

НАЦИОНАЛЬНАЯ АКАДЕМИЯ НАУК УКРАИНЫ  
НАЦИОНАЛЬНОЕ КОСМИЧЕСКОЕ АГЕНСТВО УКРАИНЫ  
Институт технической механики

**Ю. А. ПРОКОПЧУК**

**ПРИНЦИП ПРЕДЕЛЬНЫХ ОБОБЩЕНИЙ:  
методология, задачи, приложения**

Под редакцией  
д.т.н. Алпатова А. П.

Днепропетровск 2012

УДК 004:007:51-7

**Прокопчук Ю. А.**

Принцип предельных обобщений: методология, задачи, приложения. Монография. – Днепропетровск: Ин-т технической механики НАНУ и НКАУ, 2012.- 384 с.

В монографии описаны подходы к формализации системных представлений о когнитивных и метакогнитивных процессах на уровне математических моделей, компьютерных архитектур и программ. В основу предлагаемых моделей положен Принцип предельных обобщений. За счет высокого уровня обобщенности модели оказываются пригодными для описания множества схожих ситуаций, образов, а также трансляции решений в другие предметные области. Зачастую в качестве обобщенных моделей выступают параметры порядка развития сложных ситуаций и процессов. Приводятся алгоритмы построения таких моделей.

В монографии также нашли отражение теоретические и практические результаты многолетней работы автора в области создания специализированных информационных систем, ориентированных на работу с информацией и знаниями в различных предметных областях: медицине, образовании, технике, научных исследованиях.

Для специалистов и научных работников в области когнитивных технологий, интеллектуальных систем управления, а также студентов компьютерных специальностей.

Утверждено к печати Ученым советом  
Института технической механики НАН Украины и НКА Украины  
(Протокол № 6 от 19.04.2012)

Под редакцией  
д-ра технических наук, профессора А. П. Алпатова

Рецензенты:

д-р физико-математических наук, профессор Ахметшин А.М.  
д-р технических наук, профессор Пошивалов В.П.

ISBN 978-966-02-6398-7

© Прокопчук Ю. А. 2012

## ПРЕДИСЛОВИЕ

*Истина рождается как ересь, а умирает как предрассудок*  
Гегель

Как ожидается, главные события в сфере науки и технологий XXI века будут связаны с когнитивными технологиями. Когнитивные технологии – это способы и алгоритмы достижения целей, опирающиеся на данные о процессах познания, обучения, коммуникации, обработки информации человеком и животными, на представления нейронауки, на теорию самоорганизации и квантовую теорию, компьютерные информационные технологии, математическое моделирование элементов сознания, ряд других научных направлений. С помощью метакогнитивных технологий совершенствуются сами механизмы познания, мышления и управления.

Основой для развития когнитивных технологий является наличие гигантской информационно-телекоммуникационной инфраструктуры от глобального до локального уровня. Ни одна отрасль в истории человечества не знала таких темпов развития, как телекоммуникационные и компьютерные технологии. Отрасль когнитивных технологий имеет все шансы стать столь же успешной.

Когнитивные технологии позволяют на новом уровне дать ответ на те проблемы, которые поставила информационная революция – достижение нового качества управления всё более сложными процессами и создание истинно партнерских систем. Трудно переоценить вклад метакогнитивного подхода в развитие образовательных технологий.

Наиболее трудный этап развития когнитивных и особенно метакогнитивных технологий заключается в понимании и модельном воспроизведении когнитивных процессов. Одни из главных вопросов, относящихся к пониманию подобных процессов, можно сформулировать так: «Каким образом происходит природная категоризация? Как устроено феноменологическое (субъективное) пространство наблюдателя? Какие движущие силы определяют динамику процессов в этом пространстве? Как измерить размерность и сложность когнитивного пространства? Как осуществляется переход от физического мира к феноменологическому пространству? Как созревают функциональные системы когнитивно-поведенческого уровня? Как происходит целеполагание, восприятие, предсказание, предвосхищение? Как интегрируются кибернетика, синергетика и квантовая теория при описании когнитивных процессов?».

Традиционные подходы к моделированию, такие как сети формальных нейронов, экспертные системы, нечеткий и вероятностный подход, теория принятия решений и распознавания образов, теория автоматического

управления, отчасти теория самоорганизации, не смогли дать убедительных ответов на поставленные выше и другие вопросы. Так, нейронные сети в существующей интерпретации позволяют моделировать лишь «искусственный инстинкт» (хотя для многих задач этого достаточно). Значительный разрыв имеется между гуманитарными представлениями о когнитивных процессах (когнитивной психологии, философии и т.д.) и их формальными моделями. Представители точных наук не замечают или не придают значения многим определяющим когнитивным феноменам, в частности феноменам «созревания» и «дозревания» (весьма длительного) когнитивных структур, решений, образов. Подобные феномены лежат в основе творческих и научных «озарений», а также модернизации (временами радикальной) всей «картины мира» наблюдателя. Другими словами, подавляющая часть всей когнитивной (мыслительной) работы выполняется «подсознательно» и необходимо понять суть и механизм таких процессов. Нельзя забывать и о когнитивной эволюции: наряду с химическим кодом (периодической таблицей элементов) и биологическим кодом (ДНК), по всей видимости, природа изобрела и усовершенствовала «мыслительный (информационный) код», который еще предстоит открыть.

Многие сходятся во мнении, что ключевая способность человека, позволившая опередить остальные виды, заключается в умении быстро выявлять параметры порядка в разных ситуациях, следить за ними и использовать в организации когнитивно-поведенческих актов.

В настоящем исследовании предпринята попытка формализации системных представлений о когнитивных и метакогнитивных процессах на уровне математических моделей, компьютерных архитектур и программ. В основу предлагаемых моделей положен «Принцип предельных обобщений», который определяет общую методологию и парадигму исследования. Суть Принципа такова: в любой ситуации анализа и синтеза природным путем на основе самоорганизации или с помощью осмысленных (интеллектуальных) действий формируется предельно обобщенная модель ситуации, образа, явления действительности. За счет предельной обобщенности модель оказывается пригодной для описания множества других схожих ситуаций, образов, а также трансляции решений в другие предметные области. Зачастую в качестве обобщенных моделей выступают параметры порядка развития сложных ситуаций. Приводятся алгоритмы построения таких моделей.

Наряду с «Принципом предельных обобщений» важную методологическую роль играет «Принцип полимодельной дополнительности, конкурентности и отбора». Данные принципы в полной мере характеризуют естественные (природные) процессы категоризации, позволяя выявить и сохранить в памяти конкурентные наборы параметров

порядка образов, явлений и сценариев развития произвольных ситуаций действительности. В связке двух базовых принципов Принцип предельных обобщений играет роль движущей и направляющей силы самоорганизации, в то время как второй принцип создает необходимые предпосылки для возникновения самоорганизации. Второй принцип нацеливает на изучение и практическое применение всех известных математических формализмов: непрерывного, дискретного, интервального, гранулярного, фрактального, символического, нечеткого, квантового и т.д.

Представленные исследования формируют единый теоретический базис процессов извлечения и представления знаний, обеспечивают полноту методов структурирования знаний и, в конечном счете, высокую адаптивность (нежесткость) программных систем.

В книге приведены многочисленные примеры применения методов обработки и анализа данных на основе Принципа предельных обобщений при решении практических задач диагностики, прогнозирования и управления в различных предметных областях.

На протяжении многих лет автор развивал образовательные технологии на основе Принципа предельных обобщений. Результатом стали усовершенствованные программы курсов «Базы данных и знаний» «Основы проектирования компьютерных систем», «Основы проектирования интеллектуальных систем», «Экспертные системы» для бакалавров и магистров кафедры «Специализированные компьютерные системы» Украинского государственного химико-технологического университета. В последнем разделе седьмой главы приведены основные учебные темы и задания, касающиеся Принципа. В рамках лабораторных работ, курсовых и дипломных проектов создавались прототипы многих интеллектуальных и когнитивных приложений, о которых идет речь в монографии. В списке литературы имеются необходимые ссылки.

Автор выражает искреннюю благодарность своим коллегам, принимавшим участие в отдельных исследованиях, результаты которых приведены в данной книге, рецензентам, а также тем, кто оказывал дружескую поддержку автору при написании книги. Особую признательность автор выражает редактору книги профессору Алпатову А. П. за многолетнюю поддержку научных и педагогических изысканий.

## СПИСОК СОКРАЩЕНИЙ

ИИ – искусственный интеллект  
ИС – интеллектуальная система  
ППО – принцип предельных обобщений  
МПО – метод предельных обобщений  
КДС – когнитивная динамическая система  
КИС – когнитивная информационная система  
ВСС – виртуальная сплошная среда  
СР – субъективная реальность  
ПрО – предметная область  
МБкЗ – многоцелевой банк знаний  
БЗ – база знаний  
ОВЗ – область возможных значений  
ЛД – лексическое дерево  
АДД – алгоритм дифференциальной диагностики  
ИФЗ – индекс физического здоровья  
СД – семантический дифференциал  
СПУ – синдромный принцип управления  
СУ – синдромное управление  
ФС – функциональная система  
АФП – анализ формальных понятий  
ФП – формальное понятие  
ИНС – искусственная нейронная сеть  
НМС – нейроморфная сеть  
НР – нечеткий регулятор

# ГЛАВА 1.

## ФИЗИЧЕСКОЕ И ФЕНОМЕНОЛОГИЧЕСКОЕ ОПИСАНИЕ ДЕЙСТВИТЕЛЬНОСТИ

Данная глава является, по сути, введением. Ее основное назначение – показать проблематику, противоречия и сложности моделирования ситуаций действительности и когнитивных функций наблюдателя [1 – 212]. Материал главы должен послужить отправной точкой в понимании базовых принципов развиваемой когнитивной парадигмы.

### 1.1 Микроуровень и макроуровень действительности

Нечто наблюдаемое современной физикой требует измеримости. Но чтобы нечто допускало измеримость, ему следует обладать некоторым свойством инвариантности. Инвариантность предполагает возможность сохранения некоторых идеальных условий. Однако это же и означает, что требуемая стабильность должна налагаться как априорное условие возможности непосредственно теории, и вид предполагаемой нами стабильности определяет и вид воплощающей наши представления теории [126]. Таким образом, стандартная физика не способна служить средством описания некоторой автономной основы действительности. Мир стандартной физики оказывается в данном смысле человеческим построением, но в таком случае нужно активно изучать и использовать те довольно специфические способы познания, которые соответствуют миру качественного опыта [8, 23, 53, 57, 85, 105, 108, 126, 133].

На рис. 1.1 выделены два уровня действительности: *микроуровень* и *макроуровень*. Уровни разделяет условный барьер, который необходимо преодолеть.

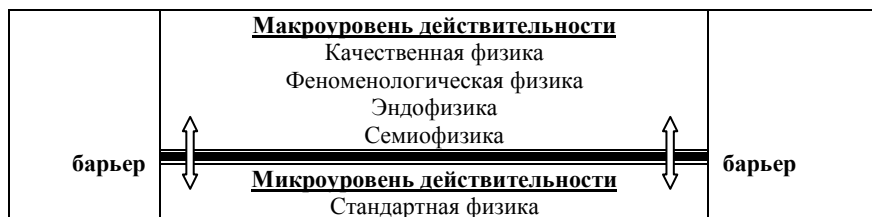


Рис.1.1 – Уровни действительности (пояснения в тексте)

Охарактеризуем кратко некоторые существующие концептуальные подходы к описанию макроуровня действительности.

В работе [8] выделяются два способа видения мира: изнутри (эндофизика) и извне (экзофизика). Эндофизика означает изучение природного мира изнутри его самого. Позиция внутреннего наблюдателя накладывает определенные ограничения на процесс наблюдения. Во-первых, нужно моделировать не только мир, но и самого наблюдателя. Во-вторых, интерфейс как промежуточная зона между внутренним наблюдателем и самим миром становится единственной реальностью, которую наблюдатель может воспринимать. Эндофизика сближается тем самым с физикой сознания и теорией мозга. То, что нам дано, что доступно нашему наблюдению – это только срез реального мира. Идея эндофизики состоит в том, что каждый наблюдатель в силу его собственных телесных и когнитивных качеств вызывает из реальности особенный ее контур и воспринимает реальность именно и только в нем.

Семиофизика [126] базируется на предположении о том, что локальное проявление качеств в качественной действительности может быть представлено как степень соответствующих интенсивностей: цвета посредством частот, горячее и холодное состояние посредством температур и т.д. Очевидно, что изменения величин интенсивности могут быть как прерывными, так и непрерывными. Ключевая идея сводится к тому, что необходимый нам формализм должен использовать как его главный примитив *качественную прерывность*, от которой следует ожидать иллюстративности в части прерывных изменений некоторым образом количественно представленных качеств.

Макроскопическая физика знает множество примеров феномена пересечения внутренних состояний системы. Они известны как критические феномены [96]. Типичными примерами служат фазовые переходы в термодинамике, где система подвергается внезапному изменению фазового состояния (например, меняясь из твердой в жидкую или из жидкой в газообразную, из магнитной в немагнитную фазу, из нормальной проводимости в сверхпроводимость и т.д.) Такие изменения происходят тогда, например, когда такой параметр, как температура пересекает точку критического значения. Можно привести примеры многих других критических феноменов: ударных волн в акустике, турбулентных перемещений в гидродинамике, деформации в теории пластичности и т.д. Все подобные феномены служат средством поддержки качественной морфологической организации феноменального мира.

Качественные структуры возникают, прежде всего, там, где некоторые мелкомодульные микроструктуры в достаточной мере равномерны для того, чтобы допустить грубозернистую морфологическую организацию, устраиваемую посредством разделений (разграничений) на макроскопическом уровне.



Одна из основных целей настоящего исследования заключается в том, чтобы разработать научную методологию, способную перебрасывать мостик от количественного представления к качественному, или между физической и феноменологической модальностями проявления действительности. Данная методология будет использована для решения задач анализа и синтеза, а также создания интеллектуальных приложений.

Объединяя вместе рациональные идеи перечисленных выше и других подходов, можно заключить, что развиваемый формализм должен опираться, в частности, на следующие положения и представления:

- выделение действительности посредством переднего и заднего планов внимания (фокусов внимания), свойство гранулированности (гранулированность позволяет эффективно проследживать низкоуровневые составляющие объектов переднего плана), свойство многокачественности;

- в фокусе внимания наблюдателя оказываются, как правило, качественные прерывности;

- существование разных уровней обобщенности любого параметра, фактора, интенсивности любого качества, а значит, и ситуации действительности в целом; существование базового («срединного») уровня обобщения [85]; отсутствие «атомов простоты» или пределов делимости материи и информации;

- каждому уровню обобщенности наилучшим образом отвечает своя группа математических методов (непрерывных, дискретных, интервальных, фрактальных, квантовых, лингвистических и т.д.);

- когнитивный процесс обладает дуализмом «локальное – нелокальное»: локальность проявляется в дискретности процесса и его составляющих, нелокальность проявляется, в частности, в неразделяемой суперпозиции смыслов, набросков, а возможно, и физических процессов (квантовые эффекты [125]);

- при восприятии действительности и синтезе управления имеет место дуализм «стабильность – нестабильность» (в частности, зоны когнитивной стабильности и нестабильности); идеи и смыслы одновременно являются и сугубо индивидуальными и коллективными, а потому - субъективными и объективными одновременно;

- рефлексия – основополагающий механизм разделения внутреннего и внешнего когнитивного пространства наблюдателя, а также основной фактор целенаправленного саморазвития;

- необходимо учитывать соотношения иерархических, кооперативных и конкурентных контуров управления в структуре адаптивной саморазвивающейся системы (целостного организма), соотношения централизации и относительной автономности в функционировании ее (его) подсистем;

- субъективная реальность (СР) – исходная форма всякого знания [57, 66, 67].

В последующих разделах главы некоторые из перечисленных исходных положений будут рассмотрены более детально.

## 1.2 Когнитивная стабильность и нестабильность

Восприятие конкретной ситуации действительности (образа, явления) может иметь неоднородную структуру. Ее можно сравнить с лоскутным одеялом, в котором каждый кусочек – зона, охваченная определенной динамической когнитивной конфигурацией [18]. Между зонами пролегают швы, в которых происходит столкновение, конфликт между различными конфигурациями. Назовем лоскуты *областями когнитивной стабильности* (области порядка), а границы стыков между ними – *областями нестабильности* или *мерцающими зонами* (рис. 1.2) (рис. взят из [18]).

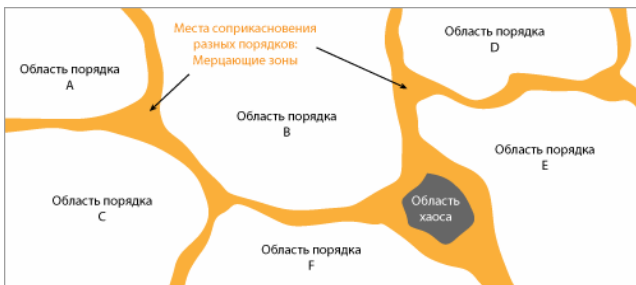


Рис. 1.2 – «Лоскутное одеяло» при восприятии реальности

Области нестабильности являются важными качественными прерывностями.

Что происходит в стыках – областях нестабильности? В них происходит конфликт и взаимодействие двух или большего числа соседних порядков. Это создает в них особую "напряженность" и сверхчувствительность. В математических сложных системах такие области представляют собой подмножества начальных, краевых или параметрических условий, минимальные вариации которых приводят к существенным изменениям в поведении системы [18, 99, 180]. Многие выдающиеся научные и культурные находки являлись там, где соприкасались различные направления науки, различные цивилизации и культуры. В рамках любого исследования имеются две области порядка: «область знания» и «область незнания». Исследователь постоянно работает на стыке этих областей.

Область хаоса на рис. 1.2 – источник инноваций и коренных перемен. Здесь встречаются разные законы, действуют смешанные языки,

рождаются новые смыслы. Отсюда актуальность методов связывания, сращивания, соединения образов, набросков решений в областях хаоса. Поэтому в поисках решений иногда целесообразно погружаться в область хаоса сознательно. В главе 7 Хаос будет представлен как первооснова «Виртуальной сплошной среды», которая в ряде аспектов может служить моделью Субъективной реальности.

Известная зрительная иллюзия, представленная на рис. 1.3, дает наглядное представление о том, что такое когнитивная стабильность и нестабильность [18, 181].



Рис. 1.3 – Бистабильный образ

Что изображено на этом рисунке?

Одним видится молодая женщина, отвернувшаяся от художника, а другим – старуха с огромным носом. После небольшой тренировки можно видеть либо женщину, либо старуху, свободно переключаясь от одного образа к другому. Однако если попытаться зафиксировать свое восприятие в какой-то промежуточной точке, начинаются трудности: наше восприятие "сопротивляется" остановке в промежуточном положении и скатывается то к образу молодой женщины, то к образу старухи. В промежуточной области между ними изображение воспринимается как лишенный смысла бесформенный набор линий и пятен. При этом сам рисунок остается неизменным, но переключение гештальта радикально изменяет смысл, который мы придаем каждой его детали.

Переключение внимания между зонами когнитивной стабильности Г. Хакен связал со сменой *параметров порядка* [180, 181].

Таким образом, в когнитивно стабильных зонах идентичность ситуации (образа, явления, понятия) оказывается хорошей, она обладает определенным и ясным смыслом и внимание легко фиксируется на ней, а в мерцающих зонах – плохой и воспринимать ситуацию как нечто цельное почти невозможно (рис. 1.4) (рис. взят из [18]).



Рис. 1.4 – Характеристика областей порядка и мерцающих зон

Общая задача, следовательно, состоит в том, чтобы на основе имеющихся данных выявить все возможные зоны когнитивной стабильности в восприятии ситуации действительности (конкретным наблюдателем или группой наблюдателей), а также выявить параметры порядка каждой зоны. Параметры порядка играют роль предельно обобщенных моделей. По мнению многих, ключевая способность наблюдателя заключается именно в умении быстро выявлять параметры порядка в разных ситуациях, следить за ними и использовать в управлении.

Метод предельных обобщений (глава 5) позволяет на основе банка тестов и эмпирических данных, описывающих прецеденты, выявить все конкурирующие наборы параметров порядка (множества формальных синдромов, каждый из которых однозначно определяет зону стабильности – исход, заключение). наброски ситуаций, образов (главы 3 и 6) характеризуют естественную смысловую многозначность восприятия действительности. Существование множества набросков разного уровня общности и, как следствие, суперпозиция смыслов позволяют предположить наличие «волновой функции» на основе орграфа набросков, которая динамически изменяется, выделяя ту или иную группу набросков в зависимости от контекста. В этом, в частности, проявляется дуализм «стабильность – нестабильность» природного восприятия любой ситуации, образа, понятия, категории. Когнитивная «волновая функция» играет важную роль в квантово-семантических эффектах.

Небольшими и экономичными операциями с параметрами порядка (детерминантами) можно достигать значительных изменений в зонах стабильности, а также «перескакивать» из одной зоны в другую (например, неблагоприятное течение процесса переводить в благоприятную фазу) [109, 110]. Благодаря этому, управление параметрами порядка может проводиться с небольшими материальными затратами, иногда скрыто и очень часто без понятного для посторонних намерения. Подобный подход

реализован в рамках Синдромного принципа управления (глава 7), что означает, по сути, «Управление параметрами порядка».

Внедрив в начале 90-х годов прошлого века информационную систему в участковой поликлинике, работавшей в рамках прогрессивной экономической модели (оплата по визитам), нам удалось в кратчайшие сроки существенно переориентировать первичную медицинскую помощь с оказания услуг в больнице на оказание услуг на дому [6]. Внедрение в середине 90-х годов в работу медико-санитарной части Никопольского завода ферросплавов автоматизированной системы контроля выполнения стандартов диспансеризации позволило в течение нескольких месяцев существенно улучшить статистику заболеваемости и повысить загрузку мощностей медсанчасти [7]. Оба эти примера свидетельствуют о способности информационных технологий в соответствующих экономических условиях экономично, быстро и радикально изменять ситуацию действительности в нужном направлении. Внедрение информационных технологий и экономическая модель выступали здесь в качестве параметров порядка [4].

Принято считать, что простых формальных правил для поиска параметров порядка, или *детерминантов* восприятия образа, ситуации действительности, не существует и только с опытом у специалиста развивается определенный навык их отыскания. Основной прикладной целью настоящего исследования как раз и является разработка таких прозрачных формальных правил (методов) отыскания детерминантов.

В роли детерминантов часто выступают качественные прерывности. Именно на них, в первую очередь, фиксируется внимание наблюдателя. Примеры прерывностей: «прогнозируемое – неожиданное» (реальное, реализовавшееся событие диаметрально противоположно прогнозируемому); «невозможное – возможное»; «незнание – знание» (такая ситуация типична для научного исследования); «фазовые переходы» («жизнь – смерть», «благоприятный режим – неблагоприятный», «устойчивый бизнес – банкротство» и т.п.).

Если разработать эффективную процедуру поиска детерминант, то вопрос о том, как их целенаправленно использовать, отпадает сам собой. Для управления развитием ситуации необходимо нейтрализовать одни детерминанты и создать другие, добиваясь, чтобы они оказывали нужное влияние на развитие ситуации. На практике детерминанты и действия с ними бывают неоднозначными, поэтому для хороших результатов и при анализе и при поиске точек воздействия необходимо опираться не на единичные детерминанты, а работать с их группами [18]. Часто несколько небольших вмешательств в разные точки системы лучше большого (по ресурсам) вмешательства в одну точку. В управлении детерминантами

(параметрами порядка) заключается суть Синдромного принципа управления сложными развивающимися системами (глава 7). В роли детерминантов выступают формальные синдромы и вероятностные закономерности (предвестники), которые определяются или с помощью метода предельных обобщений, или с помощью модели многоуровневых набросков.

### **1.3 Когнитивная самоорганизующаяся критичность**

В 1987 году был предложен механизм, согласно которому большая, находящаяся далеко от равновесия сложная система, сформированная взаимодействием множества частей, спонтанно эволюционирует к особой критической точке [193]. Поведение подобного типа было названо “самоорганизующейся критичностью” (Self-Organized Criticality, SOC). Как правило, самоорганизующуюся критичность демонстрируют масштабно-инвариантные системы со степенными функциями распределения, имеющими “длинные хвосты”. Получается, что многие нелинейные системы пребывают в устойчивой критической точке на границе хаоса и порядка. Стюарт Кауфман выдвинул гипотезу, что все живые системы находятся именно в этом режиме [200]. Данный режим гарантирует оптимальную эволюционную гибкость [165].

Развиваемый в настоящем исследовании Принцип предельных обобщений раскрывает один из ключевых механизмов индивидуального формирования критических/надкритических состояний и процессов. Критическим состояниям соответствуют предельные синдромные и вероятностные модели знаний, а также пограничные экстремальные слои набросков (модель многоуровневых набросков). Экстремальные наброски однозначно характеризуют образ и при этом имеют максимальный уровень обобщения в рамках фиксированного множества операторов агрегирования. Закритическим состояниям соответствуют многозначные (мерцающие) наброски, на основе которых формируются метафоры и ассоциативные связи. Потоки системопаттернов и системоквантов формируют «критические пути».

Критические уровни описания действительности, формируемые в рамках метода предельных обобщений, и экстремальные пограничные слои набросков можно рассматривать как границу между порядком – множеством описаний (набросков) с однозначным решением целевой задачи (например, восприятия образа, категории) и хаосом – множеством описаний (набросков), где теряется однозначность решения целевой задачи. Следовательно, если учесть открытость системы, ее диссипативность и множественность набросков с существенным разбросом по масштабу, то

предельные модели знаний и экстремальные пограничные слои набросков можно рассматривать как результат *когнитивной самоорганизующейся критичности*. Диссипация, лежащая в основе долговременной когнитивной самоорганизации, определяется, как механизмом «забывания» второстепенных деталей, так и постепенной ориентацией на работу с критическими набросками (максимально обобщенными набросками). При этом достигается минимум энергии и максимум скорости обработки данных (за счет работы с меньшим объемом данных).

Критичность моделей знаний определяется их максимальной фальсифицируемостью, а именно: если поступает какая-либо новая ситуация действительности с верифицированным заключением, которая нарушает хотя бы одну закономерность модели знаний, то рушится вся модель знаний. Может также нарушиться свойство «предельности» синдрома/закономерности при изменении онтологии (Банка тестов). В обоих случаях либо скачком осуществляется переход к другой критической модели знаний, которая не противоречит изменившейся «картине мира», либо заново запускается процесс построения всех предельных моделей знаний. Можно сказать, что предельные модели знаний и пограничный экстремальный слой набросков обладают повышенной чувствительностью к малейшему нарушению специфичности любого из синдромов модели знаний или экстремального наброска.

Состояние «на краю хаоса» (смысловой многозначности), в которое когнитивная система спонтанно переходит в процессе эволюции при решении определенного класса задач, характеризуется типичными синергетическими свойствами: согласованностью локальных процессов (например, обобщением-конкретизацией в рамках отдельных оргграфов доменов тестов) и глобальных процессов (например, применением моделей знаний и функциональных систем), интенсивным согласованным взаимодействием компонент (например, в процессе движения к предельным моделям знаний или экстремальным слоям набросков) и др. В основе механизма *когнитивной самоорганизующейся критичности* лежит естественная (забывание) и/или искусственная диссипация и потребность в минимизации когнитивной (мыслительной, вычислительной) энергии.

В начале 70-х годов прошлого века был открыт Принцип природной добиологической и биологической *гиперциклической самоорганизации* (на примере молекулярных структур) [191]. За это достижение М. Эйгену и П. Шустеру была присуждена Нобелевская премия.

Поскольку при изучении когнитивных процессов мы имеем дело с динамическими системами кооперирующихся когнитивных компонентов (структур и их физических носителей – ансамблей нейронов), наибольший интерес представляют устойчивые аттракторы внутри области

когнитивных концентраций (когнитивных многообразий). Результаты представленных исследований показывают, что на длительном интервале времени процесс формирования (предельных) когнитивных структур (аттракторов) во многих чертах напоминает гиперциклическую самоорганизацию. Это позволяет выдвигать разные концепции образования упорядоченных (предельных) когнитивных структур из неупорядоченной априорной и опытной информации на основе эволюционных алгоритмов [165]. На физическом уровне подобные концепции объясняют, как возникают устойчивые нейронные ансамбли, соответствующие когнитивным и когнитивно-поведенческим структурам (например, радикал – это нейронное воплощение синдрома, предвестника или системопаттерна, а функциональная система – это подмножество среды радикалов, т.е. нейронное воплощение модели знаний). Когнитивные, поведенческие и когнитивно-поведенческие гиперциклы удовлетворяют всем критериям интеграции информации [191]:

1. Селективная устойчивость каждого компонента (предельной) когнитивной структуры (например, синдрома, предвестника, системопаттерна, наброска, радикала) из-за успешной конкуренции с ошибочными копиями.

2. Кооперативное поведение компонентов, объединенных в новую функциональную единицу на всех уровнях организации (синдром, предвестник, набросок, системопаттерн, орграф набросков, модель знаний, функциональную систему, категорию и т.д.).

3. Успешная конкуренция этой функциональной единицы с другими, менее эффективными единицами.

Критериями отбора при одинаковой функциональности служат или больший уровень обобщенности, или меньший компонентный состав (большая редуцированность). Каким образом происходит генерирование новой информации в системе? Путем случайного выбора, например, протосиндром или протозакономерность возникают как случайная совокупность значений тестов разного уровня общности, которая обеспечивает достижение цели (решение  $Z$ -задачи, где  $Z$  – список заключений). В дальнейшем протосиндром подвергается мутации и селективному отбору. В результате этого формируются синдромы (максимальная редукция) и далее – предельные синдромы (максимальное обобщение, т.е. оптимальный продукт эволюции). Последние выдерживают конкуренцию с любыми мутантами, следовательно, на их основе формируются модели знаний и далее функциональные системы когнитивного и когнитивно-поведенческого уровней. Подобным образом формируются орграфы набросков, орграфы доменов тестов, модели знаний и функциональные системы. В этом состоит суть *когнитивного гиперцикла*



– принципа когнитивной (нейронной) самоорганизации, обуславливающего интеграцию и согласованную эволюцию системы функционально связанных когнитивных единиц и, соответственно, нейронных ансамблей.

Когнитивный гиперцикл – это средство объединения когнитивных единиц с ограниченной информацией (доменов, синдромов, предвестников, набросков, системопаттернов) в новую устойчивую систему, способную к согласованной эволюции (орграф доменов, орграф набросков, модель знаний, категорию, концепт, радикал, среду радикалов, функциональную систему, феноменологическое пространство). При этом когнитивный гиперцикл гарантирует выполнение трех необходимых условий гиперциклической самоорганизации [191]:

- поддерживает конкуренцию в мутантных распределениях всех когнитивных единиц;

- допускает существование нескольких (конкурирующих) единиц и их мутантных распределений;

- объединяет эти единицы в систему, способную к согласованной эволюции, где преимущества одного элемента могут использоваться всеми членами системы, причем система как целое продолжает интенсивно конкурировать с любой системой иного состава.

Для когнитивного гиперцикла недостаточно лишь свойства воспроизведения активности отдельных единиц. Важно системное свойство гиперциклической организации, т.е. воспроизведение согласованной активности всей когнитивной или когнитивно-поведенческой системы. Это возможно путем отбора функционально кооперирующихся партнеров через функциональные связи, которые дают либо взаимное усиление воспроизведения, либо структурную стабилизацию.

Таким образом, когнитивные, поведенческие и когнитивно-поведенческие гиперциклы – это новые классы нелинейных сетей системопаттернов (реакций, автоматизмов и т.д.), обладающих уникальными свойствами и поддающихся унифицированному математическому описанию. Зародившись, когнитивные гиперциклы эволюционируют в сторону усложнения, в частности, предельного обобщения.

Гиперциклическая самоорганизация Эйгена – Шустера – это скорее «выбор», а не «отбор», который подразумевает оптимизацию (см. предисловие редакторов перевода к [191]). «Выбор» может привести к разным вариантам развития. Когнитивный гиперцикл в ряде случаев приводит к предельным структурам (предельные синдромы и вероятностные закономерности, экстремальные слои набросков и т.д.), т.е. именно к отбору, следовательно, в этих случаях он обуславливает также самоорганизующуюся критичность.

Когнитивная гиперциклическая самоорганизация и когнитивная самоорганизующаяся критичность лежат в основе многих процессов самоорганизации в рамках феноменологического пространства наблюдателя. Они заложены и в идее построения многоуровневых набросков ситуаций, образов, моделей. Конечной целью построения орграфа набросков является формирование экстремального пограничного слоя набросков (слоя максимального уровня обобщения – абстрагирования), достаточного для решения целевых задач анализа и синтеза. В силу множественности набросков «волновая функция» когнитивной структуры никогда точно не соответствует критической точке (точке равновесия), под действием внешних и внутренних случайных факторов она претерпевает постоянные флуктуации в ее окрестностях. В этом также проявляется дуализм «стабильность – нестабильность». Эволюционная необходимость выделения экстремальных пограничных слоев набросков и других предельных когнитивных структур определяется потребностью минимизации всех видов ресурсов: памяти, энергии, времени и т.д. Чем более экономно расходуются ресурсы, тем более сложные когнитивные задачи можно решать.

Таким образом, *самоорганизующаяся критичность* – это способность системы эволюционировать в направлении достижения критичности и поддерживать себя в этом состоянии. В соответствии с этой способностью, сложные системы, включая когнитивные, сами собой переходят в критические состояния в процессе самоорганизации. То есть, критические состояния для когнитивных систем – это не исключительные моменты, а естественное направление их эволюции и внутренней динамики [193]. Когнитивная гиперциклическая самоорганизация объясняет схему интеграции и формирования функциональных систем разных уровней организации.

Важнейшей задачей интеллектуальных технологий является поиск специфических признаков, которые бы позволяли заранее узнавать, что сложная система приближается к критическому переходу. Нет необходимости обосновывать чрезвычайную актуальность данной задачи, в частности для медицины, а также предупреждения природных и техногенных катастроф. Метод предельных обобщений позволяет найти предвестники неблагоприятного развития ситуации. При приближении к критическому переходу предвестников может становиться все больше, а формальных синдромов текущей благоприятной фазы все меньше. В системе далекой от критического перехода, в силу того, что гомеостатические силы велики, предвестники или отсутствуют или их крайне мало. Задачей синдромного принципа управления является

нейтрализация неблагоприятных предвестников и увеличение числа благоприятных параметров порядка (формальных синдромов).

#### 1.4 Смысловая реальность. Предельные смыслы

В первом приближении *смысл* можно определить как отношение между субъектом и объектом или явлением действительности, которое определяется местом объекта (явления) в жизни субъекта, выделяет этот объект (явление) в образе мира и воплощается в личностных структурах, регулирующих поведение субъекта по отношению к данному объекту (явлению) [48, 89].

Изменив смысловую систему, идею, под действием которой находится наблюдатель или коллектив, можно радикально изменить его поведение и решения в одних и тех же информационных и материальных обстоятельствах. Это общий принцип когнитивного управления [18, 89].

Смысл является особой философско-психологической реальностью, несводимой к другим явлениям. Понятие смысла соотносимо и с объективной, и с субъективной, и с интерсубъективной (групповой, коммуникативной) реальностью. Смысловая реальность проявляется в многообразии своих форм и образует единство, по меньшей мере, трех сторон или аспектов ее рассмотрения [17, 58, 89]: онтологического, феноменологического и деятельностного. В настоящем исследовании онтологическая сторона смысловой реальности представлена, в частности, орграфами доменов тестов, орграфами набросков образов (модели возрастающей концентрации смысла в единице информации), предельными моделями знаний (предельная концентрация смысла на множестве прецедентов и Банка тестов), моделями категорий и концептов. Феноменологическая сторона представлена, в частности: моделями созревания протоструктур, структур и субструктур; моделями «созревания» решений и управлений; моделью когнитивной самоорганизующейся критичности; концепцией Виртуальной сплошной среды; моделью рефлексии (Баланса энтропий). Деятельностная или поведенческая сторона представлена, в частности, потоками системопаттернов и автоматизмов, системоквантами, полем импульсов.

Настоящее исследование содержит описание различных видов смысловых структур и их взаимосвязей. Так моделью минимальных смысловых структур являются орграфы набросков, орграфы доменов элементарных тестов, системопаттерны. Важнейшими моделями структур смысловой реальности являются *субструктуры высоких порядков* (субструктуры субструктур и т.д.), которые возникают под действием

высокоэнергетических импульсов. В качестве модели субъективной реальности выступает *Виртуальная сплошная среда* (глава 7).

Некоторые субструктуры представляют собой *фокусы взаимодействия* или зоны конвергенции различных видов информации [66]. На модельном уровне в качестве фокуса взаимодействия выступает радикал, а в более общем случае – системопаттерн. Гипотетическая схема биологического фокуса взаимодействия состоит из групп нервных клеток, которые соединены жесткими связями, основанными на структурных изменениях в синапсах. Благодаря такой структуре фокус объединяет и синтезирует информацию, циркулирующую в различных нейронных сетях, на основе чего достигается решение. Все сказанное дает возможность объединить данные, полученные при изучении механизмов восприятия и мышления, единым *принципом информационного синтеза* как мозговой основы возникновения нового качества в виде субъективных переживаний [67].

Функциональное значение психических переживаний состоит в том, что, представляя собой результат синтеза информации, психические феномены содержат интегрированную оценку ситуации, способствуя тем самым выбору поведенческого ответа (пример – радикал). Элементы обобщения содержатся уже в самых простых психических феноменах, таких как ощущение (пример – орграф доменов теста). Возникающие на базе синтеза информации психические явления обладают новым качеством, по сравнению с процессами физиологического уровня (пример – модели знаний, функциональные системы). Это новое качество определяет и другую логику развития, последовательности явлений, которые подчиняются иным закономерностям высшего порядка [67].

Некоторые смысловые субструктуры возникают в результате специфических процессов, например, таких как *когнитивные сингулярности*. Одна из сингулярностей представляет собой «*информационный взрыв*» – локальное (в пределах некоторого многообразия) производство гигантского объема информации, осуществляемое автоматизмами среды. В результате «информационного взрыва» образуются, например, орграфы набросков образов. Противоположным типом сингулярности является *информационный коллапс*. Примером коллапса является предельное сжатие (концентрация) смысла. Как правило, информационный коллапс следует за информационным взрывом. В результате рождаются предельные структуры, например модели знаний и экстремальные пограничные слои набросков. По сути, метод предельных обобщений можно представить как цепочку когнитивных сингулярностей: «информационный взрыв → информационный коллапс». Феноменологическая модель процесса

принятия решений также представляет собой схему «расширение → локализация» (глава 7).

К когнитивным сингулярностям можно отнести *процессы разрушения структур и субструктур*, а также разного масштаба *когнитивные катастрофы*. При появлении фальсифицирующего прецедента или нового домена какого либо теста может разрушиться модель знаний. Когнитивные катастрофы могут приводить к разрывам Виртуальной сплошной среды.

Говорить о смысловой реальности следует не только как о полиформной реальности, но и как о реальности многомерной и многоуровневой, на которую можно смотреть под разными углами зрения и с разной степенью обобщенности. При рассмотрении ее в разных аспектах получаются разные, не совпадающие между собой картины, совместить которые в непротиворечивое целостное представление позволяют модели многоуровневых набросков образов, ситуаций и модели многоуровневой морфологии феноменов действительности (главы 6, 7).

Различные стороны смысловой реальности поддаются изучению неспецифическими методами: модельными, экспериментальными, проективными, психометрическими, психосемантическими, качественно-феноменологическими. Одна из основных задач психосемантики – построение так называемого *семантического пространства*, то есть нахождение системы тех латентных факторов, в рамках которых наблюдатель «работает», так или иначе оценивая какие-либо объекты [127]. Наблюдатель, как правило, не дает себе отчета в существовании этих факторов. Семантическое пространство (с системой координат в виде семантических шкал) в социологии является *исследовательской моделью структуры индивидуального сознания*, на основе которой происходит восприятие наблюдателем объектов, их классификация, сравнение. Метод семантического дифференциала [127] направлен не только на поиск семантического пространства и анализ лежащих в его основе факторов, но и на изучение взаимного расположения объектов (концептов) в этом пространстве, то есть различий в восприятии этих объектов.

Принцип предельных обобщений позволяет существенно развить представления о семантическом пространстве. Так, в качестве системы координат семантического (феноменологического) пространства предлагается Банк тестов, содержащий орграфы доменов тестов. Другими словами, каждая семантическая шкала заменяется соответствующим орграфом доменов теста. Комбинация метода предельных обобщений и метода семантического дифференциала позволяет для каждого концепта найти все предельные синдромы и вероятностные закономерности (как индивидуальные, так и коллективные).

Работы по семантике сознания [127] обнаружили феномены смысловой (семантической) разорванности сознания. В общем случае можно говорить о дисгармоничности, конфликтности и конкурентности смысловой сферы.

### 1.5 Тесты, системопаттерны, системокванты

При изучении феноменологического пространства наблюдателя базовым предположением является предположение о существовании единого *универсального принципа структурирования информации* (этот принцип закладывался в “устройство” мышления с эволюционным возникновением самого мышления). Именно благодаря этому общему принципу происходит сопряжение (непосредственное или опосредованное) любой мыслительной структуры с любой другой, и любого уровня информационной масштабности [134 – 161].

Следствием универсального принципа структурирования информации являются модели двух взаимосвязанных базовых информационных объектов: *орграфа набросков* и *динамического системопаттерна* (или сокращенно – *системопаттерна*). Разновидностью орграфа набросков является орграф доменов элементарного теста (домен – множество значений теста со связями наследования). Формальная модель системопаттерна представима в виде:

$$f/\mu: k^1 \rightarrow k^2, \quad (1.1)$$

где  $f/\mu$  – некоторое отображение (морфизм), в котором  $\mu$  – механизм реализации отображения;  $k^1$  – исходные данные или состояние;  $k^2$  – требуемые данные или состояние (цели и критерии). Механизм  $\mu$  отвечает за системность, так как он привлекает все виды ресурсов для реализации отображения, включая поиск необходимых исходных данных (при необходимости). В дальнейшем нотация (1.1) будет конкретизирована.

С помощью системопаттернов мы будем описывать любые преобразования, движения информации, материи или энергии. Другими словами, любые реальные или предполагаемые изменения действительности, включая субъективные переживания (феноменологический уровень). С помощью системопаттернов можно описывать не только конечные процессы, но и бесконечные (с точки зрения жизненного цикла системы), например, асимптотические или триггерные.

Разновидностями системопаттернов являются формальные синдромы, вероятностные закономерности (предвестники, ассоциативные связи), реакции, импульсы, автоматизмы. Автоматизмы выполняются всегда, если они активированы (наподобие рефлексов). Выполнение остальных

системопаттернов зависит от наличия энергии (ресурсов) даже в случае их активации.

Важно отметить, что системопаттерны различного уровня организации имеют изоморфную структуру. Данный факт нацеливает на поиск единообразного пути процедурной реализации системопаттернов. Один из таких путей предлагается в главе 7 при рассмотрении вопросов, связанных с созданием Многоцелевых банков знаний и Банков математических моделей.

Взаимосвязь между оргграфом набросков и системопаттернами осуществляется, в частности, следующим образом. В оргграфе доменов теста связь между доменами представляется в виде:  $(T \rightarrow T')$ , где  $T, T'$  – домены. В общем случае связь между набросками представляется в виде:  $(P \rightarrow P')$ , где  $P, P'$  – наброски. Очевидно, любое подобное преобразование является простейшим динамическим паттерном и даже более того – является автоматизмом вычислительной среды. С другой стороны, данные (или состояния)  $k^1, k^2$  описываются с помощью элементарных тестов и набросков образов.

В 1979 году Судаковым К.В. была сформулирована гипотеза о системном квантовании поведения и жизнедеятельности [176]. Под "системоквантами" жизнедеятельности понимаются дискретные системные процессы от формирования любой потребности до ее удовлетворения. Системокванты являются своеобразными операторами динамической деятельности различных функциональных систем организма, постулированных П.К. Анохиным [11]. Деятельность системоквантов различного уровня организации направлена, прежде всего, на поддержание устойчивости различных показателей жизнедеятельности [177]. В данной работе под системоквантом понимается совокупность одновременно работающих системопаттернов (гранул действий), в частности радикалов.

Рассмотрим особенности представления операторов эволюции динамических систем с помощью системопаттернов.

Понятие *динамическая система* (ДС) включает в себя два основных компонента [99]: фазовое пространство  $\mathbf{P}$  (метрическое пространство или многообразие) и однопараметрическую непрерывную или дискретную группу (полугруппу)  $\varphi^t(\mathbf{x})$  его преобразований, т.е. для любого вектора (точки)  $\mathbf{x} \in \mathbf{P}$  и  $t \geq 0$  существует единственный вектор (точка)  $\varphi^t(\mathbf{x}) \in \mathbf{P}$ . Параметр группы  $t$  – это время. Отображение  $\varphi$  называется также *оператором эволюции* ДС. Эквивалентное обозначение –  $\varphi(\mathbf{x}, t)$ .

Поскольку в рамках развиваемого формализма вектор  $\mathbf{x}$  является вектором элементарных тестов, задаваемых оргграфами доменов, то каждая точка в исходном фазовом пространстве  $\mathbf{P}$  порождает облако точек в фазовых пространствах более высоких уровней общности. Вместо  $\mathbf{x}$

следует писать  $\{x/X\}$ , где  $X$  – домен. Таким образом, любой оператор эволюции  $\varphi$ , заданный в исходном фазовом пространстве  $\mathbf{P}$ , порождает множество системопаттернов вида:

$$f/\mu: \mathbf{x}, t \rightarrow \varphi(\mathbf{x}, t), \quad (1.2)$$

$$\{g/\mu: \{x/X_0\} \uparrow \{x/X\}, t/\Lambda_0 \uparrow t/\Lambda \rightarrow \varphi(\{x/X_0\}, t/\Lambda_0) \uparrow \{x/X\}'\}, \quad (1.3)$$

$$\{h/\mu: \{x/X\} \downarrow \{x/X_0\}, t/\Lambda \downarrow t/\Lambda_0 \rightarrow \varphi(\{x/X_0\}, t/\Lambda_0) \uparrow \{x/X\}'\}, \quad (1.4)$$

где  $\mathbf{x} = \{x/X_0\}$ ,  $t = t/\Lambda_0$  ( $X_0, \Lambda_0$  – базовые домены);  $\mu$  – механизмы реализации системопаттернов; операция ‘ $\uparrow$ ’ означает обобщение; операция ‘ $\downarrow$ ’ означает нисходящую детерминацию. Каждый системопаттерн действует в собственном фазовом пространстве  $\langle \{x/X\}, t/\Lambda \rangle$ , которые отличаются друг от друга уровнем обобщения. Системопаттерн (1.2) можно отнести к физическому уровню описания действительности, а системопаттерны (1.3) – (1.4) к феноменологическому уровню.

Иной способ построения системопаттернов динамических систем предоставляет символическая динамика [122]. При данном подходе фазовое пространство динамической системы разбивается на области таким образом, что любая ее траектория однозначно определяется последовательностью прохождения этих областей (символической последовательностью). По заданному разбиению фазового пространства строится ориентированный граф, называемый *символическим образом*. Конечные пути на графе и будут определять системопаттерны.

## 1.6 Принципы когнитивной парадигмы

Резюмируя материал главы, сформулируем базовые принципы развиваемой когнитивной парадигмы, которая лежит в основе формализма и технологий проектирования интеллектуальных приложений. Таких принципов всего два [140, 146]:

- *Принцип предельных обобщений;*
- *Принцип полимодельной дополнительной, конкурентности и отбора.*

Суть Принципа предельных обобщений (ППО) можно выразить следующим образом: среди всех допустимых моделей (решений) следует выбирать модели (решения), которые обладают максимальной общностью. Принцип предельных обобщений основан на стремлении когнитивной системы к *глобальному энергетическому минимуму и максимуму скорости обработки* (оперативной) информации: если система оперирует меньшим адекватным объемом информации, ей требуется для этого меньше энергии и она может это делать с большей скоростью.



Обсуждаемые теоретические представления созвучны с развиваемыми нейрофизиологами положениями об иерархическом принципе интеграции скоростных, медленных и сверхмедленных информационно-управляющих систем головного мозга, формируемых для обеспечения психических состояний и познавательной деятельности человека, включая вербальную ассоциативно-мыслительную деятельность [70].

Связка двух принципов позволяет на уровне моделей выразить сложность субъективной реальности через непрерывное взаимодействие относительно простых структур с использованием автоматизмов среды. Основная гипотеза состоит в том, что связка указанных принципов олицетворяет «встроенную» *оптимальность мышления*.

Ведущая роль Принципа предельных обобщений подчеркивается тем, что он вынесен в заглавие книги. Во втором принципе проявляется феномен полимодельности познания. Несмотря на важность каждого принципа в отдельности, главный смысл развиваемой когнитивной парадигмы заключается в их глубинной взаимосвязи. При совместном применении обоих принципов происходит генерирование большого числа разноуровневых описаний, набросков и, соответственно, переходных слоев (заполняется пространство смыслов). В переходных слоях происходит перестройка картины мира с одного масштаба на другой. Здесь встречаются разные законы, действуют смешанные языки, рождаются новые смыслы. Отсюда актуальность методов связывания, сращивания, соединения образов, набросков решений в переходных слоях. Наброски решений, полученные на одних уровнях общности, передаются на нижележащие уровни для детализации, конкретизации. Взаимодействия реализуются через системопаттерны.

Следствием базовых принципов являются:

- *Принцип подобия* (начиная с определенного уровня общности, разные по своей природе системы и процессы описываются схожими закономерностями);

- *Принцип инвариантности* (находятся модели знаний, наброски образов, решения, удовлетворяющие максимальному числу ситуаций действительности);

- *Принцип фальсифицируемости* (выбирать нужно ту гипотезу, которая раньше других опровергалась бы новыми данными, полученными в результате наблюдений или эксперимента, если была бы ложной);

- *Принцип «минимального обеспечения» целостной функции* (теория системогенеза);

- *Принцип минимальной длины описания* (среди всех моделей следует выбрать ту, которая позволяет описать данные наиболее коротко с учетом длины описания самой модели);

- *Принцип редукции* в задачах обнаружения закономерностей;
- *Принципы квантовой семантики* (дополнительности, нелокальности, суперпозиции смыслов);
- *Принцип когнитивного резонанса* (модельно-резонансный подход);
- *Принцип когнитивной самоорганизующейся критичности* (образование за счет диссипации и автоматизмов среды критичных категориальных структур);
- *Принцип соответствия* (язык описания сложной системы должен соответствовать характеру располагаемой о ней информации);
- *Принцип грануляции Л. Заде* (для эффективной работы с неточной информацией следует выбрать наибольший уровень грануляции, соответствующий допустимому уровню неточности);
- *Принцип Баланса энтропий* (модельно-резонансной рефлексии).

Основные идеи принципа минимальной длины описания (МДО) были сформулированы в 1960-х годах. Однако широкое распространение при решении задач автоматического анализа данных он получил только в девяностых годах [131]. В различных областях информатики этот принцип неоднократно переизобретался в более частных и конкретных формах, однако с сохранением общей идеи. Обширная литература, посвященная использованию принципа МДО, показывает его явные преимущества в получении более эффективных решений по сравнению с другими методами. При этом принцип МДО не отвергает существующие методы, а позволяет уточнить использующиеся в них критерии качества модели, что приводит к улучшению этих методов. В то же время далеко не всегда является очевидным, как именно следует применять этот принцип.

В основе теории редукции лежит давно известная зависимость расположения точек в пространстве от координат этого пространства. Такая зависимость указывает на то, что любые закономерности вида  $Y = F(X)$  существуют лишь в пространстве с определенными координатами. Принцип, основанный на предварительном формировании пространства, назван *принципом редукции* (упрощения), а соответствующая данному принципу теория – теорией редукции [36, 37]. В основе принципа редукции лежит процедура отбора таких свойств объектов, в пространстве которых достижимо простое решение задачи.

Принцип предельных обобщений, как и принцип редукции, отбрасывает несущественную информацию и выделяет только то, что способствует объединению объектов в образы или вносит различия между ними. Однако в отличие от принципа редукции, а также методов синтеза сложнейших разделяющих поверхностей в исходном пространстве огромной размерности, ППО не стремится построить одну единственную (статическую) разделяющую поверхность какой бы то ни было

размерности, а находит все наборы параметров порядка (смысловые инварианты), приводящие к требуемому результату. При таком подходе истинность всей модели (знаний) формируется эволюционно путем когнитивной самоорганизации. Процессы самоорганизации моделей в рамках ППО определяются не только появлением новых эмпирических данных, но также относительно независимым развитием банка тестов, позволяющего описать переход от физической к качественной действительности.

Смысловая нагрузка понятия резонансного взаимодействия множества моделей заключается в том, что в отличие от простой суммы (например, простого голосования решающих правил [134]), они (по аналогии с физическим смыслом резонанса волновых паттернов) коммуникативно взаимодействуют и качественно дополняют друг друга, образуя некую "открытую интерконцептуальную" систему и репрезентируя таким образом многомерное видение описываемого объекта [55].

Концептуальный гештальт "полимодельного резонанса", используемый в качестве логико-методологической основы взаимодействия различных "точек зрения", позволил сформировать новый подход, нацеленный не на отбрасывание тех или иных моделей (непрерывных, дискретных, интервальных, нечетких, лингвистических, фрактальных, грануляционных, агентных, нейросетевых и т.д.), а на превращение их в рациональные конструкты посредством выяснения смыслообразующих интервалов и границ применимости в зависимости от природы используемых познавательных средств и процедур. В такой идейный контекст вписывается представление о конкурирующих концепциях как своеобразных "предельных структурах" многомерного объекта исследования, отражающих различные его стороны, которые проявляются в разных условиях (в этом плане имеется сходство с химической теорией резонанса и ее набором так называемых "предельных структур" [55]).

С практической точки зрения, предлагаемый подход позволяет представить произвольный феномен действительности (процесс, образ, объект, ситуацию) на заданном уровне общности как результат суперпозиции множества соответствующих этому феномену частных несамодостаточных предельных структур (например, набросков в орграфе набросков). Причем на каждом уровне общности имеется свой набор предельных структур.

Таким образом, представленные в монографии исследования формируют единый теоретический базис процессов извлечения и представления знаний, обеспечивают полноту методов структурирования знаний и, в конечном счете, высокую адаптивность программных систем.

## ГЛАВА 2. ЭЛЕМЕНТАРНЫЕ ТЕСТЫ. ОРГРАФЫ ДОМЕНОВ

### 2.1 Базовые определения

Пусть  $\{\tau\}$  – множество элементарных тестов, с помощью которых описываются все факторы, обстоятельства и явления, имеющие отношение к изучаемой ситуации действительности. Элементарность теста означает, что его результат может быть представлен в виде: «тест = значение» или в эквивалентном виде «тест? значение». Мощность множества всех элементарных тестов совпадает с мощностью всех односложных вопросов, так как любой такой вопрос представим в виде «вопрос? ответ».

Конкретный результат теста  $\tau$  будем обозначать через  $\underline{\tau}$ . Результаты тестов могут выбираться (формироваться) из разных доменов (множеств значений). Для фиксации того, что в качестве множества результатов теста  $\tau$  используется домен  $T$ , будем использовать нотацию:  $\tau/T$ . Примеры нотаций:  $a/A$ ,  $\{b/B\}$  – множество тестов,  $\{c\}/C$  – множество однотипных тестов,  $\{a/A\}$  – множество значений группы тестов. Примеры:

$\{b/B\} = \{\text{ФИО}/C; \text{Пол}/C; \text{Возраст}/N; \text{АДс}/N; \text{АДд}/N; \text{Жалобы}/C\}$ ,

где  $C$  – текст;  $N$  – число; АДс и АДд – АД систолическое и диастолическое;

$\{b/B\} = \{\text{ФИО}/C? \text{Петров И.И.}; \text{Пол}/C? M; \text{Возраст}/N? 34; \text{АДс}/N? 145; \text{АДд}/N? 87; \text{Жалобы}? \text{На головную боль, на боли в суставах}\}$ .

Используя разные домены, можно управлять общностью (масштабом) результата одного и того же теста. Во многих случаях между доменами разного уровня общности могут быть заданы однозначные правила пересчета значений. Правила пересчета значений теста из одного домена в другой задают *ориентированный граф доменов*  $G(\tau) \equiv \{T \rightarrow T'\}_\tau$  а  $N_\tau$  – общее количество доменов в орграфе  $G(\tau)$ . Домен слева от стрелки в связке  $(T \rightarrow T')$  будем называть *доменом-предком (родителем)*, а справа от стрелки *доменом-потомком*. Соответственно любую связку  $(T \rightarrow T')$  можно именовать «предок – потомок». Совокупная смысловая область элементов домена–потомка полностью совпадает с совокупной смысловой областью элементов домена–предка.

*Примечание:* домен – это не просто множество элементов, а множество элементов с обобщающими связями.

Правила  $\{T \rightarrow T'\}_\tau$  задают, по сути, *эмпирическую аксиоматику*, а также *систему интерпретаций* в рамках конкретной задачи (класса задач) предметной области. Действительно, значение теста «Возраст/ $B1 = 15$ » не содержит интерпретацию, а значение теста «Возраст/ $B3 = \text{юный}$ » уже

содержит. Система интерпретаций лежит в основе «понимания» – основного феномена, отличающего мышление человека от операций компьютера.

Поскольку у любого теста может быть несколько орграфов доменов, то в некоторых случаях вместо нотации  $\tau/T$  можно применять расширенную нотацию  $\tau/G:T$ .

Без ограничений общности положим, что домены орграфа  $G(\tau)$  состоят из альтернативных элементов (точечных или атомарных элементов) и атомарный элемент любого домена однозначным образом преобразуется в атомарные элементы доменов-потомков (если они существуют). Будем также считать, что каждый домен по умолчанию может содержать *неопределенный элемент* (мерцающее значение теста), который при преобразовании также переходит в неопределенный элемент домена-потомка. Если не оговорено другое, то неопределенный элемент в любом домене будем обозначать  $u$ , следовательно, запись  $\tau/T?u$  обозначает неопределенное (мерцающее) значение теста. Неопределенные элементы играют важную роль при построении набросков ситуаций, образов, текстов.

Для любого атомарного элемента  $a$  домена  $A$  автоматически определен элемент (и терм) «не  $a$ »  $\equiv \neg a = A \setminus a$ . То же справедливо и для любого подмножества  $A' \subset A$ : автоматически определены элементы – термы «не  $A'$ »  $\equiv \neg A' = A \setminus A'$ . Таким образом, операция отрицания действует в рамках конкретного домена.

Примем, что для любой связки  $(T \rightarrow T')$  выполняется соотношение:  $|T| > |T'|$ , где  $|*|$  – мощность множества.

Орграф имеет одну *базовую вершину* – базовый домен со значениями (элементами) максимально высокого уровня точности. В базовую вершину не входит ни одна дуга орграфа. Любой домен орграфа проецируется на весь базовый домен. Орграф доменов не имеет циклов. Если имеются разные пути перехода к какой-либо вершине, то эти пути должны приводить к одному и тому же результату (не должно быть конфликтов). *Терминальные вершины* или вершины, из которых не выходит ни одна дуга, задают домены максимальных уровней общности. Орграф не может содержать двух одинаковых вершин, т.е. таких вершин, которые выдают одинаковые значения теста при одном и том же базовом значении.

Базовый домен может быть объединением конечного числа непрерывных интервалов, но любой домен-потомок всегда является дискретным. Преобразование от домена-предка к домену-потомку назовем *восходящим* преобразованием (преобразованием от частного к общему, т.е. обобщением), возможно и обратное – нисходящее преобразование (от общего к частному), которое олицетворяет *нисходящую детерминацию*. Для реализации одного из вариантов нисходящей детерминации любой

элемент дискретного домена (не базового) может быть снабжен одной или несколькими функциями принадлежности, а именно:

$$T = \{a \{\mu_a\}; b \{\mu_b\}; c \{\mu_c\}; \dots\},$$

где  $\mu$  – функции принадлежности (влияют только на нисходящую детерминацию). Функции принадлежности одного элемента отличаются друг от друга, в частности, привязкой к доменам предшественникам (вплоть до базового домена). В общем случае, нисходящая детерминация на множестве тестов описывается взаимосогласованными (калибровочными) колебательными процессами типа «маятника». Такие процессы базируются на законах гомеостатики и предполагают наличие модели системы (базы знаний). Более детально этот вопрос будет рассмотрен в разделе 2.4.

Необходимо стремиться к тому, чтобы в любом орграфе  $G(\tau)$  все терминальные вершины являлись дуадами или бинарными оппозициями. Примеры: {Истина; Ложь}, {Норма: Отклонение}, {Эффективно; Неэффективно}, {Жалобы есть; Жалоб нет}, {Благоприятный; Неблагоприятный} и т.д. Очевидно, орграф может состоять из одной базовой вершины, например: Пол/ {М; Ж}. Орграфы всех тестов, имеющих дихотомические значения, состоят из одной вершины.

На алгоритм преобразования (обобщения) значений доменов могут влиять значения других тестов. Пусть  $(T \rightarrow T) \in G(\tau)$ , а внешние тесты  $\{a/A\}$  влияют на алгоритм пересчета значений из домена  $T$  в домен  $T$ . Для фиксации данного факта будем использовать нотацию:  $T \{a/A\} \rightarrow T$ . Если результат какого либо теста из  $\{a/A\}$  неизвестен, то будем считать, что и преобразование  $T \{a/A\} \rightarrow T$  не определено, следовательно, не определен весь подграф с базовой вершиной  $T$ . Таким образом, орграф  $G(\tau | \{a/A\})$ , где  $\{a/A\}$  – множество всех используемых внешних тестов, имеет переменную структуру.

Правила вычислений элементов некоторых доменов на основе значений других тестов, в общем виде, имеют вид:  $f/\mu \underline{d}/T, \{a/A\} \rightarrow \underline{d}/T$ , где  $f/\mu$  – некоторая однозначная функция ( $\mu$  – механизм реализации функции). Данные правила будем относить к *онтологическим соглашениям*, так как сами орграфы принадлежат онтологии (задачи, предметной области). Примеры таких соглашений будут приведены ниже. Наиболее общие правила пересчета значений одних тестов в значения других тестов содержатся в базе знаний.

Примеры доменов для теста  $\tau =$  ‘Возраст’:

$V1 = [0; 100]$ ;

$V2 = \{\text{юный; молодой; средних лет; пожилой; старческий}\}$ ;

$V3 = \{\text{молодой; средних лет; пожилой}\}$ .

Примеры результатов тестов:

Возраст/  $V2 = \langle \text{пожилой} \rangle$ ; Возраст/  $V1? 87$ ; Возраст/  $V3 = \langle \text{пожилой} \rangle$ .

Отметим, что (Возраст/  $V2? \text{пожилой}$ )  $\neq$  (Возраст/  $V3? \text{пожилой}$ ). Зададим правила пересчета между доменами  $V1$ ,  $V2$  и  $V3$  следующим образом:

$V1.[0; 14] \rightarrow V2.\{\text{юный}\}$ ;  $V1.[15; 33] \rightarrow V2.\{\text{молодой}\}$ ;

$V1.[34; 55] \rightarrow V2.\{\text{средних лет}\}$ ;  $V1.[56; 70] \rightarrow V2.\{\text{пожилой}\}$ ;

$V1.[71; 100] \rightarrow V2.\{\text{старческий}\}$ ;  $V2.\{\text{юный; молодой}\} \rightarrow V3.\{\text{молодой}\}$ ;

$V2.\{\text{средних лет}\} \rightarrow V3.\{\text{средних лет}\}$ ;

$V2.\{\text{пожилой; старческий}\} \rightarrow V3.\{\text{пожилой}\}$ .

Правила пересчета  $\{V1 \rightarrow V2; V2 \rightarrow V3\}$  задают оргграф доменов для теста 'Возраст'. Обозначим оргграф через  $G_1(\text{Возраст})$ .

Если задан оргграф доменов теста, то появляется возможность автоматического вычисления значений теста в рамках всех доменов по одному, например, самому точному значению. Примеры:

Возраст/  $V1? 5 \rightarrow$  Возраст/  $V2? \text{юный} \rightarrow$  Возраст/  $V3? \text{молодой}$ ;

Возраст/  $V1? 76 \rightarrow$  Возраст/  $V2? \text{старческий} \rightarrow$  Возраст/  $V3? \text{пожилой}$ .

Преобразование множества значений:

Возраст/  $V1? [13; 20] \rightarrow$  Возраст/  $V2? \{\text{юный; молодой}\} \rightarrow$  Возраст/  $V3? \text{молодой}$ .

Результат получен путем объединения преобразований атомарных значений.

Зададим другой набор доменов для того же теста  $\tau = \text{'Возраст'}$ :

$V1 = [0; 100]$ ;

$V4 = \{[0; 17]; [18; 30]; [31; 40]; [41; 55]; [56; 60]; [61; 100]\}$ ;

$V5 = \{\text{молодой; средних лет; пожилой}\}$ ;

$V6 = \{\text{трудоспособный; нетрудоспособный}\}$ .

Правила пересчета значений из домена  $V1$  в домен  $V4$  очевидны. Зададим остальные правила (с учетом разного возраста ухода на пенсию мужчин и женщин: мужчины в 60 лет, женщины в 55):

$V4.\{[0; 17]; [18; 30]\} \rightarrow V5.\{\text{молодой}\}$ ;

$V4.\{[31; 40]; [41; 55]\} \rightarrow V5.\{\text{средних лет}\}$ ;

$V4.\{[56; 60]; [61; 100]\} \rightarrow V5.\{\text{пожилой}\}$ ;

Пол? М &  $V4.\{[18; 30]; [31; 40]; [41; 55]; [56; 60]\} \rightarrow V6.\{\text{трудоспособный}\}$ ;

Пол? Ж &  $V4.\{[18; 30]; [31; 40]; [41; 55]\} \rightarrow V6.\{\text{трудоспособный}\}$ ;

Пол? М &  $V4.\{[0; 18]; [61; 100]\} \rightarrow V6.\{\text{нетрудоспособный}\}$ ;

Пол? Ж &  $V4.\{[0; 18]; [56; 60]; [61; 100]\} \rightarrow V6.\{\text{нетрудоспособный}\}$ .

В итоге получаем второй орграф:

$$G_2(\text{Возраст} \mid \{\text{Пол}\}) \equiv \{B1 \rightarrow B4; B4 \rightarrow B5; B4 \{\text{Пол}\} \rightarrow B6\}.$$

Обратим внимание, что, несмотря на внешнее совпадение термов доменов B3 и B5, это разные домены с отличными по смыслу терминами. Следует также отметить, что если значение теста «Пол» не задано, то связка (B4 {Пол} → B6) не определена и орграф  $G_2(\text{Возраст})$  меняет структуру на  $\{B1 \rightarrow B4; B4 \rightarrow B5\}$ .

Примеры верного вычисления значений теста на основе орграфа  $G_2(\text{Возраст})$ :

Возраст/ B1? 5 → Возраст/ B4? [0; 17] → Возраст/ B5? молодой;

Пол/{М; Ж}? М & Возраст/ B4? [0; 17] → Возраст/ B6? нетрудоспособный.

Операции обобщения в рамках орграфов доменов отнесем к автоматизмам вычислительной среды (вычислительного интеллекта).

На рис. 2.1 показана структура орграфов  $G_1(\text{Возраст})$  и  $G_2(\text{Возраст})$ .

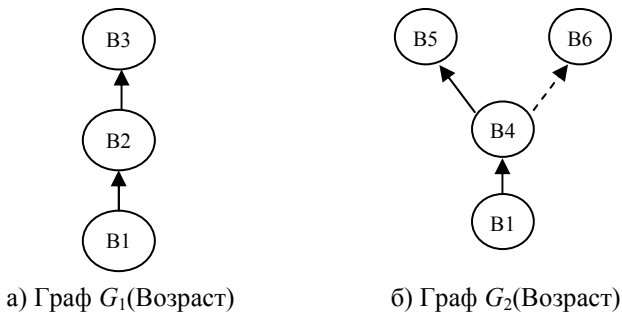


Рис. 2.1 – Структура орграфов для теста «Возраст»

Из рисунка видно, что орграф  $G_1(\text{Возраст})$  имеет линейную структуру с одной терминальной вершиной, а орграф  $G_2(\text{Возраст})$  – ветвящуюся, с двумя терминальными вершинами (пунктирная стрелка означает зависимость преобразования от внешних тестов).

Определим *операцию объединения* ‘ $\cup$ ’ разных орграфов одного и того же теста, если у них совпадают по содержанию как минимум базовые домены. В результирующем орграфе общим становится базовый домен. Обязательно объединяются и другие вершины – домены, если между элементами этих доменов имеется взаимнооднозначное соответствие и доказано, что отсутствуют конфликты (конфликт имеет место, когда разные пути вычислений приводят к разным значениям в одном домене). Пусть  $G_k(\tau \mid \{a/A\}_k)$  – разные орграфы теста  $\tau$  ( $k = 1, \dots, K$ ) с совпадающим



базовым доменом. В результате объединения получим орграф  $G(\tau)$ :

$$G(\tau | \cup_{k=1, \dots, K} [\{a/A\}_k]) = G_1(\tau | [\{a/A\}_1]) \cup \dots \cup G_K(\tau | [\{a/A\}_K]). \quad (2.1)$$

Названия вершин (доменов) в объединенном орграфе могут отличаться от названий вершин в исходных орграфах. Операция объединения орграфов коммутативна и ассоциативна.

Пусть  $G(\text{Возраст}) = G_1(\text{Возраст}) \cup G_2(\text{Возраст} | \{\text{Пол}\})$ . На рис. 2.2 показана результирующая структура орграфа  $G(\text{Возраст})$ .

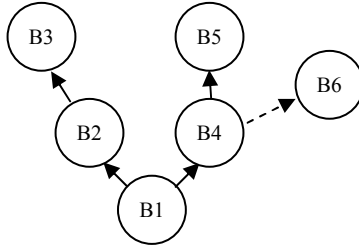


Рис. 2.2 – Структура результирующего орграфа для теста «Возраст»

Отметим, что, несмотря на внешнее совпадение, вершины B3 и B5 объединить нельзя, так как в результате объединения имел бы место конфликт значений теста. Действительно, предположим, что вершины B3 и B5 объединены (условно). Зададим значение «Возраст/ B1? 33» и вычислим значения теста в доменах B3 и B5. Получим:

Возраст/ B1? 33 → Возраст/ B4? [31; 40] → Возраст/ B5? средних лет;  
 Возраст/ B1? 33 → Возраст/ B2? молодой → Возраст/ B3? молодой.

При объединении вершин B3 и B5 имел бы место конфликт значений, а именно: (Возраст/ B5? средних лет) ≠ (Возраст/ B3? молодой).

Данный пример показывает, что семантический смысл элементов домена (термов) жестко привязан к самому домену, точнее даже к маршруту его вычисления из базового домена. В более широком контексте любой терм может стать размытым, нечетким, многозначным.

В некоторых случаях может применяться операция прививки одного орграфа к другому. Граф  $G_2(\tau)$  может быть привитым базовой вершиной к графу  $G_1(\tau)$  в вершине  $T$  с обязательным объединением других совпадающих вершин тогда и только тогда, когда совпадает состав элементов домена  $T$  и базовой вершины орграфа  $G_2(\tau)$ , совпадают их смысловые зоны (области возможных значений – ОВЗ) и нет конфликтов значений в результирующем орграфе. Граф, получаемый в результате прививки  $G_2(\tau)$  к  $G_1(\tau)$ , обозначим следующим образом:

$$G(\tau) = G_1(\tau) [T] G_2(\tau), \quad (2.2)$$

где  $T$  – вершина орграфа  $G_1(\tau)$ , к которой прививается базовая вершина орграфа  $G_2(\tau)$ . Операция прививки нескольких графов выполняется последовательно, например, прививка двух графов запишется следующим образом:

$$G(\tau) = (G_1(\tau) [T] G_2(\tau)) [T'] G_3(\tau). \quad (2.3)$$

Операция прививки не коммутативна и не ассоциативна. Очевидно, операция объединения графов является частным случаем операции прививки. Приведем пример операции прививки.

Зададим орграф  $G_3(\text{Возраст})$  следующим образом:

$D1 = \{\text{юный; молодой; средних лет; пожилой; старческий}\};$

$D2 = \{\text{юный; не юный}\}.$

$G_3(\text{Возраст}) = \{D1 \rightarrow D2\}.$

Правила преобразования:

$D1. \{\text{юный}\} \rightarrow D2. \{\text{юный}\};$

$D1. \{\text{молодой; средних лет; пожилой; старческий}\} \rightarrow D2. \{\text{не юный}\}.$

Привьем граф  $G_3(\text{Возраст})$  к графу  $G_1(\text{Возраст})$  в вершине  $B2$ . Операция прививки запишется следующим образом:

$$G''(\text{Возраст}) = G_1(\text{Возраст}) [B2] G_3(\text{Возраст}).$$

На рис. 2.3 изображен результирующий граф.

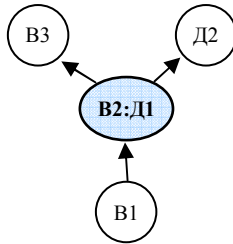


Рис. 2.3 – Операция прививки орграфа

Помимо операций объединения и прививки необходимо ввести *операцию сращивания* двух и более графов. Операция сращивания применяется к орграфам, у которых не совпадают соединяемые вершины. Для сращивания двух вершин разных графов необходимо задать связку между соединяемыми вершинами. Один из двух сращиваемых графов в обязательном порядке соединяется базовой вершиной. Пусть сращиваются два орграфа:  $G_1(\tau)$  и  $G_2(\tau)$ , причем орграф  $G_2(\tau)$  базовой вершиной  $T'$

присоединяется к вершине  $T$  орграфа  $G_1(\tau)$ . Если ОБЗ вершин  $T$  и  $T'$  совпадают, то операция сращивания выполнима. Ее можно записать следующим образом:

$$G(\tau) = G_1(\tau) [T \rightarrow T'] G_2(\tau). \quad (2.4)$$

На рис. 2.4 показан пример сращивания двух орграфов. Пунктирной стрелкой показана новая связь ( $T \rightarrow T'$ ).

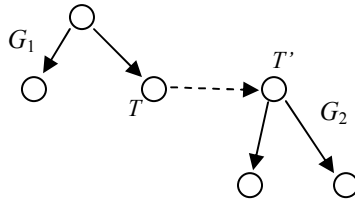


Рис. 2.4 – Операция сращивания двух орграфов

Так как орграфы доменов тестов принадлежат онтологии какой-либо задачи, то операции объединения, прививки и сращивания орграфов доменов могут быть полезны при слиянии разных онтологий (онтологическом инжиниринге).

Выше уже отмечалось, что необходимо стремиться к тому, чтобы терминальные вершины орграфа были дуадами. В этом смысле орграф  $G(\text{Возраст})$  не оптимальный. Введем *операцию расщепления вершины-домена* следующим образом: для любой дискретной вершины  $D = \{a_1, a_2, \dots, a_n\}$  такой, что  $n > 2$ , порождаются  $n$  терминальных вершин-листьев вида  $\{a_j; \text{не } a_j\} \equiv \{a_j; \neg a_j\} \equiv \{a_j; D \setminus a_j\}$ , где  $j = 1, \dots, n$ . В результирующем орграфе остаются только те вершины, которые отсутствовали в орграфе до выполнения операции расщепления. В рамках настоящего исследования будем предполагать, что имеется воспроизводимый алгоритм проверки совпадения (тождественности) вершин с учетом синонимии и разных областей отрицания (в разных доменах). Примеры типичной синонимии:

$\{\text{Жалоб на озноб нет}; \text{Жалобы на озноб}\} \equiv \{\text{Жалоб на озноб нет}; \neg \text{‘Жалоб на озноб нет’}\};$

$\{\text{Кожные покровы нормального цвета}; \text{Желтушность кожных покровов}\} \equiv \{\text{Кожные покровы нормального цвета}; \neg \text{‘Кожные покровы нормального цвета’}\}.$

Последний пример справедлив в рамках следующего орграфа теста «Желтушность кожных покровов»:  $G = \{D1 \rightarrow D2\}$ , где

$D1$  {Кожные покровы нормального цвета  $\wedge N$ ; Умеренная желтушность кожных покровов  $\wedge a$ ; Интенсивная желтушность кожных покровов  $\wedge b$ };  
 $D2$  {Кожные покровы нормального цвета  $\wedge N$ ; Желтушность кожных покровов  $\wedge a \vee b$ }.

Правила преобразования:  $D1.N \rightarrow D2.N$ ;  $D1.\{a; b\} \rightarrow D2.\{a\}$ .

Условные обозначения (алиасы) 'N', 'a', 'b' введены для упрощения записи правил преобразования. Справедливость утверждения об эквивалентности доменов вытекает из правил преобразования в рамках орграфа  $G$ , а именно:  $\{N; \neg N\} \equiv \{N; a \vee b\}$ .

Пример разных областей отрицания. Пусть  $D = \{a; b; c\}$  – исходная вершина. Новые терминальные вершины, порождаемые операцией расщепления домена, будут иметь вид:

$$D1 = \{a; b \vee c\} \equiv \{a; \neg a\} \equiv \{a; \text{не } a\};$$

$$D2 = \{b; a \vee c\} \equiv \{b; \neg b\} \equiv \{b; \text{не } b\};$$

$$D3 = \{c; a \vee b\} \equiv \{c; \neg c\} \equiv \{c; \text{не } c\}.$$

Здесь «не  $a$ », «не  $b$ » и «не  $c$ » – автоматически образованные новые элементы (термы).

Пусть  $H = \{a; b; c; d\}$ . Одна из порождаемых в результате расщепления терминальных вершин будет иметь вид:

$$H1 = \{a; b \vee c \vee d\} \equiv \{a; \neg a\} \equiv \{a; \text{не } a\}.$$

Совпадают ли вершины  $D1$  и  $H1$ ? На основе имеющейся информации определенный ответ дать нельзя. Действительно, пусть  $H \rightarrow D$  и одно из правил преобразования имеют вид:  $H.\{a\} \rightarrow D.\{a\}$ . В таком варианте вершины  $D1$  и  $H1$  совпадают, так как не происходит изменения смысловой области элемента  $a$ .

Рассмотрим другой вариант преобразования. Пусть одно из правил имеет вид:  $H.\{a; b\} \rightarrow D.\{a\}$ . В данном варианте произошло расширение смысловой области элемента  $a$ , следовательно,  $D1$  и  $H1$  отличаются.

Следует пояснить, почему при расщеплении порождаются только дуады, а не триады, например. Дело в том, что дуады относятся к базовым аспектам системы знаний, а именно – к оппозициям. Выявление оппозиций, как и неоднородностей, является фундаментом для синтеза доменов тестов. Дуады-листья позволяют выявить экстремумы реальности, полюса, дуальное переключение и т.д.

Операцию расщепления можно применять локально – лишь к одной вершине орграфа, а можно последовательно ко всем дискретным вершинам орграфа (при соблюдении условия  $|D| > 2$ ). Орграф, получаемый в результате применения операции расщепления ко всем допустимым вершинам (*операция наращивания орграфа*), назовем *структурно-*

*завершенным*. Структурно-завершенный оргграф, построенный на основе оргграфа  $G(\tau)$ , обозначим через  $G^+(\tau)$ .

**Предложение 2.1.** Для оргграфа  $G(\tau)$  его структурно-завершенный оргграф  $G^+(\tau)$  определяется в общем случае неоднозначно, но состав вершин во всех оргграфах  $G^+(\tau)$  одинаковый.

Доказательство вытекает из того факта, что операция расщепления вершины применяется последовательно ко всем дискретным доменам оргграфа, но порядок применения операции не фиксирован и разная последовательность обхода вершин оргграфа может дать разные результаты. Одинаковость состава вершин во всех оргграфах  $G^+(\tau)$  вытекает из свойств полноты порождения терминальных вершин и одновременно отсутствия избыточности вершин любого оргграфа из  $\{G^+(\tau)\}$ , т.е. совпадающих вершин.

Вершины, которые образуются в результате расщепления, и которые могут иметь одновременно несколько доменов-предков, назовем *блуждающими*. Из-за наличия блуждающих вершин операция наращивания оргграфа, в том виде, как она приведена выше (последовательное расщепление вершин), не обеспечивает получение всех истинных структурно-завершенных графов.

Построение полного множества структурно-завершенных графов предполагает первоначальное выявление всех блуждающих вершин с последующим построением всех комбинаций связей блуждающих вершин с доменами-предками. Предложение 2.1 справедливо и для полного множества (истинных) структурно-завершенных графов.

Множество всех истинных структурно-завершенных оргграфов для оргграфа  $G(\tau)$  обозначим  $\{G^+(\tau)\}$ .

Оргграф, в котором все блуждающие вершины соединены со всеми возможными доменами-предками, назовем *предельным структурно-завершенным оргграфом* (сокращенно – предельным оргграфом). Обозначим его через  $G^{++}(\tau)$ . Из определения истинного множества структурно-завершенных оргграфов следует, что предельный оргграф принадлежит  $\{G^+(\tau)\}$ .

**Предложение 2.2.** Для произвольного оргграфа  $G(\tau)$  его предельный структурно-завершенный оргграф  $G^{++}(\tau)$  является корректным и определяется единственным образом.

Утверждение доказывается от противного. Предположим, что предельный оргграф определяется не единственным образом, тогда для некоторого  $G(\tau)$  существуют, как минимум, два предельных графа:  $G_1^{++}(\tau)$  и  $G_2^{++}(\tau)$ . Оба предельных графа принадлежат  $\{G^+(\tau)\}$ , следовательно, согласно предложению 2.1 состав вершин у них одинаковый. Таким

образом, оба орграфа отличаются только некоторой связкой ( $T \rightarrow T$ ), соединяющей домен-предок с листком. Но по построению у предельного орграфа все листья связаны со всеми доменами-предками, следовательно, у другого предельного орграфа также должна быть связка ( $T \rightarrow T$ ). Пришли к противоречию, что доказывает единственность. Корректность будет показана при рассмотрении канонических орграфов, которые предназначены именно для этой цели.

Алгоритм построения предельного орграфа простой. Опишем его.

Алгоритм 2.1. - *LimitGraf*

Вход: Орграф доменов  $G(\tau)$ .

Выход: Предельный структурно-завершенный орграф  $G^{++}(\tau)$ .

Полагаем:  $G^{++}(\tau) = G(\tau)$ .

ЦИКЛ по всем дискретным вершинам  $G^{++}(\tau)$  с числом элементов больше 2.

- Для текущей вершины  $T$  выполняем операцию расщепления.
- ЦИКЛ по всем расщепленным вершинам – листьям  
ЕСЛИ текущий листок  $T$  или ребро ( $T \rightarrow T$ ) отсутствуют в  $G^{++}(\tau)$ , то выполняем:  $G^{++}(\tau) := G^{++}(\tau) \cup (T \rightarrow T)$ .

КОНЕЦ цикла.

КОНЕЦ цикла.

ВЫХОД.

Справедливо следующее утверждение.

**Предложение 2.3.** Алгоритм *LimitGraf* для произвольного орграфа  $G(\tau)$  за конечное число шагов строит предельный структурно-завершенный орграф  $G^{++}(\tau)$ .

Общее число итераций по внешнему циклу не превышает  $N_\tau$  (число вершин в  $G(\tau)$ ). Число итераций во внутреннем цикле не превышает максимального числа элементов среди всех вершин графа  $G(\tau)$ . Обозначим это число через  $M_\tau$ , следовательно, общее число итераций не превышает  $M_\tau \cdot N_\tau$ , что и доказывает конечность алгоритма. Предположим, что алгоритм строит не предельный орграф  $G^{++}(\tau)$ , а некий орграф  $G'(\tau)$ . Орграфы  $G^{++}(\tau)$  и  $G'(\tau)$  должны отличаться некоторой связкой ( $T \rightarrow T$ ), причем вершина  $T$  принадлежит и  $G^{++}(\tau)$  и  $G'(\tau)$ . Пусть связка ( $T \rightarrow T$ ) отсутствует в  $G^{++}(\tau)$  и присутствует в  $G'(\tau)$ . Вершина  $T$  может быть только листком, так как единственной операцией порождения новых вершин и/или ребер является операция расщепления, причем к листьям (новым вершинам) данная операция не применяется. Следовательно, связка ( $T \rightarrow T$ ) орграфа  $G'(\tau)$  образуется в результате расщепления вершины  $T$ , согласно алгоритму *LimitGraf*, но тогда связка ( $T \rightarrow T$ ) должна принадлежать и  $G^{++}(\tau)$  (по определению). Пришли к противоречию.

Аналогично показывается, что и  $G^{++}(\tau)$  не может содержать связок, которых нет в  $G'(\tau)$ , следовательно, они тождественны.

К какому типу непрерывному или дискретному отнести вершину В1 орграфа теста «Возраст»? Если ее отнести к дискретному типу (каждый год – отдельный элемент), то при построении структурно-завершенного орграфа операция расщепления создаст 101 новую терминальную вершину. Для большинства приложений это может оказаться избыточным. Исключить вершину В1 из списка вершин, к которым будет применена операция расщепления, можно лишь одним способом – отнести ее к типу «конструктивная непрерывность», что эквивалентно фразе «очень большое число элементов». К вершинам орграфа двух типов – «непрерывная» и «конструктивно непрерывная» – операция расщепления не применяется.

Результатом применения операции расщепления к вершине В3 являются три новые вершины:

$V7 = \{\text{молодой; средних лет} \vee \text{пожилой}\} \equiv \{\text{молодой; не молодой}\};$

$V8 = \{\text{средних лет; молодой} \vee \text{пожилой}\} \equiv \{\text{средних лет; не средних лет}\};$

$V9 = \{\text{пожилой; молодой} \vee \text{средних лет}\} \equiv \{\text{пожилой; не пожилой}\}.$

Результатом применения операции расщепления к вершине В5 являются три аналогичные вершины: В10, В11, В12. Так как вершина В6 является дихотомической, то к ней операция расщепления не применяется. После избирательного расщепления получим новый граф  $G'(\text{Возраст} \mid \{\text{Пол}\})$ . Его структура показана на рис. 2.5 (листья окрашены зеленым и желтым цветом, желтым цветом окрашена блуждающая вершина В8).

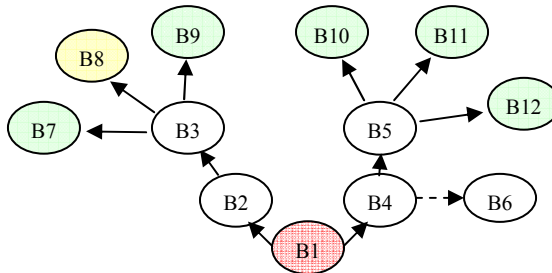


Рис. 2.5 – Структура орграфа  $G'(\text{Возраст} \mid \{\text{Пол}\})$

Орграф  $G'(\text{Возраст} \mid \{\text{Пол}\})$ , по-прежнему не является структурно-завершенным, так как остались вершины, к которым не применена операция расщепления, а именно, вершины В2 и В4 (вершину В1 отнесем к типу «конструктивная непрерывность» – красный оттенок). Применяя операцию расщепления к вершинам В2 и В4, получим дополнительно 10

терминальных вершин (4 и 6 соответственно с учетом одной блуждающей вершины). Таким образом, структурно-завершенный орграф  $G^+(\text{Возраст} \mid \{\text{Пол}\})$  будет содержать 22 вершины. Такое число вершин будут содержать все структурно-завершенные орграфы из  $\{G^+(\text{Возраст} \mid \{\text{Пол}\})\}$ . Предельный структурно-завершенный орграф  $G^{++}(\text{Возраст} \mid \{\text{Пол}\})$  будет содержать дополнительное ребро ( $B2 \rightarrow B8$ ). При рассмотрении канонических орграфов будет приведен пример  $G^+(\text{Возраст} \mid \{\text{Пол}\})$ .

Рассмотрим еще примеры. Пусть значения теста «Аппетит» выбираются из двух доменов:

Тест: *Annetum* ^A1:

D2 {Аппетит нормальный ^N; Аппетит изменен ^b}

D1 {Отвращение к пище (анорексия) ^b; Аппетит снижен ^b; Аппетит нормальный ^N; Аппетит повышен ^b}

Граф  $G(A1) = \{D1 \rightarrow D2\}$ ,

где «A1» – условное обозначение теста; «N» и «b» – условные обозначения элементов доменов, они задают однозначные схемы преобразования:  $D1.b \rightarrow D2.b$ ;  $D1.N \rightarrow D2.N$ .

На базе  $G(A1)$  построим структурно-завершенный орграф. Так как домен D2 дихотомический, то к нему операция расщепления не применяется. В результате применения операции расщепления к базовой вершине D1 появляются 4 терминальные вершины. Однако автоматически сгенерированный домен «Аппетит нормальный; Аппетит не нормальный» по смыслу совпадает с доменом D2, хотя формулировка второго термина и отличается. Автоматизмы среды должны распознавать подобные совпадения и не порождать вершины – дубликаты. Решению задач распознавания должна способствовать база синонимов термов, которая является частью онтологии предметной области. В любом случае пользователь должен убедиться в корректности результатов автоматических процедур. Таким образом, структурно-завершенный орграф  $G^+(A1)$  будет содержать 5 вершин: одну базовую и 4 терминальных. В данном примере имеем только один вариант структурно-завершенного орграфа.

Зададим другой орграф для теста «Аппетит»:

Тест: *Annetum* ^A2:

D3 {Аппетит нормальный ^N; Аппетит изменен ^a b}

D2 {Аппетит снижен ^a; Аппетит нормальный ^N; Аппетит повышен ^b}

D1 {Отвращение к пище (анорексия) ^a; Аппетит снижен ^a; Аппетит нормальный ^N; Аппетит повышен ^b; Аппетит очень высокий ^b}

Граф  $G(A2) = \{D1 \rightarrow D2; D2 \rightarrow D3\}$ .

Правила преобразования:  $D1.b \rightarrow D2.b$ ;  $D1.a \rightarrow D2.a$ ;  $D1.N \rightarrow D2.N$ ;



$D2.N \rightarrow D3.N$ ;  $D2.\{a; b\} \rightarrow D3.a$ .

Построим структурно-завершенный орграф  $G^+(A2)$ . Выше уже отмечалось, что операция отрицания связана с конкретным доменом. Поскольку  $D1.N \rightarrow D2.N \rightarrow D3.N$ , т.е. не происходит расширения смысловой зоны элемента «Аппетит нормальный», то операция расщепления, примененная к доменам D1 и D2, порождает одинаковые вершины «Аппетит нормальный; Аппетит *не* нормальный», которые совпадают с D3. Сложнее обстоит дело с элементами (термами) доменов D1, D2 – «Аппетит снижен» и «Аппетит повышен». При переходе от домена D1 к домену D2 происходит расширение смысловой зоны каждого термина, например:

$D1.\{\text{Анорексия } \wedge a; \text{Аппетит снижен } \wedge a\} \rightarrow D2.\{\text{Аппетит снижен } \wedge a\}$ .

Это приводит, например, к следующему факту: «Аппетит снижен;  $\neg_{D1}$ Аппетит снижен»  $\neq$  «Аппетит снижен;  $\neg_{D2}$  Аппетит снижен».

В итоге структурно-завершенный орграф  $G^+(A2)$  будет содержать 9 вершин.

Построим структурно-завершенный орграф для теста «Жалобы на боли в области живота»:

Тест: *Жалобы на боли в области живота*  $\wedge$  ЖБОЖ:

D2 {Болей в животе нет (норма)  $\wedge$  N; Жалобы на боли в животе  $\wedge$  b}

D1 {Болей в животе нет (норма)  $\wedge$  N; Постоянные ноющие боли в правом подреберье  $\wedge$  b; Схваткообразные боли в правом подреберье  $\wedge$  b; Опоясывающие боли в эпигастрии  $\wedge$  b; Боли в левом подреберье  $\wedge$  b}.

$G(\text{ЖБОЖ}) = \{D1 \rightarrow D2\}$ . Правила:  $D1.b \rightarrow D2.b$ ;  $D1.N \rightarrow D2.N$ .

Формально домен D1 в результате расщепления порождает 5 листьев (по числу элементов), однако вершина D2 входит в число данных вершин, поэтому структурно-завершенный орграф  $G^+(\text{ЖБОЖ})$  содержит 6 вершин – одну базовую и 5 терминальных (рис. 2.6).

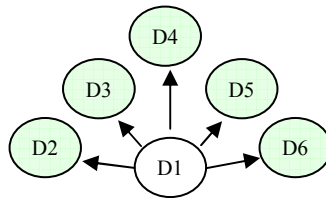


Рис. 2.6 – Структурно-завершенный орграф теста «Жалобы на боли в области живота»

Определим число вершин структурно-завершенного орграфа для теста

«Температура тела»:

*Температура тела*  $\wedge$ ТТела:

- 4 {Нормальная температура  $\wedge$ N; Ненормальная температура  $\wedge$ b c}
- 3 {Пониженная температура  $\wedge$ b; Нормальная температура  $\wedge$ N; Повышенная температура  $\wedge$ c}
- 2 {Резкая гипотермия  $\wedge$ b b1 [34.0; 35]; Гипотермия  $\wedge$ b b2 (35.0; 36); Нормальная температура  $\wedge$ N (36.0; 36.9); Субфебрильная температура  $\wedge$ c c1 (36.9; 38); Фебрильная температура  $\wedge$ c c2 (38.0; 40); Предельная температура  $\wedge$ c c3 (40.0; 42.0)}
- 1 {[34.0; 42.0]}

Граф  $G(\text{ТТела}) = \{1 \rightarrow 2; 2 \rightarrow 3; 3 \rightarrow 4\}$ .

Правила: 1.[34.0; 35]  $\rightarrow$  2.b1; 1.(35.0; 36]  $\rightarrow$  2.b2; 1.(36.0; 36.9]  $\rightarrow$  2.N;

1.(35.0; 36]  $\rightarrow$  2.c1; 1.(38.0; 40]  $\rightarrow$  2.c2; 1.(40.0; 42.0]  $\rightarrow$  2.c3;

2.b  $\rightarrow$  3.b; 2.c  $\rightarrow$  3.c; 2.N  $\rightarrow$  3.N; 3.{b; c}  $\rightarrow$  4.b; 3.N  $\rightarrow$  4.N.

Поясним обозначения. Дополнительные условные обозначения b1, b2, c1, c2, c3 к элементам второго домена добавлены лишь для компактности записи правил, хотя они могут также играть роль уникальных ссылок ('b', 'c' не могут быть такими ссылками). Условных обозначений для тестов, доменов и элементов доменов может быть сколь угодно много. Названия доменов могут быть чисто цифровыми, что облегчает реализацию орграфов с помощью баз данных и электронных таблиц.

Структура орграфа  $G(\text{ТТела})$  линейная, с одной терминальной вершиной. При расщеплении вершина «2» порождает 6 терминальных вершин, одна из которых совпадает с вершиной «4». Вершина «3» порождает 3 терминальные вершины, одна из которых также совпадает с вершиной «4». Вершина «1» относится к типу «непрерывная», поэтому операция расщепления к ней не применяется. В итоге получаем, что структурно-завершенный орграф для теста «Температура тела» содержит 11 вершин, из которых 8 терминальные.

Автоматически порождаемые в результате операции расщепления терминальные вершины, а также вершины, которые можно отнести к такому типу, названы *листьями*. Листья не нужно включать в онтологические соглашения. Листьев в орграфе может быть достаточно много и хранить их в памяти весьма неэкономно. В памяти достаточно хранить *ствол орграфа* – орграф без листьев, а листья должны порождаться автоматизмами вычислительной среды в моменты использования орграфа. Поскольку предполагается, что имеется алгоритм установления эквивалентности доменов, то этот же алгоритм позволяет однозначно отнести любую вершину к типу «листья». Для произвольного орграфа  $G(\tau)$  его ствол обозначим через  $G^{\sim}(\tau)$ .

Не все дихотомические вершины являются листьями, например, вершина В6 в орграфе  $G(\text{Возраст})$  не является листком, так как не принадлежит множеству вершин расщепления домена-предка. У орграфа, состоящего из одной дихотомической вершины, листьев быть не может. У остальных орграфов листья есть или могут порождаться.

В орграфе доменов могут быть такие бинарные вершины, которые одновременно являются и «листком» и «не листком» (имеется дуализм). Приведем пример. Определим орграф доменов теста  $X$  следующим образом:

Тест  $X$ :

3 {1 2; 3; 4};

2 {1 2; 3 4};

1 {1; 2; 3; 4}.

Граф  $G(X) = \{1 \rightarrow 2; 1 \rightarrow 3\}$ .

Правила пересчета значений доменов очевидны.

Вершина «2» графа  $G(X)$  является дуадой, но не является листком. В процессе построения предельного орграфа  $G^{++}(X)$  расщепление вершины «3» приведет к появлению листка  $\{1\ 2; \neg(1\ 2)\} = \{1\ 2; 3\ 4\}$ , который совпадает с вершиной «2». Согласно алгоритму *LimitGraf* в  $G^{++}(X)$  будет добавлено ребро  $(3 \rightarrow 2)$ . Таким образом, в орграфе  $G^{++}(X)$  вершина «2» становится *двойственной дуадой* – она одновременно и является, и не является листком.

При построении ствола орграфа доменов двойственные вершины (дуады) не удаляются, а удаляются лишь ребра к тем вершинам, листком которых является дуадная вершина. Таким образом, в стволе орграфа все двойственные дуады перестают быть двойственными.

Через  $\{(G^{\sim}(\tau))^+\}$  обозначим множество всех истинных структурно-завершенных графов, построенных на основе ствола  $G^{\sim}(\tau)$ . Предельный структурно-завершенный орграф, построенный на основе  $G^{\sim}(\tau)$ , обозначим через  $(G^{\sim}(\tau))^{++}$ .

**Предложение 2.4.** Ствол  $G^{\sim}(\tau)$  произвольного орграфа  $G(\tau)$  определяется единственным образом. Предельный структурно-завершенный орграф  $(G^{\sim}(\tau))^{++}$  является корректным и определяется единственным образом. Состав вершин всех орграфов из  $\{(G^{\sim}(\tau))^+\}$  одинаковый и, кроме того, выполняются следующие соотношения:

(i)  $\{G^+(\tau)\} \subseteq \{(G^{\sim}(\tau))^+\}$ .

(ii)  $\forall G^+(\tau) \in \{G^+(\tau)\}$  выполняется:  $(G^+(\tau))^{\sim} = G^{\sim}(\tau)$ .

(iii)  $(G^{\sim}(\tau))^{++} = G^{++}(\tau)$ .

Единственность ствола орграфа вытекает из однозначности операции порождения листьев (операция расщепления) и операции удаления листьев

(с учетом двойственных вершин). Единственность предельного орграфа вытекает из предложения 2.2. Одинаковость состава вершин всех графов из  $\{(G^{\sim}(\tau))^+\}$  вытекает из предложения 2.1. Действительно, орграфы множества  $\{(G^{\sim}(\tau))^+\}$  имеют общее основание – орграф  $G^{\sim}(\tau)$  и отличаются только связью блуждающих вершин с доменами-предками, но состав блуждающих вершин неизменный.

Соотношение (i) означает, что ствол графа порождает больше структурно-завершенных графов, чем исходный граф. Данный факт вытекает из следующих соображений. Исходный граф  $G(\tau)$  уже может содержать листья с фиксированными связями с доменами-предками (структурно-завершенные графы  $G^+(\tau)$  могут лишь добавить новые связи). От таких связей можно избавиться, только построив ствол  $G^{\sim}(\tau)$ .

Соотношение (ii) означает, что ствол любого структурно-завершенного орграфа совпадает со стволом исходного орграфа. Доказательство вытекает из соотношения (i).

Соотношение (iii) доказывается от противного. Предположим, что орграфы  $G^{++}(\tau)$  и  $(G^{\sim}(\tau))^{++}$  отличаются. Они должны отличаться некоторой связкой  $(T \rightarrow T')$ , причем вершина  $T$  принадлежит и  $G^{++}(\tau)$ , и  $(G^{\sim}(\tau))^{++}$ . Пусть связка  $(T \rightarrow T')$  отсутствует в  $G^{++}(\tau)$ . Вершина  $T$  может быть только листком, так как единственной операцией порождения новых вершин является операция расщепления, причем, к листьям (новым вершинам) данная операция не применяется. Так как  $(T \rightarrow T') \in (G^{\sim}(\tau))^{++}$ , то вершина  $T$  принадлежит к числу листьев, образующихся в результате расщепления вершины  $T$ , но тогда  $T$  должна принадлежать и  $G^{++}(\tau)$ . Пришли к противоречию. Аналогично приходим к противоречию предположив, что связка  $(T \rightarrow T')$  отсутствует в  $(G^{\sim}(\tau))^{++}$ .

Стволы орграфа и структурно-завершенного орграфа совпадают. Отсюда следует, что по стволу орграфа невозможно точно восстановить исходный орграф, можно восстановить только множество структурно-завершенных графов.

Примеры. Ствол орграфа  $G(\text{Аппетит} - A1)$  состоит из одной вершины D1. Ствол орграфа  $G(\text{Аппетит} - A2)$  состоит из связки  $\{D1 \rightarrow D2\}$ . Ствол орграфа  $G(\text{ЖБОЖ})$ , и соответственно орграфа  $G^+(\text{ЖБОЖ})$  состоит из одной вершины D1. Ствол орграфа  $G(\text{ТТела})$  состоит из связок  $\{1 \rightarrow 2; 2 \rightarrow 3\}$ .

Любой из элементов домена B4 теста «Возраст» можно разбить на несколько частей, например, первый элемент  $[0; 17]$  можно представить как обобщение элементов  $\{[0; 5]; [6; 10]; [11; 17]\}$ . Новые элементы также можно разбить на несколько частей. При любом подобном разбиении получается новый – более точный домен. *Операцию деления* произвольного домена, состоящего из интервалов, можно формализовать разными

схемами. Одни схемы являются чистыми разбиениями (например, каждый элемент домена делится на  $N$  равных частей), другие схемы предполагают отбрасывание некоторых участков (фрактальный способ деления). При отбрасывании участков должна сохраняться полнота проекции на базовый домен. На рис. 2.7 приведен фрагмент схемы равномерного разбиения произвольного интервала.

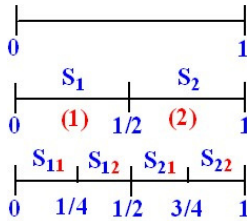


Рис. 2.7 – Пример разбиения интервала

Схемы деления могут быть *адаптивными* и зависеть от распределения эмпирических данных, например, интервал можно разделить на  $N$  частей, при этом в каждую часть должно попасть одинаковое число эмпирических данных.

Строго говоря, операция деления может применяться к любой связке  $(T \rightarrow T)$ , в которой домены  $T$ ,  $T$  – интервальные домены. Результатом последовательного деления является совокупность промежуточных доменов  $\{T_1, \dots, T_k\}$  такая, что  $T \rightarrow T_1 \rightarrow T_2 \rightarrow \dots \rightarrow T_k \rightarrow T$ . Главное условие деления: вставка промежуточных доменов должна полностью сохранить логику преобразования связки  $(T \rightarrow T)$ .

Если базовый домен непрерывный, то операция деления домена-потомка может выполняться бесконечное число раз, символизируя отсутствие пределов делимости материи и информации. Благодаря этому появляется возможность по существу неограниченного производства информации. *Схема деления*  $\Sigma_\tau$  домена-потомка для базового (непрерывного) домена является неотъемлемой частью орграфа  $G(\tau)$ . В общем случае, схема деления задается совокупностью параметров  $\{s/S\}$ , т.е.  $\Sigma_\tau = \{s/S\}_\tau$ . Одним из ключевых параметров является глубина деления – число промежуточных уровней доменов  $k$ .

Отметим важный методологический момент, который состоит в том, что операция расщепления вершин и операция деления вершин позволяют в автоматическом режиме на основе ствола орграфа доменов теста порождать потенциально бесконечное множество орграфов одного и того же теста. Ствол орграфа доменов теста играет роль памяти.

Обратной по отношению к операции деления является *операция схлопывания* последовательности промежуточных доменов. В результате схлопывания промежуточные домены  $\{T_1, \dots, T_k\}$  удаляются, остается только связка  $(T \rightarrow T')$ .

Новая *парадигма моделирования* заключается в том, что любой параметр модели и любая переменная (сила, скорость, ускорение, масса, момент инерции, риск и т.д.) являются тестами, которые задаются своими орграфами доменов. При таком подходе нет необходимости говорить о непрерывных, дискретных, интервальных, нечетких, текстовых или лингвистических переменных (о доменах с нечеткими элементами речь пойдет в следующем разделе). Орграф доменов включает все эти типы.

Значением теста может быть произвольный «текст» или упакованное упорядоченное множество (кортеж) значений определенных тестов, а именно:  $\text{text} = \langle \{ \underline{a}/T \} \rangle$ , например,  $\langle \{ \text{term?}a \} \rangle$ , где  $a$  – некоторый символ (слово) фиксированного алфавита; «*term*» – тест. Произвольное изображение или сигнал также можно представить в виде упакованного упорядоченного множества  $O = \langle \{ \underline{a}/T \} \rangle$ , например,  $\langle \{ \text{pixel?}a \} \rangle$ , где  $a$  – цвет единицы изображения, «*pixel*» – тест. Пример:

Сигнал = «001010011100001010101011101».

Важную роль играют операции распаковки значения теста и операция упаковки. *Операцию распаковки значения теста* можно изобразить следующим образом:  $\underline{a}/A \rightarrow \{ \underline{b}/B \}$  (множество или кортеж). *Операция упаковки* противоположна по смыслу:  $\{ \underline{b}/B \} \rightarrow \underline{a}/A$ . Пример распаковки результата теста «Сигнал»: Сигнал  $\rightarrow \langle v?0; v?0; v?1; \dots; v?1 \rangle$ .

При решении локальной задачи из орграфа  $G(\tau)$  может быть выделен произвольный подграф  $G'(\tau)$  с одной базовой вершиной (*операция выделения орграфа*). Препятствием для выделения могут быть лишь онтологические соглашения на маршруте до базовой вершины подграфа. В случае выделения подграфа при наличии таких соглашений может нарушиться логика вычислений. Примеры выделенных орграфов из орграфа  $G(\text{Возраст} \mid \{ \text{Пол} \})$ :

$$G'(\text{Возраст} \mid \{ \text{Пол} \}) = \{ B4 \rightarrow B5; B4 \{ \text{Пол} \} \rightarrow B6 \};$$

$$G''(\text{Возраст}) = \{ B2 \rightarrow B3 \}.$$

Если орграф  $G'(\tau)$  выделен из орграфа  $G(\tau)$ , то будем писать:  $G'(\tau) \in G(\tau)$ . Очевидно:  $G(\tau) \in G(\tau)$ .

Как известно время выхода на пенсию будут постепенно повышать, следовательно, будет меняться и верхняя граница трудоспособного возраста (55 лет, 60 лет и т.д.). В общем случае, любая граница, разделяющая элементы домена, может динамически меняться в зависимости от контекста ситуации. Следовательно, в орграфе  $G(\tau)$  там, где

это целесообразно, вместо чисел нужно использовать параметры  $\{p/P\}$ , которые также являются тестами (со своими орграфами). Итоговая спецификация орграфа имеет вид:

$$G(\tau \{ \{a/A\} \mid \{p/P\} \mid [\Sigma_r] \}) = \{T \rightarrow T\}_\tau \quad (2.5)$$

где  $\{a/A\}$  – внешние тесты для расчета значений теста;  $\{p/P\}$  – структурные параметры;  $\Sigma_r$  – схема деления домена-потомка для базового домена. Параметры могут быть взаимозависимы, что определяется онтологическими соглашениями. Структурные параметры должны быть заданы перед началом использования орграфа. Пример:

*Температура блока химических батарей* ^ ТБХБ:

D2 {Низкая ^Н [0; p1]; Средняя ^С (p1; p2]; Высокая ^В (p2; 50)};

D1 {0; 50}.

$G(\text{ТБХБ} \mid \{p1/P1; p2/P2\}) = \{D1 \rightarrow D2\}$ ,

где  $p1/P1 \equiv p1/[18; 22]$ , а  $p2/P2 \equiv p2/[28; 32]$ . Здесь «ТБХБ» – условное обозначение теста, а «Н», «С», «В» – условные обозначения элементов домена D2. Условные обозначения могут использоваться наравне с исходными названиями. Интервал [0; p1] задает правило:

D1.[0; p1]  $\rightarrow$  D2. {Низкая}.

Аналогично трактуются другие интервалы. Подобная простая нотация позволяет во многих случаях избежать явного задания правил пересчета из одного домена в другой. Конкретизация параметров выглядит следующим образом:  $G(\text{ТБХБ} \mid \{p1/P1? 20; p2/P2? 30\})$ .

Параметры  $\{p/P\}$  произвольного орграфа  $G(\tau)$  могут быть адаптивными и определяться автоматически на основе имеющихся эмпирических данных. В дальнейшем под записью  $G(\tau)$  без ограничения общности будем понимать  $G(\tau \mid \{p/P\})$ .

## 2.2 Каноническое представление орграфа доменов

Непрерывный или конструктивно непрерывный базовый домен орграфа  $G(\tau)$  обозначим  $\{\}$ . Элементы самого точного дискретного домена заменим их порядковыми номерами:  $\{1; 2; 3; \dots; N\}$ . Элементы других доменов – вершин заменим комбинациями номеров  $1, \dots, N$  в соответствии со схемами обобщения. Все домены могут содержать интервалы (в этом проявляется двойственность восприятия элементов домена). Результирующее представление орграфа назовем *каноническим*, а сам *канонический орграф* обозначим  $\underline{G}(\tau)$ . Названия доменов в каноническом орграфе можно оставить прежними, а можно изменить, например, присвоив числовые

коды. Если зависимости от внешних тестов нет, то канонический оргграф строится однозначным образом, так как ОВЗ всех вершин совпадают, и связи между вершинами четко определены. Если есть зависимость от внешних тестов, то канонические оргграфы строятся для всех значений внешних тестов. Вместо числового кода элементы первого дискретного домена можно кодировать произвольными символами (комбинациями символов), например: {a; b; c; d; ...}.

Примеры. Построим каноническое представление для оргграфа  $G_1(\text{Возраст}) = \{B1 \rightarrow B2; B2 \rightarrow B3\}$  с вершинами:

$B1 = [0; 100];$

$B2 = \{\text{юный; молодой; средних лет; пожилой; старческий}\};$

$B3 = \{\text{молодой; средних лет; пожилой}\}.$

Канонический оргграф  $\underline{G}_1(\text{Возраст}) = \{B1 \rightarrow B2; B2 \rightarrow B3\}$  представлен вершинами:

$B1 = \{\}; B2 = \{1; 2; 3; 4; 5\}; B3 = \{1\ 2; 3; 4\ 5\}.$

Канонический оргграф  $\underline{G}^+$  (ЖБОЖ) содержит вершины (рис. 2.6):

$D1 \{1; 2; 3; 4; 5\}; D2 \{1; 2\ 3\ 4\ 5\}; D3 \{2; 1\ 3\ 4\ 5\}; D4 \{3; 1\ 2\ 4\ 5\}; D5 \{4; 1\ 2\ 3\ 5\}; D6 \{5; 1\ 2\ 3\ 4\}.$

Построим каноническое представление оргграфа

$G_2(\text{Возраст} \mid \{\text{Пол}\}) \equiv \{B1 \rightarrow B4; B4 \rightarrow B5; B4 \{\text{Пол}\} \rightarrow B6\}.$

Канонический оргграф  $\underline{G}_2(\text{Возраст} \mid \{\text{Пол?М}\})$  представлен вершинами:

$B1 = \{\}; B4 = \{1; 2; 3; 4; 5; 6\}; B5 = \{1\ 2; 3\ 4; 5\ 6\}; B6 = \{2\ 3\ 4\ 5; 1\ 6\}.$

Канонический оргграф  $\underline{G}_2(\text{Возраст} \mid \{\text{Пол?Ж}\})$  представлен вершинами:

$B1 = \{\}; B4 = \{1; 2; 3; 4; 5; 6\}; B5 = \{1\ 2; 3\ 4; 5\ 6\}; B6 = \{2\ 3\ 4; 1\ 5\ 6\}.$

Как видим структура элементов вершины  $B6$  в зависимости от значения теста «Пол» изменяется.

Пусть в оргграфе доменов  $G(\tau)$  переход от базового непрерывного или конструктивно-непрерывного домена к дискретным доменам-потомкам выполняется корректно, а внешние тесты отсутствуют.

Целесообразность введения канонических представлений оргграфов доменов определяется, в частности, следующим предложением.

**Предложение 2.5.** Оргграф  $G(\tau)$  корректен тогда и только тогда, когда корректен канонический оргграф  $\underline{G}(\tau)$ .

Проверка корректности (правильности) оргграфа доменов включает пять позиций:

1. Все вершины оргграфа должны быть уникальны.
2. Все дискретные вершины должны иметь полный набор значений:  $1, \dots, N$  (смысловые области всех вершин должны совпадать).



3. В каждой связке  $(T \rightarrow T')$  должна соблюдаться преемственность элементов.
4. В каждой связке  $(T \rightarrow T')$  должно выполняться условие:  $|T'| < |T|$ .
5. Не должно быть конфликтов при разных путях обобщения.

Приведем пример нарушения преемственности элементов:

$$T \{2; 1\ 3; 4\ 5\} \rightarrow T' \{2\ 1; 3\ 4\ 5\}.$$

Заметим, что в данном примере все остальные условия корректности орграфа соблюдаются. Корректность перехода от непрерывного домена к дискретному принимается как постулат.

Канонический орграф позволяет убедиться в выполнении всех позиций, что и доказывает данное предложение. Собственно в максимальном упрощении проверки корректности орграфа и состоит главный смысл введения структуры «канонический орграф». Другое назначение состоит в упрощении анализа структурной энтропии орграфа доменов, но об этом речь пойдет в разделе 2.4.

Корректность перехода от непрерывного домена к дискретному можно определить следующим образом. Пусть  $A$  – ОВЗ (область возможных значений), определяемая базовым непрерывным или конструктивно-непрерывным доменом  $T$ , а  $B_i$  ( $i = 1, \dots, N$ ) – множество интервалов (элементов) домена-потомка  $T'$ , тогда корректность перехода  $(T \rightarrow T')$  означает выполнение следующих двух условий:

$$1. A = \cup_{i=1, N} B_i; \quad 2. \cap_{i=1, N} B_i = \emptyset.$$

**Предложение 2.6.** Корректность орграфа  $G(\tau | \{a/A\})$  полностью определяется корректностью всех канонических орграфов из множества  $\{\underline{G}(\tau | \{a/A\})\}$  при условии корректности перехода от непрерывного домена к доменам-потомкам.

Обоснованность утверждения вытекает из предложения 2.5. Пример множества канонических орграфов:

$$\{\underline{G}(\tau | \{a/A\})\} = \{\underline{G}_2(\text{Возраст} | \{\text{Пол?М}\}); \underline{G}_2(\text{Возраст} | \{\text{Пол?Ж}\})\}.$$

В ряде случаев множество  $\{\underline{G}(\tau | \{a/A\})\}$  может состоять из одного орграфа. Приведем пример соответствующего орграфа доменов:

*Индекс массы миокарда левого желудочка* ^ИММЛЖ:

D3 {Норма ^N;

Увеличение ИММЛЖ ^a b c}

D2 {Норма

Умеренное увеличение ИММЛЖ

Значительное увеличение ИММЛЖ

Резкое увеличение ИММЛЖ

D1 {[30; 400]}

^N M[30; 124] Ж[30; 109];

^a M[125; 149] Ж[110; 134];

^b M[150; 199] Ж[135; 184];

^c M[200; 400] Ж[185; 400]}

Граф  $G(\text{ИММЛЖ} \mid \{\text{Пол}\}) = \{D1 \{\text{Пол}\} \rightarrow D2; D2 \rightarrow D3\}$ .

Единственный канонический орграф  $\underline{G}(\text{ИММЛЖ} \mid \{\text{Пол? } M \underline{\vee} \text{Ж}\})$  имеет вид:  $D1 \{\} \rightarrow D2 \{N; a; b; c\} \rightarrow D3 \{N; a b c\}$ .

С помощью канонических орграфов удобно моделировать структурно-завершенные орграфы. Пусть дан канонический орграф теста «Тест\_1»:

Тест\_1 {2 {a; b d; c} 1 {a; b; d; c}}

$G(\text{Тест}_1) = \{1 \rightarrow 2\}$ .

На рис 2.8 представлены два структурно-завершенных орграфа из множества  $\{G^+(\text{Тест}_1)\}$ . Коричневым (темным) цветом обозначены вершины канонического орграфа (они же – ствол орграфа), остальные вершины-листья генерируются автоматически. Блуждающие вершины обозначены разными оттенками желтого цвета (вершины ‘c’, ‘a’).



Рис. 2.8 – Структурно-завершенные орграфы на основе канонического орграфа

Изобразим графически один из структурно-завершенных орграфов  $G^+(\text{Возраст} \mid \{\text{Пол? } M\})$ . Каноническое представление орграфа  $\underline{G}(\text{Возраст} \mid \{\text{Пол? } M\})$  имеет вид:

Возраст\_M {

B6 # B4 {b c d e; a h}

B5 {a b; c d; e h}

B4 # B1 {a; b; c; d; e; h}

B3 {1 2; 3; 4 5}

B2 {1; 2; 3; 4; 5}

B1 {}}

Запись «B6 # B4» означает связку (B4 → B6), а запись «B4 # B1» означает связку (B1 → B4). Для остальных вершин наследование идет снизу вверх. На рис. 2.9 показан пример структурно-завершенного графа  $\underline{G}^+(\text{Возраст} \mid \{\text{Пол? } M\})$ . Всего таких орграфов три. Два отличаются друг от друга расположением единственной блуждающей вершины с кодом «3» (изображена желтым цветом). Третий – предельный орграф (блуждающая

вершина связана с B2 и B3). Красным цветом изображена базовая вершина (красным цветом маркируется непрерывный или конструктивно-непрерывный тип базовой вершины).

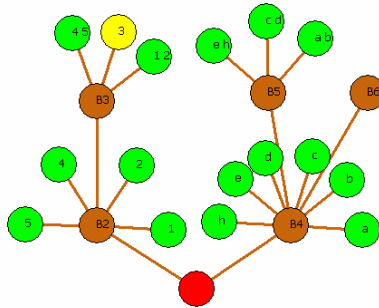


Рис. 2.9 – Структурно-завершенный орграф для теста «Возраст»

Все вершины канонического представления могут содержать интервалы, отражающие двойственность представления. Пример:

$v$  - Скорость:

V5 {a c [0; 3,0]  $\cup$  (7,0; 10,0]; b (3,0; 7,0)};

V4 {a b [0; 7,0]; c (7,0; 10,0)};

V3 {a [0; 3,0]; b c (3,0; 10,0)};

V2 {a [0; 3,0]; b (3,0; 7,0]; c (7,0; 10,0)};

V1 {[0; 10,0]}.

Орграф  $\underline{G}$ (Скорость) = {V1  $\rightarrow$  V2; V2  $\rightarrow$  V3; V2  $\rightarrow$  V4; V2  $\rightarrow$  V5}.

### 2.3 Замыкание, верхний и нижний пределы множества данных

Введем в рассмотрение функцию  $Test$ , которая произвольному множеству значений  $\{\underline{g}/T\}$  ставит в соответствие множество используемых тестов  $\{\tau\}$ . Можем записать:  $Test(\{\underline{g}/T\}) = \{\tau\}$ .

Примеры:

$Test(\{\text{ФИО}/C? \text{ Петров И.И.}; \text{Пол}/C? \text{ М}; \text{Возраст}/N? \text{ 34}; \text{АДс}/N? \text{ 145}; \text{АДд}/N? \text{ 87}\}) = \{\text{ФИО}; \text{Пол}; \text{Возраст}; \text{АДс}; \text{АДд}\}$ ;

$Test(\{\text{Возраст}/B1? \text{ 33}; \text{Возраст}/B4? \text{ [31; 40]}\}) = \{\text{Возраст}\}$ .

Введем операцию  $\{\underline{g}/T\} \nabla \{G(\tau)\}$ , означающую, что если  $\tau \in Test(\{\underline{g}/T\})$ , то  $G(\tau) \in \{G(\tau)\}$ . Другими словами, все результаты тестов интерпретируемы в рамках имеющейся онтологии.

Произвольное множество  $\{\underline{g}/T\}$  назовем *изначально неконфликтным*, если в рамках любого орграфа  $G(\tau): \tau \in Test(\{\underline{g}/T\})$  соответствующее ему

подмножество  $\{\underline{t}/T\}$  неконфликтно, т.е. в результате всех обобщений не возникает конфликта ни в одном из доменов орграфа  $G(\tau)$ .

В рамках орграфа  $G(\text{Возраст})$  множество  $\{\text{Возраст}/V1? 33; \text{Возраст}/V5? \text{молодой}\}$  изначально конфликтно, так как при обобщении значения «Возраст/ V1? 33» появляется значение «Возраст/ V5? средних лет», что приводит к конфликту.

Любой орграф  $G(\tau)$  может содержать онтологические соглашения. Для иллюстрации таких соглашений рассмотрим равномерное и прямолинейное движение. Пусть заданы орграфы трех тестов «Скорости», «Пути» и «Времени»:

$v$  - Скорость:

$V2 \{\text{Маленькая} \wedge a [0; 3,0]; \text{Средняя} \wedge b (3,0; 7,0); \text{Большая} \wedge c (7,0; 10,0)\}$

$V1 \{[0; 10,0] (s/S1/ t/T1)\}$ .

Граф  $G(\text{Скорость}) = \{V1 \rightarrow V2\}$ .

$s$  - Путь:

$S2 \{\text{Маленький} \wedge a [0;15,0]; \text{Средний} \wedge b (15,0; 40,0); \text{Большой} \wedge c (40,0;50,0)\}$

$S1 \{[0; 50,0] (v/V1 * t/T1)\}$ .

Граф  $G(\text{Путь}) = \{S1 \rightarrow S2\}$ .

$t$  - Время:

$T2 \{\text{Маленькое} \wedge a [0; 1,5]; \text{Среднее} \wedge b (1,5; 3,5); \text{Большое} \wedge c (3,5; 5,0)\}$

$T1 \{[0; 5,0] (s/S1/ v/V1)\}$ .

Граф  $G(\text{Время}) = \{T1 \rightarrow T2\}$ .

Условные обозначения и числовые интервалы в доменах  $V2$ ,  $S2$ ,  $T2$  задают однозначные правила перехода от базовых доменов. Запись « $s/S1/ v/V1$ » означает операцию деления, а запись « $v/V1 * t/T1$ » – операцию умножения.

Все орграфы данного примера содержат онтологические соглашения в базовом домене, позволяющие вычислить значения любого теста, зная результаты двух других. Онтологические соглашения в общем случае могут быть представлены в виде системопаттернов. Пример:

$f/\mu_f: s/S1, t/T1 \rightarrow v/V1; \quad h/\mu_h: v/V1, t/T1 \rightarrow s/S1; \quad g/\mu_g: s/S1, v/V1 \rightarrow t/T1,$

где  $\mu_f$ ;  $\mu_h$ ;  $\mu_g$  – механизмы реализации системопаттернов.

Произвольную совокупность орграфов  $\{G(\tau)\}$  назовем *онтологически согласованной* (в контексте конкретной задачи или предметной области), если при любых реальных изначально неконфликтных исходных данных  $\{\underline{t}/T\} \nabla \{G(\tau)\}$ , описанных в терминах базовых доменов, в результате применения всех онтологических соглашений, не выходящих за рамки  $\{G(\tau)\}$ , не возникает конфликта значений тестов (в рамках одного домена не появляются разные значения теста и не нарушаются ОВЗ). Реальность данных означает, что они удовлетворяют соотношениям модели

фундаментальных знаний (модели метазнаний). Подробнее о моделях знаний речь пойдет в главах 3 и 7.

Как легко убедиться множество  $\{G(\text{Скорость}), G(\text{Путь}), G(\text{Время})\}$  онтологически согласовано (в контексте равномерного и прямолинейного движения). Любые реальные данные, например –  $\{s/S1? 10; v/V1? 5; t/T1? 2\}$ , не приводят к конфликтам. Согласованными являются также ОВЗ всех тестов.

Введем в рассмотрение оргграф  $G(\text{Время}_2) = \{T1 \rightarrow T2\}$  следующим образом:

$t$  -  $\text{Время}_2$ :

$T2$  {Маленькое  $\wedge a$  [0; 3]; Среднее  $\wedge b$  (3; 7]; Большое  $\wedge c$  (7; 10]}

$T1$  {[0; 10,0] ( $s/S1/ v/V1$ )}.

Множество  $\{G'(\text{Скорость}), G(\text{Путь}), G(\text{Время}_2)\}$  онтологически несогласованно. Действительно, исходные изначально неконфликтные данные  $\{v/V1? 10; t/T1? 10\}$  приводят к нарушению ОВЗ базового домена теста «Путь».

Любой отдельный оргграф является онтологически согласованным по определению. Доказательство онтологической согласованности множества оргграфов  $\{G(\tau \mid \{a/A\} \mid \{p/P\} \mid \{\Sigma_d\})\}$  может оказаться нетривиальной задачей.

Если  $\{G(\tau)\}$  онтологически согласовано, то множество структурно-завершенных графов  $\{(G^{\sim}(\tau))^+ \mid G(\tau) \in \{G(\tau)\}\}$ , как правило, также онтологически согласовано.

Множество стволов оргграфов  $\{G^{\sim}(\tau) \mid G(\tau) \in \{G(\tau)\}\}$  всегда онтологически согласовано, если согласовано множество  $\{G(\tau)\}$ . Действительно, пусть  $\{\underline{t}/T\}$  – произвольное множество начальных данных, о которых идет речь в определении онтологической согласованности, а  $\{\underline{t}/T\}'$  – множество всех результирующих данных, которые получены с помощью всех онтологических соглашений на множестве графов  $\{G(\tau)\}$ . Множество  $\{\underline{t}/T\}'$  не содержит конфликтов по условию, следовательно, любое его подмножество также не содержит конфликтов. Это и доказывает онтологическую согласованность  $\{G^{\sim}(\tau) \mid G(\tau) \in \{G(\tau)\}\}$ .

В дальнейшем по умолчанию будем считать  $\{G(\tau)\}$  онтологически согласованным. Отметим, что онтологическую согласованность множества оргграфов доменов тестов может нарушить конфликтная база знаний, принимающая участие в вычислении значений тестов.

*Замыкаем* произвольного множества значений  $\{\underline{t}/T\}$  над оргграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{t}/T\})\}$  назовем множество  $\{\underline{t}/T\}^+$ , содержащее все вычисляемые значения тестов на основе оргграфов и исходного множества  $\{\underline{t}/T\}$ .

Множество  $\{\underline{t}/T\}$  назовем *неконфликтным* над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{t}/T\})\}$ , если замыкание  $\{\underline{t}/T\}^+$  над данными орграфами не содержит конфликтов значений тестов (нет нарушений ОВЗ тестов и нет конфликтных значений в рамках любого домена).

Неконфликтное множество данных не может содержать более одного значения теста в рамках любого домена. Однако это не означает, что нельзя строить замыкания для произвольных множеств  $\{\underline{t}/T\}$ .

Изначально неконфликтное множество  $\{\underline{t}/T\}$  не всегда является неконфликтным. Изначально конфликтное множество остается конфликтным всегда. Пример конфликтного множества:  $\{\underline{t}_1; \underline{t}_2\}/T$ .

Изменение состава множества графов  $\{G(\tau)\}$  может изменить характеристику конфликтности  $\{\underline{t}/T\}$  на противоположную.

*Верхним пределом* множества  $\{\underline{t}/T\}$  над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{t}/T\})\}$  назовем множество  $\{\underline{t}/T\}^{\wedge}$ , являющееся подмножеством замыкания  $\{\underline{t}/T\}^+$  и состоящее из результатов тестов максимального уровня общности (значения принадлежат доменам, образующим терминальные вершины).

*Нижним пределом* множества  $\{\underline{t}/T\}$  над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{t}/T\})\}$  назовем множество  $\{\underline{t}/T\}^{-}$ , являющееся минимальным подмножеством  $\{\underline{t}/T\}^+$ , замыкание которого содержит  $\{\underline{t}/T\}$ , т.е.  $\{\underline{t}/T\} \subset (\{\underline{t}/T\}^{-})^+$ .

Все множества  $\{\underline{t}/T\}$ ,  $\{\underline{t}/T\}^{-}$ ,  $\{\underline{t}/T\}^{\wedge}$  принадлежат  $\{\underline{t}/T\}^+$ . Очевидно, выполняется соотношение:  $\{\underline{t}/T\}^+ = (\{\underline{t}/T\}^{-})^+$  над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{t}/T\})\}$ .

Подчеркнем, что введенные выше операции замыкания, получения верхнего и нижнего пределов целиком опираются на онтологические описания и не используют возможности базы знаний. Данные операции следует отнести к автоматизмам вычислительной среды.

Во многих случаях для принятия решения достаточно информации  $\{\underline{t}/T\}^{\wedge}$ , т.е. информации максимального уровня общности. Напротив, при хранении прецедентов в базе данных достаточно хранить нижние пределы множеств результатов тестов. При извлечении прецедента из памяти с помощью автоматизмов среды возможно генерирование полного описания (замыкания множества всех имеющихся данных).

Примеры. Пусть имеется оргграф  $G(\text{Возраст})$ , представленный на рис. 2.2. Легко убедиться в справедливости следующих результатов:

$\{\text{Возраст}/\text{В1? 33}\}^+$  содержит 6 элементов;  
 $\{\text{Возраст}/\text{В2? молодой}\}^+ = \{\text{Возраст}/\text{В2? молодой}; \text{Возраст}/\text{В3? молодой}\}$   
 $\{\text{Возраст}/\text{В1? 33}\}^{\wedge} = \{\text{Возраст}/\text{В5? средних лет}; \text{Возраст}/\text{В3? молодой};$   
 $\text{Возраст}/\text{В6? трудоспособный}\};$

{Возраст/ В1? 33; Возраст/ В5? средних лет}^- = {Возраст/ В1? 33}.

Справедливо следующее утверждение.

**Предложение 2.7.** Пусть задано  $\{\underline{d}/T\}$ , тогда замыкание и верхний предел над орграфами  $\{G(\tau) \mid \tau \in Test(\{\underline{d}/T\})\}$  определяются единственным образом, а нижний предел, в общем случае, определяется не единственным образом.

Единственность замыкания доказывается от противного. Предположим, что существуют два разных замыкания  $(\{\underline{d}/T\}^+)_1$  и  $(\{\underline{d}/T\}^+)_2$ , тогда должны существовать непустые множества

$$A = (\{\underline{d}/T\}^+)_1 \setminus (\{\underline{d}/T\}^+)_2 \quad \text{и} \quad B = (\{\underline{d}/T\}^+)_2 \setminus (\{\underline{d}/T\}^+)_1,$$

иначе одно замыкание будет подмножеством другого, что противоречит определению замыкания. Так как  $\{\underline{d}/T\} \subset (\{\underline{d}/T\}^+)_1 \cap (\{\underline{d}/T\}^+)_2$ , то получается, что из  $\{\underline{d}/T\}$  одновременно и можно, и нельзя получить элементы множеств  $A$  и  $B$ . Пришли к противоречию, что доказывает единственность замыкания. Поскольку замыкание определяется единственным образом, то его подмножество – элементы терминальных вершин – определяется единственным образом, следовательно, верхний предел определяется единственным образом.

Для доказательства не единственности определения нижнего предела приведем пример.

Пусть  $\{\underline{d}/T\} = \{s/S1? 10; v/V1? 5; t/T1? 2\}$ . Как легко убедиться, множество  $\{\underline{d}/T\}$  неконфликтно над орграфами  $\{G(\tau) \mid \tau \in Test(\{\underline{d}/T\})\}$  (замыкание не содержит конфликтных значений, так как в базовых доменах выполняется условие:  $\underline{s} = \underline{v} * \underline{t}$ ). Определим нижний предел  $\{\underline{d}/T\}$  над орграфами  $\{G(\tau) \mid \tau \in Test(\{\underline{d}/T\})\} = \{G(\text{Скорость}); G(\text{Путь}); G(\text{Время})\}$ . Имеем три варианта нижних пределов:

$$\{\underline{d}/T\}^- = \{s/S1? 10; v/V1? 5\}; \quad \{\underline{d}/T\}^- = \{v/V1? 5; t/T1? 2\};$$

$$\{\underline{d}/T\}^- = \{s/S1? 10; t/T1? 2\}.$$

Любой из представленных вариантов позволяет восстановить исходное множество результатов тестов (благодаря онтологическим соглашениям). Данный пример показывает, что нижний предел определяется в ряде случаев неоднозначно.

Рассмотрим еще примеры над орграфами  $\{G(\text{Скорость}); G(\text{Путь}); G(\text{Время})\}$ .

Если  $\{\underline{d}/T\} = \{s/S1? 10; v/V1? 5\}$ , то получим единственный вариант нижнего предела над орграфами  $\{G(\tau) \mid \tau \in Test(\{\underline{d}/T\})\} = \{G(\text{Скорость}); G(\text{Путь})\}$ , который совпадает с исходными данными.

Множество  $\{\underline{d}/T\} = \{s/S1? 10; v/V1? 5; t/T2? \text{Среднее}\}$  неконфликтно. Легко убедиться, что  $\{\underline{d}/T\}^- = \{s/S1? 10; v/V1? 5\}$  и оно единственно.

Множество  $\{\underline{t}/T\} = \{s/S1? 10; v/V1? 5; t/T2? \text{Большое}\}$  конфликтно, так как при замыкании появляется, в частности, значение « $t/T2? \text{Среднее}$ », что приводит к конфликту.

Множество  $\{\underline{t}/T\} = \{s/S1? 10; v/V1? 5; t/T1? 4\}$  конфликтно, так как при замыкании появляется, в частности, значение « $s/S1? 20$ », что приводит к конфликту.

*Банком тестов* будем называть произвольное онтологически согласованное множество орграфов  $\{G(\tau)\}$ . Банк тестов является частью онтологии ПрО.

Пусть фиксирован произвольный банк тестов  $\{G(\tau)\}$ .

*Замыканием* произвольного множества значений  $\{\underline{t}/T\}$  над банком тестов  $\{G(\tau)\}$  назовем множество  $\{\underline{t}/T\}^+$ , содержащее все вычисляемые значения тестов на основе  $\{G(\tau)\}$  и исходного множества  $\{\underline{t}/T\}$ .

Множество  $\{\underline{t}/T\}$  назовем *неконфликтным* над банком тестов  $\{G(\tau)\}$ , если замыкание  $\{\underline{t}/T\}^+$  над данными орграфами не содержит конфликтов значений тестов. В силу того, что банк тестов  $\{G(\tau)\}$  онтологически согласован, реальные данные  $\{\underline{t}/T\}$  неконфликтны.

*Верхним пределом* множества  $\{\underline{t}/T\}$  над банком тестов  $\{G(\tau)\}$  назовем множество  $\{\underline{t}/T\}^{\wedge}$ , являющееся подмножеством замыкания  $\{\underline{t}/T\}^+$  и состоящее из результатов тестов максимального уровня общности (значения принадлежат доменам, образующим терминальные вершины графов из  $\{G(\tau)\}$ ).

*Нижним пределом* множества  $\{\underline{t}/T\}$  над банком тестов  $\{G(\tau)\}$  назовем множество  $\{\underline{t}/T\}^{\vee}$ , являющееся минимальным подмножеством  $\{\underline{t}/T\}^+$ , замыкание которого содержит  $\{\underline{t}/T\}$ , т.е.  $\{\underline{t}/T\} \subset (\{\underline{t}/T\}^{\vee})^+$ .

Мы сохранили обозначения для замыкания, верхнего и нижнего пределов, так как они являются, по сути, обобщениями первых.

Проиллюстрируем отличия определений замыкания и верхнего предела на следующем примере. Пусть банк тестов содержит орграфы двух тестов. Тест первый:

*Индекс левого предсердия*  $\wedge$  иЛП:

D3 {Норма  $\wedge$  N;

Увеличение иЛП  $\wedge$  a b}

D2 {Норма

$\wedge$  N [1,40; 2,10];

Умеренное увеличение иЛП

$\wedge$  a (2,10; 2,50);

Значительное увеличение иЛП

$\wedge$  b (2,50; 5,0) }

D1 {[1,40; 5,0]}.

Граф  $G(\text{иЛП}) = \{D1 \rightarrow D2; D2 \rightarrow D3\}$ ,

где D1, D2, D3 – домены; иЛП – условное обозначение теста (признак условного обозначения – знак ‘ $\wedge$ ’); N, a, b – условные обозначения



элементов доменов D2 и D3 (они же задают однозначные правила пересчета значений из домена D2 в домен D3); интервалы вида [1,40; 2,10] задают однозначные правила пересчета значений из домена D1 в домен D2.

Тест второй:

*Дилатация полости левого предсердия* (ЛП)  $\wedge$  ДПЛП:

D2 {Признаков дилатации ЛП нет  $\wedge$  N};

Признаки дилатации ЛП есть  $\wedge$  a b}

D1 {Признаков дилатации полостей нет  $\wedge$  N (иЛП/D2? N) ;

Умеренная дилатация  $\wedge$  a (иЛП/D2? a);

Значительная дилатация  $\wedge$  b (иЛП/D2? b) }.

Граф  $G(\text{ДПЛП}) = \{D1 \rightarrow D2\}$ .

Дилатация (dilatation) – увеличение или расширение какого-либо полого органа. Особенность второго теста в том, что элементы домена D1 могут быть вычислены на основе результатов первого теста.

Пусть множество исходных данных состоит из результата первого теста:  $\{\underline{a}/T\} = \{\text{иЛП/D1? } 4,0\}$ . Определим замыкания и верхние пределы согласно двум определениям.

Так как  $\text{Test}(\{\text{иЛП/D1? } 4,0\}) = \{\text{иЛП}\}$ , то замыкание и верхний предел над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{a}/T\})\}$  будут иметь следующий вид:

$\{\text{иЛП/D1? } 4,0\}^+ = \{\text{иЛП/D1? } 4,0; \text{иЛП/D2? } b; \text{иЛП/D3? } a\} \equiv \{\text{иЛП/D1? } 4,0;$

$\text{иЛП/D2? } \text{Значительное увеличение иЛП}; \text{иЛП/D3? } \text{Увеличение иЛП}\}$ ;

$\{\text{иЛП/D1? } 4,0\}^\wedge = \{\text{иЛП/D3? } a\} \equiv \{\text{иЛП/D3? } \text{Увеличение иЛП}\}$ .

Построим замыкание и верхний предел над банком тестов  $\{G(\tau)\} = \{G(\text{иЛП}), G(\text{ДПЛП})\}$ :

$\{\text{иЛП/D1? } 4,0\}^+ = \{\text{иЛП/D1? } 4,0; \text{иЛП/D2? } b; \text{иЛП/D3? } a; \text{ДПЛП/D1? } b;$

$\text{ДПЛП/D2? } a\}$ ;  $\{\text{иЛП/D1? } 4,0\}^\wedge = \{\text{иЛП/D3? } a; \text{ДПЛП/D2? } a\}$ .

Различия очевидны.

Отличия нижних пределов можно продемонстрировать над банком тестов  $\{G(\text{Скорость}); G(\text{Путь}); G(\text{Время})\}$ . Если  $\{\underline{a}/T\} = \{s/S1? 10; v/V1? 5\}$ , то получим единственный вариант нижнего предела над орграфами  $\{G(\tau) \mid \tau \in \text{Test}(\{\underline{a}/T\})\} = \{G(\text{Скорость}); G(\text{Путь})\}$ , который совпадает с исходными данными. Над банком тестов получим три варианта нижних пределов:  $\{\underline{a}/T\}^- = \{s/S1? 10; v/V1? 5\}$ ;  $\{\underline{a}/T\}^- = \{v/V1? 5; t/T1? 2\}$ ;  $\{\underline{a}/T\}^- = \{s/S1? 10; t/T1? 2\}$ .

В общем случае может быть определено замыкание, верхний и нижний пределы над моделью предметной области, которая включает банк тестов (онтологию) и модель знаний.

Пусть даны два банка тестов  $\{G(\tau)\}_1$  и  $\{G(\tau)\}_2$ . Будем писать:  $\{G(\tau)\}_1 \subset \{G(\tau)\}_2$  тогда и только тогда, когда  $\forall \tau, (G(\tau))_1 \in (G(\tau))_2$

(является подграфом), где  $(G(\tau))_1 \in \{G(\tau)\}_1$ , а  $(G(\tau))_2 \in \{G(\tau)\}_2$ . Принадлежность одного графа другому определяется с точностью до обозначений, а точнее – на уровне канонических представлений каждого орграфа. К примеру, можно положить:  $\{G(\tau)\}_2 = \{(G^-(\tau))^{++}\}_1$ .

На практике, зачастую, поступают еще проще, а именно: если граф  $G(\tau)$  линейный, то расщепляют только терминальную вершину. Рассмотрим пример на основе орграфа  $G(\text{Скорость}) \equiv (G(\tau))_1$ . Зададим орграф  $G(\text{Скорость}) \equiv (G(\tau))_2$  путем расщепления вершины  $V2$ :

$v$  - Скорость:

$V5$  {Маленькая или Средняя  $\wedge$  a b; Большая  $\wedge$  c}

$V4$  {Маленькая или Большая  $\wedge$  a c; Средняя  $\wedge$  b}

$V3$  {Маленькая  $\wedge$  a; Средняя или Большая  $\wedge$  b c}

$V2$  {Маленькая  $\wedge$  a [0; 3,0]; Средняя  $\wedge$  b (3,0; 7,0]; Большая  $\wedge$  c (7,0; 10,0]}

$V1$  {[0; 10,0] (s/S1/ t/T1)}.

Граф  $(G(\text{Скорость}))_2 = \{V1 \rightarrow V2 \rightarrow V3; V2 \rightarrow V4; V2 \rightarrow V5\}$ .

По построению  $(G(\tau))_1 \in (G(\tau))_2$ .

Построим орграф  $G(\text{Скорость}) \equiv (G(\tau))_3$ , который семантически мощнее орграфа  $(G(\tau))_1$ , но, вместе с тем,  $(G(\tau))_1$  не является его частью:

$v$  - Скорость:

$V2$  {Маленькая  $\wedge$  1 2; Средняя  $\wedge$  3; Большая  $\wedge$  4 5}

$V1\_2$  {Очень маленькая  $\wedge$  1 [0; 1,5]; Маленькая  $\wedge$  2 (1,5; 3,0]; Средняя  $\wedge$  3 (3,0; 7,0]; Большая  $\wedge$  4 (7,0; 8,5]; Очень большая  $\wedge$  5 (8,5; 10,0]}

$V1$  {[0; 10,0] (s/S1/ t/T1)}.

Граф  $(G(\text{Скорость}))_3 = \{V1 \rightarrow V1\_2 \rightarrow V2\}$ .

Орграф  $(G(\tau))_3$  отличается от орграфа  $(G(\tau))_1$  только наличием промежуточной вершины  $V1\_2$ . Остальные вершины совпадают (по ОВЗ).

Справедливо утверждение.

**Предложение 2.8** (о монотонности замыкания). Пусть дано  $\{\underline{g}/T\}$  и два банка тестов  $\{G(\tau)\}_1$  и  $\{G(\tau)\}_2$ , причем  $\{\underline{g}/T\} \nabla \{G(\tau)\}_1$ ,  $\{\underline{g}/T\} \nabla \{G(\tau)\}_2$  и  $\{G(\tau)\}_1 \subset \{G(\tau)\}_2$ , тогда справедливо соотношение:  $\{\underline{g}/T\}_1^+ \subseteq \{\underline{g}/T\}_2^+$ .

Суть предложения состоит в том, что развитие структур всех или отдельных орграфов приводит к росту объема заключений, а, следовательно, и мощности вычислительного интеллекта в целом.

Обоснованность предложения достаточно очевидна. Единственное, на что нужно обратить внимание, так это на то, что о характере конфликтности замыканий речь не идет.

В основе введенных выше понятий «изначальной неконфликтности» данных, «онтологической согласованности» орграфов и «неконфликтности (замыкания)» лежит требование невозможности появления двух значений

теста в рамках одного домена. Такую трактовку неконфликтности отнесем к *сильной неконфликтности*. В основе ряда важных автоматизмов среды могут быть операции, включающие произвольные домены разных тестов, например, интервальные операции сложения, умножения и т.д. Подобные операции приведут к нарушению требования сильной неконфликтности практически при любых исходных данных. Выход заключается в формулировке *слабой неконфликтности*, допускающей работу с интервальными данными вида  $\{\underline{\delta}/T\}$ , где  $\delta$  – интервалы.

#### 2.4 Нисходящая детерминация. Субоптимизация

Пусть задано произвольное значение теста  $\underline{\tau}/T$ , которое интерпретируется в рамках орграфа  $G(\tau)$ . *Нисходящая детерминация* задает неконфликтное множество значений  $\{\underline{\tau}/T\}$ , где  $\{T\}$  – все предшествующие  $T$  домены в рамках орграфа  $G(\tau)$ , вплоть до базового домена  $T_0$ . Очевидно, множество  $\{\underline{\tau}/T\}$  определяется не единственным образом.

Примеры. Пусть задано значение теста «Возраст/ В3? молодой» в рамках орграфа  $G_1(\text{Возраст})$ . Приведем несколько вариантов нисходящей детерминации:  $\{\underline{\tau}/T\} = \{\text{Возраст/ В1? 5; Возраст/ В2? юный}\}$ ;  $\{\underline{\tau}/T\} = \{\text{Возраст/ В1? 30; Возраст/ В2? молодой}\}$ .

Пусть задано значение теста «иЛП/Д3? Увеличение иЛП» в рамках орграфа  $G(\text{иЛП}) = \{D1 \rightarrow D2 \rightarrow D3\}$ . Вариант нисходящей детерминации:  $\{\underline{\tau}/T\} = \{\text{иЛП/Д1? 4,0; иЛП/Д2? b}\}$ .

В конкретных задачах процедура нисходящей детерминации может подчиняться дополнительным регламентирующим правилам (закономерностям), например, возможна нейросетевая реализация процедуры с обучением. Если используются элементы теории нечетких множеств (нечеткой логики), то нисходящая детерминация включает в себя дефаззификацию (на определенных уровнях). В целом, нисходящая детерминация относится к врожденным или приобретенным *автоматизмам* (вычислительной) среды.

Нисходящая детерминация играет важную роль в задачах управления, обеспечивая функционирование механизма субоптимизации.

Под *субоптимизацией* будем понимать оптимизацию на максимально высоком (предельном) уровне общности описания ситуации действительности с учетом возможностей нисходящей детерминации и критерия удовлетворительности решения (в целом).

Значения тестов на других уровнях общности определяются благодаря восходящим и нисходящим преобразованиям (автоматизмам среды). Восходящие преобразования позволяют получить обобщенную модель

управления, что обеспечивает эффективный мониторинг процесса и реализацию схем изменения стратегий управления (в случае необходимости).

Субоптимизация способна обеспечить предельно высокое быстродействие систем управления и при этом минимизировать энергетические затраты на выработку управления путем максимального подключения автоматизмов среды как в процессе восходящих преобразований (обобщения), так и в процессе нисходящей детерминации. Основным инструментом субоптимизации является комплекс взаимосвязанных методов, реализующих принцип предельных обобщений.

В самом общем случае в реализации восходящих преобразований и нисходящей детерминации ведущую роль играют прямолинейная (восходящая), вращательная (вихревая) и колебательная динамика (флуктуации) как внутри доменов, так и между доменами орграфов тестов.

Важным фактом следует считать принципиальную многозначность и неустойчивость значений тестов при нисходящей детерминации. Флуктуации стабилизируемых отрицательными обратными связями параметров порождают колебательные процессы с амплитудой либо затухающей, либо постоянной (при некоторой частоте и амплитуде колебаний, соответствующих определенной энергии детерминации), либо возрастающей до пределов устойчивости системы с ростом энергии. Снижение энергии детерминации уменьшает область спонтанно занимаемых состояний, уменьшает энтропию, снижает вероятность разрушения управления из-за внутренних флуктуаций. Уровней «флуктуации» столько же, сколько уровней элементов в иерархии систем (например, доменов в орграфе доменов), поэтому относительно каждого конкретного уровня можно говорить о существенных (различимых) для него *макрофлуктуациях* и о *микрофлуктуациях*, определяющих судьбу макрофлуктуаций на отрезках времени и при порядках величин, где макрофлуктуации могут рассматриваться как вполне детерминированные движения.

Рассмотрим примеры нисходящей детерминации в варианте, когда она реализована с помощью функций принадлежности и нечеткой логики.

Зададим тест «Прикладываемая сила» следующим образом:

F - Прикладываемая сила:

D3 {N ( $\mu_N$ ); Ze ( $\mu_{Ze}$ ); P ( $\mu_P$ )}

D2 {Nb ( $\mu_{Nb}$ ); Nm ( $\mu_{Nm}$ ); Ns ( $\mu_{Ns}$ ); Ze ( $\mu_{Ze}$ ); Ps ( $\mu_{Ps}$ ); Pm ( $\mu_{Pm}$ ); Pb ( $\mu_{Pb}$ )}

D1 {[-8; 8]}

Орграф  $G(F) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Связь между элементами домена D1 и домена D2 показана на рис. 2.10.

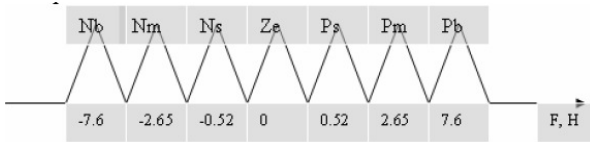


Рис. 2.10 – Переменные и функции принадлежности для «Силы»

Как следует из рисунка и спецификации доменов, каждый элемент домена D2 задается функцией принадлежности определенного (треугольного) вида. Следует отметить, что при нечетком подходе вид функции принадлежности играет важную роль в реализации процедуры нисходящей детерминации и определение вида данной функции – это отдельная непростая (оптимизационная) задача [173, 174].

Пример правила преобразования из домена D1 в домен D2:

D1.[-8; -5] → D2.{Nb ( $\mu_{Nb}$ )}.

Связь между элементами домена D2 и домена D3 зададим следующим образом:

D2.{Nb, Nm, Ns} → D3.{N ( $\mu_N$ )}; D2 {Ze} → D3.{Ze ( $\mu_{Ze}$ )};

D2 {Ps, Pm, Pb} → D3.{P ( $\mu_P$ )}.

Подчеркнем, что нечеткость в модели орграфа доменов играет роль *исключительно* при нисходящей детерминации, определенным образом влияя на выбор одного из нескольких значений при конкретизации (например, дефазсификации). При восходящем преобразовании (от частного к общему) все процессы строго детерминированы и однозначны.

На основе орграфа  $G(F)$  могут быть автоматически построены структурно-завершенные орграфы  $G^+(F)$ . Каноническое представление орграфа  $G(F)$  имеет вид:

Тест «Сила» - F:

D3 {1 2 3; Ze; 5 6 7}

D2 {1; 2; 3; Ze; 5; 6; 7}

D1 {}

$\underline{G}(F) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Так как блуждающая вершина одна (порождается элементом Ze), то имеются всего три структурно-завершенных орграфа. Два орграфа показаны на рис 2.11. Третий – предельный структурно-завершенный орграф  $G^{++}(F)$  образуется путем соединения блуждающей вершины Ze с вершинами D2 и D3. Общее число вершин в таких орграфах 12 (домен D1 непрерывный). Орграф, изображенный на рис. 2.11.a) является энтропийно

оптимальным на множестве  $\{G^+(F)\}$  (см. раздел 2.5).

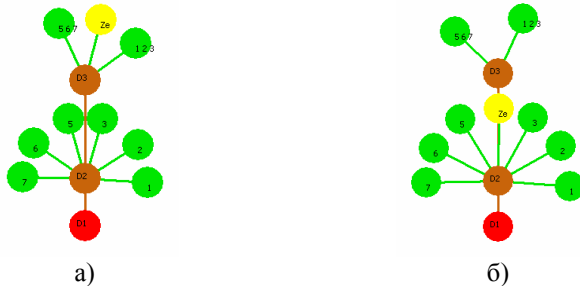


Рис. 2.11 – Структурно-завершенные оргграфы для теста «Сила»

Если в результате субоптимизации управления определено значение «F/D3? P», то значения прикладываемой силы в рамках доменов D2 и D1 определяются в результате нисходящей детерминации. Такими значениями могут быть, например, «F/D2? Pm» (выбор на основе  $\mu_P$ ) и «F/D1? 2.7» (выбор на основе  $\mu_{Pm}$ ). Прикладываемую силу могут реализовывать большое количество двигательных единиц (пример – мышца), следовательно, нисходящая детерминация должна также распределить усилие «F/D1? 2.7» по всем двигательным единицам.

Рассмотрим для конкретности управление перевернутым маятником. Зададим тесты «Угол отклонения подвеса» и «Угловая скорость отклонения подвеса» следующим образом:

$\alpha$  - Угол отклонения подвеса:

D3 {N ( $\mu_N$ ); Ze ( $\mu_{Ze}$ ); P ( $\mu_P$ )}

D2 {Nb ( $\mu_{Nb}$ ); Nm ( $\mu_{Nm}$ ); Ns ( $\mu_{Ns}$ ); Ze ( $\mu_{Ze}$ ); Ps ( $\mu_{Ps}$ ); Pm ( $\mu_{Pm}$ ); Pb ( $\mu_{Pb}$ )}

D1 {[-25; 25]}

Оргграф  $G(\alpha) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Связь между элементами домена D1 и домена D2 показана на рис. 2.12.

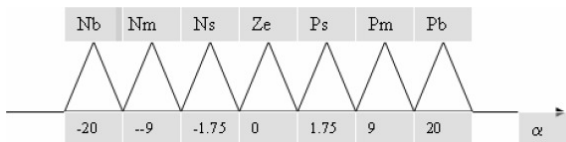


Рис. 2.12 – Лингвистические переменные и функции принадлежности для «Угла отклонения подвеса»

Для простоты взяты треугольные функции принадлежности, но, в общем случае, вид функций – это отдельный вопрос для изучения, оптимизации и

анализа «врожденных» автоматизмов [46, 56, 173, 174]. Связь между элементами домена D2 и домена D3 определяется аналогично тесту F.

$\omega$  - Угловая скорость отклонения подвеса:

D3 {N ( $\mu_N$ ); Ze ( $\mu_{Ze}$ ); P ( $\mu_P$ )}

D2 {Nb ( $\mu_{Nb}$ ); Nm ( $\mu_{Nm}$ ); Ns ( $\mu_{Ns}$ ); Ze ( $\mu_{Ze}$ ); Ps ( $\mu_{Ps}$ ); Pm ( $\mu_{Pm}$ ); Pb ( $\mu_{Pb}$ )}

D1 {[-7; 7]}

Орграф  $G(\omega) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Связь между элементами домена D1 и домена D2 показана на рис. 2.13.

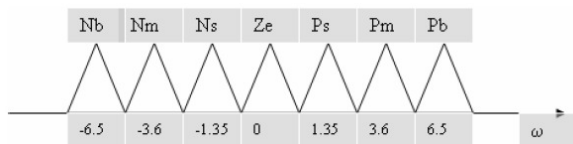


Рис. 2.13 – Лингвистические переменные и функции принадлежности для «Угловой скорости отклонения подвеса»

Связь между элементами домена D2 и домена D3 определяется аналогично тесту F.

Так же как для «Силы» для тестов «Угол» и «Угловая скорость» могут быть автоматически сгенерированы структурно-завершенные орграфы  $\{G^+(\alpha)\}$  и  $\{G^+(\omega)\}$ .

Общее число описаний, в рамках которых может проводиться субоптимизация, составляет:  $3*3*3 = 27$ . Переход с одного уровня описания на другой придает мощную внутреннюю (циклическую) динамику информационным (оптимизационным) процессам.

Пусть в результате субоптимизации на уровне  $\alpha/D2$ ,  $\omega/D2$ ,  $U/D2$  определена таблица лингвистических правил субоптимального регулятора  $\{\alpha_k/D2, \omega_k/D2 \rightarrow U_k/D2\}$ , где  $U_k$  – управление вида F. Пример типовых правил показан в таблице 2.1 (таблица взята из [173]).

Таблица 2.1 – Нечеткие правила регулятора (на уровне D2)

$\alpha_k, \omega_k \rightarrow U_k$		Угол							
		Nb	Nm	Ns	Ze	Ps	Pm	Pb	
Скорость	Nb	Nb	Nb	Nb	Nb	Ns	Ps	Pm	Ze
	Nm	Nm	Nb	Nm	Nm	Ze	Pm	Ps	
	Ns	Nb	Nb	Nm	Ns	Ze	Pm	Pm	
	Ze	Nm	Nb	Ns	Ze	Ps	Pb	Pm	
	Ps	Nm	Nm	Ze	Ps	Pm	Pb	Pb	
	Pm	Ns	Nm	Ze	Pm	Pm	Pb	Pm	
	Pb	Ze	Ns	Ps	Pb	Pb	Pb	Pb	

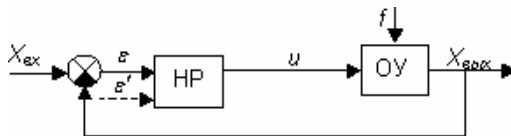
Между доменами D2 и D3 всех тестов примера имеется большой энтропийный скачок (скачок неопределенности), так как три элемента домена D2 переходят в один элемент домена D3. Для устранения такого скачка целесообразно между доменами D2 и D3 вставить промежуточные домены (операция деления домена D3).

Структура орграфов  $G(U)$ ,  $G(\alpha)$  и  $G(\omega)$  и особенно их структурно-завершенных вариантов позволяет находить субоптимальное управление более высокого уровня общности с использованием механизмов нисходящей детерминации. На качественном уровне данную возможность проиллюстрируем в рамках следующего примера.

Рассмотрим задачу управления электроприводом переменного тока. При векторном способе управления электроприводом переменного тока (например, дизель-поезда) возникает необходимость управления процессом разгона с учетом качества протекания электромагнитных процессов и оптимизации энергетических затрат. Одним из критериев качества является величина перерегулирования, которая не должна превышать некоторого максимального значения в переходном режиме. В структуру системы векторного управления входят несколько регуляторов, в частности, ПИ-регулятор скорости, который не обладает соответствующими качественными характеристиками. Методы нечеткой логики позволяют строить цифровые регуляторы, которые обеспечивают высокое быстродействие систем управления [46]. Однако остается открытым вопрос оптимизации процедуры нисходящей детерминации (на языке нечеткой логики – выбора функций принадлежности для нечетких переменных), описывающей функционирование регулятора [63, 64].

Предельно высокое быстродействие систем управления может быть получено в результате субоптимизации на основе банка тестов (параметров и переменных) с использованием процедур нисходящей детерминации.

Воспользуемся одной из типовых структур системы управления с нечетким регулятором (НР), которая представлена на рис 2.14 [46].



$X_{вх}$  – входной сигнал;  $X_{вых}$  – выходной сигнал;  $\varepsilon$  – отклонение;  
 $\varepsilon'$  – производная от отклонения;  $u$  – управляющий сигнал;  $f$  – возмущающее  
 воздействие; НР – нечеткий регулятор; ОУ – объект управления

Рис. 2.14 – Структура системы управления с нечетким регулятором



Нечеткий регулятор (фаззи-регулятор) включает три основных блока (рис. 2.15): блок фаззификации (fuzzyfication), блок формирования логического решения и блок дефаззификации (defuzzyfication).

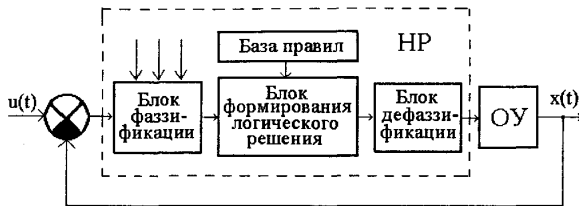


Рис. 2.15 – Функциональная схема нечеткого регулятора

Система с НР обычно устойчива в отношении изменений параметров объекта управления, что связано с нечеткой природой правил функционирования [46, 174]. Традиционные методы описания регуляторов, например, при помощи передаточных функций, для НР не подходят и не требуются. Особенности НР являются: *отсутствие динамики* в самом НР (процессы субоптимизации, в отличие от НР, характеризуются мощной внутренней динамикой информационных потоков); отсутствие «памяти»; описание процесса управления с помощью лингвистических правил.

Для управления процессами, в которых построение модели было затруднено, либо ожидалось большие изменения значений параметров в процессе функционирования системы, были предложены самоорганизующиеся нечеткие логические контроллеры или многоканальные НР [46]. Для синтеза комбинированного управления использовалось несколько параллельно включенных нечетких контроллеров.

Пусть для грубого управления некоторым процессом начальный диапазон ошибки имеет разрешение  $U_0 = \{-1000 \dots +1000\}$ , универсум  $U_0$  квантифицирован на 8 уровней (лингвистических переменных), точка перехода на тонкое управление – интервал  $U_1 = \{-100 \dots +100\}$ , и в этом универсуме  $U_1$  количество лингвистических переменных равно, например, 5 (5 уровней квантификации). Более того, для каждого уровня управления определены свои таблицы правил. Принятие решений о включении квантификатором той или иной модели управления осуществляется на основании метазнаний, полученных от эксперта.

При субоптимизации реализуется схожий подход, но без участия эксперта (с помощью эксперта могут формироваться только орграфы доменов, да и то не всегда). Суть идеи субоптимизации заключается в стремлении максимально обобщить законы управления, а конкретизацию

управления по иерархии исполнительных систем переложить на автоматизии среды (нисходящая детерминация) [136]. При этом законы (правила) управления на разных уровнях общности описания ситуации действительности и метаправила переключения этих законов определяются автоматически (без участия эксперта) на основе непрерывного самообучения.

Для описания сути субоптимизации и обобщения схемы нечеткого регулятора зададим орграфы доменов трех тестов:  $E$  – сигнала ошибки по угловой скорости вращения ротора тягового асинхронного двигателя;  $CE$  – сигнала ошибки по производной угловой скорости;  $DU$  – сигнала нечеткого управления. Областью определения или базовым доменом всех тестов является интервал относительных значений  $[-1, 1]$ .

$E$  - сигнал ошибки по угловой скорости:

D7 {Нулевой сигнал ошибки  $\wedge Z1$ ; не Нулевой сигнал  $\wedge NP1$ }

D6 {Положительный  $\wedge P1$ ; не Положительный  $\wedge ZN1$ }

D5 {Отрицательный  $\wedge N1$ ; не Отрицательный  $\wedge ZP1$ }

D4 {Нулевой сигнал ошибки  $\wedge Z1$ ; Отрицательный  $\wedge N1$ ; Положительный  $\wedge P1$ }

D3 {Нулевой сигнал ошибки  $\wedge Z1$ ; Отрицательный  $\wedge N1$ ; Большой отрицательный  $\wedge NB1$ ; Положительный  $\wedge P1$ ; Большой положительный  $\wedge PB1$ }

D2 {Нулевой сигнал ошибки  $\wedge Z1$ ; Небольшой отрицательный  $\wedge NS1$ ; Средний отрицательный  $\wedge NM1$ ; Большой отрицательный  $\wedge NB1$ ; Небольшой положительный  $\wedge PS1$ ; Средний положительный  $\wedge PM1$ ; Большой положительный  $\wedge PB1$ }

D1  $\{-1; 1\}$

Граф  $G(E) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4 \rightarrow D5; D4 \rightarrow D6; D4 \rightarrow D7\}$ .

Фрагмент правил преобразования:

D2.  $\{NS1, NM1\} \rightarrow D3. \{N1 (\mu_{N1})\}$ ; D2.  $\{PS1, PM1\} \rightarrow D3. \{P1 (\mu_{P1})\}$ ;

D3.  $\{N1, NB1\} \rightarrow D4. \{N1 (\mu_{N1})\}$ ; D3.  $\{P1, PB1\} \rightarrow D4. \{P1 (\mu_{P1})\}$ ;

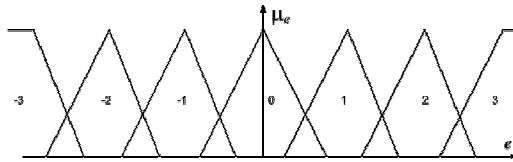
D4.  $\{P1, Z1\} \rightarrow D5. \{ZP1 (\mu_{ZP1})\}$ ; D4.  $\{N1, Z1\} \rightarrow D6. \{ZN1 (\mu_{ZN1})\}$ ;

D4.  $\{P1, N1\} \rightarrow D6. \{NP1 (\mu_{NP1})\}$ .

Отметим, что вершины D5, D6, D7 являются листьями для вершины D4. Ствол орграфа:  $G^{\sim}(E) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4\}$ . На основе ствола орграфа могут быть автоматически построены все структурно-завершенные орграфы  $\{(G^{\sim}(E))^+\}$ .

Правила преобразования из D1 в D2 задаются как в предыдущем примере – функциями принадлежности, вид которых надлежит установить [64]. На рис. 2.16 показан возможный вид функций принадлежности. Здесь 7 термов образуют покрытие терм-множества, но не его разбиение,

поскольку соседние термы пересекаются.



Гранулированные значения термов: 0 – нулевая ошибка; +1 – малая положительная ошибка; +2 – средняя положительная ошибка; +3 – большая положительная ошибка; -1 – малая отрицательная ошибка; -2 – средняя отрицательная ошибка; -3 – большая отрицательная ошибка

Рис. 2.16 – Виды функций принадлежности

*CE* - сигнал ошибки по производной угловой скорости:

D7 {Нулевой сигнал по производной ошибки  $\wedge Z2$ ; не Нулевой сигнал  $\wedge NP2$ }

D6 {Положительный  $\wedge P2$ ; не Положительный  $\wedge ZN2$ }

D5 {Отрицательный  $\wedge N2$ ; не Отрицательный  $\wedge ZP2$ }

D4 {Нулевой сигнал по производной ошибки  $\wedge Z2$ ; Отрицательный  $\wedge N2$ ; Положительный  $\wedge P2$ }

D3 {Нулевой сигнал по производной ошибки  $\wedge Z2$ ; Отрицательный  $\wedge N2$ ; Большой отрицательный  $\wedge NB2$ ; Положительный  $\wedge P2$ ; Большой положительный  $\wedge PB2$ }

D2 {Нулевой сигнал по производной ошибки  $\wedge Z2$ ; Небольшой отрицательный  $\wedge NS2$ ; Средний отрицательный  $\wedge NM2$ ; Большой отрицательный  $\wedge NB2$ ; Небольшой положительный  $\wedge PS2$ ; Средний положительный  $\wedge PM2$ ; Большой положительный  $\wedge PB2$ }

D1  $\{-1; 1\}$

$G(CE) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4 \rightarrow D5; D4 \rightarrow D6; D4 \rightarrow D7\}$ .

*DU* - сигнал нечеткого управления:

D4 {Нулевой сигнал управления  $\wedge Z3$ ; Отрицательный  $\wedge N3$ ; Положительный  $\wedge P3$ }

D3 {Нулевой сигнал управления  $\wedge Z3$ ; Отрицательный  $\wedge N3$ ; Большой отрицательный  $\wedge NB3$ ; Положительный  $\wedge P3$ ; Большой положительный  $\wedge PB3$ }

D2 {Нулевой сигнал управления  $\wedge Z3$ ; Очень малый отрицательный  $\wedge NVS3$ ; Небольшой отрицательный  $\wedge NS3$ ; Средний отрицательный  $\wedge NM3$ ; Большой отрицательный  $\wedge NB3$ ; Очень малый положительный  $\wedge PVS3$ ; Небольшой положительный  $\wedge PS3$ ; Средний положительный  $\wedge PM3$ ;

Большой положительный  $\wedge PB3$

D1  $\{-1; 1\}$

$G(DU) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4\}$ .

Фрагмент правил преобразования:

D2.  $\{NVS3, NS3, NM3\} \rightarrow D3. \{N3 (\mu_{N3})\}$ ;

D2.  $\{PVS3, PS3, PM3\} \rightarrow D3. \{P3 (\mu_{P3})\}$ .

Для тестов  $E$ ,  $CE$ ,  $DU$  могут быть автоматически сгенерированы структурно-завершенные орграфы  $(G^{\sim}(E))^+$ ,  $(G^{\sim}(CE))^+$  и  $(G^{\sim}(DU))^+$ . Отметим также большой скачок энтропии (неопределенности) при переходе от домена D2 к домену D3 в орграфе теста  $DU$  (три элемента перешли в один элемент). Скорее всего, есть смысл ввести промежуточный домен либо образовать новую ветку орграфа.

Каноническое представление стволов графов  $G^{\sim}(E)$  и  $G^{\sim}(CE)$  имеет вид:

Сигнал ошибки  $\{$

D4  $\{1\ 2\ 3; z; 5\ 6\ 7\}$

D3  $\{1; 2\ 3; z; 5\ 6; 7\}$

D2  $\{1; 2; 3; z; 5; 6; 7\}$

D1  $\{\}$ .

На рис. 2.17 изображены два структурно-завершенных орграфа из множеств  $\{(G^{\sim}(E))^+\}$  и  $\{(G^{\sim}(CE))^+\}$ . Орграф, изображенный на рис. 2.17.б) является энтропийно оптимальным (см. раздел 2.5).

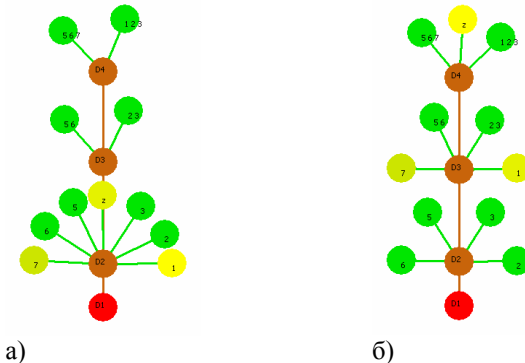


Рис. 2.17 – Структурно-завершенные орграфы для тестов «Сигнал ошибки»

Уровень общности описания задачи формируется путем комбинации доменов всех тестов. В нашем случае общее количество комбинаций

составляет:  $7*7*4 = 196$  (без построения структурно-завершенных графов). Таким образом, имеем 196 описаний, на каждом из которых может проводиться своя оптимизация. Как уже отмечалось, переход с одного уровня описания и оптимизации на другой задает мощную внутреннюю (циклическую: восходящую и нисходящую) динамику информационных процессов, что принципиально отличает субоптимизацию от НР. В рамках нечеткой логики оптимизация традиционно проводится на одном уровне, а именно:  $E/D2$ ,  $CE/D2$ ,  $DU/D2$ .

Формирование нечеткого вывода контроллером осуществляется на основании нечетких правил  $\{E/D2, CE/D2 \rightarrow DU/D2\}$ , приведенных в таблице 2.2 (одна из типовых таблиц для нечеткого регулятора). Нечеткие правила принадлежат либо онтологии (Банку тестов), либо базе знаний.

Таблица 2.2 – Нечеткие правила субоптимального регулятора [64]  
(на уровне D2)

<b><i>CE/E</i></b>	<b><i>NB1</i></b>	<b><i>NM1</i></b>	<b><i>NS1</i></b>	<b><i>Z1</i></b>	<b><i>PS1</i></b>	<b><i>PM1</i></b>	<b><i>PB1</i></b>
<b><i>NB2</i></b>	<i>NB3</i>	<i>NB3</i>	<i>NB3</i>	<i>NM3</i>	<i>NS3</i>	<i>NVS3</i>	<i>Z3</i>
<b><i>NM2</i></b>	<i>NB3</i>	<i>NB3</i>	<i>NM3</i>	<i>NS3</i>	<i>NVS3</i>	<i>Z3</i>	<i>PVS3</i>
<b><i>NS2</i></b>	<i>NB3</i>	<i>NM3</i>	<i>NS3</i>	<i>NVS3</i>	<i>Z3</i>	<i>PVS3</i>	<i>PS3</i>
<b><i>Z2</i></b>	<i>NM3</i>	<i>NS3</i>	<i>NVS3</i>	<i>Z3</i>	<i>PVS3</i>	<i>PS3</i>	<i>PM3</i>
<b><i>PS2</i></b>	<i>NS3</i>	<i>NVS3</i>	<i>Z3</i>	<i>PVS3</i>	<i>PS3</i>	<i>PM3</i>	<i>PB3</i>
<b><i>PM2</i></b>	<i>NVS3</i>	<i>Z3</i>	<i>PVS3</i>	<i>PS3</i>	<i>PM3</i>	<i>PB3</i>	<i>PB3</i>
<b><i>PB2</i></b>	<i>Z3</i>	<i>PVS3</i>	<i>PS3</i>	<i>PM3</i>	<i>PB3</i>	<i>PB3</i>	<i>PB3</i>

Правила формируются следующим образом:

1. Если  $E = NB1$  и  $CE = NS2$ , то  $DU = NB3$ ,
2. Если  $E = NM1$  и  $CE = NB2$ , то  $DU = NB3$  и т.д.

Преобразуем нечеткие правила с уровня D2 на уровень D4, оставляя для каждого сочетания значений ( $E$ ,  $CE$ ) только разные значения. Результат представим в виде таблицы 2.3.

Данные таблицы 2.3 можно представить в виде однозначных (по элементам уровня D4) правил:

$E? N1, CE? N2 \rightarrow DU? N3$ ;  $E? N1, CE? Z2 \rightarrow DU? N3$ ;

$E? Z1, CE? N2 \rightarrow DU? N3$ ;  $E? Z1, CE? Z2 \rightarrow DU? Z3$ ;

$E? Z1, CE? P2 \rightarrow DU? P3$ ;

$E? P1, CE? Z2 \rightarrow DU? P3$ ;

$E? P1, CE? P2 \rightarrow DU? P3$ .

Таблица 2.3 – Нечеткие правила субоптимального регулятора  
(на уровне D4)

$CE \setminus E$	$N1$	$N1$	$N1$	$Z1$	$P1$	$P1$	$P1$
$N2$	$N3$			$N3$	$N3$		
$N2$						$Z3$	
$N2$							$P3$
$Z2$	$N3$			$Z3$			$P3$
$P2$	$N3$			$P3$			
$P2$		$Z3$					
$P2$			$P3$				$P3$

Данные правила можно выполнять адаптивно, «подруливая» до тех пор, пока не возникнет ситуация ( $E? Z1, CE? Z2$ ). Для реализации нисходящей детерминации нужно использовать функции принадлежности соответствующих элементов. Отметим также, что даже на самом высоком уровне обобщения для однозначности управления по знаку недостаточно знать значение только одного теста.

Перевод правил на уровень D5 и D6 позволяет (автоматически) сформулировать общее метаправило: «Если знаки ошибок совпадают, то управление имеет тот же знак. Точность измерения ошибок может быть предельно низкой (с точностью до знака)».

Неоднозначность управления (по знаку) на уровне D4 имеет место только в двух случаях: ( $E/D4? N1, CE/D4? P2$ ) и ( $E/D4? P1, CE/D4? N2$ ), т.е. когда ошибки разного знака (области окрашены желтым цветом). Раскроем эти состояния на уровне D3 (таблицы 2.4 и 2.5).

Таблица 2.4 – Нечеткие правила для состояния ( $E/D4? N1, CE/D4? P2$ )  
(на уровне D3)

$CE \setminus E$	$NB1$	$N1$	$N1$
$P2$	$N3$	$N3$	$Z3$
$P2$		$Z3$	$P3$
$PB2$	$Z3$		$P3$

Таблица 2.5 – Нечеткие правила для состояния ( $E/D4? P1, CE/D4? N2$ )  
(на уровне D3)

$CE \setminus E$	$P1$	$P1$	$PB1$
$NB2$	$N3$		$Z3$
$N2$	$N3$	$Z3$	
$N2$	$Z3$	$P3$	$P3$

Как видим, уровень D3 позволяет выделить следующие однозначные (по элементам уровня D3) правила:

$E? NB1, CE? P2 \rightarrow DU? N3$ ;  $E? NB1, CE? PB2 \rightarrow DU? Z3$ ;  
 $E? N1, CE? PB2 \rightarrow DU? P3$ ;  $E? P1, CE? NB2 \rightarrow DU? N3$ ;  
 $E? PB1, CE? NB2 \rightarrow DU? Z3$ ;  $E? PB1, CE? N2 \rightarrow DU? P3$ .

Приведенные правила позволяют сформулировать еще, как минимум, два метаправила:

«Если ошибка по угловой скорости большая отрицательная, то управление не положительное (скорее отрицательное)»;

«Если ошибка по угловой скорости большая положительная, то управление не отрицательное (скорее положительное)».

Неоднозначность управления на уровне D3 также имеет место только в двух случаях (области окрашены желтым цветом): ( $E/D3? N1, CE/D3? P2$ ) и ( $E/D3? P1, CE/D3? N2$ ). Для этих состояний необходимо возвращаться на уровень D2, т.е. брать правила из таблицы 2.2.

Таким образом, в разных ситуациях управления допустима разная точность по всем тестам. Чем ближе ситуация (процесс) к конечной цели управления (устранение ошибок), тем точнее должны измеряться тесты, включая управление. Другими словами решение сложной и громоздкой задачи вычисления точных воздействий подменяется значительно более простой и гибкой стратегией адаптивного «подруливания» при сохранении требуемой точности результата [46, 174].

## 2.5 Энтропийные характеристики орграфов доменов

Под *семантическим потенциалом* (SP) некоторой (онтологической) структуры будем понимать способность данной структуры адекватно отображать все разнообразие проявлений действительности. Очевидно, чем менее дифференцированы значения теста  $\tau$  в рамках домена  $T$ , тем меньше семантический потенциал данного домена.

Не вызывает сомнения, что семантический потенциал орграфа доменов теста зависит как от мощности орграфа (числа вершин – доменов и числа элементов в доменах), так и от рациональной организации структуры орграфа (связей). В «хорошо» организованном орграфе доменов в силу эмерджентности семантический потенциал орграфа  $G(\tau) = \{T \rightarrow T\}_\tau$  превышает сумму семантических потенциалов всех составляющих вершин – доменов:

$$SP(G(\tau)) > SP(T_1) + \dots + SP(T_n). \quad (2.6)$$

Соотношение (2.6) отвечает сверхаддитивному нелинейному сложению:

«целое больше суммы составляющих его частей» или «сумма знаний, содержащихся в отдельных доменах, в орграфе доменов порождает новое знание».

Всеобщей количественной мерой неопределенности, непредсказуемости, беспорядка, хаоса, дезорганизованности вероятностных систем является энтропия [132]. Если термодинамическая энтропия есть мера неупорядоченности (или беспорядка) микрочастиц, то энтропия в широком смысле — мера неупорядоченности (беспорядка, несовершенства) объекта по любым признакам.

Энтропия орграфов доменов связана как с неопределенностью нисходящей детерминации (при движении от общего к частному), так и с обобщением (при движении от частного к общему).

Если при интеграции (объединении) энтропия системы уменьшается, это означает, что появляется новое интегративное свойство системы (орграфа), которое до объединения элементов не существовало. Так, например, если из строительных деталей собрать дом или из деталей автомобиля собрать автомобиль, то структурная энтропия этих систем уменьшится (по сравнению с суммой энтропий деталей).

Рассмотрим каноническую связку доменов ( $T \rightarrow T'$ ):

$$T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3\ 4\ 5\}.$$

Видно, что четыре элемента домена  $T$  перешли в один элемент домена  $T'$ . Произошел резкий энтропийный скачок, т.е. в домене  $T'$  по отношению к домену  $T$  резко возросла неопределенность интерпретации одного элемента и, соответственно, резко снизился семантический потенциал домена  $T'$ . Структурный энтропийный скачок в связке ( $T \rightarrow T'$ ) будет всегда, однако минимальный энтропийный скачок будет иметь место в случае, если только два элемента перейдут в один элемент (*схема преобразования «2 в 1»*), например:

$$T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3; 4; 5\};$$

$$T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3; 4\ 5\}.$$

Скачок энтропии при схеме преобразования «1 в 1» отсутствует (равен нулю). Скачок энтропии при схеме преобразования «2 в 1» будем считать равным единице. В общем случае, скачок энтропии при схеме преобразования « $N$  в 1» будем считать равным  $(N-1)$ .

Определим общий *структурный энтропийный скачок* ( $\Delta\mathcal{E}$ ) связки дискретных доменов ( $T \rightarrow T'$ ) как сумму энтропий преобразований всех элементов. Примеры энтропийных скачков между дискретными доменами:

$$\Delta\mathcal{E}(T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3; 4; 5\}) = 1;$$

$$\Delta\mathcal{E}(T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3; 4\ 5\}) = 1+1 = 2;$$



$$\Delta\mathcal{E}(T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3\ 4; 5\}) = 2;$$

$$\Delta\mathcal{E}(T \{1; 2; 3; 4; 5\} \rightarrow T' \{1; 2\ 3\ 4\ 5\}) = 3.$$

Таким образом, приращение энтропии в связке доменов «предок – потомок» – это мера множества тех элементов домена-предка, дифференциация которых теряется при переходе к домену-потомку.

Общую *структурную энтропию орграфа доменов*  $G(\tau)$  определим выражением

$$\mathcal{E}(G(\tau)) = \sum_G \Delta\mathcal{E}(T \rightarrow T'). \quad (2.7)$$

Уточним, что мы будем понимать под приращением энтропии в рамках преобразования  $(T\{\} \rightarrow T')$ , где  $T\{\}$  – базовая непрерывная или дискретно-непрерывная вершина. Если применить в чистом виде логику рассуждений для дискретных доменов, то  $\Delta\mathcal{E}(T\{\} \rightarrow T') = \infty$ , так как применяется схема преобразования «интервал значений в 1» или « $\infty$  в 1». Для конструктивности можно предложить следующую формулу:

$$\Delta\mathcal{E}(T\{\} \rightarrow T') = C + g(|T'|), \quad (2.8)$$

где  $C$  – некоторая большая положительная константа, а  $g(|T'|)$  – функция, которая уменьшается с увеличением числа элементов в домене  $T'$  (энтропия перехода с увеличением  $|T'|$  уменьшается). Конкретный вид функции  $g$  определяется требованием однозначности вычисления структурной энтропии произвольной вершины орграфа, но об этом чуть позже.

Видимо будет правильно считать, что  $C$  больше суммарной структурной энтропии дискретной части орграфа доменов, т.е.

$$C \geq \sum_G \Delta\mathcal{E}(T \rightarrow T' | T, T' \text{ - дискретные домены}). \quad (2.9)$$

Справедливы следующие предложения.

**Предложение 2.9.** Для любого орграфа доменов  $G(\tau)$  выполняются следующие соотношения:

(i) Для любой дискретной связки  $(T \rightarrow T') \in G(\tau)$  верно:

$$1 \leq \Delta\mathcal{E}(T \rightarrow T') \leq |T'| - 2. \quad (2.10)$$

(ii)  $\mathcal{E}(G(\tau)) \geq |\{T \rightarrow T'\}|$ , где  $|\{T \rightarrow T'\}|$  – множество всех дуг в орграфе.

**Предложение 2.10.** Пусть  $\underline{G}(\tau)$  – каноническое представление орграфа  $G(\tau)$ , тогда выполняются соотношения:

(i)  $\forall (T \rightarrow T') \in G(\tau)$  верно:  $\Delta\mathcal{E}(T \rightarrow T') = \Delta\mathcal{E}(\underline{T} \rightarrow \underline{T}')$ .

(ii)  $\mathcal{E}(G(\tau)) = \mathcal{E}(\underline{G}(\tau))$ .

Доказательство данных предложений не вызывает затруднений.

Пусть  $[T_1 \rightarrow T_n] = (T_1 \rightarrow T_2 \rightarrow \dots \rightarrow T_n)$  произвольная цепочка

обобщений в рамках орграфа  $G(\tau)$ , тогда энтропия пути определяется соотношением:

$$\Delta\mathcal{E}([T_1 \rightarrow T_n]) = \Delta\mathcal{E}(T_1 \rightarrow T_2) + \Delta\mathcal{E}(T_2 \rightarrow T_3) + \dots + \Delta\mathcal{E}(T_{n-1} \rightarrow T_n). \quad (2.11)$$

Из предложения 2.10 вытекает справедливость следующего выражения:

$$\Delta\mathcal{E}([T_1 \rightarrow T_n]) = \Delta\mathcal{E}([\underline{T}_1 \rightarrow \underline{T}_n]), \quad (2.12)$$

где  $\underline{T}_j$  – каноническое представление вершины  $T_j$  ( $j = 1, \dots, n$ ).

Энтропию базовой вершины-домена  $T_0$  примем равной нулю. Структурную энтропию вершины  $T$  орграфа  $G(\tau)$  определим следующим образом:

$$\mathcal{E}(T) = \Delta\mathcal{E}([T_0 \rightarrow T]), \quad (2.13)$$

где  $T_0$  – базовый домен,  $[T_0 \rightarrow T]$  – один из путей к вершине  $T$ . Выражение (2.13) показывает, что с повышением уровня общности (удаленности от базовой вершины) энтропия вершины монотонно возрастает.

Справедливы следующие предложения.

**Предложение 2.11.** Если базовая вершина орграфа  $G(\tau)$  дискретная, то структурная энтропия любой вершины орграфа определяется единственным образом и не зависит от наличия различных путей к вершине.

**Предложение 2.12.** Если базовая вершина орграфа  $G(\tau)$  непрерывная, то можно таким образом подобрать функцию  $g$  в выражении (2.8), что энтропия любой вершины орграфа будет определяться единственным образом, т.е. не будет зависеть от наличия различных путей к вершине.

Если в орграфе  $G(\tau)$  существует только одна связка ( $T \{ \} \rightarrow T$ ), то можно положить  $g = 0$ . Пусть существует несколько связок ( $T \{ \} \rightarrow T_j$ ), где  $j = 1, \dots, J$ . Без потери общности будем считать, что вершины  $T_j$  упорядочены по мощности, а именно:  $|T_1| \geq |T_2| \geq \dots \geq |T_J|$ . Положим:

$$g(|T_j|) = |T_1| - |T_j|. \quad (2.14)$$

Функция  $g$ , определяемая формулой (2.14), удовлетворяет требованию предложения 2.12. Выражение (2.8) можно переписать в следующем виде:

$$\Delta\mathcal{E}(T \{ \} \rightarrow T_j) = C + |T_1| - |T_j|. \quad (2.15)$$

Для примера зададим орграф «теста X» следующим образом:

Тест X:  $1 \{ \} \ 2 \{1; 2; 3; 4\} \ 3 \{1 \ 2; 3 \ 4\}$ .

$G(X) = \{1 \rightarrow 2 \rightarrow 3; 1 \rightarrow 3\}$ .

В данном примере вопрос заключается в энтропии вершины «3», так как имеются два пути вычисления энтропии данной вершины. Очевидно:  $|2|=4$ ,

а  $|3| = 2$ . Согласно (2.15) положим:  $\Delta\mathcal{E}(1 \rightarrow 2) = C$ ;  $\Delta\mathcal{E}(1 \rightarrow 3) = C + 2$ . Следовательно, независимо от пути  $\mathcal{E}(3) = C + 2$ .

Приведем примеры структурных энтропий.

Тест «Печеночные знаки»:

D2 {Печеночных знаков нет  $\wedge N$ ; Имеются печеночные знаки  $\wedge 1\ 2\ 3\ 4\ 5\ 6$ }

D1 {Печеночных знаков нет  $\wedge N$ ; Сосудистые звездочки  $\wedge 1$ ; Расширены подкожные вены живота  $\wedge 2$ ; Печеночные ладони  $\wedge 3$ ; Ксантелазмы  $\wedge 4$ ; Гинекомастия  $\wedge 5$ ; Атрофия мышц плечевого пояса  $\wedge 6$ }.

$G(\text{Печеночные знаки}) = \{D1 \rightarrow D2\}$ .

$\mathcal{E}(G(\text{Печеночные знаки})) = \Delta\mathcal{E}(D1 \rightarrow D2) = 5$ ;  $\mathcal{E}(D2) = 5$ .

Тест «Питание»:

D3 {Нормальное питание  $\wedge N$ ; Ненормальное питание  $\wedge 1\ 3$ }

D2 {Пониженное питание  $\wedge 1\ 2$ ; Нормальное питание  $\wedge N$ ; Повышенное питание  $\wedge 3\ 4$ }

D1 {Резко пониженное питание (кахексия)  $\wedge 1$ ; Умеренно пониженное питание  $\wedge 2$ ; Нормальное питание  $\wedge N$ ; Умеренно повышенное питание  $\wedge 3$ ; Резко повышенное питание (ожирение)  $\wedge 4$ }

$G(\text{Питание}) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

$\mathcal{E}(G(\text{Питание})) = \Delta\mathcal{E}(D1 \rightarrow D2) + \Delta\mathcal{E}(D2 \rightarrow D3) = 2 + 1 = 3$ .

$\mathcal{E}(D2) = 2$ ;  $\mathcal{E}(D3) = 3$ .

Сформулируем *принцип структурно-энтропийной оптимальности* дискретной части орграфа  $G(\tau)$ : все преобразования элементов между дискретными доменами должны принадлежать к типам «1 в 1» или «2 в 1». Следование данному принципу при проектировании орграфов доменов обеспечивает минимальные энтропийные скачки при переходе с одного дискретного уровня общности на другой.

Введем *меру несовершенства (энтропию несовершенства)* дискретной части орграфа  $G(\tau)$  или любой связки дискретных доменов ( $T \rightarrow T'$ ) как сумму фактов отклонений (0/1) схем преобразований элементов доменов от оптимальных схем. Обозначим данную меру как  $\mathcal{E}_{\text{опт}}(G(\tau))$ , а приращение меры:  $\Delta\mathcal{E}_{\text{опт}}(T \rightarrow T')$ . Индекс «опт» означает, что энтропия связана с принципом оптимальности.

*Энтропию несовершенства орграфа доменов  $G(\tau)$*  определим выражением

$$\mathcal{E}_{\text{опт}}(G(\tau)) = \sum_G \Delta\mathcal{E}_{\text{опт}}(T \rightarrow T'). \quad (2.16)$$

С помощью энтропии  $\mathcal{E}_{\text{опт}}$  можно переформулировать принцип структурно-энтропийной оптимальности орграфа доменов: орграф  $G(\tau)$  является структурно-энтропийно оптимальным (его дискретная часть)

тогда и только тогда, когда  $\mathcal{E}_{\text{опт}}(G(\tau)) = 0$ .

Примеры:

$$\Delta\mathcal{E}_{\text{опт}}(T\{1; 2; 3; 4; 5\} \rightarrow T\{1; 2\ 3; 4\ 5\}) = 0;$$

$$\Delta\mathcal{E}_{\text{опт}}(T\{1; 2; 3; 4; 5\} \rightarrow T\{1; 2\ 3\ 4; 5\}) = 1;$$

$$\Delta\mathcal{E}_{\text{опт}}(T\{1; 2; 3; 4; 5\} \rightarrow T\{1; 2\ 3\ 4\ 5\}) = 1;$$

$$\mathcal{E}_{\text{опт}}(G(\text{Питание})) = \Delta\mathcal{E}_{\text{опт}}(D1 \rightarrow D2) + \Delta\mathcal{E}_{\text{опт}}(D2 \rightarrow D3) = 0 + 0 = 0;$$

$$\mathcal{E}_{\text{опт}}(G(\text{Печеночные знаки})) = \Delta\mathcal{E}_{\text{опт}}(D1 \rightarrow D2) = 1.$$

Оргграф теста «Питание» является структурно-энтропийно оптимальным ( $\mathcal{E}_{\text{опт}} = 0$ ), а оргграф теста «Печеночные знаки» не является ( $\mathcal{E}_{\text{опт}} = 1$ ).

Рассмотрим еще один пример:

Тест «Высота стояния правого купола диафрагмы ^ВСПКД»:

D3 {Диафрагма справа на уровне VI ребра (норма) ^4; Диафрагма справа на уровне VI ребра ^1 5}

D2 {Диафрагма справа ниже уровня VI ребра ^1 2 3; Диафрагма справа на уровне VI ребра (норма) ^4; Диафрагма справа на уровне VII ребра ^5}

D1 {Диафрагма справа на уровне III ребра ^1; Диафрагма справа на уровне IV ребра ^2; Диафрагма справа на уровне V ребра ^3; Диафрагма справа на уровне VI ребра (норма) ^4; Диафрагма справа на уровне VII ребра ^5}

$$G(\text{ВСПКД}) = \{D1 \rightarrow D2; D2 \rightarrow D3\}.$$

$$\mathcal{E}(G(\text{ВСПКД})) = \Delta\mathcal{E}(D1 \rightarrow D2) + \Delta\mathcal{E}(D2 \rightarrow D3) = 2 + 1 = 3.$$

$$\mathcal{E}_{\text{опт}}(G(\text{ВСПКД})) = \Delta\mathcal{E}_{\text{опт}}(D1 \rightarrow D2) + \Delta\mathcal{E}_{\text{опт}}(D2 \rightarrow D3) = 1 + 0 = 1.$$

Справедливо предложение.

**Предложение 2.13.** Для любого подграфа доменов  $G'(\tau) \subset G(\tau)$  верно: если  $\mathcal{E}_{\text{опт}}(G(\tau)) = 0$ , то  $\mathcal{E}_{\text{опт}}(G'(\tau)) = 0$ .

Следствием данного предложения является тот факт, что если  $\mathcal{E}_{\text{опт}}(G(\tau)) = 0$ , то  $\mathcal{E}_{\text{опт}}(G^-(\tau)) = 0$ , т.е. ствол оргграфа также оптимальный.

Рассмотрим вопрос о структурно-энтропийной оптимальности структурно-завершенных оргграфов. Начнем с примера.

Вершина D3 оргграфа теста «Питание» является блуждающей и в оргграфе  $G(\text{Питание})$  она прикреплена к вершине D2. Такое расположение D3 дало следующие результаты:  $\Delta\mathcal{E}(D2 \rightarrow D3) = 1$ ;  $\Delta\mathcal{E}_{\text{опт}}(D2 \rightarrow D3) = 0$ .

Стол оргграфа  $G(\text{Питание})$  имеет вид:  $G^-(\text{Питание}) = \{D1 \rightarrow D2\}$ . При построении структурно-завершенного оргграфа  $(G^-(\text{Питание}))^+$  блуждающую вершину D3 прикрепим к вершине D1. Получим:

$$\Delta\mathcal{E}(D1 \rightarrow D3) = 3; \Delta\mathcal{E}_{\text{опт}}(D1 \rightarrow D3) = 1.$$

Видим, что такое расположение блуждающей вершины привело к значительному ухудшению всех показателей, а именно: возросла структурная энтропия и нарушилась оптимальность оргграфа.

Обобщим результаты примера.

**Предложение 2.14.** Минимум структурной энтропии в структурно-завершенном орграфе  $(G^+)$  достигается тогда и только тогда, когда блуждающие вершины прикреплены к вершинам ствола  $G^-$  с максимальной энтропией.

Пусть каноническое представление первичного орграфа имеет вид:

Тест В:

3 #1 {1; 2 3; 4 5; 6 7 8 9; 10};

2 {1; 2 3; 4; 5; 6 7; 8 9; 10};

1 {1; 2; 3; 4; 5; 6; 7; 8; 9; 10};

0 {}.

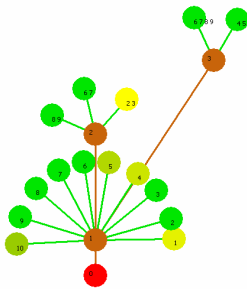
$G(B) = \{0 \rightarrow 1; 1 \rightarrow 2; 1 \rightarrow 3\}$ .

$\mathcal{E}_d(G(B)) = \Delta\mathcal{E}(1 \rightarrow 2) + \Delta\mathcal{E}(1 \rightarrow 3) = 3 + 2 = 5$

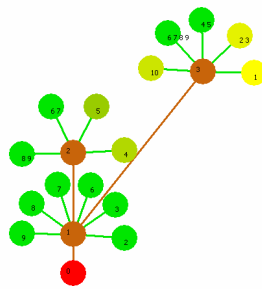
$\mathcal{E}_{\text{опт}}(G(B)) = \Delta\mathcal{E}_{\text{опт}}(1 \rightarrow 2) + \Delta\mathcal{E}_{\text{опт}}(1 \rightarrow 3) = 0 + 0 = 0; \quad \mathcal{E}(2) = 3; \mathcal{E}(3) = 5.$

$\mathcal{E}_d$  – энтропия дискретной части. Нотация «3 #1» означает связку  $(1 \rightarrow 3)$ . Для остальных вершин обобщение идет снизу вверх.

На рис. 2.18 для теста В показаны два структурно завершенных орграфа доменов с разным расположением блуждающих вершин и, соответственно, разной энтропией (оттенками желтого и зеленого обозначены блуждающие вершины, красным цветом обозначена базовая вершина, коричневым цветом обозначены дискретные вершины первичного орграфа).



а)  $\mathcal{E}_d = 109, \mathcal{E}_{\text{опт}} = 16$



б)  $\mathcal{E}_d = 91, \mathcal{E}_{\text{опт}} = 16$

Рис. 2.18 – Структурно-завершенные орграфы с разной структурной энтропией

Данный пример показывает, что, несмотря на структурно-энтропийную оптимальность орграфа  $G(\text{Тест})$ , все его структурно-завершенные орграфы не являются оптимальными. Можно говорить лишь о минимизации структурно-энтропийной не оптимальности графов  $(G^+)$ . Орграф,

изображенный на рис. 2.18.б), обладает наименьшей структурной энтропией на множестве графов  $\{(G^-)^+\}$ .

Справедливы предложения.

**Предложение 2.15.** Наибольшей структурной энтропией и энтропией несовершенства на множестве орграфов  $\{(G^-)^+\}$  обладает предельный структурно-завершенный орграф  $(G^-)^{++}$ .

**Предложение 2.16.** Если некоторый орграф из  $\{(G^-)^+\}$  обладает минимальной структурной энтропией, то он обладает также минимальной энтропией несовершенства. Обратное, в общем случае, неверно.

Тот факт, что из минимальной энтропии несовершенства не вытекает минимальность структурной энтропии, наглядно демонстрирует рис. 2.18.

Предложение 2.14 подсказывает алгоритм построения структурно-завершенного орграфа с минимальной структурной энтропией, а именно: *построение вершин – листьев (операция расщепления) необходимо начинать с вершин ствола, имеющих максимальную структурную энтропию, т.е. терминальных вершин. В последнюю очередь расщепляется базовая вершина (если она дискретная)*. Назовем этот алгоритм ОПТИМА.

**Предложение 2.17.** Алгоритм ОПТИМА обеспечивает построение структурно-завершенного орграфа доменов с минимальной структурной энтропией на множестве  $\{(G^-)^+\}$ . Решение в общем случае не единственно.

Обоснованность утверждения вытекает из Предложения 2.14. Не единственность решения вытекает из возможности наличия вершин с одинаковой максимальной энтропией. Как следует из предложения 2.16, алгоритм ОПТИМА обеспечивает построение структурно-завершенного орграфа с наименьшей энтропией несовершенства.

Приведем пример двух структурно-завершенных орграфов с минимальной энтропией. Зададим орграф «Теста А» следующим образом:

Тест А:

3 #1 {1; 2; 3; 4 5};

2 {1; 2 3; 4; 5};

1 {1; 2; 3; 4; 5}.

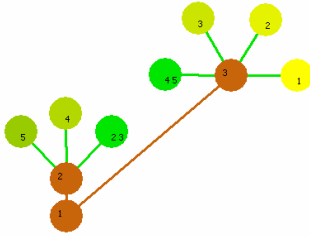
$G(A) = \{1 \rightarrow 2; 1 \rightarrow 3\}$ .

$\mathcal{E}(1) = 0; \mathcal{E}(2) = 1; \mathcal{E}(3) = 1; \mathcal{E}(G(A)) = \Delta\mathcal{E}(1 \rightarrow 2) + \Delta\mathcal{E}(1 \rightarrow 3) = 1 + 1 = 2$ .

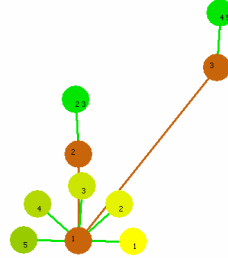
$\mathcal{E}_{\text{онт}}(G(A)) = \Delta\mathcal{E}_{\text{онт}}(1 \rightarrow 2) + \Delta\mathcal{E}_{\text{онт}}(1 \rightarrow 3) = 0 + 0 = 0$ .

На рис. 2.19.а) показан структурно-завершенный орграф с минимальной структурной энтропией, а на рис. 2.19.б) – с максимальной структурной энтропией на множестве всех деревьев (в это число не входит предельный структурно-завершенный орграф). В данном примере существует ровно два структурно-завершенных орграфа с минимальной структурной энтропией. Первый представлен на рис. 2.19.а). Другой получается из первого перемещением блуждающей вершины-листка с кодом «1» (желтая

вершина) на вершину «2».



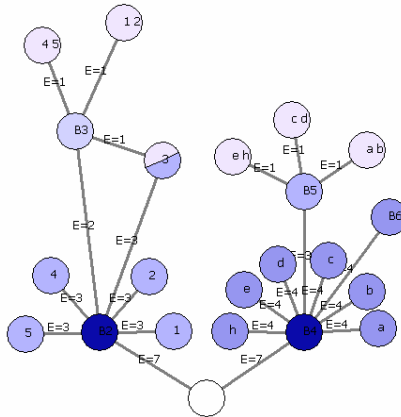
а)  $\mathcal{E} = 16, \mathcal{E}_{\text{опт}} = 7$



б)  $\mathcal{E} = 21, \mathcal{E}_{\text{опт}} = 7$

Рис. 2.19 – Структурно-завершенные орграфы с разной энтропией

Структурно-энтропийное многообразие, связанное с орграфом, удобно отображать в монохромном цвете. Базовую вершину будем отображать белым цветом, так как ее энтропия равна нулю. Интенсивность закрашивания произвольной вершины  $T$  пропорциональна величине  $E = \Delta\mathcal{E}(T \rightarrow T')$ . Если вершина имеет несколько предков, то она может быть разделена на соответствующее число секторов, каждый из которых имеет свою интенсивность закрашивания. Потомки непрерывной вершины имеют максимальную интенсивность закрашивания. На рис. 2.20 – 2.22 показаны примеры структурно-энтропийных многообразий.



$\mathcal{E}(G^{++}) = 54; \mathcal{E}_{\text{опт}} = 12$

Рис. 2.20 – Энтропийное многообразие для теста «Возраст»

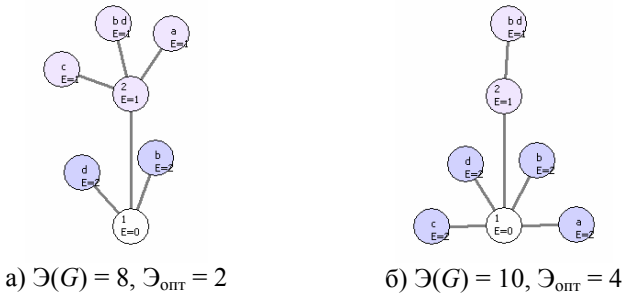


Рис. 2.21 – Энтропийные многообразия различных структурно-завершенных орграфов

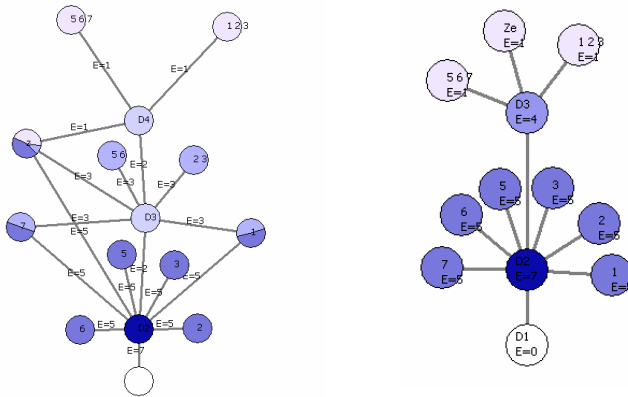


Рис. 2.22 – Энтропийные многообразия тестов «Сигнал ошибки» и «Сила»

Рассмотрим вопрос о движущей силе энтропийных преобразований или другими словами: «Для чего нужно вводить и использовать «грубые» домены, если есть базовые домены?» Такой движущей силой является жесточайший лимит всех видов ресурсов и, прежде всего – энергии (вычислительной, «психической») и времени. Следствием экономии энергии является относительно небольшая оперативная память, в рамках которой осуществляется выработка оперативного управления в условиях жестких временных ограничений. Недостаток энергии, времени и памяти с одной стороны и потенциальная бесконечность описания действительности (мощность доменов тестов), с другой стороны, определяют неизбежность оперирования высокэнтропийной информацией (доменами) по



большинству тестов, участвующих в описании действительности. Точность измерения тестов должна быть достаточна для решения целевой задачи и не более того. В этом суть Принципа предельных обобщений.

Таким образом, стремлению к точности данных, т.е. понижению энтропии, противостоит стремление к уменьшению ресурсов, что ведет к повышению энтропии. Возникает проблема баланса энтропии (точности) и ресурсов (энергии). То же касается баланса управляемости и автономности. Не следует стремиться управлять всем. Что-то надо передавать для самоуправления и самореализации. На этом принципе будут строиться функциональные системы когнитивно-поведенческого уровня (см. главу 7).

Под *информационным управлением* понимается механизм, когда управляющее воздействие носит неявный, косвенный, информационный характер и объекту управления (ОУ) дается определенная информационная картина, ориентируясь на которую он как бы самостоятельно выбирает линию своего поведения, или организованность [117]. Именно такой подход применяется при субоптимизации управления с учетом автоматизмов нисходящей детерминации.

Антиэнтропийная направленность формирования автоматизмов нисходящей детерминации позволяет выявить важную закономерность этого процесса.

В процессе нисходящей детерминации на этапе  $T.a \rightarrow T.\{b_1; \dots; b_k\}$  реализуется приращение *информационной энтропии* следующего вида:

$$\Delta \mathcal{E}_{\text{инф}}(T.a \rightarrow T.\{b_1(w_1); \dots; b_k(w_k)\}) = - \sum_{i=1,k} w_i \ln w_i, \quad (2.17)$$

где  $w_i$  – вероятность выбора элемента  $b_i$  ( $i = 1, \dots, k$ ). Если одна из вероятностей  $w$  равна единице, а все остальные равны нулю, то неопределенность в информации отсутствует и  $\Delta \mathcal{E}_{\text{инф}} = 0$ . С другой стороны, информационная энтропия процесса нисходящей детерминации получает максимальное приращение, когда все  $w_i$ , равны  $1/k$ , т.е. имеется максимальная неопределенность с выбором значения домена-предка.

Суммарная информационная энтропия нисходящей детерминации складывается из приращений информационной энтропии на всех участках преобразований.

Таким образом, главной целью создания автоматизмов нисходящей детерминации является максимальное снижение приращений информационной энтропии вида (2.17). Отметим, что структурная энтропия орграфа доменов и информационная энтропия нисходящей детерминации коррелируют между собой очень слабо. Только в случае, если для всех нисходящих преобразований  $\{T \rightarrow T'\}$  все  $w_i$  равны  $1/k$ , можно говорить о прямой связи между структурной и информационной энтропиями.

## 2.6 Созревание протоструктур тестов

Важность рассмотрения вопросов «созревания» различных когнитивных структур вытекает из важности изучения динамики и трансформации смысловых структур и систем. Одно из существенных следствий постановки и разработки такой проблемы – создание оснований для масштабного доформального и формального изучения смысловых процессов, включая творчество, принятие решений и управление. Выделяя в качестве основных смысловых процессов смыслообразование и смыслоосознание, мы стремимся содержательно раскрыть их механизмы, показывая, что за динамикой этих процессов стоят автоматизмы среды, такие как грануляция, обобщение, самоорганизация, естественный отбор и другие.

Само понятие «созревание» является частным случаем понятия «развитие» и представляет сложный интегральный процесс, включающий как прогресс, так и регресс. Действительно изменения, происходящие в предметах и явлениях в процессе развития, имеют две тенденции: стремление к усложнению организации и одновременно к упрощению. Тенденция к усложнению равносильна накоплению информации, а тенденция к упрощению означает уменьшение информации и накопление энтропии. Следовательно, созревание когнитивной структуры (протоструктуры) – это одновременно и усложнение, и упрощение структуры, но при этом доминирующими являются тенденции усложнения, т.е. накопления информации и снижения энтропии.

В каждом классе природных, технических, материальных, когнитивных систем преимущественно развиваются те системы и структуры, которые при данной совокупности внутренних и внешних условий достигают минимального уровня *Баланса энтропий* – минимального совокупного уровня несовершенств в шкалах «несовершенство – совершенство» (максимум КПД, надежности, производительности, долговечности, спроса; минимум стоимости, эксплуатационных издержек, энергии управления, отказов и т.п.). Общий текущий Баланс энтропий устанавливается в результате рефлексии, развития, самоорганизации внутренних процессов и естественного отбора лучших структур и схем процессов.

Таким образом, требование перманентного (ситуативного) снижения общего Баланса энтропий является тем принципом оптимальности, который задает текущее *характеристическое направление* вектору созревания–изменения когнитивных структур. Ситуативность означает неравновесность процесса созревания: то, что является удовлетворительным (зрелым) в одной ситуации действительности, в другой ситуации может оказаться не вполне удовлетворительным и,

следовательно, требующим изменения, адаптации, т.е. нового витка созревания. Характеристическое направление задается потоком внутренней энергии, стимулирующим определенные виды автоматизмов среды.

Под «структурой» будем понимать «зрелое» образование, удовлетворяющее (неравновесному) принципу оптимальности (Балансу энтропий). Все промежуточные формы укладываются в следующую цепочку:

Хаос → Хаотический ансамбль → Протоструктуры → Структуры

В первом приближении *Хаос* – очень большое число полностью изолированных и поэтому не зависящих друг от друга структур, находящихся в различных динамических состояниях (структуры, протоструктуры, субструктуры, обломки структур и т.д.). Хаос поддерживается непрерывным и всеобъемлющим действием автоматизмов среды.

*Хаотический ансамбль* – это очень большое число полностью изолированных, не зависящих друг от друга движений (трансформаций) по характеристическому направлению находящихся в различных динамических состояниях элементов когнитивного пространства. Смыслом движений является уменьшение энтропии в ансамбле (в энтропийном многообразии ансамбля) за счет возникновения определенных протоструктур и структур.

Первоначально в хаотическом ансамбле созревают протоструктуры. К протоструктурам элементарного теста относятся: прототесты, протодомены, протографы доменов.

В качестве прототеста может выступать любое объективное или субъективное качество, например, любое сходство между объектами (тепло, твердость, доброта и т. д.). Если в прототесте не обнаруживаются дополнительные измерения, то он переходит в тест.

Протодоменом прототеста является выделенная субъектом неоднородность качества. С помощью грануляции (автоматизм среды) осуществляется деление неоднородности на различающиеся между собой части – гранулы. Под гранулой понимается группа элементов, объединяемых неразличимостью, сходством, близостью (т.е. отношениями, обладающими, по крайней мере, свойствами симметричности и рефлексивности). Элементы гранул берутся из хаотического ансамбля. Более детально вопросы грануляции – дискретизации областей значений теста будут рассмотрены в главе 4.

На множестве всех протодоменов между парами протодоменов могут возникать ассоциативные и обобщающие связи (вначале на уровне элементов - гранул, затем между протодоменами). Так возникают

протографы протодоменов прототестов. Дальнейшие процессы грануляции и обобщения на всех уровнях строительства (созревания) с целью минимизации структурной энтропии и удовлетворения принципу оптимальности приводят к возникновению структур – тестов, доменов, оргграфов доменов.

На рис 2.23 приведены последовательные фазы созревания оргграфа доменов теста: вначале созревает ствол графа (коричневые вершины), затем листья (зеленые и желтые вершины).

В главе 5 при рассмотрении метода предельных обобщений будет показано, что созревание лишь одной новой вершины оргграфа доменов какого-либо теста (одного теста из сотен или тысяч) может привести к обрушению модели знаний и возникновению другой. Подобную катастрофическую реорганизацию «картины мира» можно интерпретировать как внезапное «озарение».

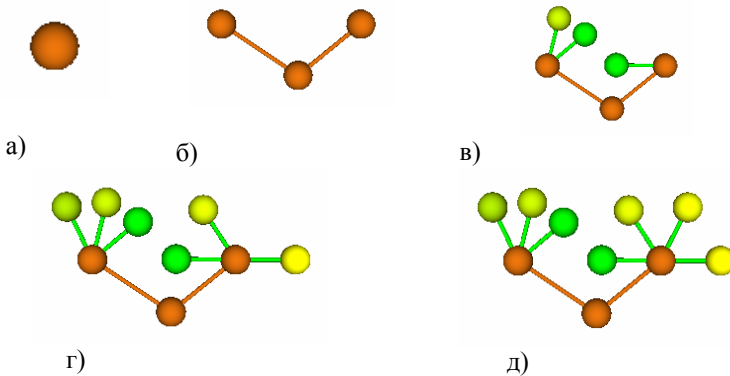


Рис. 2.23 – Динамика созревания оргграфа доменов теста

В главе 6 при рассмотрении вопросов квантовой семантики будет введена операция деления соседней пары элементов домена и на основе этой операции рассмотрена иная концепция созревания оргграфа доменов. В рамках этой модели созревание начинается с оппозиции «качество+ – качество-», при этом допускается существование нескольких базовых доменов. Данная модель имеет большую эволюционную обоснованность.

В онтологической теории интеллекта устройство интеллекта описывается с позиций трех базовых категорий: «ментальные структуры», «ментальное пространство», «ментальная репрезентация» [184]. *Ментальные структуры* – это устойчивая форма ментального опыта в виде системы психических образований, которые в условиях познавательного контакта субъекта с действительностью обеспечивают возможность

поступления информации о происходящих событиях и ее преобразование, а также управление процессами переработки информации и избирательность интеллектуальной деятельности.

Орграфы доменов тестов с полным основанием могут быть отнесены к базовым ментальным структурам. Совокупность орграфов доменов всех тестов обеспечивает сравнение сенсорных сигналов, описанных в терминах базовых доменов, со сведениями, извлеченными из памяти, что, как можно предполагать, и лежит в основе перехода физиологического процесса на уровень психических, субъективно переживаемых явлений. В сочетании с методом предельных обобщений и моделью многоуровневых набросков ситуаций/образов эти представления могут служить формальной моделью *гипотезы информационного синтеза*. Нейрофизиологический вариант подобной гипотезы был развит в работах А. М. Иваницкого [67].

В целом от развитости структур орграфов доменов тестов напрямую зависят когнитивные способности наблюдателя (интеллектуального агента). Ключевая гипотеза состоит в том, что развивающийся Банк тестов  $\{G(\tau)\}$  может выступать в качестве *системы координат феноменологического пространства* (рис. 2.24).

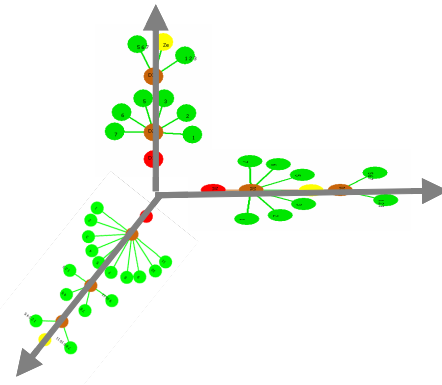


Рис. 2.24 – Система координат феноменологического пространства

Непрерывное развитие Банка тестов как в количественном, так и в качественном отношении обуславливает неравновесность критических когнитивных структур, включая предельные модели знаний. В последующих главах контекст задачи, включающий Банк тестов, будет играть решающую роль.

### ГЛАВА 3. МОДЕЛЬ ОПИСАНИЯ СИТУАЦИЙ ДЕЙСТВИТЕЛЬНОСТИ

Рассматриваются ситуации, возникающие в естественных предметных областях. Это могут быть организационно-технические, чрезвычайные, военные, политические, экономические, криминальные, экологические, клинические, учебные и иные ситуации.

#### 3.1 Свойства тестов. Лексика. Механизмы реализации тестов

Глава 2 была целиком посвящена изучению схем обобщения произвольного теста. Однако кроме схем обобщения элементарный тест, как и любая сущность, может обладать и другими свойствами, зависящими от предметной области. Множество всех свойств теста  $\tau$  обозначим: СВОЙСТВА( $\tau$ ). Приведем примеры описания свойств.

В работе [134] для предметной области «Клиническая медицина» были приведены основные свойства сущности «Элементарный диагностический тест (Д-ТЕСТ)»:

СВОЙСТВА( $\tau$ . ЭЛЕМЕНТАРНЫЙ Д-ТЕСТ) = {  
 Название теста;  
 Эквивалентные названия теста;  
 Коды теста по классификаторам (HL7, LOINC, ICD,...);  
 Описание теста;  
 Условия применимости (пол, возраст, клиническая картина);  
 Клиническое значение;  
 Лексика  $\{LT(\tau)\}$ ;  
 Орграфы доменов  $\{G(\tau)\}$ ;  
 Механизмы реализации (для каждого домена);  
 Метрологические требования (для каждого механизма);  
 Интерпретация теста (для элементов домена);  
 Рекомендации по  $\{\tau\}$  (для элементов домена)}.

Как видим, орграф доменов является лишь одним из свойств любого элементарного теста. Примеры описания свойств тестов (использован синтаксис лексических деревьев [4, 134]):

ЖЕЛТУХА: Есть; Нет. {Желтуха

*Эквивалентные названия* {

Желтая окраска кожных покровов, склер и слизистых}

*Клиническое значение* {

Повышенный гемолиз (распад эритроцитов)  
 Патология печеночных клеток  
 Нарушение эвакуации желчи  
*Механизмы реализации* {  
 Субъективная оценка.  
 Концентрация билирубина в крови? → Тест? (Пр\_756)}  
*Рекомендации по тестам* {  
 Если *Есть*, то выполнить АДД\_Желтуха.  
 Если *Есть*, то выполнить {Размеры печени / {Норма; Увеличены}}?;  
 Характер края печени / {Твердый; Мягкий; Бугристый} ?; ... }  
 } Желтуха  
 #АДД – алгоритм дифференциальной диагностики  
 ХАРАКТЕР КРАЯ ПЕЧЕНИ: Твердый; Мягкий; Бугристый. {КрайПечени  
*Условия применения* {  
 Размеры печени? Увеличены}  
*Интерпретация* {  
 Если *Твердый* & (Желтуха? *Есть*), то {↑◇Цирроз печени}.  
 Если *Мягкий* & (Желтуха? *Есть*), то {↑◇Гепатит}.  
 Если *Бугристый* & (Желтуха? *Есть*), то {↑◇Рак печени; ↑◇Цирроз  
 печени}}  
 } КрайПечени  
 #символ ‘↑◇’ означает повышение вероятности (возможности).  
 ЛИХОРАДКА: *Есть*; Нет. {Лихорадка  
*Эквивалентные названия* {  
 Повышение температуры выше 37.0°C}  
*Клиническое значение* {  
 Является следствием синтеза пирогенных веществ в организме при  
 любом воспалительном процессе, интоксикациях и нарушениях  
 терморегуляции}  
*Механизмы реализации* {  
 Субъективная оценка  
 Температура тела / T1? → Лихорадка? (Пр\_327)  
 Температура тела / T2? → Лихорадка? (Пр\_328)  
 Температура тела / T3? → Лихорадка? (Пр\_329) }  
 } Лихорадка

Комментарий 1: Синтаксис лексических деревьев выбран с целью демонстрации одного из простых (для пользователя) и вместе с тем достаточно универсального варианта процедурной реализации свойств. Не

представляет трудностей переписать свойства в XML формате, но тогда описание потеряет свою наглядность для пользователя. На процедурном уровне орграфы доменов также могут быть представлены с помощью лексических деревьев (см. главу 4).

Комментарий 2: Пр\_N – некоторые правила из базы умений БЗ ИС. Структура правил одинакова: Вход; Выход; Описание назначения; Тело правила. Правила представляют собой механизмы  $\mu$  реализации системопаттернов вида  $f/\mu: \{J_a \underline{a}/A\} \rightarrow \{J_b \underline{b}/B\}$ . Подобные правила могут входить и в соглашения орграфа доменов  $G(\tau)$ .

Очень важно, чтобы механизмы реализации включали все доступные способы определения значений тестов. Пусть, например, в механизме присутствует расчетная схема *Температура тела/T1?*  $\rightarrow$  *Лихорадка?*. Это означает следующее: если в системе каким-либо образом определен результат теста *Температура тела/T1?*, то с помощью автоматизмов среды будут определены результаты всех связанных тестов, в частности, *Лихорадка /{Есть; Нет}?*, *Температура тела /T2,T3?* и т.д. Использование свойства транзитивности при построении замыкания (каскада правил преобразования) позволяет избавиться от необходимости перечислять для каждого домена все возможные механизмы.

Пример теста с доменом «Текст» (произвольный текстовый блок):

**Тест:** Заболевание, протекающее с рвотой в сочетании с сухим и (или) обложенным языком/Текст?

*Механизмы реализации:*

1. АДД «Рвота в сочетании с сухим и (или) обложенным языком»
2. Лексическое дерево {Рвота\_в\_сочет\_с\_сух\_яз}.
3. {Головная боль/{Есть; Нет}; Начало и течение болезни /{Острое; Хроническое}; Понос /{Есть; Нет}; Напряжение брюшной стенки /{Есть; Нет}; Боль при пальпации живота /{В эпигастрии; диффузная}}  $\rightarrow$  Тест? (Пр\_1205) .
4. Консультация гастроэнтеролога.

Свойство «Рекомендации по  $\{\tau\}$ » позволяет реализовать сколь угодно сложные алгоритмы управления (диагностические, прогностические, лечебные). Пример описания данного свойства для произвольного теста  $\tau$  со значениями  $\{\underline{z}_1, \dots, \underline{z}_n\}$ :

Если  $\underline{z}_1 [\wedge \{\underline{z}'\}_1]$ , то рекомендуется выполнить  $\{\tau\}_1$ ;

...

Если  $\underline{z}_n [\wedge \{\underline{z}'\}_n]$ , то рекомендуется выполнить  $\{\tau\}_n$ .

В приведенной схеме  $\{\underline{z}'\}$  означают результаты либо ранее проведенных тестов, либо тестов, входящих в одну батарею (агрегат) с данным тестом.



Зачастую,  $\{\underline{\tau}\}$  характеризуют особенности клинической картины, такие как пол, возраст, наличие беременности, сопутствующих заболеваний и т.д.

Свойство «Интерпретация теста» можно реализовать с помощью набора импликаций:

$$\underline{\tau}_1 [\wedge \{\underline{\tau}\}_1] \rightarrow \{J_d \underline{d}/D\}_1; \dots; \underline{\tau}_n [\wedge \{\underline{\tau}\}_n] \rightarrow \{J_d \underline{d}/D\}_n,$$

где  $J_d$  – оператор оценки истинности. Указанный набор импликаций участвует в процессе построения замыкания произвольного (заданного) множества значений тестов. В некоторых случаях можно использовать упрощенную систему модальностей [134]:

- $\underline{\tau} \rightarrow \square \underline{d}/D$  – сильная положительная связь (если имеет место  $\underline{\tau}$ , то всегда имеет место  $\underline{d}/D$ );
- $\underline{\tau} \rightarrow \square \neg \underline{d}/D$  – сильная отрицательная связь (если имеет место  $\underline{\tau}$ , то всегда отсутствует  $\underline{d}/D$ );
- $\underline{\tau} \rightarrow \uparrow \diamond \underline{d}/D$  – слабая положительная связь (если имеет место  $\underline{\tau}$ , то возрастает вероятность  $\underline{d}/D$ );
- $\underline{\tau} \rightarrow \downarrow \diamond \underline{d}/D$  – слабая отрицательная связь (если имеет место  $\underline{\tau}$ , то уменьшается вероятность  $\underline{d}/D$ ).

Пример.

**Тест:** Боль в эпигастрии / {Есть; Нет}?

*Интерпретация:*

Если *Есть*  $\wedge$  (Пол? М), то  $\{\square d_1; \square d_2; \uparrow \diamond d_3; \downarrow \diamond d_4; \dots\}$ .

Если *Есть*  $\wedge$  (Пол? Ж)  $\wedge$  (Беременность? Есть), то  $\{\uparrow \diamond d_1; \uparrow \diamond d_2; \dots\}$ .

Если *Есть*  $\wedge$  (Возраст/В3? Пожилой), то  $\{\uparrow \diamond d_3; \dots\}$ .

Если *Нет*, то  $\{\square \neg d_1; \downarrow \diamond d_2; \downarrow \diamond d_3; \square d_4; \dots\}$ .

*Рекомендации:*

Если *Есть*, то выполнить {Интенсивность?, Характер?, Иррадиация?, Чем купируется?, ...}.

Если *Есть*, то выполнить АДД\_Боль\_в\_эпиг.

Модальности могут накапливаться по мере выполнения батареи тестов, например, если в результате выполнения предыдущих тестов имела место модальность  $\uparrow n \diamond d$ , а значение  $\underline{\tau}$  связано с  $d$  слабой отрицательной связью, то в итоге будем иметь  $\downarrow \uparrow n \diamond d = \uparrow n-1 \diamond d$ . Модальность  $J \in [\downarrow n \diamond; \uparrow n \diamond]$  является элементарным тестом со своим орграфом доменов.

Приведем пример из предметной области «Химия» [15, 16, 24]. Пусть  $\tau = \langle \text{ХИМ\_ЭЛЕМЕНТ} \rangle$ . Положим:

$D1 = \{H, Li, Be, B, C, N, O, F, Ne, He, Na, Mg, Al, Si, P, Ca, \dots\}$ ;

$D2 = \text{ОДЗ(ХИМ\_ГРУППА)} = \{Ia, Ib, IIa, IIb, IIIa, IIIb, IVa, IVb, Va, Vb, VIa, VIb, VIIa, VIIb, VIIIa, VIIIb\}$ ;

$D3 = \{I, II, III, IV, V, VI, VII, VIII\}$ .

$G(\text{ХИМ\_ЭЛЕМЕНТ}) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

$\text{СВОЙСТВА } (\tau \text{ ХИМ\_ЭЛЕМЕНТ}) = \{\text{АТОМНЫЙ\_НОМЕР, АТОМНЫЙ\_ВЕС, ХИМ\_ГРУППА, ХИМ\_ПЕРИОД, \dots, G(\tau)\}$ .

Рассмотрим теперь варианты формирования значения  $\tau T$  для случая, когда  $T$  – «Текст», т.е. не является числовым интервалом или совокупностью элементов. Примеры подобных тестов: «Жалобы», «Анамнез», «Объективно», «Консультация врача» и т.д. В простейшем случае значение формируется путем обычного набора текста. В рамках госпитальных информационных систем было предложено формировать значение таких тестов с помощью лексических деревьев [4], что значительно ускоряет процедуру фиксации заключения. Подробное рассмотрение этого вопроса выходит за рамки настоящего исследования, поэтому приведем лишь фрагмент лексического дерева:

#### ЖАЛОБЫ НА

|постоянную; приступообразную| !головную боль {головная\_боль

|умеренную; интенсивную|

головную боль

|возникающую; усиливающуюся| при

повышении давления

физической нагрузке

перемене положения |головы; тела|

< ? >

в течение \_\_\_\_\_ |лет; года; мес; дн|}

|постоянное; приступообразное| !головокружение {головокружение}

Формирование значения теста (заключения) на основе лексического дерева осуществляется с использованием лексического процессора [4]. Любой тест  $\tau$  может иметь несколько лексических деревьев  $\{LT(\tau)\}$ . С помощью лексического дерева  $LT(\tau)$  формируется значение теста на уровне базового домена орграфа доменов. В дальнейшем данное значение может быть многократно обобщено с помощью специальных операторов обобщения, например тех, что позволяют создать орграф набросков  $G_s(\tau)$ . Подробнее этот вопрос будет рассматриваться в разделе 3.5.

Лексические деревья позволяют не только составить протокол, но и реализовать произвольный древовидный алгоритм дифференциальной диагностики или прогнозирования. Пример записи алгоритма дифференциальной диагностики с использованием синтаксиса лексических деревьев [111, 134]:

Острая боль в эпигастральной области {АДД\_Боль\_в\_эпиг

СВЯЗЬ БОЛИ ИЛИ ЕЕ ЭКВИВАЛЕНТОВ С ПРИЕМОМ ПИЩИ?

Есть [
   
     Н.в. Патология органов брюшной полости / А.Нет
   
     Возможна ИБС на фоне заболеваний органов брюшной полости]
   
 Нет {А.Нет
   
     СВЯЗЬ БОЛИ ИЛИ ЕЕ ЭКВИВАЛЕНТОВ С ФИЗ.НАГРУЗКОЙ
   
     ИЛИ ПСИХОЭМОЦ. СТРЕССАМИ В НАМНЕЗЕ?
   
     Есть {
   
         КОРОНАРОЛИТИЧЕСКАЯ ЗАВИСИМОСТЬ?
   
         Есть [
   
             Ds: Стенокардия / ЧД\_294]
   
         Нет [
   
             Ds: Инфаркт миокарда / ЧД\_295/ РП\_ИнфМиокарда]}
   
     Нет {
   
         ЧУВСТВО РЕЗКОЙ СЛАБОСТИ НА ФОНЕ И ПОСЛЕ
   
         ПРЕКРАЩЕНИЯ БОЛИ?
   
         Есть [
   
             Ds: Инфаркт миокарда /АДД\_ЭКГ / ЧД\_295]
   
         Нет [
   
             Патология органов брюшной полости /АДД\_БолиЖивоте]
   
     }}

Данный пример наглядно показывает, что лексическое дерево содержит набор тестов с описанием множеств значений (доменов). Следовательно, из ЛД при необходимости можно извлечь список используемых тестов, а именно:  $LT(\tau) \rightarrow \{p/P\}$ . Зная значения  $\{p/P\}$ , можно установить значение  $\tau$ . Другими словами, лексика является частью механизмов реализации.

Для удобства определения состава тестов в лексике, включая алгоритмы дифференциальной диагностики, прогнозирования и оптимизации управления, введем специальную функцию

$$\text{СОСТАВ}(LT(\tau)) = \{p/P\}.$$

Как видим, механизм установления значения элементарного теста (механизм реализации) может быть связан с выполнением значительной совокупности вспомогательных тестов, что влечет за собой расход разного рода ресурсов, включая время. В целом можно говорить о том, что реализация теста предполагает выполнение определенного когнитивно-поведенческого паттерна (или нескольких паттернов одновременно).

Любой тест  $\tau$  может иметь один или несколько механизмов реализации, информационные модели которых обозначим  $\{\mu\}_\tau$ . При выполнении любого теста  $\tau$  с помощью механизма реализации  $\mu$  должны быть соблюдены определенные метрологические требования (условия), которые обозначим через  $c_{\tau\mu}$ . Приведем пример:

ГЛЮКОЗА КРОВИ:  $c_{\tau\mu} = \{\text{анализ проводится натощак; в течение последних 24 часов не было приема лекарственных средств, влияющих на уровень глюкозы в крови}\}$ .

На основе приведенных выше примеров можно составить общее описание свойств произвольного теста  $\tau$ .

$$\begin{aligned} \text{СВОЙСТВА}(\tau) = \{ & \hspace{15em} (3.1) \\ & \{a/A\} \quad \text{– Свойства – тесты;} \\ & \{\underline{u}/U\} \quad \text{– Условия применимости } \tau, \\ & \{LT(\tau)\} \quad \text{– Лексика;} \\ \text{СОСТАВ}(LT(\tau)) & \text{– Список тестов, используемых в лексике;} \\ & \{Gs(\tau)\} \quad \text{– Орграфы набросков значений теста;} \\ & \{G(\tau)\} \quad \text{– Орграфы доменов;} \\ & \{G(\Delta\tau)\} \quad \text{– Орграфы доменов приращений теста;} \\ & \{f/\mu_f: \{b/B\} \rightarrow \underline{d}/T\} \quad \text{– Механизмы реализации;} \\ & \forall \mu_\tau \rightarrow c_{\tau\mu} \quad \text{– Метрологические требования;} \\ & \{g/\mu_g: \underline{d}/T, \{\underline{c}/C\} \rightarrow J_d \underline{d}/D\} \quad \text{– Интерпретация;} \\ & \{h/\mu_h: \underline{d}/T, \{\underline{r}/R\} \rightarrow \{\underline{d}/T\}'\} \quad \text{– Рекомендации по тестам.} \end{aligned}$$

Здесь  $f/\mu_f$ ,  $g/\mu_g$ ,  $h/\mu_h$  – некоторые известные системопаттерны со своими механизмами реализации;  $\Delta\tau$  – производный тест, позволяющий выявлять качественные прерывности значений теста  $\tau$  (см. раздел 3.2). Орграфы набросков рассматриваются в разделе 3.5 и главе 6. Можно записать:

$$\begin{aligned} & \forall \tau \rightarrow \text{СВОЙСТВА}(\tau); \\ & \text{Банк элементарных тестов} = \{\tau, \text{СВОЙСТВА}(\tau)\}. \end{aligned}$$

В идеале Банк тестов должен быть замкнут относительно всех ссылок, содержащихся в (3.1).

На основе Банка тестов могут быть реализованы разнообразные вычислительные сети и среды. Более подробно этот вопрос будет затронут при рассмотрении архитектуры Многоцелевых банков знаний (глава 7).

Очень часто необходимость проведения одних тестов зависит от результатов других тестов. Так, если «ГОЛОВНАЯ БОЛЬ? отсутствует», то больше никаких уточняющих тестов проводить не нужно. Однако, если «ГОЛОВНАЯ БОЛЬ? имеется», то необходимо уточнить локализацию боли, ее характер, периодичность, которые также являются тестами. Возникает проблема процедурного представления взаимосвязанной совокупности орграфов доменов тестов.

В самом простом случае, если орграф доменов теста «ТЕСТ» состоит из одного базового домена, то при наличии уточняющих тестов на

процедурном уровне его можно представить следующим образом:

*ТЕСТ*: элем1 [{Уточн. тесты}];...; элемK [{Уточн. тесты}], (3.2)

где 'элем<sub>k</sub>' – k-й элемент базового домена; [{Уточн. тесты}] – необязательное множество уточняющих тестов каждого элемента домена. Каждый уточняющий тест, в простейшем случае, имеет ту же структуру.

В качестве примера приведем онтологическое описание теста «ГОЛОВНАЯ БОЛЬ» с использованием синтаксиса лексических деревьев (на основе первичных материалов [188]):

ГОЛОВНАЯ БОЛЬ: {Гол\_боль  
отсутствует;  
имеется {

ЛОКАЛИЗАЦИЯ: лобная область; теменная область;...

ХАРАКТЕР: тупая; острая; жгучая;...

ПЕРИОДИЧНОСТЬ: постоянная; приступообразная.

ИРРАДИАЦИЯ: отсутствует; в глаза справа; в глаза слева;...

УСЛОВИЯ ВОЗНИКНОВЕНИЯ: [Услов\_возникн]

ПРОДОЛЖИТЕЛЬНОСТЬ/ ДТ

ВРЕМЯ ВОЗНИКНОВЕНИЯ: после сна; ночью; в первой половине дня; во второй половине дня.

ЧАСТОТА ВОЗНИКНОВЕНИЯ/ ЧВ

ИНТЕНСИВНОСТЬ: нестерпимая; очень резкая; резкая; сильная;...

УСИЛЕНИЕ: в вертикальном положении тела; в горизонтальном положении тела; при переходе из одного положения в другое; при ходьбе;...

УМЕНЬШЕНИЕ: после сна; после отдыха; после приема анальгетиков;...

СОПРОВОЖДАЕТСЯ: {Сопровожд}

}} Гол\_боль

[Услов\_возникн

умственная нагрузка; физическая нагрузка; стрессовые ситуации; ломка привычных стереотипов; яркий свет; громкие звуки; резкие запахи; ...]

{Сопровожд

ОБЩЕМОЗГОВЫЕ СИМПТОМЫ: вялость; бессонница; сонливость; тошнота; рвота; головокружение; шум в ушах.

НАРУШЕНИЯ ЧУВСТВИТЕЛЬНОСТИ: {

РАССТРОЙСТВО ЗРЕНИЯ: {

СТОРОНА: справа; слева.

ХАРАКТЕР: сверкающие точки; сверкающие шары; огненные фигуры; молниеподобные вспышки; мелькающие зигзаги; диплопия}

ДРУГИЕ...}

**МОТОРНЫЕ НАРУШЕНИЯ:** нарушение речи; частичная моторная афазия; полная моторная афазия; атаксия; мышечные спазмы.

**ВЕГЕТАТИВНЫЕ НАРУШЕНИЯ:** полиурия; полидипсия; боли в животе; усиленная перистальтика кишечника.

**АЛЛЕРГИЧЕСКИЕ ПРОЯВЛЕНИЯ:** крапивница; отек Квинке; астма}.

В приведенном примере [Услов\_возникн] – домен, описание которого вынесено за рамки лексического дерева (на этот домен могут ссылаться и другие тесты); {Сопровожд} – ссылка на другое лексическое дерево – составной тест (таких вложенных ссылок может быть сколь угодно много); DT и ЧВ – базовые домены орграфов тестов. Подчеркнем, что любой тест может быть представлен не одним доменом, а орграфом.

Для корректного отражения логики уточняющих тестов во множестве результирующих данных можно использовать следующую нотацию:

$$\{\underline{t}/T\} = \{\underline{t}_1/T_1 [\{\underline{t}/T\}_1]; \dots; \underline{t}_n/T_n [\{\underline{t}/T\}_n]\}, \quad (3.3)$$

где  $\{\underline{t}/T\}_j$  – множество результатов уточняющих тестов к  $j$ -му тесту  $\underline{t}/T_j$ . Уровней уточнения может быть сколь угодно много. Пример:

$\{\underline{t}/T\} = \{\text{Головная боль? имеется} \{\text{Локализация? лобная область; Характер? тупая; ...}\}; \text{ЧСС? 77; Онемение? имеется} \{\text{Тип? тактильное; Интенсивность? умеренная; ...}\}; \text{Снижение зрения? имеется} \{\text{Глаз? правый} \{\text{Характер? вдаль; ...}\}; \text{Глаз? левый} \{\text{Характер? вблизи; ...}\}\}\}.$

Не представляет труда онтологическое описание теста представить на языке XML или другом специализированном языке. Выбор синтаксиса лексических деревьев определяется его предельной простотой, наглядностью и одновременно гибкостью [4].

### 3.2 Тест «Время». События

Пусть  $t/A$  – выделенный тест, означающий время (абсолютное).  
Примеры доменов  $A$ :

$A1 = \text{'Дата: время'}$ ;

$A2 = \text{'Дата: \{утро, день, вечер, ночь\}'}$ ;

$A3 = \text{'Дата'}$ ;

$A4 = \text{'Месяц, Год'}$ ;

$A5 = \text{'Год'}$ .

Возможны и другие домены. Конкретное значение времени согласно принятой нотации будем обозначать:  $\underline{t}/A$ . Примеры значений времени:

$\underline{t}/A1 = \text{'10.01.09: 08.30'}$ ,  $\underline{t}/A2 = \text{'10.01.09: утро'}$ ,  $\underline{t}/A3 = \text{'10.01.09'}$ ,

$t/A4 = \text{'январь, 2009'}$ ,  $t/A5 = \text{'2009'}$ .

Вполне очевиден способ задания пересчета значений из одного домена в другой (с большим номером), т.е. определен оргграф доменов:

$$G(t) = \{A1 \rightarrow A2 \rightarrow A3 \rightarrow A4 \rightarrow A5\}.$$

Уточнения требует лишь связка  $A1 \rightarrow A2$ . Определить правила пересчета времени для данной связки можно, например, на основе нечетких множеств (термы 'утро', 'день', 'вечер', 'ночь' можно рассматривать как нечеткие множества). Более простой способ задается следующими правилами:

'00.00 – 6.00' → 'ночь'; '6.00 – 10.00' → 'утро';  
'10.00 – 18.00' → 'день'; '18.00 – 24.00' → 'вечер'.

Пример другого оргграфа доменов  $G(t)$ :

$A1 = \text{'Дата: время'}$ ;

$A2 = \{1; 2; \dots; 24\}$ ;

$A3 = \{1; 2; \dots; 7\}$ .

$G(t) = \{A1 \rightarrow A2; A1 \rightarrow A3\}$ .

Правила пересчета:

$A1.t \rightarrow A2.Hour(t)$  – текущий час;

$A1.t \rightarrow A3.DayOfWeek(t)$  – текущий день недели.

Определим два элементарных события:

$$e_1 = \langle \underline{t}/T, \underline{t}/A \rangle \text{ и } e_2 = \langle \neg \underline{t}/T, \underline{t}/A \rangle. \quad (3.4)$$

Событие  $e_1$  означает, что в момент времени  $\underline{t}/A$  тест  $\underline{t}/T$  имел значение  $\underline{t}/T$ . Событие  $e_2$  означает, что в момент времени  $\underline{t}/A$  тест  $\underline{t}/T$  не имел значение  $\underline{t}/T$ . Положим:  $e_1 = \neg e_2$ ,  $e_2 = \neg e_1$ . Примеры элементарных событий:

$\langle \text{Кашель}/\{\text{Имеется; Отсутствует}\} ? \text{Имеется}, \underline{t}/A1 ? 23.02.08: 12.45 \rangle$ ,

$\langle \neg \text{'Le крови'}/\{\text{Снижено; Норма; Повышено}\} ? \text{Норма}, \underline{t}/A3 ? 23.02.08 \rangle$ ;

$\langle \text{Ток солнечных батарей}/\{\text{Низкий; Рабочий}\} ? \text{Низкий}, \underline{t}/A1 ? 13.05.09: 09.15 \rangle$ ;

$\langle \text{Оценка}/\{1; 2; 3; 4; 5\} ? 4, \underline{t}/A3 ? 2.07.10 \rangle$ .

$\langle \text{Риск инвестиций}/\{\text{Высокий; Средний; Низкий}\} ? \text{Высокий}, \underline{t}/A5 ? 2008 \rangle$ ;

$\{c(t)\} = \{c, t\}$  – произвольный временной ряд.

Любое событие может быть поименованным.

Шаблоном события назовем кортеж  $\langle \underline{t}/T, \underline{t}/A \rangle$ . Шаблоны событий задаются, как правило, перед началом мониторинга ситуации действительности.

Событие  $\langle \underline{t}/T, \underline{t}/A \rangle$  можно представить как результат теста: «Значение  $\underline{t}/T$  в момент  $\underline{t}/A = \underline{t}$ ». Подобная трактовка позволяет упростить многие

определения.

Если для теста  $\tau$  определен оргграф доменов  $G(\tau)$ , а для времени оргграф  $G(t)$ , то событие  $\langle \underline{t}/T, \underline{t}/A \rangle$ , описанное с помощью базовых доменов, автоматически порождает множество событий  $\{\langle \underline{t}/T', \underline{t}/A' \rangle\}$  общей численностью  $N_\tau \times N_t$ , где  $N_\tau$  – общее количество доменов в оргграфе  $G(\tau)$ , а  $N_t$  – общее количество доменов в оргграфе  $G(t)$ . Этот факт имеет важное значение для выстраиваемого формализма.

На основе элементарных событий можно определить *составные события* следующим образом:

$$e = \langle \underline{t}/T \rangle, \underline{t}/A \rangle \text{ или } e' = \langle \{[-]\} \underline{t}/T \rangle, \underline{t}/A \rangle.$$

Для составных событий эквивалентны записи:

$$e = \{e_1, e_2, \dots, e_n\} = \langle \underline{t}/T, \underline{t}/A \rangle = \langle \{ \underline{t}/T \}, \underline{t}/A \rangle.$$

Однако обратная операция – разделение составных событий на независимые элементарные события – не всегда возможна, так как составные события часто объединяют *батарею тестов*, т.е. те тесты, которые неразрывно связаны между собой.

Примеры составных событий: жалобы (с указанием времени фиксации жалоб), осмотр (с указанием времени осмотра), анализ крови, параметры ЭКГ, АД, пульс и т.д. Так для АД и пульса можем записать:

АД1 =  $\langle \text{АДс}/N? 145, \text{АДд}/N? 93, \underline{t}/A1? 13.05.09: 12.45 \rangle$ ;

Пульс2 =  $\langle \text{Частота}/N? 70, \text{Ритмичность}/D1? \text{ритмичный}, \underline{t}/A2? 23.02.08: \text{утро} \rangle$ ;

Исследование пульса на лучевых артериях ( $\underline{t}/A2? 23.02.08: \text{утро}$ ): {  
ЧАСТОТА/ $N$ : 86;

РИТМИЧНОСТЬ/ $D1$  = {ритмичный; аритмичный}: ритмичный;

НАПРЯЖЕНИЕ/{среднего напряжения; твердый; мягкий}: твердый;

НАПОЛНЕНИЕ/{среднего наполнения; полный; пустой}: полный;

ФОРМА/{быстрый; скачущий; медленный}: быстрый;

ДЕФИЦИТ/{имеется; отсутствует}: отсутствует}.

Приведенные примеры объединяют батареи тестов, которые всегда выполняются пакетом. Заметим, что «Исследование пульса на лучевых артериях» следует трактовать как название составного события (батареи тестов). Каждый тест батареи тестов, как и любой тест, задается своим оргграфом. В приведенных выше примерах оргграфы тестов состоят из одного домена.

Если для разных времен  $\underline{t}/A$  повторяется состав тестов  $\{ \underline{t}/T \}$ , то составной тест  $\langle \{ \underline{t}/T \}, \underline{t}/A \rangle$  можно рассматривать как *кадр*. Примером может служить последовательность изображений (видеоряд, видеопоток),



где каждое изображение – это составной тест (с каждым пикселем связано либо одно число – код цвета, либо три числа - RGB).

Произвольная совокупность тестов (событий) может образовать поименованный *блок тестов* (событий). В рамках онтологии имя блока тестов или событий должно быть уникальным.

Для многих тестов, кроме их результатов, важной дополнительной (служебной) информацией являются, в частности, следующие параметры – тесты:

- место проведения теста (МестоПровТеста?);
- кем проводился тест (КемПровТест?);
- стоимость теста (СтоимТеста?);
- условия проведения теста (УсловПровТеста?);
- метрологические факторы (Метрология?);
- время фиксации теста в БД (ВремяФиксТестаБД?);
- достоверность результата теста (оператор - тест J).

Если через  $\{p/P\}$  обозначить дополнительные параметры – уточняющие тесты, то расширенная трактовка элементарных событий примет вид составных событий:

$$e_1 = \langle \underline{t}/T, \{p/P\}, \underline{t}/A \rangle \text{ и } e_2 = \langle \neg \underline{t}/T, \{p/P\}, \underline{t}/A \rangle. \quad (3.5)$$

Пример расширенного описания составного события:

АД1 = <АДс/Д2? [140; 150]; АДд/Д2? [85; 90]; МестоПровТеста/С1? Дом; КемПровТест/С2? Врач СМП; Метрология/М3? Выполнена;  $\underline{t}/A1?$  13.05.09: 12.45>.

Если событие локализовано к какой-либо точке пространства  $\{x/X, y/Y, z/Z\}$ , то можно записать:  $e = \langle \underline{t}/T, \{x/X, y/Y, z/Z\}, \underline{t}/A \rangle$ .

*Протяженным событием* назовем событие  $e = \langle \{t/T\}, \underline{\delta}/A \rangle$ , где  $\underline{\delta}$  – произвольный временной интервал (возможно, не односвязный). Для интервалов определена операция объединения:

$$e = \&_i \langle \{t/T\}, \underline{\delta}_i/A \rangle = \langle \{t/T\}, \underline{\delta}/A \rangle, \text{ где } \underline{\delta} = \&_i \underline{\delta}_i. \quad (3.6)$$

Пример протяженного события:

<Креатинин / {мг / 100 мл}? 1.3, [12.02.09; 14.02.09]/A3>,

где  $\underline{\delta} = [12.02.09; 14.02.09]$  (в терминах домена A3).

Единственным ограничением на описание любого события является недопустимость одновременного присутствия двух элементарных тестов  $e_1$  и  $e_2$  таких, что  $e_1 = \neg e_2$ .

Если для любого момента времени  $t$  в рассматриваемом периоде наблюдения значение некоторого теста  $\underline{t}/T$  не меняется, то такой результат назовем *фактом*. Так, например, перенесенные в прошлом операции или

травмы являются фактами, заболевания родителей (наследственность) являются фактами и т.д.

*Жизненным циклом* произвольного теста не являющегося фактом назовем упорядоченное по  $t$  множество  $\{<\underline{t}/T, \underline{t}/A> \mid \tau/T \text{ фиксировано, } A \text{ фиксировано}\}$ .

Автоматизмами вычислительной среды из любого события  $<\underline{t}/T, \underline{t}/A>$  формируется связанное событие-факт: *«Имело место  $\underline{t}/T$ »*. Событие-факт позволяет устранить привязку значения теста к конкретному времени, что упрощает во многих случаях установление формальной эквивалентности между совокупностями событий.

Пусть  $J$  – оператор оценки истинности/возможности той или иной информации. Если информация  $x$  имеет истинность  $\beta$ , то будем использовать нотацию:  $J_\beta x$  или  $J_\beta \underline{t}/T$ . С возрастанием уровня общности результата теста, как правило, меняется и значение истинности результата.

Оценка истинности  $J$  так же, как любой тест, может выражаться с помощью шкал разного уровня общности – от числовой шкалы ( $[0; 1]$ ) до лингвистической шкалы следующего вида:

- “абсолютно достоверно” (полная определенность);
- “скорее всего” (крайне высокая степень определенности);
- “весьма вероятно” (высокая степень определенности);
- “противоречивые или спорные сведения” (практически полная неопределенность);
- “мало вероятно” (высокая степень определенности);
- “сомнительно, но не исключено” или “крайне мало вероятно”, т.е. почти полное отрицание (очень высокая степень определенности);
- “абсолютно невозможно”, т.е. полное отрицание (полная определенность).

В общем случае будем предполагать, что для  $J$  задан оргграф доменов  $G(J)$ . С учетом оператора  $J$  произвольное событие можно представить в виде:

$$e = <J_\tau \underline{t}/T, \{J_p \underline{p}/P\}, J_t \underline{t}/A>. \quad (3.7)$$

Считается, что функция сознания в чем-то аналогична процессу дифференцирования, т.е. вычислению первой производной. Определим скорость изменения значения произвольного теста  $\tau$ .

Пусть известны два события  $<\underline{t}/T, \underline{t}/A>$  и  $<\underline{t}'/T, \underline{t}'/A>$  таких, что  $\underline{t}/A < \underline{t}'/A$ . Если выполняется условие:

$$|\underline{t}' - \underline{t}|/A' \leq \Delta_t/A', \quad (3.8)$$

где  $\Delta_t/A'$  – параметр, зависящий от теста, то определим новое событие

$$e' = \langle (\underline{\tau} \uparrow \underline{\tau}')/T, (\underline{t}' - \underline{t})/A', \underline{t}/A \rangle, \quad (3.9)$$

которое характеризует скорость изменения значения теста (упрощенное дифференцирование). Оператор изменения « $\uparrow$ » зависит от теста. Во многих случаях удобно ввести производный тест « $\Delta \tau$ », положив:  $\Delta \underline{\tau} T' = (\underline{\tau}' - \underline{\tau})/T'$ . Выражение (3.9) перепишется следующим образом:

$$e' = \langle \Delta \underline{\tau} T', \Delta \underline{t}/A', \underline{t}/A \rangle. \quad (3.10)$$

Для теста  $\Delta \tau$  должен быть определен оргграф  $G(\Delta \tau)$ , относящийся к свойствам теста  $\tau$ . Параметры  $\{\Delta \tau/A'\}$  также принадлежат к числу свойств теста  $\tau$ . Приведем примеры:

$\langle$ Креатинин? 1.3, 12.02.11/A3 $\rangle$ ,  $\langle$ Креатинин? 1.7, 13.02.11/A3 $\rangle$ ;

$e' = \langle$ Креатинин? 1.3 $\uparrow$ 1.7, сутки, 12.02.11/A3 $\rangle$ ;

$e' = \langle \Delta$ Креатинин? 0.4, сутки, 12.02.11/A3 $\rangle$ .

$\langle$ Адс? 130, 12.02.11:12.00/A1 $\rangle$ ,  $\langle$ Адс? 155, 12.02.11:15.00/A1 $\rangle$ ;

$e' = \langle$ Адс? 130 $\uparrow$ 155, 3 часа, 12.02.11:12.00/A1 $\rangle$ ;

$e' = \langle \Delta$ Адс? 25, 3 часа, 12.02.11:12.00/A1 $\rangle$ .

$e' = \langle \Delta$ Адс/D2? Большое,  $\Delta \underline{t}/A2$ ? Малое, 12.02.11:12.00/A1 $\rangle$ .

Не всегда изменение теста можно выразить в числовом виде. Пример:

$\langle$ МОКРОТА.Цвет? белая, 17.04.11/A3 $\rangle$ ;

$\langle$ МОКРОТА.Цвет? желтоватая, 18.04.11/A3 $\rangle$ ;

$e' = \langle$ МОКРОТА.Цвет? белая $\uparrow$ желтоватая, сутки, 17.04.11/A3 $\rangle$ ;

$e' = \langle \Delta$ МОКРОТА.Цвет/{Есть; Нет}? Есть, сутки, 17.04.11/A3 $\rangle$ .

С помощью производного события «Изменение» можно устанавливать качественные прерывности в динамике тестов. Такой «рекурсивный» подход открывает новое пространство признаков (тестов), явно не присутствующих в ПрО.

### 3.3 Структуры и их свойства

Реконструкция сложного явления, феномена, ситуации включает в себя построение законченных абстрактных качественно-смысловых форм – структур разного уровня общности. Основной тезис раздела заключается в следующем: любая сущность, явление, феномен, процесс или объект ПрО полностью описывается совокупностью элементарных тестов и системопаттернов на их основе, которые укладываются в определенную структуру (информационный объект). Другими словами, элементарный тест и системопаттерн являются теми кирпичиками, с помощью которых

строятся информационные образы всех структур.

Приведем примеры из химии. Структурная формула химического вещества обозначается кортежем, элементами которого, в свою очередь, являются кортежи – компоненты формулы; элементами последних являются имя, обозначающее химический элемент, одно из имен *Ox* или *Red* и два целых числа – степень окисления элемента и его индекс [15, 16]. Примеры формул веществ:

$$S = \langle\langle S, \text{Red}, 0, 1 \rangle\rangle; \quad O_2 = \langle\langle O, \text{Ox}, 0, 2 \rangle\rangle;$$

$$SO_2 = \langle\langle S, \text{Red}, 4, 1 \rangle, \langle O, \text{Ox}, -2, 2 \rangle\rangle.$$

Как видим, каждое химическое вещество определяется совокупностью элементарных тестов. Через структуру «Химическое вещество» можно определить более сложную субструктуру – «Химический процесс» (структуру из структур) [15]:

$$(v: \text{МОМЕНТЫ\_ВРЕМЕНИ})$$

$$\text{ХИМИЧ\_ПРОЦЕСС}(v) = \langle\{\langle\langle S, \text{Red}, 0, 1 \rangle\rangle, \langle\langle O, \text{Ox}, 0, 2 \rangle\rangle\},$$

$$\{\langle\langle S, \text{Red}, 4, 1 \rangle, \langle O, \text{Ox}, -2, 2 \rangle\rangle, \langle\langle O, \text{Ox}, 0, 2 \rangle\rangle\}\rangle.$$

Комментарий: на первом шаге процесса имеются вещества *S* и *O<sub>2</sub>*, на втором шаге – вещества *SO<sub>2</sub>* и *O<sub>2</sub>*. Как и химическое вещество, химический процесс полностью определяется совокупностью элементарных тестов.

Событие является примером простейшей структуры. Структуры «Текст» и «Изображение» представляются кортежами тестов:

$$\text{Текст} = \langle\{ \text{Символ? } a \} \rangle; \quad \text{Изображение} = \langle\{ \text{Pixel? } a \} \rangle.$$

Определим функцию СОСТАВ(), которая для произвольной структуры *Str* выдает список определяющих элементарных тестов (главных качеств):

$$\forall Str, \text{СОСТАВ}(Str) = \{ \tau/T \}.$$

Справедлива запись:  $Str(\{ \tau/T \}) \Leftrightarrow \text{СОСТАВ}(Str) = \{ \tau/T \}$ . Необходимо иметь в виду, что один и тот же тест  $\tau$  может входить в структуру с разными доменами:  $\tau/T$  и  $\tau/T'$  и это будут разные вхождения. Примеры:

$$\text{СОСТАВ}(\text{Событие } X) = \{ \tau/T, t/L \};$$

$$\text{СОСТАВ}(\text{Событие } Y) = \{ \tau/T, \{p/P\}, t/L \};$$

$$\text{СОСТАВ}(LT(\tau)) = \{p/P\};$$

$$\text{СОСТАВ}(\text{АН\_КРОВИ}) = \{ \text{Лейкоциты; Эритроциты; Гемоглобин; СОЭ; ...} \};$$

$$\text{СОСТАВ}(\text{ПАСПОРТ\_ДАННЫЕ}) = \{ \text{ФИО; Пол; Дата рождения; Серия паспорта; Номер паспорта; ...} \}.$$

Заметим, что не любая совокупность тестов образует структуру. Между тестами структуры должны существовать явные или скрытые отношения.

Так тесты структуры «Анализ крови» должны удовлетворять формуле крови. В структуре «Паспортные данные» нельзя произвольно заменить значение какого-либо теста (не будет реального субъекта с новыми параметрами).

Рассмотрим особенности описания структуры «Клинический диагноз». Пусть *Str* – «Язва желудка». Определим следующие тесты: J – эндоскопические проявления заболевания; L – локализация; K – наличие инвазии Нр; O – осложнения; F – клинические стадии развития перфоративной язвы; H – виды стеноза; P – пенетрация; I – эндоскопические признаки активности кровотечения по Forest; T – степень тяжести кровопотери. Формула диагноза «Язва желудка» [80]:

[J<sub>x</sub>] язва [L<sub>x</sub>] желудка [, K<sub>x</sub>] [, осложненная O<sub>x</sub> &].

Квадратные скобки в формуле означают необязательность элемента. Знак & означает возможность повторения блока данных. Подставив допустимые элементы из доменов характеристик, получим синтаксически и семантически правильное описание клинического диагноза. Если характеристики представлены орграфами доменов, то можно говорить о формуле разного уровня общности. В базе данных (базе прецедентов) структуру можно хранить с помощью *кодowego представления*, которое для язвы желудка может иметь следующий вид:

K25 ([J<sub>x</sub>]; [L<sub>x</sub>]; [K<sub>x</sub>]; [O<sub>1</sub>:I<sub>x</sub>, T<sub>x</sub>]; [O<sub>2</sub>:F<sub>x</sub>]; [O<sub>3</sub>:P<sub>x</sub>]; [O<sub>4</sub>:H<sub>x</sub>]),

где K25 – код язвы желудка по Международному классификатору болезней (МКБ). Любой элемент [\*] может отсутствовать. Вместо индексов *x* необходимо подставить конкретные числа.

Зная формулу и кодовое представление диагноза можно однозначно воспроизвести текстовое описание клинического диагноза. Примеры кодовых представлений и текстовых описаний диагнозов:

K25 (J2; L5; K1) Рубцующая язва пилорического отдела желудка, ассоциированная с Нр;

K25.4 (O1:I1, T3) Язва желудка, осложненная кровотечением (Forest 1a – активное струйное), кровопотеря тяжелой степени.

Через структуру «Клинический диагноз» можно определить субструктуру «Патологический процесс» (структуру из структур). Пример описания патологического процесса:

(р: ПАЦИЕНТЫ) (v: МОМЕНТЫ\_ВРЕМЕНИ)

ПАТОЛОГ\_ПРОЦЕСС (р, v) = <{K25.5 (O2: F1); K80.0 (M3)}, {<>}, ...>.

Комментарий: на первом шаге процесса у пациента имеются два заболевания (язва желудка, осложненная перфорацией (O2) в стадии шока

(F1) и острый калькулезный холецистит (K80.0), гангренозная форма (M3)), на втором шаге – одно заболевание и т.д.

Любая структура, являясь сущностью, обладает собственными свойствами, которые задаются функцией СВОЙСТВА(*Str*). Примеры свойств:

СВОЙСТВА (Событие *Y*) =  $\{p/P\}$ ;

СВОЙСТВА (*Str*: ХИМ\_ВЕЩЕСТВА) = {ФОРМУЛА, МОЛЯРНАЯ\_МАССА, ПРОСТОЕ\_ВЕЩЕСТВО, КИСЛОТА, ОКСИД, ...};

СВОЙСТВА (*Str*: ХИМ\_РЕАКЦИИ) = {*Тип реакции* / {Обратимая; Необратимая}?, *Фазовое состояние* / {Гетерогенная; Гомогенная}?, ...}.

Описания структур могут содержать онтологические соглашения. К онтологическим соглашениям относятся все виды определений, включая формулы. Примеры онтологических соглашений из химии [15, 16].

(e: ХИМ\_ЭЛЕМЕНТЫ)

ЩЕЛОЧНОЗЕМЕЛЬНЫЙ(e)  $\Leftrightarrow$  ХИМ\_ГРУППА(e) = Па.

Легко видеть, что данное соглашение является частью орграфа доменов:  $G(\text{ХИМ\_ЭЛЕМЕНТ}) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Если *тело* состоит из одного вещества, то масса тела равна произведению массы одной молекулы на число молекул в теле:

(v: *тела*)

$\text{масса}(v) = \text{масса\_молекулы}(\text{вещество\_тела}(v)) \cdot \text{число\_молекул}(v)$ .

Данное соглашение, как и любое другое, можно представить с помощью одной из разновидностей системопаттерна  $f/\mu$ :  $\{J_a \underline{a}/A\} \rightarrow \{J_b \underline{b}/B\}$ .

Примеры онтологических соглашений клинической медицины:

"Заключение *d* имеет место тогда и только тогда, когда совокупность результатов тестов  $\{\underline{g}/T\}$  удовлетворяет фиксированному множеству предикатов  $\{P\}$ ", т.е.  $d \Leftrightarrow \{P(\{\underline{g}/T\}) = \text{Истина}\}$ ;

$d = \text{«Воспаление среднего уха»} \Leftrightarrow \{\text{Боль в ухе? Есть; Шум в ухе? Есть; Ухудшение слуха? Есть; Гиперемия и выпуклость барабанной перепонки? Есть}\}$ .

Данное соглашение можно представить с помощью системопаттерна  $f/\mu$ :  $\{\underline{g}/T\} \rightarrow d/D$ .

В общем случае будем считать, что онтологические соглашения задаются экспертами, а не формируются автоматически на основе эмпирических данных.

Описание концепта (произвольного понятия) с помощью онтологических соглашений имеет следующий вид:

$$[Тип:] Концепт \{ \underline{т}/T_1 [\{ \underline{т}/T \}_1]; \dots; \underline{т}/T_n [\{ \underline{т}/T \}_n] \}, \quad (3.11)$$

где  $[\{ \underline{т}/T \}_j]$  – уточняющие тесты (могут отсутствовать); *Тип* – означает тип концепта (может быть опущен), например, *Ds* – диагноз. Уточняющие тесты, в свою очередь, также могут содержать уточняющие тесты. В итоге концепт может представлять собой множественную иерархию значений тестов. Для фиксации значений тестов могут использоваться любые домены орграфов доменов тестов (не только базовые домены). Пример:

*Ds: Острая пневмония* {Кашель (отсутствует; имеется): имеется; Частота дыхания (увеличена; норма; снижена): увеличена; Одышка (отсутствует; имеется): имеется {Характер одышки (смешанная; экспираторная; инспираторная): смешанная}; ...}.

В приведенном примере тест «Характер одышки» является уточняющим тестом для результата теста «Одышка: имеется». В разделе 3.1 был приведен пример концепта «Головная боль». Подобным образом можно описать все медицинские тесты.

В рамках модели ПрО целесообразно фиксировать множество концептов и банк тестов:

*Банк концептов* {[Тип:] Концепт}; *Банк тестов* { $G(\tau)$ }.

В рамках клинической медицины в качестве Банка концептов можно рассматривать, например, онтологические описания диагнозов {*Ds*} (совокупность диагностических критериев, выраженных с помощью тестов) следующего вида:  $Ds_1 \{ \underline{т}/T \}_1; \dots; Ds_N \{ \underline{т}/T \}_N$ . Примеры:

*Ds: Эпилепсия* {Тахикардия: имеется; Головокружение: отсутствует; Головная боль: имеется; ...};

*Ds: Острая пневмония (типичная)* {Синдром интоксикации: *имеется*; Синдром уплотнения легочной ткани: *имеется*; Синдром дыхательной недостаточности: *имеется*; Катаральный синдром: *имеется*};

*Ds: Острый бронхит* {Катаральный синдром: *имеется*; Синдром интоксикации: *отсутствует*; Синдром уплотнения легочной ткани: *имеется*; Синдром дыхательной недостаточности: *отсутствует*};

*Ds: Панкреатит* {Б/х\_крови: АМГ↑, АсАТ↑, АlАТ↑, ДДГ↑, ЛДГЗ↑; Б/х\_мочи: АМГ↑; ...}.

Такие концепты как «Острая пневмония» и «Острый бронхит» представлены на максимальном уровне общности (синдромном).

Главный путь формализации свойств таких структур как «образ», «ситуация», «эмпирическое понятие» – это поиск закономерностей на множестве эмпирических данных. Примерами таких закономерностей

являются формальные синдромы и вероятностные закономерности (предвестники). Множество закономерностей при определенных условиях образует модель знаний. Детально закономерности и модели знаний будут рассмотрены в главе 5. Поясним лишь суть данных конструктов, так как с их помощью осуществляется естественная категоризация и описываются свойства многих когнитивных структур.

Пусть  $\Omega(\{\tau/T\}, Z) = \{\alpha\{\underline{t}/T\}, \underline{z}/Z\}$  – база прецедентов (образов, процессов или ситуаций действительности), представленных на уровне общности  $\{\tau/T\}$ , с известными заключениями  $z \in Z$ , где  $Z = \{1, \dots, N\}$  – домен теста «Заключение» (образует полную группу несовместных заключений). Будем говорить, что множество  $\Omega$  не содержит *конфликтов* на уровне общности  $\{\tau/T\}$ , если нет двух ситуаций с разными заключениями, но совпадающими значениями тестов.

*Формальным синдромом* (или просто *синдромом*) будем называть неизбыточную совокупность значений тестов, позволяющую однозначно установить заключение. Произвольный синдром представим в виде:

$$S = (\{\underline{t}/T\} \rightarrow \underline{z}/Z), \quad \nu = \nu(S),$$

где  $\{\underline{t}/T\}$  – значения тестов;  $\underline{z}/Z$  – заключение;  $\nu$  – вес синдрома. Вес синдрома  $S$  пропорционален количеству прецедентов с заключением  $\underline{z}/Z$ , содержащих синдром  $S$ . Вес синдрома принадлежит к числу свойств структуры «Синдром». При определенных условиях множество синдромов образует синдромную модель знаний  $\{S\}$  (см. главу 5).

*Предвестниками* некоторого события  $\underline{z}/Z$  на множестве тестов  $\{a/A\}$  назовем правила вида

$$R = (\{\underline{t}/T\} \rightarrow \underline{z}/Z), \quad p(R) \geq p^*, \quad \nu_R = \nu(R), \quad \{\tau/T\} \subseteq \{a/A\},$$

где  $p(R)$  – вероятность получения заключения  $\underline{z}/Z$  при условии  $\{\underline{t}/T\}$ ;  $p^*$  – порог (например, 0.9);  $\nu_R$  – вес правила, пропорциональный количеству прецедентов с заключением  $\underline{z}/Z$ , отвечающих правилу  $R$ . Предвестники отличаются от формальных синдромов вероятностным характером зависимости. При определенных условиях множество предвестников образует вероятностную модель знаний  $\{R\}$ . Синдромы и предвестники являются разнovidностью системопаттерна.

С помощью структур «База прецедентов», «Формальный синдром», «Предвестник», «Модель знаний», «Орграф набросков» определяются свойства такой важной когнитивной структуры как «Образ» (см. главу 6):

$$\begin{aligned} \text{СВОЙСТВА } (P: \text{ОБРАЗ}) = \{ & \\ & \{\tau, \text{СВОЙСТВА}(\tau)\} && \text{- Банк тестов;} \\ & \Omega(\{\tau/T\}, Z) = \{\alpha\{\underline{t}/T\}, \underline{z}/Z\} && \text{- База прецедентов;} \end{aligned}$$



- $\{S = (\{\underline{t}/T\} \rightarrow \underline{z}), \nu = \nu(S)\}$  - Формальные синдромы;  
 $\{R = (\{\underline{t}/T\} \rightarrow \underline{z}/Z), p(R) \geq p^*\}$  - Предвестники;  
 $\{Gs(P)\}$  - Орграфы набросков}.

С помощью системопаттернов  $\{f/\mu: \{J_a \underline{a}/A\} \rightarrow \{J_b \underline{b}/B\}\}$  и Банка тестов описываются такие макроструктуры, как «Банк математических моделей», «Многоцелевой банк знаний», «Среда радикалов», «Функциональная система», «Модель ПрО» и другие (см. главу 7).

В работе [134] приведена информационная структура концепта «Развернутое заключение», который может быть применим к широкому классу ситуаций действительности, в частности, клинических ситуаций. Набор свойств данного концепта выглядит следующим образом:

- СВОЙСТВА( $d/D$ : РАЗВЕРНУТОЕ ЗАКЛЮЧЕНИЕ) = {
- $\{\tau$ . СВОЙСТВА( $\tau$ )
  - $F(\{a/A\})$
  - $d/D \Leftrightarrow \{P(\{\underline{t}/T\}) = \text{Истина}\}$
  - $\Omega(\{\underline{t}/T\}, Z) = \{\alpha(\{\underline{t}/T\}, \underline{z}/Z)\}$
  - $\{S = (\{\underline{t}/T\} \rightarrow \underline{z}), \nu = \nu(S)\}$
  - $\{R = (\{\underline{t}/T\} \rightarrow \underline{z}/Z), p(R) \geq p^*\}$
  - $\{Gs(d/D)\}$
  - $\{\{\underline{t}/T\}_k \rightarrow J_d d/D\}$
  - $\{f/\mu: \{\underline{t}/T\}_n \rightarrow \{J_\tau \underline{t}/T\}_m\}$
  - $\{d/D \rightarrow J_k \{\underline{t}/T\}_k; d/D \rightarrow \{J_\tau \underline{t}/T\}\}$
  - $\{\underline{t}/T\} \rightarrow \{d/D \mid d/D \rightarrow \diamond\{\underline{a}/A\} \& \{\underline{a}/A\} \nabla \{\underline{t}/T\}\}$
  - $\{d/D \wedge \{\underline{t}/T\}_p \rightarrow \{J_{d'} d'/D'\}\}$
  - $\{d/D \wedge \{\underline{t}/T\}_p \rightarrow \{r/R\}\}$

где  $\{\underline{a}/A\} \nabla \{\underline{t}/T\}$  означает совместимость тестов  $\{\underline{a}/A\}$  и  $\{\underline{t}/T\}$ . Примеры формул заключения были рассмотрены выше.

Таким образом, свойства любой информационной структуры полностью определяются элементарными тестами и системопаттернами.

### 3.4 Наброски ситуаций, образов

Обозначим через  $\alpha_s$  произвольный набросок ситуации действительности или образа  $\alpha$ . Причем  $\alpha \in \{\alpha_s\}$ . Объем информации, содержащийся в наброске  $\alpha_s$ , обозначим  $|\alpha_s|$ . Выполняется соотношение:  $\forall \alpha_s \neq \alpha, |\alpha_s| < |\alpha|$ . Другими словами, наброски отличаются друг от друга степенью детализации описания образа/ситуации  $\alpha$ .

Пусть  $\{F\}$  – фиксированное множество сжимающих отображений (системопаттернов) таких, что любой набросок  $\alpha_s$  получается из  $\alpha$  в

результате последовательного применения серии отображений из  $\{F\}$ . Можно записать:  $\alpha_s = F^v(\dots F^1(\alpha)) = F^{(v)}(\alpha)$ . К каждому наброску делается попытка применить все отображения из  $\{F\}$ . Потребуем для  $\{F\}$  выполнения следующих условий:

- (а)  $\forall F \in \{F\}, \forall \alpha_s \in \{\alpha_s\}$ , если  $F$  применимо к  $\alpha_s$ , то  $|F(\alpha_s)| < |\alpha_s|$ .
- (б)  $\forall F, F' \in \{F\}, \forall \alpha_s \in \{\alpha_s\}, |F(\alpha_s)| \approx |F'(\alpha_s)|$ .
- (с)  $\forall \alpha \exists K: \{F^{(K)}(\alpha)\} \neq \emptyset$ , но либо  $\neg \exists F^{(K+1)}(\alpha)$ , либо  $\{F^{(K+1)}(\alpha)\} = \emptyset$ .

Множество набросков  $\alpha_s$  таких, что либо  $\forall F \in \{F\}, F(\alpha_s) = \emptyset$ , либо к ним нельзя применить ни одно отображение из  $\{F\}$  назовем *финитными набросками* и обозначим  $\{\alpha^*\}$ . Как правило, для класса ситуаций  $\{\alpha\}$  множество финитных набросков конечно. Примером конечного множества финитных набросков могут служить клинические диагнозы.

*Слоем набросков* назовем множество  $G_{S_v}(\alpha) = \{\alpha_s | \alpha_s = F^{(v)}(\alpha)\}$ , где  $v = 0, 1, \dots, K$ . Последний  $K$ -й слой состоит исключительно из финитных набросков. Исходная ситуация  $\alpha$  относится к 0-му слою набросков.

*Орграфом набросков* произвольного образа/ситуации действительности  $\alpha$  назовем структуру следующего вида:

$$Gs(\alpha) = \cup_v \{\alpha_s | \alpha_s = F^{(v)}(\alpha)\} = \cup_v G_{S_v}(\alpha), \quad (3.12)$$

где  $v = 0, 1, \dots, K_\alpha$ . Число слоев  $K_\alpha$  является константой образа/ситуации  $\alpha$  (при фиксированном множестве отображений  $\{F\}$ ).

Как уже отмечалось ранее, орграф доменов теста  $G(\tau)$  является частным случаем орграфа набросков  $Gs(\tau)$ . Любой домен является наброском базового домена.

Построение орграфа набросков — высшая ступень превращения сложного в простое, умения выделять существенное. Это результат глубокого проникновения в ситуацию и постижения главного в ней, решающего, того, что определяет ход событий. Процедура построения  $Gs(\alpha)$  в полной мере реализует принцип полимодельной дополнителности и конкуренции, о котором шла речь в первой главе.

Системопаттерны множества  $\{F\}$  могут быть реализованы экспертами, нейроморфной средой, мультиагентной средой и т.д. Особенно важный случай представляет множество  $\{F\}$ , которое можно отнести к автоматизмам вычислительной среды (вычислительного интеллекта). В этом случае орграф  $Gs(\alpha)$  строится всегда независимо от решаемой задачи.

Пусть  $\Omega(Z) = \{\alpha(\underline{z}/Z)\}$  — база прецедентов (ситуаций действительности, образов), относящихся к определенному классу из  $Z$  ( $Z$ -задача). В качестве  $Z$  может выступать полный набор несовместных заключений, например, диагнозов, прогнозов, управлений или кодов ситуаций.

Если набросок  $\alpha_s$  содержит в себе важнейшие черты ситуации  $\alpha$ , то

данный факт отобразим нотацией  $\alpha_s \nabla \alpha$ . Принадлежность ситуации к тому или иному классу  $\underline{z}/Z$  является важнейшей характеристикой ситуации. Если набор важнейших черт ситуации  $\alpha$  обозначить через  $\{\underline{a}/A\}_\alpha$ , то контекстом задачи построения набросков назовем  $X = \langle \{\underline{a}/A\}_\alpha, \{F\} \rangle$ . Частным случаем является контекст  $Z$ -задачи:  $X = \langle \Omega(Z), \{F\} \rangle$ .

Экстремальным пограничным слоем набросков орграфа набросков  $G_s(\alpha)$  контекста  $X$  назовем множество набросков

$$M_X(\alpha) = \{\alpha_s | \alpha_s \in G_s(\alpha) \& \alpha_s \nabla \alpha \& (\forall F \in \{F\}, \neg(F(\alpha_s) \nabla \alpha))\}. \quad (3.13)$$

Изменение контекста  $X$ , скорее всего, приведет к изменению состава экстремального пограничного слоя. Наброски из  $M_X(\alpha)$  будем называть *экстремальными* или *критическими*. Наброски набросков из  $M_X(\alpha)$  назовем *закритическими*, а наброски, предшествующие экстремальному слою, назовем *докритическими*. Закритические наброски в рамках  $Z$ -задачи служат основой метафор и ассоциаций, так как принадлежат нескольким образам/ситуациям с разными заключениями  $\underline{z}/Z$ .

**Предложение 3.1.** В рамках орграфа  $G_s(\alpha)$  наброски экстремального пограничного слоя  $M_X(\alpha)$  содержат минимальный объем информации достаточный для отражения важнейших черт ситуации/образа  $\alpha$ , в частности, однозначной идентификации ситуации/образа из  $\Omega$  в рамках  $Z$ -задачи.

Забегая вперед, отметим, что на основе орграфов набросков прецедентов и Банка тестов могут быть построены предельные синдромные модели знаний, которые содержат конкурирующие наборы параметров порядка, т.е. меньший объем информации, чем экстремальные наброски (см. главы 5 и 6).

Пример орграфа набросков с экстремальным пограничным слоем показан на рис. 3.1.

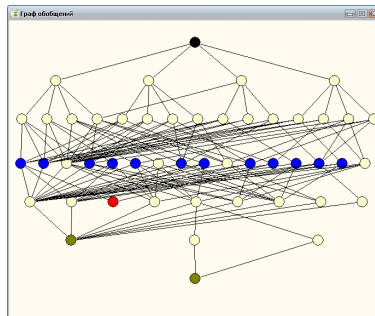


Рис. 3.1 – Пример орграфа набросков с экстремальным слоем ( $K_\alpha = 6$ )

Исходная ситуация – черная вершина в верхней части рисунка; красная и зеленые вершины внизу – финитные наброски; синие вершины в 3-м слое – наброски экстремального пограничного слоя. Множество  $\{F\}$  содержит четыре отображения.

В качестве исходного образа может выступать вся база прецедентов  $\Omega(Z)$  или любой ее фрагмент. За значимую характеристику  $\underline{a}/A$  можно принять отсутствие конфликта данных.

Приведем пример множества отображений  $\{F\}$ . Пусть любая ситуация действительности  $\alpha$  описывается с помощью банка тестов  $\{G(\tau)\}$ . Без потери общности будем считать, что каждый тест входит в описание ситуации только один раз. Элементарный акт обобщения ( $T \rightarrow T'$ ) в рамках орграфа  $G(\tau)$  обозначим  $F_\tau$ .

Множество  $\{F_\tau \mid \tau \in \{\tau, G(\tau)\}\}$  удовлетворяет требованиям (а) – (с), предъявляемым к отображениям из  $\{F\}$ . Действительно, условие (а) выполняется, так как любое обобщение приводит к потере части информации. Условие (б) выполняется, так как элементарное обобщение значения лишь одного теста незначительно уменьшает объем информации обо всей ситуации. Другими словами, коэффициенты сжатия всех  $F_\tau$  можно считать сопоставимыми. Покажем, что выполняется также и условие (с).

Пусть  $[T_1 \rightarrow T_n] = (T_1 \rightarrow T_2 \rightarrow \dots \rightarrow T_n)$  произвольная цепочка обобщений в рамках орграфа  $G(\tau)$ . Длину максимальной цепочки обобщений в произвольном орграфе  $G(\tau)$  обозначим  $L_\tau$ .

**Предложение 3.2.** Если  $\{F\} = \{F_\tau \mid \tau \in \{\tau, G(\tau)\}\}$ , то выполняется соотношение:

$$\forall \alpha(\{\tau\}), K_\alpha \leq \sum_{\{\tau\}} L_\tau \quad (3.14)$$

Предложение 3.2 доказывает условие (с).

Множество терминальных вершин в орграфе  $G(\tau)$  обозначим  $V_\tau$ . Общее количество набросков в орграфе  $G_s(\alpha)$  и общее количество финитных набросков дает следующее предложение.

**Предложение 3.3.** Если  $\{F\} = \{F_\tau \mid \tau \in \{\tau, G(\tau)\}\}$ , то выполняются соотношения:

$$\forall \alpha(\{\tau\}), |G_s(\alpha)| = \prod_{\{\tau\}} |G(\tau)|, \quad (3.15)$$

$$\forall \alpha(\{\tau\}), |\{\alpha^*\}| = \prod_{\{\tau\}} |V_\tau|, \quad (3.16)$$

где  $|G_s(\alpha)|, |G(\tau)|$  - число вершин в орграфах (при условии, что каждый тест входит в описание  $\alpha$  один раз).

Из (3.16) следует, что если в орграфах доменов всех тестов всего одна терминальная вершина, то орграф набросков содержит лишь один

финитный набросок.

Количество набросков в экстремальном пограничном слое  $M_X(\alpha)$  определяется количеством критических описаний, формируемых методом предельных обобщений (глава 5).

Принципиально важным является тот факт, что отображения из  $\{F_i\}$  относятся к автоматизмам вычислительной среды, поэтому оргграф набросков также формируется с помощью автоматизмов. К автоматизмам из  $\{F_i\}$  следует добавить еще два автоматизма: первый автоматизм – формирование синдромной модели знаний для любого наброска; второй автоматизм – формирование предельной синдромной модели знаний на основе оргграфа набросков. Оба эти автоматизма являются неотъемлемой частью метода предельных обобщений и будут рассмотрены в главе 5.

Приведем пример оргграфа набросков. Объект – Компьютерная фирма. Решается вопрос о приеме программиста на работу. Возможны три заключения:  $Z = \{1; 2; 3\}$  и, соответственно, три производных множества заключений:

$$Z1 = \{1; -1\} = \{1, 4\}; \quad Z2 = \{2; -2\} = \{2, 5\}; \quad Z3 = \{3; -3\} = \{3, 6\}.$$

Вопрос о приеме на работу решается на основании трех тестов:

- образование;
- владение языками программирования;
- знание английского языка.

Тесты задаются оргграфами доменов:

Образование  $\wedge t1 \{$

2 { Высшее  $\wedge c$ ; Не высшее  $\wedge a b$  }

1 { Среднее  $\wedge a$ ; Неполное высшее  $\wedge b$ ; Высшее  $\wedge c$  }

Владение языками программирования  $\wedge t2 \{$

2 { ООП  $\wedge a$ ; Несколько видов  $\wedge b c$  }

1 { ООП  $\wedge a$ ; ООП + WEB  $\wedge b$ ; ООП + WEB + низкоуровневый  $\wedge c$  }

Знание английского  $\wedge t3 \{$

2 { Слабое  $\wedge a$ ; Среднее и выше  $\wedge b c$  }

1 { Слабое  $\wedge a$ ; Чтение и перевод  $\wedge b$ ; Свободно  $\wedge c$  }.

Все оргграфы доменов имеют одинаковую структуру:  $G = \{1 \rightarrow 2\}$ . Схемы обобщения в оргграфах очевидны:

t1: 1.c  $\rightarrow$  2.c; 1.{a, b}  $\rightarrow$  2.a;

t2 и t3: 1.a  $\rightarrow$  2.a; 1.{b, c}  $\rightarrow$  2.b.

Пусть множество прецедентов  $\Omega$  содержит 11 случаев, которые представлены в таблице 3.1 в формате (t1/1, t2/1, t3/1), т.е. в базовых доменах. В таблице 3.2 приведены сведения о прецедентах на всех уровнях обобщения.

Таблица 3.1 – БД  $\Omega$ 

t1	t2	t3	Z
a	a	a	1
b	b	b	3
c	c	a	2
a	b	c	1
a	c	b	1
b	c	a	2
b	a	c	2
c	a	b	1
c	b	a	3
c	b	b	2
b	a	b	2

Таблица 3.2 – Расширенная БД

t1		t2		t3					
1	2	1	2	1	2	Z	Z1	Z2	Z3
a	a	a	a	a	a	1	1	5	6
b	a	b	b	b	b	3	4	5	3
c	c	c	b	a	a	2	4	2	6
a	a	b	b	c	b	1	1	5	6
a	a	c	b	b	b	1	1	5	6
b	a	c	b	a	a	2	4	2	6
b	a	a	a	c	b	2	4	2	6
c	c	a	a	b	b	1	1	5	6
c	c	b	b	a	a	3	4	5	3
c	c	b	b	b	b	2	4	2	6
b	a	a	a	b	b	2	4	2	6

В таблице 3.3 приведены все комбинации доменов для трех тестов и для каждой комбинации указана конфликтность базы прецедентов (0/1). Напомним, что конфликт будет иметь место тогда и только тогда, когда найдутся два прецедента с одинаковыми значениями тестов, но с разными заключениями. Все конфликты помечены символом «С». Символом «В» помечены экстремальные наброски (наброски экстремального пограничного слоя). Любое обобщение экстремального наброска приводит к конфликту.

Таблица 3.3 – Результирующая таблица набросков

Домены			Z		Z1		Z2		Z3	
t1	t2	t3	0/1	A,B,C	0/1	A,B,C	0/1	A,B,C	0/1	A,B,C
1	1	1	0	A	0	A	0	A	0	A
1	1	2	0	B	0	A	0	B	0	B
1	2	1	1	C	0	A	1	C	1	C
2	1	1	0	B	0	B	0	B	0	B
1	2	2	1	C	0	B	1	C	1	C
2	1	2	1	C	1	C	1	C	1	C
2	2	1	1	C	1	C	1	C	1	C
2	2	2	1	C	1	C	1	C	1	C

0 – нет конфликта; 1 – конфликт; А – докритическое описание (набросок); В – критическое описание; С – закритическое описание.

Список всех комбинаций доменов определяет *орграф набросков* базы прецедентов  $G_S(\Omega)$  (рис. 3.2). На рис. 3.3 показан орграф набросков 4-го

сверху прецедента из таблицы 3.1 –  $Gs(\alpha_4)$ . Индексы сверху значений тестов показывают домен.  
1 2 2

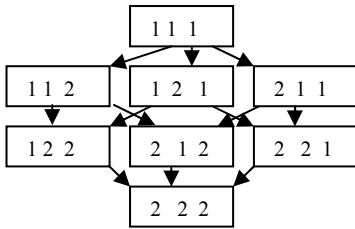


Рис. 3.2 – Схема орграфа набросков  $Gs(\Omega)$

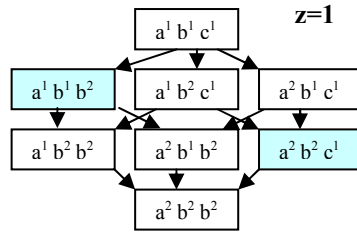


Рис. 3.3 – Орграф набросков прецедента  $Gs(\alpha_4)$

В данном примере орграф набросков любого прецедента, как и  $Gs(\Omega)$ , содержит 8 набросков:  $|Gs(\alpha)| = \prod_{i \in \tau} |G(\tau)| = 2 \cdot 2 \cdot 2 = 8$ . Финитный набросок всего один:  $|\{\alpha^*\}| = \prod_{i \in \tau} V_{\sigma} = 1 \cdot 1 \cdot 1 = 1$ . Слоев набросков четыре: 0 -3. Соответственно  $K=3$ .

Два закрашенных наброска на рис. 3.3 являются критическими набросками, относящимися к пограничному слою  $M_X(\alpha_4)$ , где  $\alpha_4$  – 4-й прецедент в базе прецедентов  $\Omega$ . Они являются уникальными для прецедента  $\alpha_4 = (t1/1/?a; t2/1/?b; t3/1/?c; z/1)$ , но любое их обобщение приведет к совпадению по тестам с другими прецедентами из  $\Omega$ .

На рис. 3.4 цветом выделены экстремальные пограничные слои набросков  $M_Z(\Omega)$  и  $M_{Z1}(\Omega)$  (соответствуют символам «В» в таблице 3.3). Здесь  $\Omega$  выступает в качестве образа.

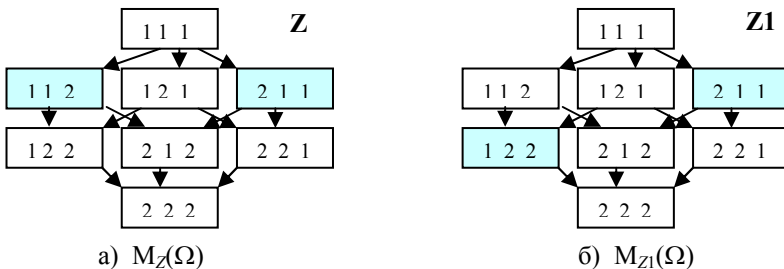


Рис. 3.4 – Экстремальные пограничные слои набросков

Таким образом, орграф набросков разрешает диалектическое противоречие «Единое – многое». Рассмотрение орграфов набросков

ситуаций, образов, моделей будет продолжено в главе 6.

### 3.5 Модель развития ситуации. Импульсы. Описание ситуации

Произвольную ситуацию действительности обозначим через  $\alpha$ , а множество ситуаций действительности (прецедентов) обозначим через  $\Omega = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$ . По умолчанию будем предполагать, что все ситуации множества  $\Omega$  описаны с помощью одного и того же банка тестов  $\{G(\tau)\}$ , т.е. в рамках одной онтологии. В общем случае, если  $A$  и  $B$  – два разных наблюдателя, то каждый из них имеет собственную онтологию –  $\{G(\tau)\}_A$  и  $\{G(\tau)\}_B$  соответственно и, следовательно, разные форматы описаний одного и того же множества ситуаций действительности:  $\Omega_A$  и  $\Omega_B$ . В связи с этим возникают нетривиальные задачи обмена опытом, передачи знаний, выработки общей точки зрения и т.д.

В простейшем случае эмпирическую модель описания произвольной ситуации действительности на основе онтологии можно представить как объединение событий:

$$\alpha = \alpha(\{ \langle J_\tau \underline{t}/T, \{ J_p \underline{p}/P \}, J_t \underline{t}/A \rangle \}). \quad (3.17)$$

С такой моделью наблюдатель сталкивается при первом знакомстве с ситуацией. Она представляет собой лишь набор разрозненных фактов, событий. Важно подчеркнуть, что необходимых тестов для описания ситуации (феномена) действительности вначале может и не быть. Адекватная система тестов должна быть создана в процессе изучения неоднородностей потока данных (типов и форм изменчивости ситуации). Концептуально речь идет о «созревании» орграфов доменов тестов на основе прототестов, протодоменов, протографов.

В процессе изучения ситуации необходимо ответить на четыре главных вопроса:

- Как известные факты, события, характеристики ситуации связаны между собой?
- Каков прогноз развития ситуации, и каковы движущие силы развития ситуации?
- Каковы параметры порядка, предельные смыслы, задающие смысловую регуляцию ситуации?
- Каковы аналоги или сходные ситуации, каковы ситуации-цели?

Ответ на первый вопрос может быть представлен в виде модели развития ситуации (локальной модели знаний), связывающей все известные факты, характеристики и события. Такую модель для ситуации  $\alpha$  можно представить в виде множества системопаттернов:



$$k_\alpha = \{f/\mu: \{J_a \underline{a}/A\} \rightarrow \{J_b \underline{b}/B\}\}_\alpha \cup P_k, \quad (3.18)$$

где  $f/\mu$  – отображения с известным механизмом реализации  $\mu$ ,  $P_k$  – правила композиции отображений. Наличие модификатора истинности значений тестов указывает на приближительный характер некоторых отображений.

Модель (3.18) раскрывает «физику явлений». Другими словами, устанавливает универсальные отношения и системные атрибуты, раскрывающие базовые (фундаментальные) механизмы системогенеза.

Описание «Физику явления» чрезвычайно важно при анализе техногенных аварий или катастроф, экономических кризисов, экологических катастроф, социальных кризисов, заболеваний и т.д. Примерами реализации локальных моделей знаний могут служить когнитивные карты [1, 98].

В клинической медицине модель (3.18) описывает, в частности, этиологию и патогенез развития заболевания (развития клинической ситуации), а также функциональные связи между характеристиками и параметрами системы (организма). Приведем пример реализации функциональных связей в задаче «Влияние адреналина на гликогенез» (по материалам [171]). Введем следующие тесты:

$$\begin{aligned} \tau_1 &= \text{«Продукция адреналина»}; \tau_2 = \text{«Активность аденилатциклазы»}; \\ \tau_3 &= \text{«Содержание АТФ»}; \tau_4 = \text{«Содержание АМФ»}; \\ \tau_5 &= \text{«Содержание фосфорилазы В»}; \tau_6 = \text{«Содержание фосфорилазы»}; \\ \tau_7 &= \text{«Содержание гликогена»}; \tau_8 = \text{«Содержание глюкозы»}. \end{aligned}$$

Каждый из тестов задается своим орграфом доменов  $G(\tau)$ . В итоге имеем банк тестов  $\{G(\tau_1), \dots, G(\tau_8)\}$ .

На уровне базовых доменов D1 всех тестов цепочка функциональных связей представляется множеством отображений:

$$\begin{aligned} k_\alpha &= \{f_1/\mu_1: \tau_1 \rightarrow \tau_2; f_2/\mu_2: \tau_2 \rightarrow \tau_3; \dots; f_7/\mu_7: \tau_7 \rightarrow \tau_8\} \cup P_k \equiv \{\tau_1 (+a) \rightarrow \tau_2; \\ \tau_2 (-b) &\rightarrow \tau_3; \tau_3 (-c) \rightarrow \tau_4; \tau_4 (-d) \rightarrow \tau_5; \tau_5 (-e) \rightarrow \tau_6; \tau_6 (-g) \rightarrow \tau_7; \\ \tau_7 (-k) &\rightarrow \tau_8\} \cup P_k. \end{aligned}$$

В скобках приведены структурные коэффициенты влияния (знак показывает направление влияния – в сторону уменьшения или повышения). Запись  $x(-a) \rightarrow y$  в аналитической форме означает:  $y = -a \cdot x$ , т.е. таким способом задается механизм реализации отображения  $f/\mu$ .

Правила композиции  $P_k$  реализуют в данном примере процедуру аналитического транзитивного замыкания, что позволяет получить отображение

$$\tau_1/D1 (+abcdegk) \rightarrow \tau_8/D1$$

или в развернутом виде:

Продукция адреналина (+*abcdefgk*) → Содержание глюкозы.

Знак структурного коэффициента показывает, что с увеличением синтеза адреналина содержание глюкозы в крови возрастает, что в действительности и наблюдается [171].

Каждый структурный коэффициент  $p$  является тестом  $p/P$ , поэтому замыкание на логическом уровне (с каким-то механизмом  $\mu$ )

$$f/\mu: (\tau/T)_1 \rightarrow (\tau/T)_8$$

на процедурном уровне запишется следующим образом:

$$(\tau/T)_1 (\Pi_{j=1,N} (p/P)_j) \rightarrow (\tau/T)_8.$$

Двойственность описания функциональных или причинно-следственных связей проявляется в том, что эти связи могут быть представлены либо в локальной модели знаний  $k_\alpha$ , либо в виде онтологических соглашений в орграфах доменов тестов. Пример онтологического соглашения для теста  $\tau_2$ :

Тест  $\tau_2$ : D3 {...} D2 {...} D1 {[0; 1] (a ·  $\tau_1/D1$ )}.

$$G(\tau_2) = \{D1 \rightarrow D2 \rightarrow D3\}.$$

Аналогичным образом записываются онтологические соглашения для остальных тестов. Отметим, что для орграфов доменов тестов может быть выбрана одна из стандартных схем (см. главу 4).

Если предположить, что орграфы всех тестов содержат по 3 домена, то замыкание  $(\tau_1/D1)^+$  над банком тестов  $\{G(\tau_1), \dots, G(\tau_8)\}$  будет содержать 24 значения восьми тестов разного уровня общности. Те же 24 значения могут быть получены на основе банка тестов и локальной модели знаний  $k_\alpha$ , например:

$$\{f_1/\mu_{1,3}: \tau_1/D3 \rightarrow \tau_2/D3; f_2/\mu_{2,2}: \tau_2/D2 \rightarrow \tau_3/D2; \dots; f_7/\mu_{7,3}: \tau_7/D3 \rightarrow \tau_8/D3\}.$$

Вернемся к рассмотрению общих вопросов анализа и синтеза модели ситуации действительности.

Ответ на вопрос о дальнейшем развитии ситуации и ее движущих силах дает *поле импульсов*:

$$\{Imp: \{J_a \underline{a}/A\}, \underline{e}/E \rightarrow \{b/B\}\}_t, \quad \sum_{\{Imp\}} \underline{e}/E \leq \underline{e}^*/E, \quad (3.19)$$

где  $Imp_t$  – импульс в момент времени  $t$ ;  $\{J_a \underline{a}/A\}$  – установленные факты, события;  $\{b/B\}$  – факты и события, которые планируется достичь;  $\underline{e}/E$  – выделяемая энергия (ресурсы) на достижение цели;  $\underline{e}^*/E$  – общая доступная энергия (ресурсы).

Направления импульсов и параметры  $\{e/E\}$  являются, по сути, средствами управления развитием ситуации в момент времени  $t$ . Необходимо отметить, что импульсы могут существенно различаться не только по энергии, но и по времени выполнения. Ввиду ограниченности общих ресурсов импульсы могут конфликтовать между собой (в борьбе за ресурсы). Чем больше конфликтов, тем больше энергии (ресурсов) расходуется впустую.

Одни импульсы могут строиться на основе системопаттернов модели знаний и для них известен механизм реализации  $\mu$ . Для других импульсов механизм реализации требуется создать (такие импульсы определяют «желания», которые не подкреплены «умениями»). Работа по созданию механизмов относится к метакогнитивным технологиям.

Импульсы играют важную роль в концепции Виртуальной сплошной среды – ВСС (глава 7). Импульсы с большой энергией формируют *характеристические направления*, в рамках которых происходит наиболее интенсивное формирование хаотических ансамблей протоструктур и дальнейшее созревание протоструктур, в частности, прототестов и протопаттернов (механизмов реализации импульсов).

Импульс, энергия которого максимальна и значительно превосходит все остальные импульсы, называется *резонатором*. В области действия резонатора формируются субструктуры высоких порядков (субструктуры субструктур и т.д.). В рамках теории интеллекта субструктуры высоких порядков можно отождествить с «сознанием», тогда как остальные импульсы и автоматизмы формируют «подсознательные» энергоинформационные потоки. Принципиально важным является тот факт, что импульсы «живут» до тех пор, пока у них имеется хоть какая-то энергия, пусть даже самая малая. Такой слабый импульс не может быть в области «сознания» (т.е. порождать субструктуры высоких порядков), но он будет влиять на процессы в «подсознании», питаясь фоновой энергией. В процессе созревания структур ВСС может созреть механизм реализации импульса и тогда произойдет «озарение» – импульс получит фокус внимания (дополнительную энергию). Подобная схема описывает, по сути, процесс созревания решения, образа.

Выделение импульсов и установление их конфигурации позволяет говорить о *фронте развития ситуации* в текущий (заданный) момент времени.

Ответ на вопрос о параметрах порядка, предельных смыслах, задающих смысловую регуляцию ситуации, представим в виде множества  $\{\pi\}_\alpha$ . Таким образом,  $\{\pi\}_\alpha$  – это параметры порядка, детерминанты, предельные смыслы, когнитивные инварианты (каноны, стереотипы и т.д.), принципы, исследовательская программа, которые функционируют как внутренние

регуляторные механизмы, воздействующие на протекание самих процессов деятельности и познания ситуации действительности. Параметры  $\{\pi\}$  выступают в качестве "среды", в которую погружен исследователь и посредством которой он вступает в контакт с открываемой и создаваемой им естественной и искусственной реальностью.

Временная локализация смысловых ориентиров может быть представлена в виде множества протяженных событий:

$$\{\pi\} = \{ \langle J_\tau \underline{d}T, \{ J_p \underline{p}/P \}, J_t \underline{\delta}A \rangle \}_\pi. \quad (3.20)$$

Таким образом, параметры  $\{\pi\}$  вводят системообразующие принципы, меры и механизмы самоорганизации, завершающие оформление ситуации как феномена действительного мира.

Символический образ орбиты ситуации  $\alpha$  длительностью  $\delta A$  в пространстве тестов  $\{a/A\}$  определим выражением:

$$\begin{aligned} Orb_\alpha(\{a/A\}, \delta A) = \cup_\nu \{ \langle \omega/W, \delta A, \{p/P\} \rangle_{i=1,N} \}_\nu, \\ \forall \nu \sum_{i=1,N} (\delta_i/A)_\nu = \delta A, \end{aligned} \quad (3.21)$$

где  $\nu$  – индекс схемы разбиения пространства на ячейки (схемы обобщения);  $N_\nu$  – общее число ячеек, через которые проходит траектория за время  $\delta A$  при  $\nu$ -ом разбиении;  $(\omega_i/W)_\nu$  – имена (или адреса) ячеек  $\nu$ -го пространства разбиения (схемы обобщения);  $(\delta_i/A)_\nu$  – длительность непрерывного пребывания траектории в ячейке  $(\omega_i/W)_\nu$ ;  $(\{p/P\}_i)_\nu$  – характеристики траектории в ячейке  $(\omega_i/W)_\nu$ , например, такие как плотность заполнения, информационное отражение и т.д. Как и ранее все тесты определяются своими орграфами доменов. Пространство тестов  $\{a/A\}$  задает тематический аспект.

В определенном смысле выражение (3.21) описывает *последовательную временную орбиту (траекторию)* развития ситуации  $\alpha$  в рамках заданного аспекта  $\{a/A\}$ . В рамках аспекта  $\{b/B\}$ , скорее всего, будет другая орбита. Совокупность всех аспектов формирует множество «точек зрения» на развитие ситуации. Аспект включает в себя и контекст развития ситуации.

Орбита ситуации в рамках фиксированного аспекта формирует «Историю развития ситуации». В отличие от аморфных событий история имеет топологическую структуру. Смежность ячеек орбиты подчеркивает пространственную, временную, идеологическую и логическую связь смысловых единиц ситуации.

Линеаризация одновременных действий (импульсов) и перестановка мотивов (параметров  $\{\pi\}$ ) превращают естественный порядок истории в более или менее отличающийся от него искусственный порядок. Новая топологическая структура приводит к новой упорядоченности

(эквивалентности) смысловых элементов ситуации по отношению к их исходной позиции. Вместе с тем, благодаря аспекту пространство развития ситуации организовано не в чисто топологическом смысле, но приобретает символическую валентность, когда признак принадлежности смысловой единицы к тому или иному фрагменту орбиты тематически актуализирован.

Таким образом, орбиту или историю следует рассматривать как объективный эвристический инструмент, обеспечивающий методику перехода от отдельных частей (фрагментов) ситуации к их смысловому целому.

В медицине, рассматривая клиническую ситуацию, говорят об анамнезе жизни (истории жизни) и анамнезе заболевания (истории развития заболевания).

В итоге когнитивную модель описания ситуации действительности в фиксированный момент времени можно представить следующим образом:

$$\begin{aligned} \text{СВОЙСТВА}(\alpha, \text{СИТУАЦИЯ ДЕЙСТВИТЕЛЬНОСТИ}) = \{ & \quad (3.22) \\ \{\tau, \text{СВОЙСТВА}(\tau)\} & \quad - \text{Банк тестов;} \\ \{<J_\tau \underline{a}/T, \{J_p \underline{p}/P\}, J_i \underline{t}/A>\} & \quad - \text{События;} \\ \{f/\mu: \{J_b \underline{b}/B\} \rightarrow \{J_a \underline{a}/A\}\}_\alpha \cup P_k & \quad - \text{Модель знаний (локальная);} \\ \{Imp: \{J_a \underline{a}/A\}, \underline{e}/E \rightarrow \{b/B\}\}_i & \quad - \text{Импульсы;} \\ \{Str(\{\underline{a}/T\})\} & \quad - \text{Структуры;} \\ \{\pi\}_\alpha & \quad - \text{Параметры порядка, смыслы;} \\ Orb_\alpha(\{a/A\}, \delta/A) & \quad - \text{Орбита ситуации;} \\ \{Gs(\alpha)\} & \quad - \text{Орграфы набросков}. \end{aligned}$$

Данная модель должна быть дополнена ресурсными ограничениями (возможно с выделением каждого вида ресурса, например, временного, финансового, вычислительного, агентного):

$$\sum_{\{Imp\}} \underline{e}/E \leq \underline{e}^*/E, \quad (3.23)$$

а также онтологическими соотношениями и автоматизмами среды, в частности:

$$\begin{aligned} \forall \tau \rightarrow \text{СВОЙСТВА}(\tau); \\ \forall Str \rightarrow \text{СВОЙСТВА}(Str); \\ \forall \tau, G(\tau | [\{a/A\}] | [\{p/P\}] | [\Sigma_\tau]) = \{T \rightarrow T'\}; \\ \forall \tau, G(\tau) \rightarrow \underline{G}(\tau), \exists(G(\tau)), \exists_{\text{онт}}(G(\tau)); \\ \forall \tau, G(\tau) \rightarrow \{G^+(\tau)\}, G^{++}(\tau), G^-(\tau), \{(G^-(\tau))^+\}; \\ \forall \{\underline{a}/T\} \rightarrow \{\underline{a}/T\}^+, \{\{\underline{a}/T\}\}, \{\{\underline{a}/T\}\}^\wedge; \\ \forall \tau, \{\underline{a}\} \rightarrow G(\tau) \text{ (созревание)}. \end{aligned}$$

Последние три строчки описывают, по сути, автоматизмы вычислительной среды, которые функционируют постоянно в процессе решения любой задачи анализа и синтеза.

*Моделью естественной предметной области* назовем четверку  $A = \langle O, k, \Omega, \pi \rangle$ , где  $O$  – онтология ПрО;  $k$  – модель знаний ПрО;  $\Omega = \{\alpha\}$  – множество ситуаций (прецедентов) ПрО;  $\pi$  – принципы оптимальности и смысловые регуляторы процессов ПрО, например, процессов принятия решений и управления. Модель знаний  $k$  включает в себя, в частности, автоматизмы и локальные модели знаний  $\{k_{\alpha}\}$ , где  $\alpha \in \Omega$ , а также метазнания (глава 7). Рефлексия над системой знаний  $\{k_{\alpha}\}$  приводит к выходу за ее пределы и порождению нового знания.

Модель ситуации (3.22) в сочетании с моделью ПрО отражает *понимание* эволюции ситуации  $\alpha$  некоторым познающим субъектом *Agent* (индекс –  $Ag$ ), если она содержит в себе ответы на потенциальные вопросы-тесты  $\{q/Q\}_{Ag}$ . Модель ситуации  $\alpha$  является *полной* для агента  $Ag$  тогда и только тогда, когда  $\exists (g/\mu)_{Ag}: \alpha \rightarrow \{q/Q\}_{Ag}$ , где  $(g/\mu)_{Ag}$  – инструменты добычи информации, а  $\{q/Q\}_{Ag}$  – ответы на вопросы агента  $Ag$ . Поскольку у каждого агента своя модель предметной области  $A_{Ag}$ , то и полнота модели ситуации может быть разной.

### 3.6 Информационные множества. Множества обобщения

Под *информационным множеством*  $I_{\gamma}(\alpha, \{\tau/T\})$  для ситуации  $\alpha$  и аспекта  $\{\tau/T\}$  будем понимать совокупность всех реальных или гипотетических ситуаций действительности, удовлетворяющих заданным критериям истинности  $\gamma$  и совместимых с ситуацией  $\alpha$  на уровне общности  $\{\tau/T\}$ . Можно записать:

$$I_{\gamma}(\alpha, \{\tau/T\}) = \{\beta \mid \{\tau/T\}_{\beta} = \{\tau/T\}_{\alpha} \ \& \ \gamma(\beta) = true\}, \quad (3.24)$$

где  $\gamma(\ )$  – оператор оценки истинности информации. Разные ситуации  $\beta \in I_{\gamma}(\alpha)$  могут формироваться разными агентами. В самом общем случае ситуации  $\beta$  из информационного множества охватывают прошлое, настоящее и будущее развитие ситуации  $\alpha$ .

В ряде приложений оператор  $\gamma$  может осуществлять проверку принадлежности ситуаций  $\beta$  к множеству прецедентов  $\Omega$ . В таком случае будем использовать следующую нотацию:

$$I_{\Omega}(\alpha, \{\tau/T\}) = \{\beta \mid \{\tau/T\}_{\beta} = \{\tau/T\}_{\alpha} \ \& \ \beta \in \Omega\}.$$

Обратим внимание на то, что в силу универсальности понятия теста определение информационного множества включает также конструкцию

события. Действительно событие  $\langle \underline{t}/T, \underline{t}/A \rangle$  можно представить как результат теста: «Значение  $\tau/T$  в момент  $\underline{t}/A = \underline{t}$ ». Если такое представление, по каким-либо причинам неудобно, то можно использовать развернутое определение информационного множества, а именно:

$$I_{\gamma}(\alpha, \{\langle \tau/T, \underline{t}/A \rangle\}) = \{\beta \mid \{\langle \underline{t}/T, \underline{t}/A \rangle\}_{\beta} = \{\langle \underline{t}/T, \underline{t}/A \rangle\}_{\alpha} \ \& \ \gamma(\beta) = true\}. \quad (3.25)$$

В определение информационного множества при необходимости можно включить модели развития ситуаций, а именно:

$$I_{\gamma}(\alpha, \{\tau/T\}, k) = \{\beta \mid \{\underline{t}/T\}_{\beta} = \{\underline{t}/T\}_{\alpha}, k_{\alpha} \subseteq k_{\beta}, \gamma(\beta) = true\}. \quad (3.26)$$

Информационным множеством будет являться восстановление любого  $N$ -мерного образа объекта по  $M$ -мерной проекции, где  $M < N$ .

Информационное множество может быть получено путем замены в ситуации  $\alpha$  точечных событий на совместимые с ними протяженные события, а именно:

$$\alpha = \alpha(\{\langle J_{\tau} \underline{t}/T, J_{\tau} \underline{t}/A \rangle\}) \rightarrow I(\alpha) = \{\beta = \beta(\{\langle J_{\tau} \underline{t}/T, J_{\delta} \underline{t}/A \rangle\})\}.$$

Относительно простой способ построения информационного множества состоит в автоматической генерации скоростей изменений тестов (для выявления качественной прерывности):

$$\alpha = \alpha(\{\langle \underline{t}/T, \underline{t}/A \rangle\}) \rightarrow I(\alpha) = \alpha \cup \{\langle \Delta \underline{t}/T, \Delta \underline{t}/A', \underline{t}/A \rangle\}.$$

Более сложная задача состоит в генерации заданных классов событий и гипотез (закономерностей) на основе имеющихся событий. Приведем пример.

Пусть ситуация  $\alpha$  описывается множеством событий  $\{c(t)\} \equiv \{\langle c, t \rangle\}$ , где  $c(t)$  – значение временного ряда в момент времени  $t$ . Разный масштаб времени зададим с помощью функций  $Hour(t)$  – текущий час и  $DayOfWeek(t)$  – текущий день недели (1- понедельник, 2- вторник и т.д.).

Пусть значения временного ряда  $c(t)$  обобщаются с помощью орграфа доменов  $G(c)$ , а значения времени  $t$  обобщаются с помощью орграфа  $G(t)$ . Структура орграфа  $G(t)$  приведена в разделе 3.1.

Для построения информационного множества ситуации  $\alpha$  определим новый класс событий следующим образом:

$$\langle (c(t-i) * c(t-j))/C, Hour(t) = k \rangle \equiv \langle (c(t-i) * c(t-j))/C, t/A2? k \rangle;$$

$$\langle (c(t-i)/C * q/Q, Hour(t) = k \rangle \equiv \langle (c(t-i)/C * q/Q, t/A2? k \rangle;$$

$$\langle (c(t-i) * c(t-j))/C, DayOfWeek(t) = k \rangle \equiv \langle (c(t-i) * c(t-j))/C, t/A3? k \rangle;$$

$$\langle (c(t-i)/C * q/Q, DayOfWeek(t) = k \rangle \equiv \langle (c(t-i)/C * q/Q, t/A3? k \rangle;$$

где  $c/C$  – значение временного ряда соответствующего уровня общности; \* – один из символов множества  $\{\neq, \leq, \geq, <, >, =\}$ ;  $i, j, k$  – произвольные допустимые параметры;  $q$  – параметр, задаваемый орграфом доменов  $G(q)$ . Конструируемая ситуация  $\beta \in I_\gamma(\alpha)$  дополнительно к  $\{c(t)\}$  включает события указанного вида.

Построим локальную модель знаний ситуации  $\beta$ , содержащую вероятностные закономерности возникновения *аномальных событий*, например, нарушений динамики ряда.

Правило  $R=A_1 \& \dots \& A_n \rightarrow A_0$  является *вероятностной закономерностью* [39], если для любого правила  $A_{i1} \& \dots \& A_{ik} \rightarrow A_0$  такого, что  $\{A_{i1}, \dots, A_{ik}\} \subset \{A_1, \dots, A_n\}$ , условная вероятность

$$0 < p(A_0 | A_{i1} \& \dots \& A_{ik}) < p(A_0 | A_1 \& \dots \& A_n) = p(R).$$

Вероятностные неравенства, входящие в определение вероятностной закономерности, проверяются на данных с использованием статистических критериев.

Правила выявления аномальных событий, в частности правила нарушения динамики ряда, формируются на основе множества вероятностных закономерностей следующим образом [39]. Для каждого правила  $R$  берется его отрицание:  $\neg R = A_1 \& \dots \& A_n \rightarrow \neg A_0$ . Если  $p(R) > 0,9$ , то конъюнкция  $A_1 \& \dots \& A_n \& \neg A_0$  будет являться признаком аномального события, а правило

$$R' = A_1 \& \dots \& A_n \& \neg A_0 \rightarrow Abnormal \quad (3.27)$$

будет прогнозировать аномальную ситуацию с вероятностью  $p(R)$ . Обнаруженные вероятностные закономерности  $J_r R$  и  $J_r R'$  ( $J_r = p(R)$ ) помещаются в локальную базу знаний ситуации  $\beta$ .

На языке тестов правила запишутся следующим образом:

$$R(\underline{p}/P): \{\underline{a}/A\} \rightarrow \underline{b}/B; \quad R'(\underline{p}/P): \{\underline{a}/A\} \& \neg \underline{b}/B \rightarrow Abnormal \quad (3.28)$$

или

$$\begin{aligned} R(\underline{p}/P): \{\langle \underline{a}/A, \underline{t}/\Lambda \rangle\} &\rightarrow \langle \underline{b}/B, \underline{t}'/\Lambda \rangle; \\ R'(\underline{p}/P): \{\langle \underline{a}/A, \underline{t}/\Lambda \rangle\} &\& \langle \neg \underline{b}/B, \underline{t}'/\Lambda \rangle \rightarrow Abnormal, \end{aligned} \quad (3.29)$$

где  $p/P$  – вероятность, рассматриваемая как элементарный тест со своим орграфом доменов  $G(p)$ . Пример орграфа  $G(p)$ .

$p$  - Вероятность:

P3 {Маленькая  $\wedge$ 1; Средняя  $\wedge$ 2 3 4; Высокая  $\wedge$ 5}

P2 {Очень маленькая  $\wedge$ 1 [0; 0,1]; Маленькая  $\wedge$ 2 (0,1; 0,35]; Средняя  $\wedge$ 3 (0,35; 0,65]; Высокая  $\wedge$ 4 (0,65; 0,9]; Очень высокая  $\wedge$ 5 (0,9; 1]}



$P1 \{[0; 1]\}$

Граф  $G(p) = \{P1 \rightarrow P2 \rightarrow P3\}$ .

Приведем примеры возможных закономерностей:

1. Закономерность нормальной динамики (АД – артериальное давление, P1 – базовый домен теста «Вероятность»):

$(t/A2? 15) \& (АД(t-6) > АД(t)) \rightarrow (АД(t-6) > АД(t-3)), p/P1 = 0,970$ .

Данная закономерность отражает тот факт, что если давление падает в течение дня, то оно падает монотонно.

Правило, предсказывающее нарушение:

$(t/A2? 15) \& (АД(t-6) > АД(t)) \& \neg (АД(t-6) > АД(t-3)) \rightarrow Abnormal, p/P1=0,970$ .

2. Закономерность нормальной динамики:

$(t/A3 = 2) \& (c(t-5)/C < q/Q) \rightarrow (c(t-1)/C < q/Q), p/P3 = \text{«Высокая»}$ .

Правило, предсказывающее нарушение:

$(t/A3 = 2) \& (c(t-5)/C < q/Q) \& \neg (c(t-1)/C < q/Q) \rightarrow Abnormal, p/P3 = \text{«Высокая»}$ .

Эффективный путь поиска вероятностных закономерностей максимального уровня общности предоставляет метод предельных обобщений (глава 5).

Описанный выше подход естественным образом обобщается на задачу поиска детерминантов (синдромов) и предвестников некоторого события на основе многомерного временного ряда (задача прогнозирования критического или аномального события). Конкретизируем данную задачу.

Рассмотрим многомерный временной ряд (ВР), представляющий собой совокупность нескольких одномерных ВР, каждый из которых описывает изменение во времени значения какого-либо количественного или качественного признака  $\{\varphi(t)\}_i \equiv \{<\varphi, t>\}_i$ , характеризующего объект исследования. Многомерный ВР представляет собой ситуацию  $\alpha$ . Естественнно предполагать, что наступлению интересующего события предшествует некая (неизвестная) комбинация значений (неизвестных уровней общности) входных признаков (тестов) или производных тестов, событий, которую назовём *синдромом* (синонимы – детерминант, явление). Будем в этом случае говорить, что синдром *иницирует* данное событие. Синдромов может быть несколько. *Предвестником* назовём комбинацию значений только тех тестов, которые существенно связаны с наступлением события (по типу правил  $R$  и  $R'$ ) и могут быть использованы для его прогнозирования. Предвестников также может быть несколько.

В рамках развиваемого формализма первоначально строятся орграфы доменов для каждого признака –  $\{G(\varphi_i)\}$ . Строится также орграф доменов теста «Интервал поиска -  $\Delta t$ » –  $G(\Delta t)$ , определяющий масштабные уровни интервала времени до наступления события. Далее на множестве всех описаний  $\beta = \langle \{<\varphi//\Phi, t/A>\}_1, \dots, \{<\varphi//\Phi, t/A>\}_n \rangle$ , Событие –  $1/0 >_{\Delta t/A}$  находится экстремальный пограничный слой набросков, а также синдромные и вероятностные модели знаний, включая предельные модели (см. главу 5). В главе 7 дополнительно будут рассмотрены алгоритмы выявления закономерностей в процессе познания действительности.

Важными средствами построения информационного множества являются эмпирический оператор эволюции (глава 7), а также индуктивное развертывание ситуации на основе замыканий значений тестов.

Пусть запись  $\beta:\alpha$  означает, что ситуация  $\beta$  является обобщением ситуации  $\alpha$ , а  $\alpha$  рассматривается как конкретизация ситуации  $\beta$ . Обобщение будем связывать с заменой тестов  $\{a/A\}_\alpha$  на более общие тесты  $\{b/B\}_\beta$ . Для такой замены в базе знаний  $k$  должен существовать системопаттерн вида  $f/\mu: \{a/A\} \rightarrow \{b/B\}$ , с помощью которого осуществляется преобразование значений одной группы тестов в значения другой группы тестов.

Переход от первичных данных ситуации  $\alpha$  к нечетким или интервальным множествам является примером построения обобщения  $\beta$ .

Главным свойством операции обобщения является ее необратимость, т.е. из  $\alpha$  можно получить  $\beta$ , но из  $\beta$  нельзя получить  $\alpha$ . Модель иерархической структуры мира представима в виде:  $U = \{\alpha_1: \alpha_2: \dots: \alpha_n\}$ .

Множеством обобщения ситуации/образа  $\alpha$  назовем множество  $G_\gamma(\alpha)$  следующего вида:

$$G_\gamma(\alpha) = \{\beta \mid \beta:\alpha \ \& \ \gamma(\beta) = true\}. \quad (3.30)$$

Пусть, например, ситуация/образ  $\alpha$  описывается с помощью базовых доменов  $N$  тестов, орграфы которых содержат  $M_i$  доменов (без потери общности примем, что каждое вхождение теста – это отдельный тест). Общее число различных описаний действительности, получаемых из  $\alpha$  и обобщающих  $\alpha$ , составляет  $M_1 \times M_2 \times \dots \times M_N - 1$ . Положим  $\gamma(\beta) = true$ , если  $\beta$  является одним из таких описаний. Роль отображений  $\{f/\mu\}$  играют восходящие преобразования в рамках орграфов доменов. Таким образом, получили точную оценку мощности множества  $G_\gamma(\alpha)$  для данного примера, а именно:

$$|G_\gamma(\alpha)| = M_1 \times M_2 \times \dots \times M_N - 1.$$

В данном примере  $G_s(\alpha) = \alpha \cup G_\gamma(\alpha)$ . Однако далеко не всегда множество

обобщения связано с орграфом набросков.

Приведем пример обобщения из области анализа электроэнцефалограмм (ЭЭГ) [59]: ставится задача найти отличия в ЭЭГ биологических объектов, подвергавшихся определенным видам воздействия, от ЭЭГ этих же биологических объектов в обычной обстановке (фоновых ЭЭГ). Множество всех ЭЭГ образуют ситуацию действительности  $\alpha$ . Для упрощения анализа строится обобщение  $\alpha$ . Сначала производится сглаживание ЭЭГ по методу скользящего среднего. Затем участки сигнала с положительной первой производной заменяются на 1, остальные на 0. Таким образом, вместо исходного сигнала обработке подвергается последовательность, состоящая из 0 и 1. В более «тонком» варианте анализа для величины производной необходимо ввести орграф доменов с несколькими уровнями общности (в исходном варианте всего 2 уровня, причем второй уровень дихотомический – {1; 0}). Разные преобразования ЭЭГ дают разные ситуации  $\beta$ , при этом  $\beta:\alpha$ . В итоге строится множество  $G_\gamma(\alpha)$ , содержащее ситуации  $\beta$  с разным уровнем детальности описания. Искомые отличия дают синдромные модели знаний, которые находятся с помощью метода предельных обобщений (см. главу 5). Дополнительные возможности дает сочетание метода предельных обобщений и модели многоуровневых набросков  $\{Gs(\beta)\}$ .

Важной прикладной задачей построения множества обобщения является автоматическое многоуровневое реферирование ситуации действительности, т.е. построение набросков ситуации, затем набросков набросков и т.д. На этом принципе построена модель многоуровневых набросков ситуации, образа, текста (глава 6).

Примеры множеств обобщения дает символическая динамика. Часто фазовое пространство динамической системы удается разбить на области таким образом, что любая ее траектория однозначно определяется последовательностью прохождения этих областей. Символический образ представляет собой ориентированный граф, построенный по заданному разбиению фазового пространства, что позволяет использовать хорошо разработанные алгоритмы на графах [122]. Примеры множеств обобщения дают гранулярные вычисления и интервальная математика.

*Множеством обобщения для группы ситуаций или образов  $\{\alpha\}$  назовем множество  $G_\gamma(\{\alpha\})$  следующего вида:*

$$G_\gamma(\{\alpha\}) = \{\beta \mid \beta:\{\alpha\} \equiv (\beta:\alpha \ \forall \alpha \in \{\alpha\}) \ \& \ \gamma(\beta) = true\}. \quad (3.31)$$

Следует иметь в виду, что, как правило,  $G_\gamma(\{\alpha\}) \neq \cup_{\alpha \in \{\alpha\}} G_\gamma(\alpha)$ .

Эффективным средством формирования множества обобщений для произвольной совокупности прецедентов  $\Omega$  с известным заключением является метод предельных обобщений. В качестве  $\{\beta\}$  выступают

предельные модели знаний (главы 5, 6) и экстремальные слои набросков.

Введем *операцию детализации* значений теста  $\tau$  на основе орграфа доменов  $G(\tau)$ :

$$(\underline{\tau})/T' \downarrow_G (\underline{\tau})/T. \quad (3.32)$$

Данная операция позволяет по заданному множеству значений  $(\underline{\tau})/T$  получить полное множество более точных значений  $(\underline{\tau})/T'$ . Пример на основе орграфа  $G_1(\text{Возраст}) = \{B1 \rightarrow B2 \rightarrow B3\}$  из главы 2:

Возраст/ B2? {юный; молодой}  $\downarrow_G$  Возраст/ B3? молодой;

Возраст/ B1? [0; 33]  $\downarrow_G$  Возраст/ B3? молодой.

Для маркировки выборок объектов ПрО удобно использовать понятие «обобщенный код множества». Кортеж  $K = \langle \{\underline{\tau}_1\}/T_1; \dots; \{\underline{\tau}_n\}/T_n \rangle$ , где  $T_i \in G(\tau_i)$ , будем называть *обобщенным кодом множества*

$$M(T'_1, \dots, T'_n) = A_1 \times A_2 \times \dots \times A_n,$$

где  $A_i = (\underline{\tau}_i)/T'_i$ ,  $T'_i \in G(\tau_i)$ , если  $(\underline{\tau}_i)/T'_i \downarrow_G (\underline{\tau}_i)/T_i$  ( $i = 1, \dots, n$ ).

Очевидно, один и тот же обобщенный код  $K$  представляет множества  $M(\{T'\})$  разного уровня общности. Представление подмножеств универсума  $U$  обобщенными кодами позволяет выполнять над ними все теоретико-множественные операции по формальным правилам, посредством определенных действий над обобщенными кодами.

С помощью обобщенных кодов удобно описывать высокоуровневые запросы к БД, например, пусть требуется выбрать всех пациентов с параметрами: {пожилой, высокая температура}. Соответствующий обобщенный код будет выглядеть так:  $K = \langle \text{Возраст/B3? пожилой; Температура тела/ T4? высокая} \rangle$ . В результат поиска войдут медицинские карты тех пациентов, которые удовлетворяют коду  $K$ . С использованием обобщенных кодов может быть реализован интеллектуальный интерфейс.

### 3.7 Сходство ситуаций действительности

Пусть ситуации  $\alpha$  и  $\beta$  описываются с помощью одного банка тестов  $\{G(\tau)\}$ . Пример фрагмента банка тестов «Клиническая картина пневмонии» (составлен по материалам [188]):

$$\{G(\tau)\} = \{$$

**КАШЕЛЬ:** отсутствует, имеется {

*Характер:* сухой, влажный непродуктивный, влажный продуктивный.

*Громкость и тембр:* короткий, тихий, мягкий, осторожный, громкий, надсадный, грубый, мучительный, беззвучный, лающий, сиплый, «блеяние козы».

*Время появления:* утренний, вечерний, ночной, в течение всего дня.

*Ритм:* постоянный, периодический, покашливание, приступообразный, удушательный.

*Усиление кашля:* на правом боку, на левом боку, при глубоком вдохе, при циклоническом типе погоды, в дни геомагнитных бурь}

**МОКРОТА:** отсутствует, имеется {

*Количество за сутки:* <числовые значения>.

*Характер:* слизистая, серозная, гнойная, слизисто-гнойная, серозно-гнойная, прожилки крови, в виде кровяных сгустков, “ржавая”.

*Отхаркиваемость:* трудноотделяемая, хорошо отходит.

*Цвет:* бесцветная, белая, серая, желтоватая, зеленоватая, коричневая, розовая.

*Запах:* отсутствует, затхлый, гнилостный, зловонный, нестерпимо-отвратительный, горелого мяса.

*Консистенция:* густая, вязкая, жидкая, тягучая, стекловидная.}

**БОЛЬ В ГРУДНОЙ КЛЕТКЕ:** отсутствует, имеется {

*Локализация:* по ходу трахеи, по передней поверхности грудной клетки, в правой половине грудной клетки, в левой половине грудной клетки, по задней поверхности грудной клетки.

*Иррадиация:* отсутствует, в область живота, в спину, в плечо, в шею.

*Характер:* тупая, острая, колющая, ноющая, внезапно сильная, покалывание, чувство тяжести, интенсивная, слабо интенсивная.

*Усиление:* при дыхании, при кашле, в положении на правом боку, в положении на левом боку, при сдавлении грудной клетки, при наклоне в большую сторону.

*Связана:* с дыханием, с кашлем, с изменением положения тела, с глубоким вдохом, с движением.

*Распространенность:* локальная, разлитая} ...}.

Все тесты представлены базовыми доменами. Не представляет труда для каждого основного или дополнительного теста ввести обобщающие домены.

Примеры описания двух клинических случаев на основе банка тестов “Клиническая картина пневмонии”:

$\alpha = \{КАШЕЛЬ? \text{ имеется } \{Характер? \text{ влажный продуктивный. Громкость и тембр? короткий, тихий. Время появления? утренний. Ритм? постоянный. Усиление кашля? на правом боку.}\}. МОКРОТА? \text{ имеется } \{Количество за сутки? 120 \text{ мл. Характер? “ржавая”}. Отхаркиваемость? трудноотделяемая. Цвет? серая. Запах? затхлый. Консистенция? густая}\}. БОЛЬ В ГРУДНОЙ КЛЕТКЕ? \text{ отсутствует}\};$

$\beta = \{ \text{КАШЕЛЬ? имеется} \{ \text{Характер? влажный непродуктивный. Громкость и тембр? мучительный. Время появления? в течение всего дня. Ритм? постоянный. Усиление кашля? при глубоком вдохе.} \}. \text{МОКРОТА? имеется} \{ \text{Количество за сутки? 200 мл. Характер? слизисто-гнояная. Отхаркиваемость? трудноотделяемая. Цвет? желтоватая. Запах? затхлый. Консистенция? густая} \}. \text{БОЛЬ В ГРУДНОЙ КЛЕТКЕ? имеется} \{ \text{Локализация? по ходу трахеи. Иррадиация? отсутствует. Характер? слабо интенсивная. Усиление? при кашле, в положении на правом боку. Связана? с кашлем. Распространенность? локальная} \} \}.$

Насколько данные ситуации схожи?

Прежде чем ввести меру сходства уточним маркировку ссылки на элементарный тест. Ссылка на тест может отражать иерархию вложенности тестов:

Тест1.Тест2...ТестN? *Значение*.

Подобная запись означает, что определены все вложенные цепочки тестов:

Тест1? *Значение*

Тест1.Тест2? *Значение*

Тест1.Тест2...ТестN? *Значение*

или в более общей нотации

$$\text{Тест1?} \text{Знач1.} \text{Тест2?} \text{Знач2} \dots \text{ТестN?} \text{Значение.} \quad (3.33)$$

Будем говорить, что данная маркировка задает *детализацию N-го уровня*.

Выделим в ситуации  $\alpha$  все элементарные тесты согласно общей маркировке:

$\alpha = \{$

КАШЕЛЬ? имеется.

КАШЕЛЬ.Характер? влажный продуктивный.

КАШЕЛЬ.Громкость и тембр? короткий, тихий.

КАШЕЛЬ.Время появления? утренний.

КАШЕЛЬ.Ритм? постоянный.

КАШЕЛЬ.Усиление кашля? на правом боку.

МОКРОТА? имеется.

МОКРОТА.Количество за сутки? 120 мл.

МОКРОТА.Характер? “ржавая”.

МОКРОТА.Отхаркиваемость? трудноотделяемая.

МОКРОТА.Цвет? серая.

МОКРОТА.Запах? затхлый.

МОКРОТА.Консистенция? густая.

БОЛЬ В ГРУДНОЙ КЛЕТКЕ? отсутствует}.

В ситуации  $\beta$  оставим только те тесты, которые имеются в ситуации  $\alpha$ .

$\beta = \{$

*КАШЕЛЬ? имеется*

*КАШЕЛЬ.Характер? влажный непродуктивный.*

*КАШЕЛЬ.Громкость и тембр? мучительный.*

*КАШЕЛЬ.Время появления? в течение всего дня.*

*КАШЕЛЬ.Ритм? постоянный.*

*КАШЕЛЬ.Усиление кашля? при глубоком вдохе*

*МОКРОТА? имеется.*

*МОКРОТА.Количество за сутки? 200 мл.*

*МОКРОТА.Характер? слизисто-гнояная.*

*МОКРОТА.Отхаркиваемость? трудноотделяемая.*

*МОКРОТА.Цвет? желтоватая.*

*МОКРОТА.Запах? затхлый.*

*МОКРОТА.Консистенция? густая.*

*БОЛЬ В ГРУДНОЙ КЛЕТКЕ? имеется}.*

Курсивом выделены тесты, значения которых совпадают с соответствующими тестами ситуации  $\alpha$ . Всего совпадают значения 6 тестов из 14. Подчеркнем, что все тесты представлены базовыми доменами.

В рамках детализации 1-го уровня ситуации  $\alpha$  и  $\beta$  будут иметь вид:

$\alpha = \{КАШЕЛЬ? \text{ имеется}; МОКРОТА? \text{ имеется}; БОЛЬ В ГРУДНОЙ КЛЕТКЕ? \text{ отсутствует}\};$

$\beta = \{КАШЕЛЬ? \text{ имеется}; МОКРОТА? \text{ имеется}; БОЛЬ В ГРУДНОЙ КЛЕТКЕ? \text{ имеется}\}.$

В рамках максимального уровня обобщения ситуаций совпадают значения 2-х тестов из 3.

Для элементарного теста  $\tau$  справедливо следующее правило: если  $\underline{\tau}$  – произвольное значение теста, то элементарным тестом будет также конструкция

$\tau.\underline{\tau} / \{Истина; Ложь\}?$

Более того, если  $w(\{\underline{\tau}\})$  – произвольная логическая формула, составленная из результатов теста  $\tau$ , то элементарным тестом будет также конструкция

$\tau.w(\{\underline{\tau}\}) / \{Истина; Ложь\}?$

Домен  $\{Истина; Ложь\}$  эквивалентен доменам  $\{Есть; Нет\}$ ,  $\{Да; Нет\}$ ,  $\{Имеется; Отсутствует\}$  и т.п. Примеры:

Об-но.Веки? *Гиперемия;*

Об-но.Веки.Гиперемия? *Есть*;  
 Жалобы? *Боль при глотании*;  
 Жалобы? *Изжога отсутствует*;  
 Жалобы.[*Боль при глотании*  $\wedge$   $\neg$ *Изжога*]? *Истина*.

Подобные трансформации записей результатов тестов необходимо учитывать при построении мер сходства.

Различных мер сходства огромное количество и мы не ставим перед собой цель охватить какую-либо их часть. Для нас важна иллюстрация общего подхода к определению подобных мер на основе развиваемого формализма. К принципиальным положениям относятся два факта: 1) любая мера сходства или различия ситуаций является элементарным тестом со своим орграфом доменов; 2) сходство ситуаций существенно зависит от уровня общности описания ситуаций.

Определим базовый домен теста «*Сходство двух ситуаций/образов*» в аспекте  $\{\tau/T\}$  с помощью разновидности коэффициента Джаккарда:

$$Sim(\alpha, \beta | \{\tau/T\}) = |\{\underline{\tau/T}\}_{\alpha} \cap \{\underline{\tau/T}\}_{\beta}| / |\{\tau/T\}_{\alpha} \cup \{\tau/T\}_{\beta}| \quad (3.34)$$

либо

$$Sim(\alpha, \beta | \{\tau/T\}) = |\{\underline{\tau/T}\}_{\alpha} \cap \{\underline{\tau/T}\}_{\beta}| / |\{\tau/T\}_{\alpha} \cap \{\tau/T\}_{\beta}|. \quad (3.35)$$

В первом варианте в расчет принимаются все тесты. Во втором варианте в расчет принимаются только те тесты, которые имеются в каждой ситуации или каждом образе (для примеров будем использовать именно этот вариант). Совпадение названий тестов определяется в смысле (3.33).

Аспект  $\{\tau/T\}$  может касаться любых фрагментов модели ситуации (3.22). Можно говорить о тематическом сходстве по актуализированным тематическим признакам. Тематическое сходство ситуаций распадается на сходство параметров порядка, событий, действий и т.д.

**Предложение 3.4** Сходство, определяемое выражениями (3.34) и (3.35), рефлексивно, симметрично и, кроме того, для любых ситуаций  $\alpha, \beta$  выполняется:

$$0 \leq Sim(\alpha, \beta | \{\tau/T\}) \leq Sim(\alpha, \beta | \{\tau/T\}') \leq 1, \quad (3.36)$$

где  $\{\tau/T\}'$  задают меньший уровень детализации, чем  $\{\tau/T\}$  в рамках Банка тестов  $\{G(\tau)\}$ .

Из утверждения следует, что с повышением уровня общности сходство ситуаций может только возрастать. Из соотношения (3.36) следует также, что в рамках заданного банка тестов  $\{G(\tau)\}$  существует предел сходства двух ситуаций:

$$Sim(\alpha, \beta | \{\tau/T\})^{\wedge} = \max_{\{\tau/T\}} Sim(\alpha, \beta | \{\tau/T\}), \quad (3.37)$$



где  $\{\tau/T\}^\wedge$  – домены максимального уровня общности в рамках  $\{G(\tau)\}$ .

Пусть  $G_s(\alpha)$  и  $G_s(\beta)$  орграфы набросков ситуаций  $\alpha$  и  $\beta$ . Пусть  $\alpha_s \in G_s(\alpha)$ ,  $\beta_s \in G_s(\beta)$ ,  $\{\alpha_s\}_v$  и  $\{\beta_s\}_v$  – адекватные слои набросков, а  $\{\alpha^*\}$  и  $\{\beta^*\}$  – множества финитных набросков ситуаций. Можно ожидать выполнения следующих соотношений ( $\forall v$ ):

$$Sim(\alpha, \beta) \leq \max \{Sim(\alpha_s, \beta_s) | \alpha_s \in \{\alpha_s\}_v, \beta_s \in \{\beta_s\}_v\} \leq \max_{\{\alpha^*\}, \{\beta^*\}} Sim(\alpha^*, \beta^*).$$

Легко убедиться, что максимальное сходство двух модельных клинических ситуаций  $\alpha$  и  $\beta$  будет иметь место в рамках детализации 1-го уровня и определяется значением

$$Sim(\alpha, \beta | 1\text{-й уровень детализации}) = Sim(\alpha^*, \beta^*) = 2/3.$$

Детализация 1-го уровня является в данном примере ничем иным, как финитными набросками.

Сходство модельных ситуаций в рамках полной онтологии «Клиническая картина пневмонии» определяется величиной

$$Sim(\alpha, \beta | \{G(\tau)\}) = 6/14 = 3/7.$$

Таким образом, получили диапазон изменения сходства двух ситуаций в зависимости от уровня детализации –  $[3/7; 2/3]$ .

Орграф  $G(\text{Сходство})$  можно задать, например, следующим образом:

Тест «Сходство»:

D3 {Несхожи  $\wedge$  1 2 3; Схожи  $\wedge$  4 5; Идентичны  $\wedge$  6}

D2 {Абсолютно несхожи  $\wedge$  1 [0; 0]; Практически несхожи  $\wedge$  2 (0; 0,2]; Несхожи  $\wedge$  3 (0,2; 0,5]; Незначительно схожи  $\wedge$  4 (0,5; 0,8]; Практически схожи  $\wedge$  5 (0,8; 1); Идентичны  $\wedge$  6 [1; 1]}

D1 {[0; 1]};

Орграф  $G(\text{Сходство}) = \{D1 \rightarrow D2 \rightarrow D3\}$ .

Данное представление не единственно. Более того, существует бесконечное множество вариантов орграфа  $G(\text{Сходство})$ . В частности, для каждого элемента доменов D2 и D3 могут быть заданы функции принадлежности, определяющие нисходящую детерминацию. На основе любого варианта орграфа  $G(\text{Сходство})$  можно построить множество структурно-завершенных орграфов  $\{G^+\}$ .

Подобие двух ситуаций в некотором заданном аспекте можно определить следующим образом: ситуации  $\alpha$  и  $\beta$  подобны в аспекте  $\{\tau/T\}$ , если выполняется соотношение  $\{\tau/T\}_\beta = \{\tau/T\}_\alpha$

В отличие от теста «Сходство» тест «Подобие» является ассоциативным (в рамках фиксированного аспекта). Орграф теста «Подобие» состоит из одного дихотомического домена:

D1 {Подобны; Не подобны}.

Предложенный подход к определению сходства, подобия (эквивалентности) вводит в процесс познания действительности необозримое множество корреляций. Эквивалентность  $\alpha$  и  $\beta$  может в одном случае иметь характер сходства, а в другом, когда акцентирован другой признак (аспект), характер оппозиции. Эта множественная соотносимость ситуаций образует основу для чрезвычайно сложной смысловой регуляции действительности.

Рассмотрим простой пример определения сходства ситуаций на разных уровнях общности. Пусть заданы орграфы доменов трех тестов «Температура тела», «Возраст» и «Диагноз».

$\tau_1$  – *Температура тела*:

T1 = [34,40] градусов;

T2 = {[34, 35], (35, 36.5), [36.5, 36.8], (36.8, 37.5), [37.5, 40]};

T3 = {пониженная; нормальная; повышенная; высокая};

T4 = {нормальная, ненормальная}.

$G(\tau_1) = \{T1 \rightarrow T2 \rightarrow T3 \rightarrow T4\}$ .

Правила преобразования:

T2. {[34, 35], (35, 36.5)}  $\rightarrow$  T3. "пониженная";

T2. {[36.5, 36.8]}  $\rightarrow$  T3. "нормальная";

T2. {(36.8, 37.5)}  $\rightarrow$  T3. "повышенная"; T2. {[37.5, 40]}  $\rightarrow$  T3. "высокая";

T3. "нормальная"  $\rightarrow$  T4. "нормальная";

T3. {пониженная; повышенная; высокая}  $\rightarrow$  T4. "ненормальная";

$\tau_2$  – *Возраст*:

B1 = [0...100];

B2 = {молодой, средних лет, пожилой, старческий}.

$G(\tau_2) = \{B1 \rightarrow B2\}$ .

Правила преобразования:

B1.[0...33]  $\rightarrow$  B2. "молодой"; B1.[34...59]  $\rightarrow$  B2. "средних лет";

B1.[60...69]  $\rightarrow$  B2. "пожилой"; B1.[70...100]  $\rightarrow$  B2. "старческий".

$d$  – *Диагноз*:

D = {DS<sub>1</sub>; DS<sub>2</sub>}.

Приведем фрагмент базы данных медицинских карт (DS – диагноз). Исходные данные (фрагменты карт) зададим в виде формы «T1-B1». Каждая строка – это конкретный клинический случай. С помощью орграфов  $G(\tau_1)$  и  $G(\tau_2)$  сформируем обобщенные описания и сохраним их в формах «T3 - B1», «T3 - B2», «T4 - B2», «T2 - B2», «T4 - B1».

Форма «Т1 – В1»		
$\tau_1/T1$	$\tau_2/B1$	d
36.0	12	DS <sub>1</sub>
36.7	87	
37.2	32	DS <sub>2</sub>
39.0	50	

Форма «Т3 – В1»		
$\tau_1/T3$	$\tau_2/B1$	d
Пониженная	12	DS <sub>1</sub>
Нормальная	87	
Повышенная	32	DS <sub>2</sub>
Высокая	50	

Форма «Т3 – В2»		
$\tau_1/T3$	$\tau_2/B2$	d
Пониженная	Молодой	DS <sub>1</sub>
Нормальная	Старчeck.	
Повышенная	Молодой	DS <sub>2</sub>
Высокая	Сред. лет	

Форма «Т4 – В2»		
$\tau_1/T4$	$\tau_2/B2$	d
Ненорм.	Молодой	DS <sub>1</sub>
Нормальная	Старчeck.	
Ненорм.	Молодой	DS <sub>2</sub>
Ненорм.	Сред. лет	

Форма «Т2 – В2»		
$\tau_1/T2$	$\tau_2/B2$	d
(35, 36.5)	Молодой	DS <sub>1</sub>
[36.5, 36.8]	Старчeck.	
(36.8, 37.5)	Молодой	DS <sub>2</sub>
[37.5, 40]	Сред. лет	

Форма «Т4 – В1»		
$\tau_1/T4$	$\tau_2/B1$	d
Ненорм.	12	DS <sub>1</sub>
Нормальная	87	
Ненорм.	32	DS <sub>2</sub>
Ненорм.	50	

Сходства между группами клинических ситуаций {1, 2} и {3, 4} на уровне «Т1-В1» нет, т.е.  $Sim(\alpha_i, \alpha_j | \{\tau_1/T1, \tau_2/B1, d\}) = 0$ , где  $i = 1,2; j = 3,4$ . Минимальное сходство есть внутри каждой из групп:

$$Sim(\alpha_1, \alpha_2 | \{\tau_1/T1, \tau_2/B1, d\}) = 1/3.$$

Аналогичная ситуация и на уровне «Т3 – В1». Легко убедиться в следующих результатах:

$$Sim(\alpha_3, \alpha_4 | \{\tau_1/T4, \tau_2/B1, d\}) = 2/3; Sim(\alpha_1, \alpha_3 | \{\tau_1/T4, \tau_2/B2, d\}) = 2/3.$$

Последний пример интересен еще и тем, что он демонстрирует единственный конфликт, а именно: имеют место разные заключения при совпадающих значениях тестов. При рассмотрении метода предельных обобщений описание «Т4 – В2» будет отнесено к разряду *закритических*, т.е. содержащих конфликт. На масштабных уровнях «Т2 – В2» и «Т3 – В2» сходство между ситуациями одинаково и не превышает 1/3.

Таким образом, сходство между некоторыми клиническими ситуациями при повышении уровня общности изменяется от нуля до 2/3, иногда даже приводя к конфликту данных.

## ГЛАВА 4. КОНФИГУРАТОРЫ ТЕСТОВ

В главе 2 было введено понятие ориентированного графа доменов произвольного теста. Орграф доменов – это математический объект. Процедурная реализация данного объекта может быть выполнена разными способами. Один из них – представление орграфа доменов в виде конфигулятора. На самом деле во второй и третьей главе мы уже использовали неявно данное представление. Теперь дадим его точное формальное описание, а также рассмотрим способы автоматического формирования конфигуляторов тестов в разных ПрО.

### 4.1 Определение конфигулятора. Примеры

Общую схему конфигуляторов с использованием синтаксиса лексических деревьев можно представить следующим образом:

$$\begin{aligned} \text{Тест } [^{\wedge}\text{Тест}...] \{ \# \text{Тест}X... \} & \{ & (4.1) \\ \text{Dom}_1 [^{\wedge}\text{Dom}_1...] \{ \# \text{Dom}X... \} & \{ ; ; \} \\ \text{Dom}_2 [^{\wedge}\text{Dom}_2...] \{ \# \text{Dom}Y... \} & \{ ; ; \} \\ \dots & \\ \text{Dom}_N [^{\wedge}\text{Dom}_N...] & \{ ; ; \} \}, \end{aligned}$$

где ‘Тест’ – название теста; ‘ $\wedge$ Тест...’ – список условных обозначений теста; ‘#ТестX...’ – список ссылок на более общие тесты; ‘Dom<sub>K</sub>’ – название K-го домена (произвольный терм, состоящий из букв и/или чисел); ‘ $\wedge$ Dom<sub>K</sub>...’ – список условных обозначений K-го домена; ‘#DomX...’ – ссылка на домены предки; { ; ; } – список элементов домена со связями обобщения. Каждый элемент домена может иметь собственный список обозначений, которые также играют роль символов групп обобщения. Элементы доменов могут содержать дополнительные конструкции, которые обеспечивают однозначность вычислительных схем в зависимости от внешних тестов.

Порядок размещения доменов в конфигуляторе – сверху вниз и слева направо – означает рост точности значений теста за счет большей детализации (увеличения числа элементов). В упорядоченной последовательности доменов метки элементов любого домена явно задают однозначные правила перерасчета значений из домена-предка в домен-потомок, размещенный выше или слева. Конфигураторы тестов, как и орграфы доменов, будем обозначать *G*. Примеры конфигуляторов:

*Уровень гемоглобина*  $\wedge$ Hb {

$D4$  {Нормальный уровень  $Hb \wedge N$ ; Ненормальный уровень  $Hb \wedge a b$ }  
 $D3$  {Анемия  $\wedge a$ ; Нормальный уровень  $Hb \wedge N$ ; Повышенный уровень  $Hb \wedge b$ }  
 $D2$  {Резко выраженная анемия  $\wedge a [20; 50]$ ; Выявленная анемия  $\wedge a (50; 90)$ ;  
 Умеренная анемия  $\wedge a (90; 120)$ ; Нормальный уровень  $Hb \wedge N (120; 160)$ ;  
 Уровень  $Hb$  умеренно повышен  $\wedge b (160; 170)$ ; Повышенный уровень  $Hb$   
 $\wedge b (170; 180)$ ; Резко повышенный уровень  $Hb \wedge b (180; 200)$  }  
 $D1$  {[20; 200]}  
 $G(Hb) = \{D1 \rightarrow D2 \rightarrow D3 \rightarrow D4\}$ .

*Температура блока химических батарей*  $\wedge TBXB$  {  
 $5 \# 2$  {Средняя и ниже  $\wedge C H$ ; Выше среднего  $\wedge B$ }  
 $4 \# 2$  {Ниже среднего  $\wedge H$ ; Средняя и выше  $\wedge C B$ }  
 $3$  {Средняя  $\wedge C$ ; Ниже или выше среднего  $\wedge H B$ }  
 $2$  {Низкая  $\wedge H [0; 20]$ ; Средняя  $\wedge C (20; 30)$ ; Высокая  $\wedge B (30; 50)$ }  
 $1$  {[0; 50]} }

$G(TBXB) = \{1 \rightarrow 2 \rightarrow 3; 2 \rightarrow 4; 2 \rightarrow 5\}$ .

*Интенсивность (Писк, Скорость, Ускорение, Сила,...)* { $6 \# 3$  {Small  $\wedge 1$ ; Not Small  $\wedge 3$  }  
 $5 \# 3$  {Medium  $\wedge 3$ ; Not Medium  $\wedge 4$  }  
 $4$  {Large  $\wedge 4$ ; Not Large  $\wedge 3$  }  
 $3$  {Small  $\wedge 2$ ; Medium  $\wedge 3$ ; Large  $\wedge 4$  }  
 $2$  {Very small  $\wedge 1 [0; 15]$ ; Small  $\wedge 2 (15; 40)$ ;  
 Medium  $\wedge 3 (40; 60)$ ; Large  $\wedge 4 (60; 85)$ ; Vary large  $\wedge 5 (85; 100)$ }  
 $1$  {[0; 100]} }

$G(\text{Интенсивность}) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4; 3 \rightarrow 5; 3 \rightarrow 6\}$ .

При всей лаконичности записи каждый конфигурактор содержит все необходимые сведения для однозначной интерпретации схем преобразований одних доменов в другие. В этом и заключается главный смысл введения понятия «конфигуратор теста».

Как и исходные объекты (орграфы доменов тестов) конфигураторы чаще всего содержат экспертные знания и являются, по сути, моделью индуктивного обобщения результатов теста. Для числовых тестов ключевым моментом является переход от непрерывного интервала к дискретному разбиению (фазовый переход от бесконечности к конечности). Подобный переход может быть выполнен разными способами, что, безусловно, отражается на результатах моделирования. Важно отметить, что любой числовой тест всегда имеет множественное дискретное представление (интервальное, символическое). Кроме того, между базовым и следующим уровнем конфигуратора числового теста может быть образована интервальная иерархия по типу разбиения или физического фрактала.

С точки зрения логики конфигураторы представимы в виде совокупности эмпирических аксиом, формирующих основу логической

эмпирической теории предметной области. В аксиомы вложена система интерпретаций ПрО, т.е. семантический уровень (числовые данные не содержат интерпретацию, интервальные данные содержат лишь частично). Интерпретируемость данных и результатов обработки данных в системе понятий предметной области является необходимым условием получения полезного результата.

Конфигураторы тестов – это не только набор эмпирических аксиом, но также и набор операций (непрерывных, интервальных, нечетких, сетевых, лингвистических, фрактальных), которые можно производить со значениями тестов разного уровня общности (разных сортов) с точки зрения рассматриваемой предметной области.

Реализовать конфигураторы для конкретной задачи можно, например, с помощью электронной таблицы [144, 159].

## **4.2 Способы формирования доменов и конфигураторов**

### *4.2.1 Квалиметрический подход*

Особенности подхода рассмотрим на примере задачи расчета Индекса Физического Здоровья (ИФЗ). Несмотря на огромное число исследований, данная задача не теряет свою актуальность [120]. При этом нас будет интересовать не сама задача как таковая, а способы формирования конфигураторов тестов, используемых в данной задаче, и возможности применения конфигураторов для построения моделей знаний разного уровня общности.

Разные по масштабу диапазоны норм у контролируемых показателей организма на разных этапах онтогенеза требуют разработки специальной системы их количественной оценки. Кроме того, необходим поиск приемлемого способа интегральной оценки соматического здоровья человека на этапах онтогенеза, используя заданный комплекс функционально-адаптационных характеристик целого организма и его отдельных систем [43]. Метод предельных обобщений, используя Банк тестов и базу прецедентов, предлагает эффективный на практике способ интегральной оценки, который заключается в построении предельных моделей знаний (синдромных и вероятностных).

Одним из вариантов структуризации диапазона норм для всех медико-физиологических показателей (во всех возрастных группах) является деление диапазона на  $N$  оценочных зон, для которых установлены определенные границы. Приемлемым для практики вариантом может быть, например,  $N=6$ . Рассмотрим для конкретности именно этот вариант.

Статистический анализ результатов измерений показателей организма

выявил, что все они подчиняются закону распределения, который близок к нормальному, и это позволяет использовать стандартные отклонения в качестве критерия для установления границ оценочных зон. Один из вариантов границ представлен на рис. 4.1 [120].

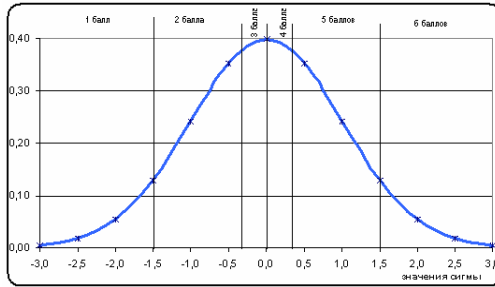


Рис. 4.1 – Модель структурирования диапазона значений для всех контролируемых показателей организма

Как видно из рисунка, зона от среднегруппового значения показателя (медианы) и до точки  $+0,3$  сигмы получает «удовлетворительную» оценку и по балльно-рейтинговой шкале 4 балла. Следующий за ней интервал от  $+0,31$  до  $+1,5$  сигмы оцениваются «хорошо» и в 5 баллов, а зона более  $+1,5$  сигмы получает оценку «отлично» и 6 баллов. Аналогичны сигмальные интервалы и в сторону уменьшения от среднего значения показателя с соответствующими оценками: «ниже среднего» = 3 балла; «плохо» = 2 балла; «очень плохо» = 1 балл.

На основе статистического анализа и выбранной экспертной модели стандартных отклонений должны быть структурированы диапазоны нормы всех показателей в  $K$  возрастных группах (отдельно для мужчин и женщин) с присвоением им балльно-рейтинговых оценок, что позволит сравнивать значения разных показателей как одного человека, так и у людей различного возраста и пола (на практике значение  $K$  колеблется от 10 до 20). С возрастом меняется не только модель нормы, представленная на рис. 4.1, но также информативность отдельных показателей [43]. Синдромы модели знаний автоматически учитывают разную информативность при расчете ИФЗ.

Ниже представлен типовой конфигуратор произвольного показателя, который отвечает модели, изображенной на рис. 4.1:

$$\begin{aligned} \text{Показатель } \{ & \\ & 6 \# 3 \{ \text{Хороший} \wedge 5; \text{не Хороший} \wedge 1 \ 3 \} \\ & 5 \# 3 \{ \text{Удовлетворительный} \wedge 3; \text{не Удовлетворительный} \wedge 1 \ 5 \} \end{aligned} \quad (4.2)$$

- 4 { Плохой  $\wedge^1$ ; не Плохой  $\wedge^3$  5 }  
 3 { Плохой  $\wedge^1$  2;  
 Удовлетворит.  $\wedge^3$  4;  
 Хороший  $\wedge^5$  6 }  
 2 { Очень плохой  $\wedge^1$  [p1; p2];  
 Плохой  $\wedge^2$  (p2; p3);  
 Ниже среднего  $\wedge^3$  (p3; p4);  
 Удовлетворит.  $\wedge^4$  (p4; p5);  
 Хороший  $\wedge^5$  (p5; p6);  
 Отличный  $\wedge^6$  (p6; p7) }  
 1 { [p1; p7] }},

где  $p_1 = -3$ ;  $p_2 = -1,5$ ;  $p_3 = -0,3$ ;  $p_4 = 0$ ;  $p_5 = 0,3$ ;  $p_6 = 1,5$ ;  $p_7 = 3$ . Значения параметров  $\{p_1, \dots, p_7\}$  относятся к базовым онтологическим соглашениям модели здоровья. При необходимости могут быть использованы другие разбиения. Помимо указанных выше параметров онтологические соглашения должны содержать параметры нормального распределения (среднее и сигму) по каждому показателю.

Банк показателей (тестов) модели здоровья  $\{G(\tau)\}$  может содержать производные тесты, которые вычисляются на основе первичных тестов.

Орграф  $G(\text{Показатель})$  имеет вид:

$$G(\text{Показатель}) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4; 3 \rightarrow 5; 3 \rightarrow 6\}.$$

Домены 4, 5, 6 образуют листья орграфа. Ствол орграфа имеет вид:

$$\tilde{G}(\text{Показатель}) = \{1 \rightarrow 2 \rightarrow 3\}.$$

Каноническое представление орграфа  $\underline{G}(\text{Показатель})$  имеет вид:

$$\begin{aligned} \text{Показатель } \{ \\ & 6 \#3 \{1 \ 2 \ 3 \ 4; 5 \ 6\} \\ & 5 \#3 \{3 \ 4; 1 \ 2 \ 5 \ 6\} \\ & 4 \{1 \ 2; 3 \ 4 \ 5 \ 6\} \\ & 3 \{1 \ 2; 3 \ 4; 5 \ 6\} \\ & 2 \{1; 2; 3; 4; 5; 6\} \\ & 1 \{\}. \end{aligned}$$

На основе ствола конфигуратора  $\tilde{G}(\text{Показатель})$  с помощью автоматизмов среды может быть построен единственный структурно-завершенный орграф  $(G^-)^+ = G^+(\text{Показатель})$ , который представлен на рис. 4.2. Как и раньше базовая непрерывная вершина, отвечающая домену 1, изображена красным цветом (самая нижняя вершина), зеленым цветом (светлые вершины) изображены листья, коричневым цветом изображены вершины ствола орграфа. На рис. 4.3 показано энтропийное многообразие, связанное с орграфом доменов  $G^+(\text{Показатель})$ .



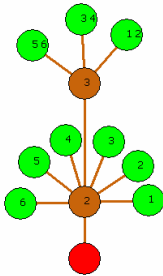


Рис. 4.2 – Оргграф  $G^+$ (Показатель)

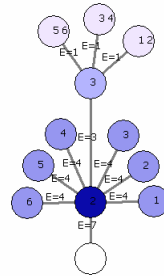


Рис. 4.3 – Структурная энтропия = 30 (дискретной части);  $\mathcal{E}_{\text{опт}} = 6$

В некоторых исследованиях может применяться упрощенный конфигуриатор, структура которого показана ниже:

$$\text{Показатель2} \{ \tag{4.3}$$

- 5 #2 {Выше нормы ^3; Норма и ниже ^1 2}
- 4 #2 {Ниже нормы ^1 Норма и выше ^2 3}
- 3 { Норма ^2; не Норма ^1 3}
- 2 { Ниже нормы ^1 [q1; q2];  
Норма ^2 (q2; q3);  
Выше нормы ^3 (q3; q4)}
- 1 {[q1; q4]};

$$G(\text{Показатель2}) = \{1 \rightarrow 2; 2 \rightarrow 3; 2 \rightarrow 4; 2 \rightarrow 5\}.$$

На рис. 4.4 показан оргграф  $G(\text{Показатель2})$ , который совпадает со своим структурно-завершенным оргграфом. На рис. 4.5 показано энтропийное многообразие, связанное с оргграфом доменов  $G^+(\text{Показатель2})$ .

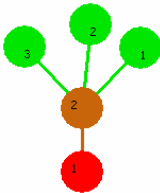


Рис. 4.4 – Оргграф  $G^+$ (Показатель2)

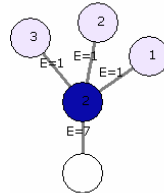


Рис. 4.5 – Структурная энтропия = 3 (дискретной части);  $\mathcal{E}_{\text{опт}} = 0$

Если конфигуриатор строится на основе разбиения абсолютной шкалы показателя (на 6 диапазонов), то схема (4.2) должна быть изменена следующим образом:

$$\begin{aligned}
 & \text{Показатель } \{ \\
 & 6 \#3 \{ \text{Хороший } \wedge 5; \text{ не Хороший } \wedge 1 \ 3 \} \\
 & 5 \#3 \{ \text{Удовлетворительный } \wedge 3; \text{ не Удовлетворительный } \wedge 1 \ 5 \} \\
 & 4 \{ \text{ Плохой } \wedge 1; \text{ не Плохой } \wedge 3 \ 5 \} \\
 & 3 \{ \text{ Плохой } \wedge 1 \ 2; \\
 & \quad \text{Удовлетворит. } \wedge 3 \ 4; \\
 & \quad \text{Хороший } \wedge 5 \ 6 \} \\
 & 2 \{ \text{ Очень плохой } \wedge 1 \ M [p1; p2] \ Ж [p1; p2']; \\
 & \quad \text{Плохой } \wedge 2 \ M (p2; p3) \ Ж (p2'; p3'); \\
 & \quad \text{Ниже среднего } \wedge 3 \ M (p3; p4) \ Ж (p3'; p4'); \\
 & \quad \text{Удовлетворит. } \wedge 4 \ M (p4; p5) \ Ж (p4'; p5'); \\
 & \quad \text{Хороший } \wedge 5 \ M (p5; p6) \ Ж (p5'; p6'); \\
 & \quad \text{Отличный } \wedge 6 \ M (p6; p7) \ Ж (p6'; p7) \} \\
 & 1 \{ [p1; p7] \} \},
 \end{aligned}
 \tag{4.4}$$

где М/Ж означают пол. В отличие от схемы (4.2) в схеме (4.4) параметры  $\{p1, \dots, p7, p2', \dots, p6'\}$  должны задаваться для каждого показателя. Как видим, для вычисления значений теста «Показатель» по схеме (4.4) необходима предварительная фиксация значения теста «Пол».

В таблице 4.1 в качестве примера показана общая схема представления базовых онтологических соглашений конфигураторов тестов в модели физического здоровья и работоспособности подростков-мальчиков (таблица составлена по материалам [120]).

Таблица 4.1 – Фрагмент базовых онтологических соглашений конфигураторов тестов в модели физического здоровья и работоспособности (Пол?М; Возраст? 13-14)

Показатели здоровья	Оценка / балл					
	Оч.плохо 1	Плохо 2	Ниже ср. 3	Удовл. 4	Хорошо 5	Отлично 6
Пульс (ЧСС) в покое, уд/мин	> 96	88 – 96	82 – 87	79 – 81	72 – 78	< 72
Жизненная емкость легких, мл/кг	< 45	45 – 49	50 – 54	55 – 59	60 - 64	> 64
Задержка дыхания, сек	< 30	30 – 35	36 – 44	45 – 54	55 – 69	> 69
Мышцы брюшн. пресса, раз	< 11	11 – 13	14 – 16	17 – 19	20 – 25	> 25
...	...	...	...	...	...	...
ИФЗ, балл	< 1,3	1,3 – 2,0	2,1 - 3,0	3,1 - 4,0	4,1 - 5,0	5,1 - 6,0

Такая же таблица должна быть построена и для подростков-девочек.

Совокупность подобных таблиц для всех возрастных групп позволяет построить конфигурации показателей для любого возраста и пола (схема 4.4 дополняется тестом «Возраст/В2»).

Последняя строка таблицы 4.1 отображает интегральный индекс физического здоровья в баллах универсальной рейтинговой шкалы ИФЗ, который одновременно является индикатором физической работоспособности (трудоспособности) человека и показателем его биологического возраста. Орграф  $G(\text{ИФЗ})$  также может определяться схемой (4.2) с соответствующей коррекцией значений параметров:  $p_1 = 0$ ;  $p_2 = 1,2$ ;  $p_3 = 2,0$ ;  $p_4 = 3,0$ ;  $p_5 = 4,0$ ;  $p_6 = 5,0$ ;  $p_7 = 6,0$ .

Если конфигурации строятся по упрощенной схеме (4.3), то аналог таблицы 4.1 будет содержать всего две колонки: «Показатели здоровья» и «Норма».

Конфигурации показателей-тестов модели физического здоровья и работоспособности человека представляют собой Банк тестов  $\{G(\tau)\}$ .

Рассмотрим следующий сценарий применения конфигураций. Пусть дана база прецедентов  $\Omega = \{\alpha(\{z/T\}, z/Z)\}$ , где  $Z = \{1, \dots, N\}$ . Прецеденты содержат также тесты «Возраст» и «Пол». Заключение  $z$  можно рассматривать в качестве кода ИФЗ, при этом  $N = 6$ . Конфигурации тестов  $\{\tau\}$  определяются одной из схем (4.2) – (4.4). Далее с помощью метода предельных обобщений (глава 5) находятся предельная синдромная и вероятностная модели знаний. Синдромная модель знаний содержит все конкурирующие наборы параметров порядка для определения ИФЗ. Параметры порядка отражают целесообразный уровень детализации значений тестов в рамках соответствующих конфигураций.

Формальные синдромы, относящиеся к зоне «плохо» и «очень плохо», рассматриваются в качестве «слабых звеньев» организма и, при некоторых условиях, как факторы риска развития соответствующей функциональной недостаточности.

Минимизация числа показателей, снижение требований к их точности, а в конечном итоге – значительное упрощение расчетов ИФЗ, позволит существенно расширить сферу применения модели физического здоровья и работоспособности. Для практики это может оказаться решающим.

#### 4.2.2 Конфигурации с вычислениями на основе соглашений

Рассмотрим следующую модельную задачу. Требуется установить наличие или отсутствие артериальной гипертензии на основе параметров, выявляемых при суточном мониторингировании артериального давления (АД). Задачу необходимо решить с помощью операции замыкания  $\{z/T\}^+$  над банком тестов  $\{G(\tau)\}$  (операция замыкания была рассмотрена в главе 2).

В качестве исходных данных будем использовать четыре параметра:

- Уровень среднего систолического АД днем;
- Уровень среднего систолического АД ночью;
- Уровень среднего диастолического АД днем;
- Уровень среднего диастолического АД ночью.

Указанные параметры вычисляются на основе двух суточных графиков – систолического и диастолического АД (не представляет сложности параметры АДс и АДд также ввести в состав банка тестов).

Ниже приведены конфигураторы банка тестов [159].

*Уровень среднего систолического АД днем* ^САДд {

D1 {САДд в норме ^N;

САДд не в норме ^a b c d}

D2 {САДд снижен ^a [50; 89];

САДд в норме ^N [90; 134];

САДд повышен I степени ^b [135; 154];

САДд повышен II степени ^c [155; 174];

САДд повышен III степени ^d [175; 250]}

D3 { [50; 250] }

$G(\text{САДд}) = \{D3 \rightarrow D2 \rightarrow D1\}$ ;  $G^-(\text{САДд}) = \{D3 \rightarrow D2\}$ .

Пример структурно-завершенного орграфа  $G^+(\text{САДд})$  показан на рис. 4.6.

Энтропийные характеристики орграфа  $G^+(\text{САДд})$  показаны на рис. 4.7.

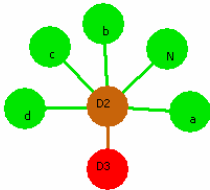


Рис. 4.6 – Орграф  $G^+(\text{САДд})$

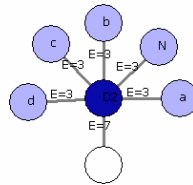


Рис. 4.7 – Структурная энтропия = 15 (дискретной части);  $\Theta_{\text{опт}} = 5$

*Уровень среднего систолического АД ночью* ^САДн {

D1 {САДн в норме ^N;

САДн не в норме ^a b c d}

D2 {САДн снижен ^a [40; 79];

САДн в норме ^N [80; 119];

САДн повышен I степени ^b [120; 139];

САДн повышен II степени ^c [140; 159];

САДн повышен III степени ^d [160; 250]}

D3 { [40; 250] }

$$G(\text{САДн}) = \{D3 \rightarrow D2 \rightarrow D1\}; G^-(\text{САДн}) = \{D3 \rightarrow D2\}.$$

*Уровень среднего систолического АД ^САД {*

- D1 {САД в норме ^N;  
САД не в норме ^a b c d}
- D2 {САД снижен ^a;  
САД в норме ^N;  
САД повышен ^b c d }
- D3 {САД снижен ^a (САДд/D2? а или САДн/D2? а);  
САД в норме ^N (САДд/D2? N и САДн/D2? N);  
САД повышен I степени ^b (max (САДд/D2?; САДн/D2?) = b);  
САД повышен II степени ^c (max (САДд/D2?; САДн/D2?) = c);  
САД повышен III степени ^d (max (САДд/D2?; САДн/D2?) = d)} }

$$G(\text{САД}) = \{D3 \rightarrow D2 \rightarrow D1\}; G^-(\text{САД}) = \{D3 \rightarrow D2\}.$$

*Уровень среднего диастолического АД днем ^ДАДд {*

- D1 {ДАДд в норме ^N;  
ДАДд не в норме ^a b c d}
- D2 {ДАДд снижен ^a [35; 59];  
ДАДд в норме ^N [60; 84];  
ДАДд повышен I степени ^b [85; 94];  
ДАДд повышен II степени ^c [95; 104];  
ДАДд повышен III степени ^d [105; 150] }
- D3 {[35; 150]} }

$$G(\text{ДАДд}) = \{D3 \rightarrow D2 \rightarrow D1\}; G^-(\text{ДАДд}) = \{D3 \rightarrow D2\}.$$

*Уровень среднего диастолического АД ночью ^ДАДн {*

- D1 {ДАДн в норме ^N;  
ДАДн не в норме ^a b c d}
- D2 {ДАДн снижен ^a [25; 44];  
ДАДн в норме ^N [45; 69];  
ДАДн повышен I степени ^b [70; 89];  
ДАДн повышен II степени ^c [90; 99];  
ДАДн повышен III степени ^d [100; 150]}
- D3 {[25; 150]} }

$$G(\text{ДАДн}) = \{D3 \rightarrow D2 \rightarrow D1\}; G^-(\text{ДАДн}) = \{D3 \rightarrow D2\}.$$

*Уровень среднего диастолического АД ^ДАД {*

- D1 {ДАД в норме ^N;  
ДАД не в норме ^a b c d}
- D2 {ДАД снижен ^a;  
ДАД в норме ^N;

ДАД повышен  $\wedge b \text{ c } d$ ;  
 D3 {ДАД снижен  $\wedge a$  (ДАД<sub>д</sub>/D2? а или ДАД<sub>н</sub>/D2? а);  
 ДАД в норме  $\wedge N$  (ДАД<sub>д</sub>/D2? N и ДАД<sub>н</sub>/D2? N);  
 ДАД повышен I степени  $\wedge b$  ( $\max$  (ДАД<sub>д</sub>/D2?; ДАД<sub>н</sub>/D2?)= b);  
 ДАД повышен II степени  $\wedge c$  ( $\max$  (ДАД<sub>д</sub>/D2?; ДАД<sub>н</sub>/D2?)= c);  
 ДАД повышен III степени  $\wedge d$  ( $\max$  (ДАД<sub>д</sub>/D2?; ДАД<sub>н</sub>/D2?)= d)}  
 $G(\text{ДАД}) = \{D3 \rightarrow D2 \rightarrow D1\}$ ;  $G^-(\text{ДАД}) = \{D3 \rightarrow D2\}$ .

Артериальная гипертензия  $\wedge AG$  {

D1 {AG нет  $\wedge N$ ;  
 AG есть  $\wedge b \text{ c } d$ }  
 D2 {AG нет  $\wedge N$  ( $\max$  (САД/D3?; ДАД/D3?) = N);  
 AG I степени  $\wedge b$  ( $\max$  (САД/D3?; ДАД/D3?) = b);  
 AG II степени  $\wedge c$  ( $\max$  (САД/D3?; ДАД/D3?) = c);  
 AG III степени  $\wedge d$  ( $\max$  (САД/D3?; ДАД/D3?) = d)}

$G(AG) = \{D2 \rightarrow D1\}$ ;  $G^-(AG) = \{D2\}$ .

Пример структурно-завершенного орграфа  $G^+(AG)$  показан на рис. 4.8. Энтропийные характеристики орграфа  $G^+(AG)$  показаны на рис. 4.9.

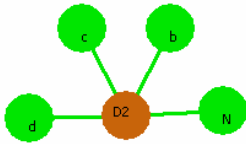


Рис. 4.8 – Орграф  $G^+(AG)$

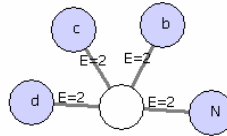


Рис. 4.9 – Структурная энтропия = 8;  
 $\mathcal{E}_{\text{орг}} = 4$

При построении конфигураторов использовалось следующее очевидное соглашение: диапазоны разбиения области возможных значений параметра задают естественный порядок на множестве термов и алиасов первого дискретного домена. Указанный порядок позволяет применять функцию  $\max$  (и другие подобные функции) к произвольным термам.

В результате замыкания исходных данных над банком тестов  $\{G(\tau)\} = \{G(\text{САД}_д); G(\text{САД}_н); G(\text{ДАД}_д); G(\text{ДАД}_н); G(\text{САД}); G(\text{ДАД}); G(AG)\}$  получим:

$\{\text{САД}_д?; \text{САД}_н?; \text{ДАД}_д?; \text{ДАД}_н?\}^+ = \{\text{САД}_д?; \text{САД}_н?; \text{ДАД}_д?; \text{ДАД}_н?; \text{САД}?; \text{ДАД}?; \text{AG}?\}$ .

Как видим тест «AG» содержится в замыкании, что и требовалось условиями задачи.

#### 4.2.3 Технологический аудит и субоптимизация

Технологический аудит и субоптимизация на основе Принципа предельных обобщений позволяют решать, в частности, следующие практические задачи [52, 75, 136]:

- получение новых системных знаний об изучаемом объекте;
- определение конкурирующих наборов управляемых входных параметров, которые достаточны для управления процессом с заданным качеством;
- анализ управляемости процесса (оценка возможностей существующей системы управления обеспечить в каждой реализации процесса заданное качество его функционирования);
- прогноз значений выходных параметров по известным значениям входных параметров, в частности, оперативная оценка значений выходных параметров, лабораторный анализ которых требует значительных затрат времени;
- синтез новых химических соединений и новых композиционных материалов (построение математических моделей зависимости между строением определенного класса химических соединений и комплексом их потребительских свойств);
- определение требуемой точности измерения входных параметров.

Рекомендации, полученные в результате аудита и субоптимизации, могут включать [52, 136]:

- оптимизацию действующих технологических процессов по технологическим, экономическим и экологическим критериям или комплексам этих критериев;
- определение научно и технологически обоснованных требований к качеству сырья, при заданных требованиях к качеству функционирования конкретного технологического процесса;
- технологию обеспечения заданного качества функционирования процесса при существенных вариациях состава сырья;
- технологию обеспечения необходимой технологической гибкости, с целью получения продукта, удовлетворяющего индивидуальным требованиям различных заказчиков.

Действующие технологические процессы на предприятиях, как объекты управления многомерны, нелинейны, обладают многими внутренними обратными связями, сложными нелинейными взаимосвязями между входными и выходными параметрами. Для них характерен высокий уровень шумов (ошибки измерений, недостаточная представительность отбираемых проб, низкая точность аналитического контроля) и не наблюдаемость некоторых существенных входных параметров (отсутствие

контроля ряда показателей состава и свойств сырья, методик анализа и средств измерений). Изучение этих процессов математическими методами осложняется трудностью использования на реальных объектах результатов исследований, полученных на лабораторных и полупромышленных установках [183].

Тем не менее, оптимизация сложных технологических процессов невозможна без наличия математических моделей, обеспечивающих с приемлемой точностью решение задачи синтеза оптимального или субоптимального управления и реализации этого управления на существующих или модернизируемых для этих целей системах управления.

Модель технологического процесса строится на основе эмпирических данных. Каждая реализация процесса рассматривается как отдельный прецедент. Для построения модели все реализации технологического процесса разбивают на  $N$  классов. Как и раньше  $Z = \{1, \dots, N\}$  – множество классов-заклучений по комплексному показателю эффективности.

В простейшем случае  $Z$  может содержать только 2 класса: “хорошие” опыты, в которых обеспечиваются заданные ограничения по себестоимости, потребительским, энергетическими и другим параметрам, и “плохие” опыты, в которых процесс или выходной продукт хотя бы по одному параметру не соответствовал заданным ограничениям. Конечная цель построения математической модели заключается в нахождении всех конкурирующих наборов параметров порядка, которые обеспечивали бы получение продукта, попадающего в класс “хороших” опытов.

Если  $Z$  содержит более двух классов (заклучений), то автоматизмами среды порождаются еще  $N$  множеств заклучений:  $Z_1 = \{1, -1\}$ ,  $Z_2 = \{2, -2\}$ , ...,  $Z_N = \{N, -N\}$ . В совокупности множества  $Z, Z_1, Z_2, \dots, Z_N$  образуют оргграф-конфигуратор заклучения  $z$ :

$$G(z) = \{Z \rightarrow Z_1; Z \rightarrow Z_2; \dots; Z \rightarrow Z_N\}.$$

Более сложные схемы конфигураторов заклучений могут быть такими:

$$z \{Z' \{1\ 2; 3\ 4; 5\ 6\} Z \{1; 2; 3; 4; 5; 6\}\},$$

где обобщение идет справа налево. Конфигуратор  $G(z) = \{Z \rightarrow Z'\}$ . Пример:

*Эффективность*  $\wedge z \{$

$Z' \{$  Низкая  $\wedge^1 2$ ; Средняя  $\wedge^3 4$ ; Высокая  $\wedge^5 6\}$

$Z \{$  Очень низкая  $\wedge^1$ ; Низкая  $\wedge^2$ ; Ниже среднего  $\wedge^3$ ; Средняя  $\wedge^4$ ; Высокая  $\wedge^5$ ;

Очень высокая  $\wedge^6\}\}$ .

Единственный структурно-завершенный оргграф  $G^+(\text{Эффективность})$  показан на рис. 4.10. Всего он содержит 11 вершин. Энтропийные характеристики оргграфа  $G^+(\text{Эффективность})$  показаны на рис. 4.11.



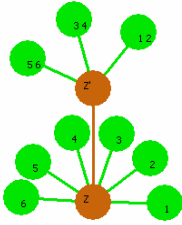


Рис. 4.10 – Орграф  $G^+$  (Эффективность)

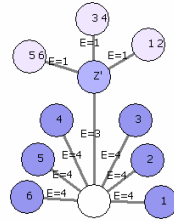


Рис. 4.11 – Структурная энтропия = 30;  
 $\mathcal{E}_{\text{онт}} = 6$

Если исходный показатель эффективности  $z$  непрерывный, то конфигурактор может иметь следующую структуру:

Эффективность  $\wedge z$  {

3 {Низкая  $\wedge 1$  2; Средняя  $\wedge 3$  4; Высокая  $\wedge 5$  6}

2 {Очень низкая  $\wedge 1$  [0; p1]; Низкая  $\wedge 2$  (p1; p2); Ниже среднего  $\wedge 3$  (p2; p3);  
Средняя  $\wedge 4$  (p3; p4); Высокая  $\wedge 5$  (p4; p5); Очень высокая  $\wedge 6$  (p5; 1)}

1 {[0; 1]}

Схемы конфигуракторов тестов  $\{z\}$  могут разные. Приведем несколько простых схем.

Схема 1. Обозначим регламентный (или рекомендуемый) интервал значений  $X_i$  через [p1, p2].

$X_i$  {2 {Ниже нормы  $\wedge 1$  [Xmin; p1]; Норма  $\wedge 2$  [p1, p2]; Выше нормы  $\wedge 3$  (p2; Xmax]} 1 {[Xmin; Xmax]}

Схема 2. Разобьем регламентный интервал значений  $X_i$  на две части, а именно: [p1, p2] = [p1, q]  $\cup$  (q, p2]. Конфигуратор представим в виде:

$X_i$  {3 {Ниже нормы  $\wedge a$ ; Норма  $\wedge b$  c; Выше нормы  $\wedge d$ } 2 {Ниже нормы  $\wedge a$  [Xmin; p1]; Норма\_1  $\wedge b$  [p1, q]; Норма\_2  $\wedge c$  (q, p2]; Выше нормы  $\wedge d$  (p2; Xmax]} 1 {[Xmin; Xmax]}

Схема 3. Примем интервал [Xmin; Xmax] за 100% и разобьем его на 5 частей. Получим следующий конфигурактор:

$X_i$  {6 #3 {Small  $\wedge 1$ ; Not Small  $\wedge 3$  4} 5 #3 {Medium  $\wedge 3$ ; Not Medium  $\wedge 1$  4} 4 {Large  $\wedge 4$ ; Not Large  $\wedge 1$  3} 3 {Small  $\wedge 1$  2; Medium  $\wedge 3$ ; Large  $\wedge 4$  5} 2 {Very small  $\wedge 1$  [0; 15]; Small  $\wedge 2$  (15; 40); Medium  $\wedge 3$  (40; 60); Large  $\wedge 4$  (60; 85); Vary large  $\wedge 5$  (85; 100]} 1 {[0; 100]}

Схема 4. Объединим схемы 2 и 3 в один конфигурактор. Результирующий предельный структурно-завершенный орграф  $G^{++}$

представлен на рис. 4.12.

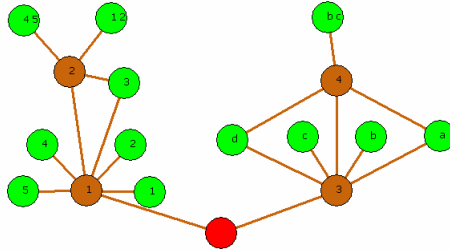


Рис. 4.12 – Предельный структурно-завершенный орграф Схемы 4

Схема 4 показывает, как разные схемы могут быть объединены в один конфигурактор. В разделе 4.3 будет рассмотрена принципиально иная схема построения конфигурактора, названная «Штрихкодом». Она также может быть интегрирована в единый конфигурактор. Построение структурно-завершенных орграфов значительно расширяет возможности обобщения, т.е. построения синдромной модели знаний, которая собственно и является математической моделью технологического процесса.

Математическая модель технологического процесса в виде предельной синдромной и вероятностной модели знаний строится в рамках контекста  $\langle \Omega = \{\alpha\{\underline{z}/T\}, \underline{z}/Z\}, \{G(\tau)\}, G(z) \rangle$ , где  $Z = \{1, \dots, N\}$ ;  $\{\tau\}$  – тесты, с помощью которых описываются прецеденты (реализации процесса).

Синдромы моделей знаний имеют вид

$$S = (\{\underline{z}/T\} \rightarrow \underline{z}/Z), \nu = \nu(S),$$

где  $\{\underline{z}/T\}$  – значения тестов;  $\underline{z}/Z$  – заключение;  $\nu$  – вес синдрома;  $\nu(S)$  – функция вычисления веса. Вес синдрома  $S$  пропорционален количеству опытов (прецедентов) с заключением  $\underline{z}/Z$ , содержащих синдром  $S$ . Синдромы описывают избыточные зависимости комплексного выходного параметра  $z$  от взаимного влияния входных параметров. Синдромы могут быть предельными по уровню общности (см. главу 5).

На основе полного перечня синдромов для каждого заключения могут быть сформированы минимальные (предельные) модели знаний, содержащие синдромы с минимальным весом  $\{S\}$ .

Помимо синдромов речь также может идти о вероятностных закономерностях  $\{R\}$  и сопряженных предельных вероятностных закономерностях, которые являются ранними предвестниками событий, например, неблагоприятного течения процесса [136].

Контекст задачи  $\langle \Omega(Z), \{G(\tau)\}, G(z) \rangle$  должен быть таким, чтобы некоторые из предельных синдромов можно было отнести к параметрам

порядка благоприятного и неблагоприятного течения изучаемого процесса. В таком случае предельная синдромная модель знаний будет содержать конкурентные наборы параметров порядка, а управление технологическим процессом и технологический аудит заключаются в синдромной диагностике и задании целевых синдромов (параметров порядка), отвечающих благоприятному режиму (выходу на этот режим и его стабилизации). При этом желательно нейтрализовать предвестники развития неблагоприятного режима, которые содержатся в предельной вероятностной модели знаний. Такой подход реализует синдромный принцип управления сложными процессами любой природы (см. главу 7).

Минимальные (предельные) синдромные модели знаний позволяют определить перечень управляемых параметров, которые минимально достаточны (в рамках выбранной модели знаний) для реализации целевой задачи. На основе минимальных моделей знаний проводят оценку возможностей существующей системы управления обеспечить в каждой реализации технологического процесса заданное качество его работы; определяют нормы технологического режима по всем входным параметрам, обеспечивающим снижение себестоимости выпускаемой продукции при условии соблюдения требований по качеству, выбросам вредных веществ в окружающую среду и другим ограничениям.

Модели знаний  $\{S\}$  позволяют учесть изменения качества сырья и других неуправляемых входных параметров. Таким образом, в зависимости от различных комбинаций в составе сырья находят соответствующую модель знаний и, следовательно, необходимое управление технологическим процессом. В результате может быть обеспечено получение высококачественного продукта при использовании некондиционного сырья.

Использование предложенных методов позволяет существенно повысить показатели выхода и качества продукции, увеличить производительность, сократить расходные нормы по сырью и энергоресурсам в действующих производствах химической, биотехнологической, металлургической и других областях промышленности. Например, технологический аудит на металлургическом заводе может быть задействован для оценки возможностей [52]: повышения выхода годной стали на сталеплавильном переделе; снижения потерь металла в стружку для отделочных линий прокатного цеха; снижения потерь металла в стружку на участке заготовки; увеличения выхода чугуна по загруженному железу и т.д.

Таким образом, для проведения технологического аудита необходима следующая информация:

- возможные и регламентные значения выхода продукта и расходных

норм по сырьевым, энергетическим и экологическим показателям;

- таблица экспериментального материала, каждая строка которой содержит значения загрузок сырья, полупродуктов и выходных показателей в одной реализации изучаемого технологического процесса;

- фактическая калькуляция себестоимости за период наблюдения;

- производительность за период наблюдения (например, годовая).

Для построения конфигураторов тестов и моделей знаний технологического процесса исходная информация представляется в виде трех таблиц. Первая таблица содержит перечень входных параметров, информацию об их минимальных и максимальных значениях, регламентных значениях и единицах измерения. Пример представления перечня входных параметров приведен в таблице 4.2 [56, 75].

Таблица 4.2 – Фрагмент перечня входных параметров при технологическом аудите доменной печи (оценка резерва по удельному расходу кокса)

№пп	Наименование входного параметра $X_i$ , ед.изм.	$X_{min} - X_{max}$	Рекомендуемый режим
1	ЖРЧШ, т/сут	3057,0 – 4457,0	4103,0 – 4457,0
2	Нагрузка по коксу, т/сут	935,0 – 1357,0	1255,5 – 1357,0
3	Влажность кокса, %	6,15 – 10,9	8,475 – 10,900
4	Расход воздуха, м <sup>3</sup> /мин	2112,0 – 2510,0	2414,5 – 2510,0
5	Давление дутья, атм	2,17 – 2,66	2,425 – 2,505
6	Температура дутья, °С	992,0 – 1130,0	1092,5 – 1130,0
7	Общая влажность дутья, г/м <sup>3</sup>	4,5 – 17,1	10,7 – 17,1

Первая таблица позволяет в автоматическом режиме сформировать конфигураторы входных параметров.

Вторая таблица содержит информацию о перечне выходных показателей, единицах их измерения и допустимых значениях. Данные таблицы служат для определения интегрального показателя качества процесса (или эффективности)  $z$ . Пример перечня выходных показателей приведен в таблице 4.3.

Таблица 4.3 – Пример перечня выходных показателей

$Y_i$	Выходная переменная $Y_i$ , ед. изм.	Допуст. значение $Y_i$
1	Выход продукта, %	$Y_1 > 93.0$
2	Массовая доля основного вещества, %	$Y_2 > 90.5$
3	Содержание примеси А, %	$Y_3 < 0.15$
...	...	...
k	Температура плавления продукта, °С	$Y_k > 145$
Комплексный критерий $z$		$G(z)$

Если выходной показатель один, то он совпадает с критерием  $z$ .

Третья таблица содержит информацию о значениях входных и выходных переменных в  $m$  реализациях изучаемого процесса, т.е. базу прецедентов  $\Omega(Z)$  (пример представления – таблица 4.4).

Таблица 4.4 – Реализации изучаемого процесса

NN	Значения входных переменных					Значения выходных переменных				
	X1	X2	X3	Xi	Xn	Y1	Y2	Yj	Yk	Z
1	-	-	-	-	-	-	-	-	-	-
...	...	...	...	...	...	...	...	...	...	...
m	-	-	-	-	-	-	-	-	-	-

Результатом задания конфигуракторов входных и выходных параметров является построение расширенной таблицы исходных данных (таблица 4.5). Выделенная строка содержит имена всех доменов – вершин орграфа каждого параметра.

Таблица 4.5 – Расширенная таблица исходных данных

NN	Значения входных переменных						Значения комплексного выходного показателя				
	X1		...	Xn			z				
	1	2	...	1	2	3	Z	Z1	Z2	...	Z N
1	-	-	...	-	-	-	-	-	-	...	-
...	...	...	...	...	...	...	...	...	...	...	...
m	-	-	...	-	-	-	-	-	-	...	-

Пример расширенной таблицы исходных данных приведен в разделе 3.4. На основе расширенной таблицы исходных данных строятся синдромные и вероятностные модели знаний. Формальные процедуры построения моделей знаний рассматриваются в главе 5.

#### 4.2.4 Семантический дифференциал

Опишем кратко суть метода семантического дифференциала. Семантический дифференциал (СД) – один из проективных методов социологии, опирающийся на достижения психосемантики, был разработан группой американских психологов во главе с Ч. Осгудом в 1952 году. Он применяется в исследованиях, связанных с восприятием и поведением человека, с анализом социальных установок и личностных смыслов. Метод СД является комбинацией метода контролируемых ассоциаций и процедур шкалирования. СД позволяет выявить бессознательные ассоциативные связи между объектами в сознании людей [127].

Метод СД позволяет найти систему латентных факторов, в рамках

которых индивид оценивает объекты. По существу, семантическое пространство – исследовательская модель структуры индивидуального сознания, и задача состоит в определении того, где в этом пространстве находится изучаемый объект.

Во многих исследованиях тестируемые объекты оцениваются по ряду *бимодальных*  $(2N+1)$ -балльных шкал, полюса которых обычно задаются вербально при помощи антонимов: хороший – плохой, теплый – холодный, активный – пассивный и т.п. Предполагается, что человек способен оценить изучаемый объект, соотнося интенсивность внутреннего переживания по поводу объекта с заданной оценочной шкалой. Деления шкалы фиксируют различные степени интенсивности данного качества объекта. Шкалы, коррелирующие между собой, группируются в независимые факторы, образующие семантическое пространство.

Пример бимодальных шкал для оценки сущностей «мужчина» и «женщина»:

X1. светлый – темный	X8. прямой – кривой
X2. низкий – высокий	X9. энергичный – скованный
X3. явный – скрытый	X10. большой – маленький
X4. реальный – идеальный	X11. отважный – трусливый
X5. безопасный – опасный	X12. быстрый – медленный
X6. мягкий – жесткий	X13. активный – пассивный
X7. острый – тупой	X14. резкий – монотонный

Наиболее часто применяются три разбиения бимодальных шкал: на 21 балл –  $[-10; 10]$ , на 7 баллов –  $[-3; 3]$  и на 3 балла –  $[-1; 1]$ . Каждая шкала является, по сути, элементарным тестом, следовательно, приведенные выше разбиения естественным образом задают схемы обобщения и, соответственно, конфигураторы тестов. Приведем пример возможной схемы конфигуратора ( $X_i$  – код или название  $i$ -ой бимодальной шкалы):

$$X_i \{$$

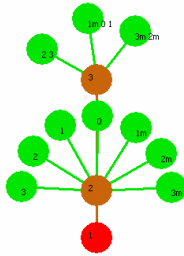
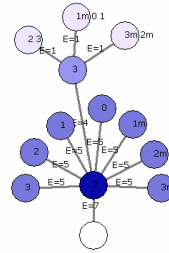
$$3 \{-1 \wedge 3m \ 2m; 0 \wedge 1m \ 0 \ 1; 1 \wedge 2 \ 3\}$$

$$2 \{-3 \wedge 3m \ [-10; -7]; -2 \wedge 2m \ [-6; -4]; -1 \wedge 1m \ [-3; -2]; 0 \wedge 0 \ [-1; 1]; 1 \wedge 1 \ [2; 3]; 2 \wedge 2 \ [4; 6]; 3 \wedge 3 \ [7; 10]\}$$

$$1 \{[-10; 10]\}$$

$$G(X_i) = \{1 \rightarrow 2 \rightarrow 3\}.$$

На рис.4.13 показан единственный структурно-завершенный оргграф  $G^+(X_i)$ , а на рис. 4.14 показаны его энтропийные характеристики. Принято, что домен «1» оргграфа  $G(X_i)$  является конструктивно-непрерывным (для него не строятся листья).

Рис. 4.13 – Орграф  $G^+(X_i)$ Рис. 4.14 – Структурная энтропия = 42  
(дискретной части);  $\mathcal{E}_{\text{опт}} = 8$ 

Наряду с вербальными, разработаны также невербальные семантические дифференциалы, где в качестве полюсов шкал используются графические оппозиции, живописные картины, фотопортреты. Список кодов всех образов формирует базовый домен конфигуратора. Обобщение образов по тем или иным аспектам приводит к появлению новых доменов.

В исследованиях также используются *монополярные шкалы*, с помощью которых объекты оцениваются по выраженности одного свойства: насколько объект хороший, насколько теплый и т.п. В случае бимодальных шкал респондент оценивает, где находится для него объект «А» по шкале «здоровый – больной», а при унимодальных шкалах оценивается насколько свойство «здоровье» присуще объекту «А». Применение унимодальных шкал вызвано тем, что зачастую антонимичные прилагательные на самом деле не являются полными противоположностями.

Результаты измерений по унимодальным шкалам удобно представить на лепестковой диаграмме. Пусть  $X_i$  ( $i = 1, \dots, 16$ ) – замеры в первичных унимодальных шкалах, а  $z_i$  – те же замеры, но в нормированной шкале (от 0 до 1). Данные двух замеров приведены в таблице 4.6, а их отображение на лепестковой диаграмме представлено на рис. 4.15.

Таблица 4.6 – Данные замеров

$z_i$	I-й замер	II-й замер	$z_i$	I-й замер	II-й замер
1	0,25	0,5	9	0,5	0,5
2	0,5	0,5	10	0,5	0,5
3	0,5	0,5	11	0,5	0,5
4	0,5	0,5	12	0,5	0,5
5	0,25	0,5	13	0,5	0,5
6	0,25	0,5	14	0,25	0,5
7	0,25	0,5	15	0,25	0,5
8	0,25	0,5	16	0,25	0,5



Рис. 4.15 – Отображение результатов 2-х замеров на базовом уровне общности

Конфигуратор для нормированного показателя  $z_i$  можно задать, например, следующей схемой:

$z_i$  {  
 3 {Низкий  $\wedge$ 1; Средний  $\wedge$ 2 3 4; Высокий  $\wedge$ 5}  
 2 {Низкий  $\wedge$ 1 [0; 0,2]; Ниже среднего  $\wedge$ 2 (0,2; 0,4]; Средний  $\wedge$ 3 (0,4; 0,6];  
 Выше среднего  $\wedge$ 4 (0,6; 0,8]; Высокий  $\wedge$ 5 (0,8; 1]}  
 1 {[0; 1]}

$G(z_i) = \{1 \rightarrow 2 \rightarrow 3\}$ .

На рис.4.16 показаны некоторые структурно-завершенные орграфы  $G^+(z_i)$ .

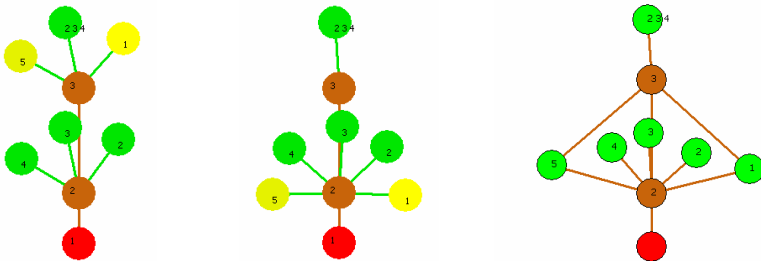


Рис. 4.16 – Структурно-завершенные орграфы  $G^+(z_i)$

Не всегда целесообразно в подобных исследованиях переходить к непрерывным шкалам. Приведем пример конфигуратора для показателя «Адаптивность»:



Адаптивность  $\wedge A$  {

A3 {не Норма  $\wedge 1$  2; Норма  $\wedge 5$ }

A2 {Дезадаптация  $\wedge 1$ ; Дискомфорт  $\wedge 2$  3 4; Норма  $\wedge 5$ }

A1 {Дезадаптация  $\wedge 1$ ; Внутриличностный конфликт  $\wedge 2$ ; Стресс  $\wedge 3$ ;  
Дискомфорт  $\wedge 4$ ; Норма  $\wedge 5$ }

$G(A) = \{A1 \rightarrow A2 \rightarrow A3\}$ ,  $G(A) = \{A1 \rightarrow A2\}$ .

В маркетинговых исследованиях признанным инструментом исследования имиджа корпорации, торговой марки или товара являются *денотативные шкалы*, которые не всегда состоят только из прилагательных-антонимов, но представляют собой, как правило, словосочетания, фразы, которые выражают ожидания, характеристики товара, как негативные, так и положительные. Аналогичные товары разных компаний-производителей можно оценивать по шкале «стоит своих денег», банки – по уровню надежности, выгоды и т.п.

Подобные шкалы возникают в задачах моделирования социальной, педагогической и медицинской реабилитации, как отдельных индивидуумов, так и целых групп. Примеры социально-педагогических шкал (тестов) при реабилитации детей «группы риска»: X1 – состояние регуляторных систем; X2 – психоэмоциональное состояние; X3 – уровень умственного развития; X4 – наличие положительных жизненных планов; X5 – развитие полезных знаний, навыков, интересов и т.д. Примером конфигуратора для подобных тестов может служить конфигуратор  $G(A)$ . На рис. 4.17 показана запись «Оценка степени индивидуального развития ребенка», разработанной нами электронной «Психолого-педагогической, медико-социальной карты учащегося» [5].

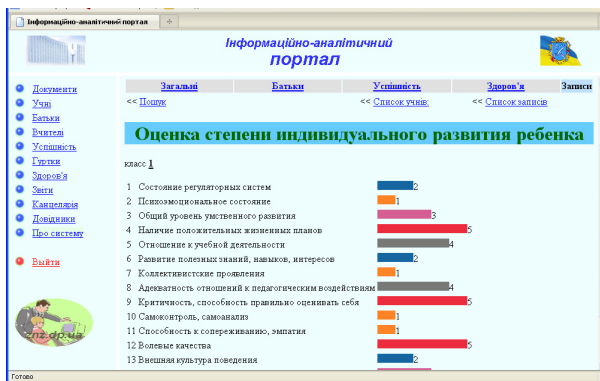


Рис. 4.17 – Пример формы «Оценка индивидуального развития ребенка»

Степень выраженности у ребенка перечисленных показателей социального развития оценивается по пятибалльной шкале, при этом «5» баллов соответствует понятию «Норма», оценка «3» является пороговым значением и соответствует нижней границе возрастной нормы.

Рассмотрим особенности формирования латентных факторов с использованием конфигураторов тестов и синдромных моделей знаний. Пусть дано  $n$  шкал  $X_1, \dots, X_n$  и, соответственно,  $n$  конфигураторов:  $G(X_1), \dots, G(X_n)$ . Список респондентов зададим множеством  $Y_1$ , а список изучаемых понятий или объектов зададим множеством  $Y_2$ . На множествах  $Y_1$  и  $Y_2$  определено множество заключений  $Z = \{1, 2, \dots, N\}$ . В простейшем случае имеется взаимнооднозначное соответствие между  $Z$  и  $Y_2$ . Будем считать, что для заключения  $z$  задан конфигурактор  $G(z)$ , в котором  $Z$  играет роль базового домена.

Исходные данные исследования в базовых доменах по всем шкалам представимы в виде таблицы 4.7. Таблица является базой прецедентов  $\Omega$ .

Таблица 4.7 – Исходное представление результатов исследований по модифицированному методу семантического дифференциала

Шкалы					Ключ		Z
X1	X2	X3	Xi	Xn	Y1	Y2	
-	-	-	-	-	-	-	-
...	...	...	...	...	...	...	...
-	-	-	-	-	-	-	-

Комбинация значений  $(y_1, y_2)$  является уникальным ключом (идентификатором) соответствующей записи.

Важно отметить, что для получения значимых результатов достаточно относительно небольших выборок – 30–50 человек, в связи с тем, что единицей анализа является не респондент, а оценки, которые он выставляет объектам. Учитывая, что каждый из 30–50 респондентов оценивает 7–10 объектов по 15–25 шкалам, то общий объем выборки оказывается вполне достаточным для построения корректных моделей знаний.

Результатом задания конфигураторов для шкал  $X_i$  и заключения  $z$  является автоматическое построение расширенной таблицы исходных данных (пример – таблица 4.5). На основе расширенной таблицы исходных данных строятся (предельные) синдромные и вероятностные модели знаний. Каждый синдром содержит связанный набор латентных факторов. Для каждого понятия выделяется директивная зона синдромов и зона вероятности. Детальный алгоритм построения синдромных и вероятностных моделей знаний рассматривается в главе 5.

Для приведенного выше примера 14 шкал и конфигуратора  $G(X_i)$  общее

пространство описаний разного уровня общности включает  $3^{14}$  описаний. Если вместо  $G(Xi)$  использовать структурно-завершенные орграфы  $G^+(Xi)$ , то пространство описаний будет включать  $13^{14}$  описаний.

В главе 3 описана процедура построения орграфа набросков. Данная процедура позволяет построить индивидуальный (для каждого индивида) орграф набросков любого понятия (на основе обобщений в рамках конфигураторов). Терминальные вершины орграфов по всем тестам задают финитные наброски. В рамках любого орграфа набросков находится экстремальный пограничный слой набросков (для фиксированной цели).

Процедура выполнения модифицированной методики семантического дифференциала в рамках конкретного исследовательского проекта будет состоять из следующих этапов:

1. Формирование и тестирование списка прилагательных, утверждений, для описания тестируемых объектов. От выбранных признаков зависит тот уровень осознанности, на котором респондент будет оценивать измеряемый объект.

2. Задание конфигулятора для каждой шкалы и заключения  $z$ .

3. Проведение исследования. Представление результатов исследования в виде таблицы 4.7. Автоматическое формирование расширенной таблицы исходных данных.

4. Построение предельных синдромных и вероятностных моделей знаний. Минимальные модели знаний позволяют, в частности, отыскать минимальные наборы координат смысла, по которым можно восстановить все объекты. Одновременно в рамках каждого синдрома устанавливается необходимая степень детальности по каждой координате смысла (шкале).

Как показывают исследования, у каждого человека обнаруживается многообразие сравнительно малой размерности, в небольшой окрестности которого лежат почти все субъективные ощущения. С помощью модифицированной СД-методики можно строить классификации людей и группировать их по критерию сходства восприятия определенных объектов, т.е. фактически составлять описание, «портрет» целевых групп.

#### 4.2.5 Когнитивные карты

Когнитивная карта – математическая модель, позволяющая формализовать описание сложной системы или проблемы и выявить структуру причинно-следственных связей между элементами системы (составляющими проблемы), а также оценить последствия в результате воздействия на эти элементы или изменения характера связей [1, 98].

Формально, когнитивная карта – это взвешенный ориентированный граф  $G = (X, A)$ , в котором  $X = \{x_1, \dots, x_n\}$  – множество вершин,

взаимнооднозначно соответствующих множеству базисных факторов,  $A$  – множество дуг, отражающих факт непосредственного влияния факторов. Значение фактора  $x_i$  указывается в числовой или лингвистической шкале (табл. 4.8, столбцы 1 и 3). Влияние  $a_{ij}$  фактора  $x_i$  на фактор  $x_j$  также указывается в числовой или лингвистической шкале (табл. 4.8, столбцы 2 и 3). Ниже значения фактора и связи будут представлены в виде конфигураторов.

Таблица 4.8 – Значения лингвистических переменных

Лингвистические значения факторов	Лингвистические значения связей	Численные значения
не изменяется	не влияет	0
очень слабо растет (падает)	очень слабо усиливает (ослабляет)	0,2 (-0,2)
слабо растет (падает)	слабо усиливает (ослабляет)	0,4 (-0,4)
умеренно растет (падает)	умеренно усиливает (ослабляет)	0,6 (-0,6)
сильно растет (падает)	сильно усиливает (ослабляет)	0,8 (-0,8)
очень сильно растет (падает)	очень сильно усиливает (ослабляет)	1 (-1)

Когнитивная карта является исходным статическим представлением (отображением) связей между факторами, существующими в исследуемой ситуации. Разрешение проблем целенаправленного развития, возникающих в слабоструктурированных ситуациях, требует построения динамической имитационной модели и получения на ее основе новых знаний о структуре и динамике исследуемой ситуации. Модель сложной ситуации строится следующим образом: на основе знаний об исследуемой ситуации составляется уравнение для каждого базисного фактора, включающее набор всех факторов, непосредственно влияющих на этот фактор (этот набор выбирается по когнитивной карте ситуации) вместе с функциональными характеристиками соответствующих влияний.

Моделирование основано на сценарном подходе и представляет собой процесс передачи управлений по вершинам графа модели. *Сценарий* – набор тенденций, характеризующих саморазвитие ситуации в начальный момент, вектора целей развития, вектора управлений (управляющих воздействий), характеризующих комплексы мероприятий, влияющих на развитие ситуации, и системы наблюдаемых параметров (факторов), иллюстрирующих тенденции процессов в ситуации.

Модель может исследоваться по трем основным направлениям [98]:

1. Прогноз саморазвития ситуации (без внешнего воздействия на процессы в ситуации);
2. Прогноз развития ситуации с выбранным вектором управлений

(прямая задача);

3. Синтез вектора управлений для достижения необходимого направления развития ситуации (обратная задача).

Пусть задан  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t))$  - вектор тенденций всех факторов в момент  $t$ . Динамика изменений этих значений во времени определяется видом модели, выбираемой для описания взаимодействия непосредственных влияний, одновременно приходящих на каждый базисный фактор от «соседних» факторов [1].

Конфигураторы для значений факторов и влияний их друг на друга строятся на основании таблицы 4.8. Возможные схемы конфигураторов приведены ниже:

*Фактор  $x_i$*  {

4 {Падает <sup>1</sup> 4; Не изменяется <sup>6</sup>; Растет <sup>7</sup> 9}

3 {Сильно падает <sup>1</sup> 2 3; Падает <sup>4</sup> 5; Не изменяется <sup>6</sup>; Растет <sup>7</sup> 8; Сильно растет <sup>9</sup> 10 11}

2 {Очень сильно падает <sup>1</sup> [-1; -0,8]; Сильно падает <sup>2</sup> [-0,8; -0,6]; Умеренно падает <sup>3</sup> [-0,6; -0,4]; Слабо падает <sup>4</sup> [-0,4; -0,2]; Очень слабо падает <sup>5</sup> [-0,2; 0]; Не изменяется <sup>6</sup> [0; 0]; Очень слабо растет <sup>7</sup> (0; 0,2]; Слабо растет <sup>8</sup> (0,2; 0,4]; Умеренно растет <sup>9</sup> (0,4; 0,6]; Сильно растет <sup>10</sup> (0,6; 0,8]; Очень сильно растет <sup>11</sup> (0,8; 1]}

1 {[-1; 1]}

$G(\text{Фактор } x_i) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4\}$ .

*Влияние  $a_{ij}$*  {

4 {Ослабляет <sup>1</sup> 4; Не влияет <sup>6</sup>; Усиливает <sup>7</sup> 9}

3 {Сильно ослабляет <sup>1</sup> 2 3; Ослабляет <sup>4</sup> 5; Не влияет <sup>6</sup>; Усиливает <sup>7</sup> 8; Сильно усиливает <sup>9</sup> 10 11}

2 {Очень сильно ослабляет <sup>1</sup> [-1; -0,8]; Сильно ослабляет <sup>2</sup> [-0,8; -0,6]; Умеренно ослабляет <sup>3</sup> [-0,6; -0,4]; Слабо ослабляет <sup>4</sup> [-0,4; -0,2]; Очень слабо ослабляет <sup>5</sup> [-0,2; 0]; Не влияет <sup>6</sup> [0; 0]; Очень слабо усиливает <sup>7</sup> (0; 0,2]; Слабо усиливает <sup>8</sup> (0,2; 0,4]; Умеренно усиливает <sup>9</sup> (0,4; 0,6]; Сильно усиливает <sup>10</sup> (0,6; 0,8]; Очень сильно усиливает <sup>11</sup> (0,8; 1]}

1 {[-1; 1]}

$G(\text{Влияние } a_{ij}) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4\}$ .

На основе интервалов 2-го домена могут быть заданы функции принадлежности (см. главу 2). Структурно-завершенные орграфы доменов тестов «Фактор» и «Влияние», включая предельный орграф, показаны на рис. 4.18.

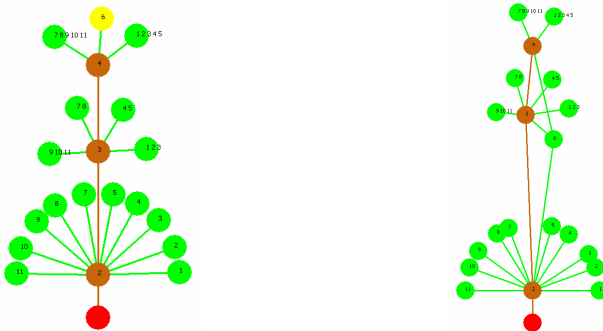
а) Орграф доменов  $G^+$ б) Предельный орграф доменов  $G^{++}$ 

Рис. 4.18 – Орграфы доменов тестов «Фактор» и «Влияние»

Общий вид когнитивной модели  $k$ , построенной на основе когнитивной карты и банка тестов  $\{G(x)\}$ ,  $\{G(a)\}$ , может быть следующим:

$$k = \{f/\mu. \{<x/X, a/A>\} \rightarrow x'/X\}, \quad (4.5)$$

где  $f/\mu$  – системопаттерн ( $\mu$  – механизм реализации);  $x'/X$  – изучаемый фактор;  $\{<x/X, a/A>\}$  – «соседние» факторы с коэффициентами влияния.

Помимо модели (4.5), для любого  $Z$ -фактора строится математическая модель в виде предельной синдромной и вероятностной модели знаний в рамках контекста  $\langle \Omega = \{\alpha\{x/X, z/Z\}\}, \{G(x)\}, G(z) \rangle$ , где  $\{x\}$  – тесты, с помощью которых описываются прецеденты  $\alpha\{x/X, z/Z\}$  (см. главу 5).

#### 4.2.6 Управление агротехнологиями

*Задачи управления продуктивностью и качеством сельхозкультур.* Рассматриваются задачи управления продуктивностью (урожайностью) и качеством сельскохозяйственных культур путем выбора и применения оптимальной агротехнологии. В качестве параметров управления могут рассматриваться, например, следующие факторы (тесты) [93]: поставленная цель (максимальное количество или максимальное качество продукции); вид севооборота; ротация севооборота; предшественники; система удобрений; сорта; способы посева; применение гербицидов и регуляторов роста; защита растений от вредителей и болезней; глубина и способ обработки почвы; сроки сева; нормы высева (млн. всхожих семян на 1 га).

Решаются две основные задачи:

1. Прогнозирование того, какие хозяйственные результаты наиболее вероятны на данном виде почв, с данными предшественниками при

условии применения имеющихся в распоряжении агротехнологий.

2. Поддержка принятия решений по выбору управляющих воздействий, т.е. консультирование по вопросам о том, какие виды почв, предшественники и агротехнологии должны быть, чтобы можно было рассчитывать с определенной уверенностью на заданный хозяйственный результат.

Задача поддержки принятия решений является обратной по отношению к задаче прогнозирования: если в задаче прогнозирования требуется по значениям факторов определить в какие состояния может под их влиянием перейти объект управления, то в задаче поддержки принятия решений, наоборот, по заданному будущему состоянию объекта управления (обычно целевому) определить значения факторов, которые с наибольшей вероятностью переведут объект управления в это состояние. Вторая задача получает разрешение в рамках синдромного принципа управления (глава 7).

Практическая значимость решения этих задач для хозяйств очевидна, т.к. достоверное прогнозирование результатов применения и научно-обоснованные рекомендации по рациональному выбору агротехнологий позволят повысить экономическую эффективность хозяйствования.

Новизна предлагаемого способа решения задач состоит в построения предельных синдромных и вероятностных моделей знаний на основе интеллектуальной обработки ретроспективных данных, отражающих фактический опыт выращивания исследуемых сортов, с использованием банка тестов. Предельный синдром – это фактически набор параметров порядка (в синергетическом смысле), определяющий эволюцию ситуации и обеспечивающий подстройку всех остальных параметров (принцип подчинения).

Для примера рассмотрим влияние различных агротехнологий на урожайность пшеницы и ячменя в конкретных агрометеорологических условиях. Для формулировки целей будем использовать тест «Урожайность» (Y1) и тест «Качество» (Y2). Варианты простейших конфигураторов приведены ниже:

*Урожайность*  $\wedge Y1 \{2 \{Низкая \wedge Н НС; Средняя \wedge С; Высокая \wedge ВС В\} 1 \{Низкая \wedge Н; Ниже средней \wedge НС; Средняя \wedge С; Выше средней \wedge ВС; Высокая \wedge В\}\}$

*Качество*  $\wedge Y2 \{1 \{Рядовая; Ценная; Сильная; Твердая; Семенная\}\}$ .

В качестве классификационных шкал (заклучения Z) могут быть выбраны следующие комбинации тестов: Урожайность (точное значение); Урожайность (интегральная оценка); Качество; Культура + сорт + урожайность; Культура + сорт + качество; Культура + урожайность (округленная); Культура + качество.

В качестве факторов-тестов, обуславливающих переход объекта управления в требуемые состояния, выберем следующие:

1. Предшественники
2. Уровень плодородия
3. Обработка почвы
4. Удобрение
5. Защита растений

Зададим конфигураторы тестов следующим образом:

*Культура*  $\wedge X_1$  {2 {Пшеница  $\wedge$ ОП; Ячмень  $\wedge$ Я ОЯ ЯЯ} 1 {Озимая пшеница  $\wedge$ ОП; Озимый ячмень  $\wedge$ ОЯ; Яровой ячмень  $\wedge$ ЯЯ}}

*Сорт*  $\wedge X_2$  {1 {-  $\wedge$ 0; Сорт 1  $\wedge$ 1; Сорт 2  $\wedge$ 2; Сорт 3  $\wedge$ 3; Сорт 4  $\wedge$ 4}}

*Предшественники*  $\wedge X_3$  {1 {-  $\wedge$ 0; Кукуруза  $\wedge$ К; Многолетние травы  $\wedge$ Т; Сахарная свекла  $\wedge$ С}}

*Уровень плодородия*  $\wedge X_4$  {1 {Высокий  $\wedge$ В; Средний  $\wedge$ С; Низкий  $\wedge$ Н}}

*Обработка почвы*  $\wedge X_5$  {1 {Безотвальная  $\wedge$ Б; Отвальная  $\wedge$ О; Поверхностная  $\wedge$ П}}

*Удобрения*  $\wedge X_6$  {2 {Имеется  $\wedge$ И МС ОС ОМ; Отсутствует  $\wedge$ О} 1 {Минеральная система  $\wedge$ МС; Органическая система  $\wedge$ ОС; Органическо-минеральная  $\wedge$ ОМ; Отсутствует  $\wedge$ О}}

*Защита растений*  $\wedge X_7$  {1 {«Бактороденцид 3 кг/га; Фосфид цинка т.п. (5%) 4 кг/га»  $\wedge$ 1; ...  $\wedge$ 2}}

*Урожайность (%)*  $\wedge Y_1$  {

4 {0.0 – 40.0  $\wedge$ 1 3; 40.0 – 80.0  $\wedge$ 5 7; 80.0 – 100.0  $\wedge$ 9}

3 {0.0 – 20.0  $\wedge$ 1 2; 20.0 – 40.0  $\wedge$ 3 4; 40.0 – 60.0  $\wedge$ 5 6; 60.0 – 80.0  $\wedge$ 7 8; 80.0 – 100.0  $\wedge$ 9 10}

2 {0.0 – 10.0  $\wedge$ 1 [0,0; 10,0]; 10.0 – 20.0  $\wedge$ 2 (10,0; 20,0); 20.0 – 30.0  $\wedge$ 3 (20,0; 30,0); 30.0 – 40.0  $\wedge$ 4 (30,0; 40,0); 40.0 – 50.0  $\wedge$ 5 (40,0; 50,0); 50.0 – 60.0  $\wedge$ 6 (50,0; 20,0); 60.0 – 70.0  $\wedge$ 7 (60,0; 70,0); 70.0 – 80.0  $\wedge$ 8 (70,0; 80,0); 80.0 – 90.0  $\wedge$ 9 (80,0; 90,0); 90.0 – 100.0  $\wedge$ 10 (90,0; 100,0)}

1 {{0; 100}}.

Примечание: урожайность каждой культуры берется в относительных единицах.

*Качество*  $\wedge Y_2$  {1 {Рядовая  $\wedge$ 1; Ценная  $\wedge$ 2; Сильная  $\wedge$ 3; Твердая  $\wedge$ 4; Семенная  $\wedge$ 5}}

$Z = \{0; 1\}$ , где  $z = 1$ - целевой класс.

Приведем примеры целевых классов ( $z = 1$ ):

Культура + сорт + урожайность: Озимая пшеница+Сорт\_2+{70.0-80.0};

Культура + сорт + качество: Озимая пшеница+Сорт\_2+Ценная;



Урожайность/4? 80.0 – 100.0.

На основе конфигураторов тестов могут быть построены структурно-завершенные орграфы доменов. На рис. 4.19а представлен структурно-завершенный орграф для тестов «Культура» (X1), «Уровень плодородия» (X4) и «Обработка почвы» (X5); на рис. 4.19б – орграф тестов «Предшественники» (X3) и «Удобрения» (X6); на рис. 4.19в – орграф тестов «Сорт» (X2) и «Качество» (Y2).

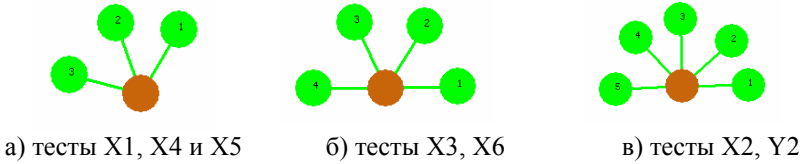
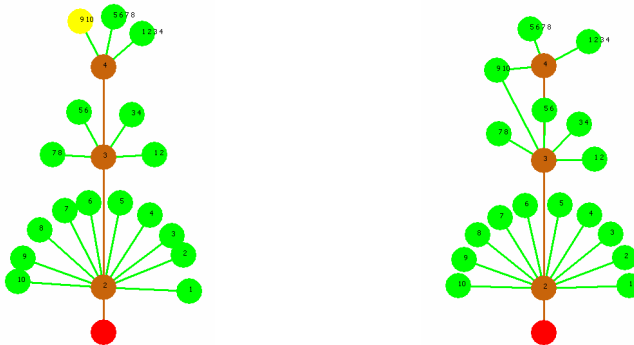


Рис. 4.19 – Орграфы  $G^+$  разных тестов

Структурно-завершенные орграфы доменов теста «Урожайность» (Y2), включая предельный орграф, показаны на рис. 4.20. Канонический орграф теста «Урожайность» имеет следующий вид (домен «1» - конструктивно-непрерывный):

*Урожайность* {4 {1 2 3 4; 5 6 7 8; 9 10} 3 {1 2; 3 4; 5 6; 7 8; 9 10} 2 {1; 2; 3; 4; 5; 6; 7; 8; 9; 10} 1 {} }.



а) Орграф доменов  $G^+$

б) Предельный орграф доменов  $G^{++}$

Рис. 4.20 – Орграфы доменов теста «Урожайность»

В таблице 4.9 приведен фрагмент базы прецедентов. В качестве значений тестов в базу прецедентов помещены метки элементов базовых доменов.

Таблица 4.9 – Фрагмент базы прецедентов

Факторы							Результаты		
X1	X2	X3	X4	X5	X6	X7	Y1	Y2	Z
ОП	2	К	В	Б	МС	1	65	2	1
ОЯ	1	К	С	О	ОС	1	54	2	0
ОП	2	Т	В	П	ОМ	1	67	2	1
ОП	1	С	Н	П	О	1	44	3	0
ЯЯ	3	Т	С	О	ОС	1	57	1	0

На основе таблицы 4.9 с помощью банка тестов создается расширенная таблица, которая для каждого фактора содержит столько колонок, сколько доменов в конфигураторе данного фактора. С помощью расширенной таблицы и банка тестов находятся (предельные) синдромные и вероятностные модели знаний (глава 5). Модели знаний служат основой для решения первой и второй задач.

Задача прогнозирования состоит в том, чтобы по планируемым к применению или уже применяемым агротехнологиям спрогнозировать наиболее вероятные хозяйственные результаты. Приведем пример Анкеты распознаваемой ситуации:

*Анкета ситуации* {ПРЕДШЕСТВЕННИКИ/1? Многолетние травы; УРОВЕНЬ ПЛОДОРОДИЯ/1? Средний; ОБРАБОТКА ПОЧВЫ/1? Безотвальная; УДОБРЕНИЕ/1? Отсутствует; ЗАЩИТА РАСТЕНИЙ/1? «Бактороденцид 3 кг/га : Фосфид цинка т.п. (5%) 4 кг/га»}.

По Анкете ситуации и фиксированной синдромной модели знаний определяются все имеющиеся синдромы. Синдромы могут указывать либо на  $z = 1$ , либо на  $z = 0$ . Если обнаружатся синдромы, которые указывают на разные заключения, то либо анализируемая ситуация не соответствует реальности, либо база прецедентов не релевантная (содержит не все значимые прецеденты), другими словами, анализируемая ситуация фальсифицирует выбранную модель знаний.

Если нас интересует вопрос о том, какие необходимо применить агротехнологии, чтобы получить очень высокую урожайность озимой пшеницы сорта\_2, то необходимо сформировать соответствующее  $Z$  и с помощью метода предельных обобщений построить синдромные модели знаний. Далее определяются достижимые целевые синдромы, которые затем реализуются. Классификационная шкала  $Z$  формируется так ( $z = 1$ ):

Культура + сорт + урожайность: Озимая пшеница+Сорт\_2+{90.0-100.0}.

Отметим, что если бы в исходных данных были указаны предшественники не только прошлого года, но и позапрошлого, и 3, 4, и более лет назад, то предложенная модель позволила бы исследовать их

влияние на количество и качество урожая зерновых, а также выявить и использовать законы севооборота для автоматизированного прогнозирования и поддержки принятия решений в растениеводстве.

Модель может быть расширена также за счет явного учета локальных (местных) особенностей регионов и даже отдельных хозяйств. Ожидается, что эти уточненные "локализованные", "адаптированные" знания будут более адекватны, а значит и более эффективны на практике, чем знания, не учитывающие местную специфику [93].

*Агрометеорологические задачи выбора микрозон и культур для выращивания.* Это задачи выявления причинно-следственных закономерностей между метеофакторами и результатами выращивания, прогнозирования периодов фенофаз по динамике метеопараметров, а также задачи триангуляции и пространственной интерполяции [93].

Основной задачей является прогнозирование количественных и качественных результатов выращивания различных культур в заданных точках (микрозонах). На основе решения данной задачи для различных культур и точек выращивания (полигонов, микрозон) могут быть решены следующие две подзадачи [93]:

1. Определен потенциал выращивания заданной культуры в различных микрозонах.
2. Определен потенциал выращивания различных культур в заданной микрозоне.

Результаты решения 1-й подзадачи подаются в наглядной и удобной для принятия решений форме геоинформационной визуализации в виде зональных географических карт, на которых зоны с одинаковым потенциалом для выращивания некоторой культуры закрашены одним условным цветом [93].

Будем считать, что биологический потенциал географической точки определяется плодородием почв и многолетней динамикой метеопараметров в этой точке. Влияние агротехнологий в данном исследовании не учитывается, хотя не представляет сложностей учесть и эти факторы (конфигураторы некоторых факторов приведены выше). Предполагается, что погода изменяется медленно, генотип культур и плодородие почв еще медленней.

Решение основной задачи базируется на гипотезе о том, что закономерности влияния динамики метеопараметров на переход культуры от одной фенофазы к другой и на количественные и качественные результаты ее выращивания определяются только плодородием почв и генотипом данной культуры. Если данная гипотеза верна, то закономерности влияния факторов среды на результаты выращивания данной культуры, выявленные на основе обработки эмпирических данных,

полученных за определенный период времени в одних точках, применимы для прогнозирования результатов выращивания этой культуры и в других точках (с известным уровнем плодородия почв).

Известно [93], что влияние одних и тех же метеопараметров на количественные и качественные результаты выращивания культуры зависит не только от самих значений этих метеопараметров, но в огромной степени – от *фенофазы*, в течение которой эти параметры действовали. Поэтому вся метеоинформация должна обрабатываться с привязкой к периодам фенофаз, характеризующим темп биологического времени для различных культур, а не просто – к физическому (календарному) времени.

Для плодовых культур выделяют 13 фенофаз [93]: 1 – Набухание плодовых почек; 2 – Распускание плодовых почек; 3 – Развертывание первых листьев; 4 – Обособление бутонов; 5 – Цветение; 6 – Конец цветения; 7 – Формирование плодов; 8 – Окончание роста побегов; 9 – Вызревание древесины; 10 – Съемная спелость; 11 – Осеннее расцвечивание листьев; 12 – Листопад; 13 – Вторичный рост побегов. Сроки начала и окончания фенофаз являются тестами со своими конфигураторами и измеряются в целочисленной шкале [1; 365] (отсчет идет от 1-го января текущего года). Все даты также привязываются к этой шкале.

Предположим, что разработан банк тестов и наполнена база прецедентов с использованием банка тестов. В этом случае решение основной задачи состоит из двух этапов.

На 1-м этапе выявляются причинно-следственные зависимости между метеоусловиями, плодородием почв и результатами выращивания сельскохозяйственных культур. Другими словами, строятся синдромные модели знаний для разных множеств заключений  $Z$ .

На 2-м этапе знания о выявленных причинно-следственных зависимостях (формальных синдромах) используются для прогнозирования количественных и качественных результатов выращивания заданной культуры в конкретном пункте выращивания.

В рамках 1-го этапа решаются две подзадачи.

*Задача 1.1:* На основе обработки эмпирической информации об условиях и результатах выращивании различных культур в конкретных микрорайонах выявить причинно-следственные закономерности влияния плодородия почв и метеопараметров, действующих в течение определенных фенотипических фаз развития, на количественные и качественные результаты выращивания.

Исходные данные:

- банк тестов;
- характеристики всех микрорайонов, включая плодородие почв (каждой микрорайону присваивается уникальный код);

- фактические периоды начала и окончания фенофаз различных культур в конкретных микрозонах (по шкале [1; 365]); орграфы набросков периодов фенофаз различных культур;

- фактические посуточные значения метеопараметров в конкретных микрозонах (по шкале [1; 365]); орграфы набросков динамики метеопараметров по различным качественным характеристикам.

Результат решения – синдромные модели знаний для количественных и качественных результатов выращивания различных культур. Синдромные модели знаний показывают список значимых для прогноза тестов, а также достаточный уровень общности (точности) по каждому тесту –  $\{t/T\}$ .

Фактические периоды фенофаз по всем микрозонам, всем сортам и всем годам хранятся в таблице вида 4.10 (по шкале [1; 365]).

Таблица 4.10 – Фактические периоды фенофаз

Зона	Год	Сорт	Фенофаза		
			№	Начало	Окончание
1	2007	Джонатан	1	83	90
1	2007	Джонатан	2	94	100
1	2007	Джонатан	3	104	110
1	2007	Джонатан	4	114	120
1	2007	Джонатан	5	120	124
...	...	...	...	...	...

Начало и окончание любой фенофазы по любому сорту (или группе сортов) является тестом со своим конфигуратором. Приведем фрагмент такого конфигуратора:

*Джонатан\_Фенофаза\_1\_Начало* {2 {...} 1 {[75; 90]} }.

Для потенциальных точек выращивания, по которым необходимо прогнозировать количественные и качественные результаты, периоды начала и окончания фенофаз для различных культур, как правило, неизвестны (синдромная модель показывает, какие периоды фенофаз действительно важны для прогноза). Поэтому возникает следующая задача.

*Задача 1.2:* на основе обработки эмпирической информации об условиях и динамике фенотипического развития различных культур в конкретных микрозонах выявить причинно-следственные зависимости влияния динамики метеопараметров на сроки начала и окончания фенофаз для этих культур. Исходные данные те же, что и в задаче 1.1.

Результат решения – синдромные модели знаний, позволяющие с минимальной погрешностью определять сроки наступления и окончания фенофаз для различных культур как в фактических, так и в предполагаемых

точках их выращивания по известному плодородию почв и известной динамике метеопараметров в этих точках.

Задача 1.2 так же, как и задача 1.1, включает в качестве вспомогательных задачи триангуляции и пространственной интерполяции.

Когда задачи 1-го этапа решены, то решение основной задачи сводится к выполнению следующих шагов (2-й этап):

1. Определение трех ближайших к заданной точке метеостанций по географическим координатам этой точки и метеостанций (триангуляция). Для метеостанций сетка нерегулярная, а для точек выращивания, как правило, регулярная (квадратная с равным шагом ячейки). Определение динамики метеопараметров в заданной точке путем взвешивания значений метеопараметров в трех ближайших метеостанциях с учетом расстояний и рельефа местности (пространственная интерполяция).

2. Прогнозирование наиболее вероятных сроков начала и окончания фенофаз для заданной культуры в заданной точке по известному плодородию почв и известной динамике метеопараметров (задача 1.2). Прогнозировать нужно только те периоды фенофаз, которые фигурируют в синдромной модели знаний.

3. Прогнозирование количественных и качественных результатов выращивания культуры по известным значениям метеопараметров, действующих в прогнозируемые периоды фенофаз (результат решения задачи 1.2), и известным закономерностям влияния метеопараметров на результаты выращивания (результат решения задачи 1.1). Прогнозирование осуществляется на основе синдромной модели знаний.

Для описания действия метеофакторов применяются, в частности, следующие тесты [93]: Температура воздуха средняя (°C); Температура воздуха максимальная (°C); Температура воздуха минимальная (°C); Температура почвы средняя (°C); Температура почвы максимальная (°C); Температура почвы минимальная (°C); Влажность относительная средняя; Влажность относительная минимальная; Влажность абсолютная; Количество осадков днем; Количество осадков за ночь; Количество осадков за сутки; Атмосферное давление на уровне метеостанции; Атмосферное давление на уровне моря; Облачность всего; Облачность низкая; Ветер – азимут; Ветер – скорость; Атмосферные явления.

Общий вид конфигуратора для числовых тестов может быть следующим (в относительных единицах):

$$\begin{aligned}
 & \text{Показатель } \{ \hspace{15em} (4.6) \\
 & 3 \{ 0 - p_2 \wedge 1 \ 2; p_2 - p_4 \wedge 3 \ 4; p_4 - 1 \wedge 5 \ 6 \} \\
 & 2 \{ 0 - p_1 \wedge 1 \ [0; p_1]; p_1 - p_2 \wedge 2 \ (p_1; p_2); p_2 - p_3 \wedge 3 \ (p_2; p_3); p_3 - p_4 \wedge 4 \ (p_3; p_4); \\
 & \quad p_4 - p_5 \wedge 5 \ (p_4; p_5); p_5 - 1 \wedge 6 \ (p_5; 1) \} \\
 & 1 \{ \{ 0; 1 \} \}
 \end{aligned}$$

$G(\text{Показатель}) = \{1 \rightarrow 2 \rightarrow 3\}$ .

В качестве «0» принимается минимальное значение параметра (MIN), а в качестве «1» максимальное (MAX). Минимальные и максимальные значения хранятся в специальной таблице (таблица 4.11). Параметры  $p_1 \dots p_5$  задаются экспертом или рассчитываются автоматически.

Таблица 4.11 – Интервалы числовых метеофакторов

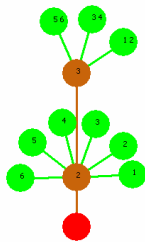
Показатель	MIN	MAX
Температура воздуха средняя (°C)	-5	20
Температура почвы средняя (°C)	-2	18
Влажность относительная средняя	50	86
Количество осадков днем	0	5
Ветер – скорость	2	10
...	...	...

Конфигуратор для теста «Ветер - азимут»:

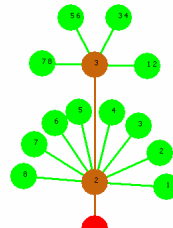
*Ветер – азимут* {

- 3 {Северный - Северо-западный ^1 2; Западный - Юго-западный ^3 4; Южный - Юго-восточный ^5 6; Восточный - Северо-восточный ^7 8}
- 2 {Северный ^1 [0; 45]; Северо-западный ^2 (45; 90]; Западный ^3 (90; 135]; Юго-западный ^4 (135; 180]; Южный ^5 (180; 225]; Юго-восточный ^6 (225; 270]; Восточный ^7 (270; 315]; Северо-восточный ^8 (315; 360]}
- 1 {[0; 360]} }

На рис. 4.21а представлен структурно – завершенный оргграф доменов числовых тестов «Показатель». На рис. 4.21б представлен структурно-завершенный оргграф доменов теста «Ветер – азимут».



а) тесты «Показатель»



б) тест «Ветер - азимут»

Рис. 4.21 – Оргграф доменов числовых тестов

Конфигуратор теста «Атмосферные явления»:

*Атмосферные явления* {1 {Штиль; Дождь; Ливневый дождь; Морось; Снег; Ливневый снег; Снежная крупа; Снежные зерна; Ледяная крупа; Ледяной

дождь; Град; Мокрый снег; Ливневый мокрый снег; Роса; Иней; Гололед; Зернистая изморось; Кристаллическая изморось; Гололедица; Туман; Просвечивающийся туман; Дымка; Гроза; Зарница; Шквал}}.

Фиксация любых метеофакторов осуществляется с помощью когнитивной конструкции «событие» (глава 3). Примеры (*t/A* - время):

<Температура воздуха средняя/1? 15; *t/A2?* 12.10.2011>;

<Атмосферные явления /1? Дождь, Шквал; *t/A2?* 12.10.2011>;

<Ветер - азимут/3? Южный - Юго-восточный; *t/A2?* 04.06.2012>;

<Показатель/3? р4 – 1; *t/A3?* 11.2012>;

<Фенофаза\_1\_Начало/1? 83; Сорт/1? Джонатан; Зона? 1; *t/A4?* 2007>.

*Задачи совершенствования методов селекции подсолнечника.* Рассматриваются следующие задачи [93]:

Задача 1: выявление причинно-следственных зависимостей между фенотипическими признаками подсолнечника и его хозяйственными свойствами.

Задача 2: прогнозирование хозяйственных свойств растений подсолнечника на основе анализа их фенотипических признаков.

Задача 3: поддержка принятия решений по отбору растений для селекции на основе анализа фенотипических признаков.

Основой решения всех трех задач является синдромная модель знаний, которая строится на основе банка тестов и базы прецедентов. Каждый синдром представляет собой набор параметров порядка, следовательно, разные синдромы обеспечивают конкуренцию параметров порядка. В базу прецедентов заносятся отдельные растения (фенотипические признаки и хозяйственные свойства).

Задача принятия решений является обратной по отношению к задаче прогнозирования. Если при прогнозировании по фенотипическим признакам определяются возможные хозяйственные свойства, то при принятии решений, наоборот, по заданным хозяйственным свойствам определяются фенотипические признаки, которыми должны обладать растения, имеющие эти свойства. Синдромные модели знаний непосредственно используются при отборе растений подсолнечника для формирования следующего поколения при селекции: выбираются растения с фенотипическими признаками, указанными в том или ином подходящем синдроме.

Заключения в базе прецедентов (*z*) формируются на основе списка хозяйственных свойств, а именно: урожайность (*г*); сборы масла (*г*); средняя масличность (%); натура (*г/л*). Для конфигураторов хозяйственных свойств (тестов) используем схему (4.6) только в абсолютных единицах:



*Урожайность (г)* {3 {менее 500 ^1 2; [500, 700) ^3 4; 700 и более ^5 6}  
 2 {менее 400 ^1 [200; 400); [400, 500) ^2 [400; 500); [500, 600) ^3 [500; 600);  
 [600, 700) ^4 [600; 700); [700, 800) ^5 [700; 800); 800 и более ^6 [800;  
 900]}  
 1 {[200; 900]}}

*Сборы масла (г)* {3 {менее 250 ^1 2; [250, 350) ^3 4; 350 и более ^5 6}  
 2 {менее 200 ^1 [100; 200); [200, 250) ^2 [200; 250); [250, 300) ^3 [250; 300);  
 [300, 350) ^4 [300; 350); [350, 400) ^5 [350; 400); 400 и более ^6 [400;  
 500]}  
 1 {[100; 500]}}

*Средняя масляничность (%)* {  
 3 {менее 51 ^1 2; [51, 53) ^3 4; 53 и более ^5 6}  
 2 {менее 50 ^1 [45; 50); [50, 51) ^2 [50; 51); [51, 52) ^3 [51; 52); [52, 53) ^4  
 [52; 53); [53, 54) ^5 [53; 54); 54 и более ^6 [54; 56]}  
 1 {[45; 56]}}

*Натура (г/л)* {3 {менее 440 ^1 2; [440, 460) ^3 4; 460 и более ^5 6} 2 {менее  
 430 ^1 [400; 430); [430, 440) ^2 [430; 440); [440, 450) ^3 [440; 450); [450,  
 460) ^4 [450; 460); [460, 470) ^5 [460; 470); 470 и более ^6 [470; 490]} 1  
 {[400; 490]}}

На рис. 4.21а) представлен единственный структурно-завершенный оргграф доменов хозяйственных свойств кукурузы.

В качестве фенотипических признаков подсолнечника примем следующие признаки [93]: Вес семян с 1 корзины (г); Время посева; Высота растения (см); Облиственность (шт); Диаметр корзины (см). Для конфигураторов фенотипических признаков используем преимущественно схему (4.6) или близкую к ней:

*Вес семян с 1 корзины (г)* {  
 3 {менее 70 ^1 2; [70, 90) ^3 4; 90 и более ^5 6}  
 2 {менее 60 ^1 [50; 60); [60, 70) ^2 [60; 70); [70, 80) ^3 [70; 80); [80, 90) ^4  
 [80; 90); [90, 100) ^5 [90; 100); 100 и более ^6 [100; 110]}  
 1 {[50; 110]}}

*Время посева* {1 {Позднезимний (декабрь); Ранневесенний (февраль);  
 Оптимальный}}

*Высота растения (см)* {  
 3 {менее 110 ^1 2 3; [110, 170) ^4 5; 170 и более ^6 7}  
 2 {менее 50 ^1 [40; 50); [50, 80) ^2 [50; 80); [80, 110) ^3 [80; 110); [110, 140)  
 ^4 [110; 140); [140, 170) ^5 [140; 170); [170, 200) ^6 [170; 200); 200 и  
 более ^7 [200; 220]}  
 1 {[40; 220]}}

*Облиственность (шт)* {3 {менее 32 ^1 2; 32 - 35 ^3 4; более 35 ^5}  
 2 {менее 30 ^1 [20; 29]; 30 - 31 ^2 [30; 31]; 32 - 33 ^3 [32; 33]; 34 - 35 ^4 [34;  
 35]; более 35 ^5 [36; 50]}  
 1 {{20; 50}}}

*Диаметр корзины (см)* {3 {менее 21 ^1 2 3; [21, 23] ^4 5; 23 и более ^6 7}  
 2 {менее 19 ^1 [16; 19]; [19, 20] ^2 [19; 20]; [20, 21] ^3 [20; 21]; [21, 22] ^4  
 [21; 22]; [22, 23] ^5 [22; 23]; [23, 24] ^6 [23; 24]; 24 и более ^7 [24; 26]}  
 1 {{16; 26}}}

#### 4.2.7 Оценка научно-технической эффективности проектов

Приведем основные виды эффектов при разработке и реализации научно-технических проектов и программ [129].

1. Экономический эффект состоит в получении экономических результатов от научно-технических проектов как в целом для страны, регионов, так и для каждого производственного субъекта. Для оценки экономической эффективности проектов используются следующие, признанные в мировой практике, показатели: чистый дисконтированный доход; период окупаемости; коэффициент выгод-затрат; индекс доходности; коэффициент финансовой автономности проекта.

2. Научно-технический эффект отражает повышение научно-технического уровня посредством улучшения параметров техники и технологий, создания новых технологических средств производства материалов и видов продукции. Для научных разработок – открытие новых природных явлений, новых законов и закономерностей в природе.

3. Социальный эффект характеризует изменения условий жизни и деятельности человека и проявляется в улучшении характера и условий труда, повышении жизненного уровня населения, улучшении бытовых условий, расширении возможностей духовного развития личности, улучшении состояния окружающей среды. Оценка социальной эффективности проектов проводится по следующим показателям: увеличение рабочих мест в отрасли и государстве; повышение уровня квалификации персонала; развитие малого и среднего бизнеса; влияние на деятельность общественных и молодежных организаций; улучшение здоровья населения; улучшение качества жизни населения; влияние на образование и подготовку кадров.

4. Суммарный эффект разделяется на две составляющие:

- прямой (непосредственный) эффект;
- косвенный (опосредованный) эффект.

Количественную оценку показателей научно-технической эффективности можно проводить с помощью следующего банка тестов

«Эффективность» (на базовом уровне тесты измеряются в баллах от 1 до 10):

*Научно-технический уровень*  $X_1$  {

- 3 {Отвечает мировому уровню  $^{10}$  9; Отвечает отечественному уровню  $^6$  4 2; Низкий уровень  $^1$ }
- 2 {Превышает лучшие мировые аналоги  $^{10}$  [10; 10]; Отвечает мировому уровню  $^9$  [7; 9]; Ниже лучших мировых аналогов  $^6$  [5; 6]; Превышает лучшие отечественные аналоги  $^4$  [3; 4]; Отвечает отечественному уровню  $^2$  [2; 2]; Ниже отечественного уровня  $^1$  [1; 1]}
- 1 {[1; 10]}

*Перспективность*  $X_2$  {

- 3 {Перспективный проект  $^8$  5; Малоперспективный проект  $^1$  2}
- 2 {Проект первоочередной важности  $^8$  [8; 10]; Важный проект  $^5$  [5; 7]; Полезный проект  $^2$  [2; 4]; Неперспективный проект  $^1$  [1; 1]}
- 1 {[1; 10]}

*Масштаб использования*  $X_3$  {

- 3 {Мировой рынок  $^9$ ; Внутренний рынок  $^6$  3 1}
- 2 {Мировой рынок  $^9$  [9; 10]; Народное хозяйство  $^6$  [6; 8]; Область (регион)  $^3$  [3; 5]; Отдельное предприятие  $^1$  [1; 2]}
- 1 {[1; 10]}

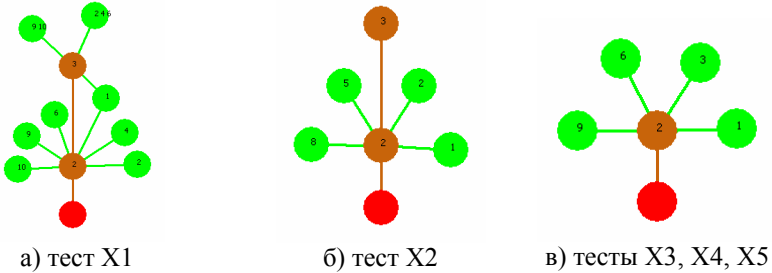
*Международная кооперация*  $X_4$  {

- 3 {Кооперация предполагается  $^9$  6 3; Кооперация не предполагается  $^1$ }
- 2 {Включение в международные программы  $^9$  [9; 10]; Соответствует международным соглашениям  $^6$  [6; 8]; Привлекаются иностранные инвестиции  $^3$  [3; 5]; Кооперация не предполагается  $^1$  [1; 2]}
- 1 {[1; 10]}

*Уверенность в достижении целей проекта*  $X_5$  {

- 3 {Большая  $^9$ ;  
Средняя и ниже  $^6$  3 1}
- 2 {Большая  $^9$  [9; 10];  
Средняя  $^6$  [5; 8];  
Слабая  $^3$  [3; 4];  
Очень слабая  $^1$  [1; 2]}
- 1 {[1; 10]}

На рис. 4.22 приведены предельные структурно-завершенные оргграфы доменов всех тестов: а)  $X_1$  – «Научно-технический уровень»; б)  $X_2$  – «Перспективность»; в)  $X_3$  – «Масштаб использования»,  $X_4$  – «Международная кооперация»,  $X_5$  – «Уверенность в достижении целей» (домен «1» - конструктивно-непрерывный).

Рис. 4.22 – Орграфы доменов  $G^+$  банка тестов «Эффективность»

Обозначим все показатели (тесты) научно-технической, экономической и социальной эффективности через  $X_1, \dots, X_n$ . Соответственно имеем банк тестов  $\{G(X_1), \dots, G(X_n)\}$ . Пусть  $Z = \{0, 1\}$  – список заключений, где  $z = 1$  – «проект принят», а  $z = 0$  – «проект отклонен». Сформируем базу прецедентов  $\Omega(\{T\}, Z) = \{\alpha(\{T\}, z/Z)\}$  из всех проектов, которые подвергались экспертизе (с известными исходами  $z \in Z$ ). В доменах «1» база проектов имеет вид таблицы 4.12 ( $N_0$  – номер проекта).

Таблица 4.12 – Фрагмент базы проектов

№	Показатели научно-техн. эффективности					Показатели эконом. эфф-ти	Показатели соц. эфф-ти		Z
	X1	X2	X3	X4	X5		...	Xn	
1	5	8	8	6	3	...	...	...	0
2	3	6	9	10	2	...	...	...	1
3	9	4	7	5	9	...	...	...	1

На основе базы прецедентов и банка тестов находятся предельные вероятностные и синдромные модели знаний. Последние применяются для анализа приоритетов экспертов, а также построения систем поддержки принятия решений (автоматизированной экспертизе проектов).

### 4.3 Метод «Штрихкода» квантования значений

От развитости Банка тестов  $\{G(\tau)\}$  зависит корректность предельных моделей знаний. Метод штрихкода в сложных ситуациях анализа позволяет автоматически строить приемлемые орграфы  $G(\tau)$  [149]. Под сложной ситуацией анализа понимается такая ситуация, когда имеющийся Банк тестов не позволяет найти корректную модель знаний.

### 4.3.1 Описание метода

Пусть дано конечное множество пар наблюдений

$$W = \{ \langle x, z \rangle \mid x \in R, z \in \{1, \dots, N\} \}.$$

Из множества  $W$  построим упорядоченное множество кортежей  $W'$  следующим образом:

1. Упорядочиваем множество  $W$  по возрастанию  $x$  и удаляем все совпадающие кортежи.

2. Объединяем все совпадающие по  $x$  пары наблюдений в один кортеж. Получаем:  $W' = \{ \langle x, z_1, \dots, z_p \rangle_j \mid \text{все } x \text{ различны} \} \equiv \{ \langle x, Z \rangle_j \mid \forall j \in \{1, \dots, k-1\}, x_j < x_{j+1}; Z_j = \{z_1, \dots, z_p\}_j \}$ ,  $|W'| = k$ .

Разобьем  $W'$  на минимальное число подмножеств  $W_j = \{ \langle x, Z \rangle_j \mid j = 1, \dots, m \}$  таким образом, что выполняются два условия:

1. Если  $\langle x, Z \rangle \in W_j$  и  $\langle x', Z' \rangle \in W_j$ , то  $Z = Z'$ .

2. Если для кортежей  $\langle x, Z \rangle$  и  $\langle x', Z' \rangle$  справедливо:  $Z = Z'$ , но при этом существует кортеж  $\langle x'', Z'' \rangle$  такой, что  $x < x'' < x'$  и  $Z'' \neq Z$ , то кортежи  $\langle x, Z \rangle$  и  $\langle x', Z' \rangle$  принадлежат разным  $W_j$ .

**Предложение 4.1.** Для любого множества наблюдений  $W$  разбиение  $W' = \cup_{j=1, m} W_j$  строится единственным образом (с точностью до нумерации).

Набор  $Z_j$  будем называть типом  $W_j$ . Разные  $W_j$  ( $j = 1, \dots, m$ ) могут иметь одинаковый тип. Удобно все типы пронумеровать произвольным образом от 1 до  $n$ , где  $n$  – максимальное число различных типов. Ясно, что  $n \leq m$ . Эквивалентные представления  $W''$ :  $W'' = \{ \langle x, p \rangle \}$ , где  $p$  – тип. Если  $p_j$  – тип  $W_j$ , то можно записать:

$$W'' = \cup_{j=1, m} W_j(p_j). \quad (4.7)$$

Разобьем отрезок  $[x_{min}, x_{max}] = [x_l, x_k]$  на непересекающиеся интервалы  $A_i$  ( $i = 1, \dots, m$ ) таким образом, чтобы в каждый интервал попали все  $x$  из  $W_i$ . Некоторые интервалы могут быть вырожденными, т.е. состоять из одной точки. Очевидно, разбиение  $\{A_i\}$  определяется не единственным образом. На разбиение  $\{A_i\}$  могут накладываться дополнительные ограничения, например, требование целочисленности интервалов.

*Метод штрихкода* для произвольного множества наблюдений  $W$  строит разные виды разбиений  $\{A_i\}$  с учетом имеющихся ограничений. Он включает в себя три этапа. На первом этапе строится множество  $W'$ . На втором этапе строится разбиение  $W''$ . На третьем этапе строится искомое разбиение  $A = \{ \langle A, p \rangle_j \mid j = 1, \dots, m \}$ , где  $p_j$  – тип, совпадающий с типом  $W_j(p_j)$ .

Рассмотрим пример. Пусть дано конечное множество пар наблюдений

$W = \{ \langle x, z \rangle \} = \{ \langle 17, 15; 1 \rangle, \langle 20, 4; 1 \rangle, \langle 22, 45; 1 \rangle, \langle 14, 1; 1 \rangle, \langle 13; 1 \rangle, \langle 15; 8 \rangle, \langle 17; 1 \rangle, \langle 19; 4 \rangle, \langle 14; 3 \rangle, \langle 15; 2 \rangle, \langle 19; 5 \rangle, \langle 12; 2 \rangle, \langle 16; 2 \rangle, \langle 12; 1 \rangle, \langle 15; 2 \rangle, \langle 17; 2 \rangle, \langle 16; 2 \rangle, \langle 19; 2 \rangle, \langle 21; 2 \rangle, \langle 13; 7 \rangle, \langle 14; 2 \rangle, \langle 17; 2 \rangle, \langle 17; 2 \rangle, \langle 18; 2 \rangle, \langle 21; 2 \rangle, \langle 14; 2 \rangle, \langle 18; 2 \rangle, \langle 14; 1 \rangle, \langle 18; 1 \rangle, \langle 13; 1 \rangle, \langle 15; 4 \rangle, \langle 17; 5 \rangle, \langle 19; 3 \rangle, \langle 12; 2 \rangle, \langle 16; 1 \rangle, \langle 17; 1 \rangle, \langle 20; 7 \rangle, \langle 21; 5 \rangle, \langle 21; 6 \rangle, \langle 21; 4 \rangle, \langle 22; 8 \rangle \}$ .

Преобразуем множество  $W$  во множество  $W'$ :

$W' = \{ \langle x, z_1, \dots, z_p \rangle \} = \{ \langle 12; 1, 2 \rangle, \langle 13; 1, 7 \rangle, \langle 14; 1, 2, 3 \rangle, \langle 14, 1; 1 \rangle, \langle 15; 2, 4, 8 \rangle, \langle 16; 1, 2 \rangle, \langle 17; 1, 2, 5 \rangle, \langle 17, 15; 1 \rangle, \langle 18; 1, 2 \rangle, \langle 19; 2, 3, 4, 5 \rangle, \langle 20; 7 \rangle, \langle 20, 4; 1 \rangle, \langle 21; 2, 4, 5, 6 \rangle, \langle 22; 8 \rangle, \langle 22, 45; 1 \rangle \} = \cup_{j=1, \dots, 15} W_j(p_j)$ .

Возможное разбиение  $A$  интервала  $[12; 22,45]$  приведено в таблице 4.13.

Таблица 4.13 – Возможное разбиение по методу штрихкода

Интервалы $A_j$	$Z_j$	Тип $p_j$	$j$
[12; 13)	1,2	1	1
[13; 14)	1,7	2	2
[14; 14,1)	1,2,3	3	3
[14,1; 15)	1	4	4
[15; 16)	2,4,8	5	5
[16; 17)	1,2	1	6
[17; 17,15)	1,2,5	6	7
[17,15; 18)	1	4	8
[18; 19)	1,2	1	9
[19; 20)	2,3,4,5	7	10
[20; 20,4)	7	8	11
[20,4; 21)	1	4	12
[21; 22)	2,4,5,6	9	13
[22; 22,45)	8	10	14
[22,45; 22,45]	1	4	15

Разбиение (4.7) графически удобно отображать в виде последовательности цветных полосок (такое представление и дало название методу). Совпадающие типы отображаются одним цветом и имеют одну высоту (рис. 4.23)

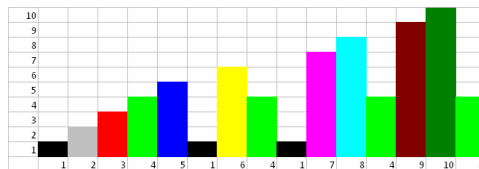


Рис. 4.23 – Разбиение «Штрихкод»

Построим (автоматически) конфигуратор на основе разбиения  $A$ . Типы  $p$  будут играть роль алиасов (меток) элементов первого дискретного домена. На следующих уровнях попарно объединяются соседние типы и так до тех пор, пока на очередном уровне не окажется 2 – 3 элемента.

*Тест* {

4 {1 - 4<sup>1</sup> 3; 5 - 8<sup>5</sup> 7; 9 - 10<sup>9</sup>}

3 {1 - 2<sup>1</sup> 2; 3 - 4<sup>3</sup> 4; 5 - 6<sup>5</sup> 6; 7 - 8<sup>7</sup> 8; 9 - 10<sup>9</sup> 10}

2 {12 - 13<sup>1</sup> [12; 13]; 13 - 14<sup>2</sup> [13; 14]; 14 - 14,1<sup>3</sup> [14; 14,1]; 14,1 - 15<sup>4</sup> [14,1; 15]; 15 - 16<sup>5</sup> [15; 16]; 16 - 17<sup>6</sup> [16; 17]; 17 - 17,15<sup>6</sup> [17; 17,15]; 17,15 - 18<sup>4</sup> [17,15; 18]; 18 - 19<sup>1</sup> [18; 19]; 19 - 20<sup>7</sup> [19; 20]; 20 - 20,4<sup>8</sup> [20; 20,4]; 20,4 - 21<sup>4</sup> [20,4; 21]; 21 - 22<sup>9</sup> [21; 22]; 22 - 22,45<sup>10</sup> [22; 22,45]; 22,45<sup>4</sup> [22,45; 22,45]}

1 {[12; 22,45]}.

$G(\text{Тест}) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4\}$ .

Возможно, на 3-м уровне целесообразнее было объединить элементы таким образом, чтобы уравнивать их суммарные «веса» (длины интервалов). На последующих уровнях также можно уравнивать суммарные веса элементов. В ряде практических задач мы так и поступали (см. раздел 4.3.3). Наконец, в рамках одного конфигулятора можно реализовать несколько способов объединения элементов, получив «ветвистый», а не линейный конфигуратор.

Легко убедиться в том, что если все пары наблюдений из  $W$  имеют одно и то же заключение  $z$  (например,  $z = 1$ ), то разбиения не происходит, т.е.  $A = \{<x_{min}, x_{max}, p = 1>\}$ .

Одному и тому же штрихкоду (рис. 4.23) могут соответствовать разные разбиения  $A$  интервала  $[x_{min}, x_{max}]$ . Приведем алгоритмы некоторых разбиений  $A$ .

#### **Левый штрихкод.**

*Вход:* Множество наблюдений  $W$ .

1. Строим множества  $W'$  и  $W'' = \{<x, p>\}$ .

2. Полагаем  $A_1 = [x_{min}, ?) = [x_1, ?)$ , тип -  $p_1$ ,  $W_1(p_1) = <x, p >_1$ .

Требуется определить правую границу интервала  $A_1$ .

2.1 Берем следующее наблюдение  $<x, p >_2$ . Если  $p_2 = p_1$ , то  $A_1$  поглощает точку  $x_2$  ( $W_1 = W_1 \cup <x, p >_2$ ) и так до тех пор, пока не встретится кортеж  $<x, p >_j$  с  $p_j$  отличным от  $p_1$ .

2.2 Полагаем  $A_1 = [x_1, x_j)$  и  $A_2 = [x_j, ?)$ , тип -  $p_j$ ,  $W_2(p_j) = <x, p >_j$ .

3. Шаг 2 выполняем для  $A_2$  и далее до полного исчерпания множества  $W''$ .

*Выход:* Разбиение  $A$ .

Пример левого штрихкода приведен в таблице 4.13.

*Правый штрихкод* отличается от левого направлением обработки кортежей множества  $W$ , а именно: справа налево (на Шаге 2 полагаем:  $A_1 = (?, x_k)$ ).

*Срединный штрихкод* отличается от левого правилом определения границы интервала, а именно, на шаге 2.2 полагаем:  $A_1 = [x_1, (x_{j-1} + x_j)/2)$  и  $A_2 = [(x_{j-1} + x_j)/2, ?)$ . Срединный штрихкод не применяется для целочисленных интервалов.

Смысл разбиения непрерывного или конструктивно-непрерывного интервала по методу штрихкода заключается в стремлении гарантировать отсутствие конфликтов в базе прецедентов хотя бы на одном дискретном уровне (при отсутствии конфликтов на непрерывном уровне). Сформулируем соответствующее утверждение.

Пусть  $\forall \tau, G(\tau) = \{Di \rightarrow Dj\}$ , причем  $D1$  – базовый непрерывный или конструктивно-непрерывный (целочисленный) домен, а  $D2$  – первый дискретный домен. Пусть  $A_\tau = \{<A, p>_{j=1,m}\}_\tau$  – некоторое штрихкодое разбиение домена  $D1$  произвольного теста  $\tau$ . Положим

$$D2 = \{A_1 \wedge p_1 A_1; \dots; A_m \wedge p_m A_m\}. \quad (4.8)$$

Пример домена  $D2$  показан в рамках конфигуратора  $G(\text{Тест})$ .

Пусть  $\Omega(\{\tau/T\}, Z) = \{\alpha(\{\tau/T\}, z)\}$  множество ситуаций действительности с известными исходами  $z \in Z$ . Все ситуации описаны с помощью банка тестов  $\{G(\tau) = \{Di \rightarrow Dj\}\}$ . Множество  $\Omega$  не содержит конфликтов на заданном уровне общности, если нет двух ситуаций с разными исходами, но совпадающими значениями тестов.

**Предложение 4.2.** Если множество прецедентов  $\Omega(\{\tau/D1\}, Z)$  не содержит конфликтов, то его обобщение – набросок  $\Omega(\{\tau/D2\}, Z)$  также не содержит конфликтов.

Предложение 4.2 гарантирует существование моделей знаний, в синдромы которых не входят значения  $\{\tau/D1\}$ .

### 4.3.2 Пример «Ураганы»

Пусть имеются данные о координатах – Долготе – Longitude и Широте – Latitude - для 37 циклонов, достигающих силы урагана, по двум классификациям циклонов – *Baro* и *Trop*. Модельный набор данных (фрагмент приведен в таблице 4.14) использовался для целей иллюстрации в работе Elsner, Lehmiller и Kimberlain (1996), авторы которой исследовали различия между бароклиническими и тропическими циклонами в Северной Атлантике. Данный пример широко используется для демонстрации методов распознавания и принятия решений.



Таблица 4.14 – Фрагмент таблицы «Ураганы»

LONGITUD	LATITUDE	CLASS
59,00	17,00	BARO
59,50	21,00	BARO
60,00	12,00	BARO
60,50	16,00	BARO
61,00	13,00	BARO
61,00	15,00	BARO
61,50	17,00	BARO
61,50	19,00	BARO
62,00	14,00	BARO
63,00	15,00	TROP
63,50	19,00	TROP
64,00	12,00	TROP
...	...	...

Обозначим класс *Baro* как  $z = 1$ , а класс *Trop* как  $z = 2$  и приведем полные данные по долготе и классу в виде массива  $W$ :

$$W = \{ \langle x, z \rangle \} = \{ \langle 59; 1 \rangle, \langle 59,5; 1 \rangle, \langle 60; 1 \rangle, \langle 60,5; 1 \rangle, \langle 61; 1 \rangle, \langle 61; 1 \rangle, \langle 61,5; 1 \rangle, \langle 61,5; 1 \rangle, \langle 62; 1 \rangle, \langle 63; 2 \rangle, \langle 63,5; 2 \rangle, \langle 64; 2 \rangle, \langle 64,5; 2 \rangle, \langle 65; 2 \rangle, \langle 65; 2 \rangle, \langle 65; 2 \rangle, \langle 65,5; 2 \rangle, \langle 65,5; 2 \rangle, \langle 66; 2 \rangle, \langle 66; 2 \rangle, \langle 66; 2 \rangle, \langle 66,5; 2 \rangle, \langle 66,5; 2 \rangle, \langle 66,5; 2 \rangle, \langle 67; 2 \rangle, \langle 67,5; 2 \rangle, \langle 68; 1 \rangle, \langle 68,5; 1 \rangle, \langle 69; 1 \rangle, \langle 69; 1 \rangle, \langle 69,5; 1 \rangle, \langle 69,5; 1 \rangle, \langle 70; 1 \rangle, \langle 70,5; 1 \rangle, \langle 71; 1 \rangle, \langle 71; 1 \rangle \}.$$

Преобразуем множество  $W$  во множество  $W'$ :

$$W' = \{ \langle x, z_1, \dots, z_p \rangle \} = \{ \langle 59; 1 \rangle, \langle 59,5; 1 \rangle, \langle 60; 1 \rangle, \langle 60,5; 1 \rangle, \langle 61; 1 \rangle, \langle 61,5; 1 \rangle, \langle 62; 1 \rangle, \langle 63; 2 \rangle, \langle 63,5; 2 \rangle, \langle 64; 2 \rangle, \langle 64,5; 2 \rangle, \langle 65; 2 \rangle, \langle 65,5; 2 \rangle, \langle 65,5; 2 \rangle, \langle 66; 2 \rangle, \langle 66,5; 2 \rangle, \langle 66,5; 2 \rangle, \langle 67; 2 \rangle, \langle 67,5; 2 \rangle, \langle 68; 1 \rangle, \langle 68,5; 1 \rangle, \langle 69; 1 \rangle, \langle 69,5; 1 \rangle, \langle 70; 1 \rangle, \langle 70,5; 1 \rangle, \langle 71; 1 \rangle \}.$$

$$W' = W_1(p_1 = 1) \cup W_2(p_2 = 2) \cup W_3(p_3 = 1),$$

где

$$W_1(p_1 = 1) = \{ \langle 59; 1 \rangle, \langle 59,5; 1 \rangle, \langle 60; 1 \rangle, \langle 60,5; 1 \rangle, \langle 61; 1 \rangle, \langle 61,5; 1 \rangle, \langle 62; 1 \rangle \};$$

$$W_2(p_2 = 2) = \{ \langle 63; 2 \rangle, \langle 63,5; 2 \rangle, \langle 64; 2 \rangle, \langle 64,5; 2 \rangle, \langle 65; 2 \rangle, \langle 65,5; 2 \rangle, \langle 65,5; 2 \rangle, \langle 66; 2 \rangle, \langle 66,5; 2 \rangle, \langle 67; 2 \rangle, \langle 67,5; 2 \rangle \};$$

$$W_3(p_3 = 1) = \{ \langle 68; 1 \rangle, \langle 68,5; 1 \rangle, \langle 69; 1 \rangle, \langle 69,5; 1 \rangle, \langle 70; 1 \rangle, \langle 70,5; 1 \rangle, \langle 71; 1 \rangle \}.$$

В таблице 4.15 приведены результаты разбиений по левому, правому и срединному штрихкоду.

Таблица 4.15 – Результаты разбиений по методу штрихкод (долгота)

$j$	$A_j$	тип $p_j$	$Z_j$
Левый штрихкод			
1	[59; 63)	1	1
2	[63; 68)	2	2
3	[68; 71,5]	1	1
Правый штрихкод			
1	(67,5; 71,5]	1	1
2	(62; 67,5]	2	2
3	[59; 62]	1	1
Срединный штрихкод			
1	[59; 62,5)	1	1
2	[62,5; 67,75)	2	2
3	[67,75; 71,5]	1	1

На рис. 4.24 показан график разбиения штрихкод, который одинаков (с точностью до нумерации) для всех типов разбиения  $A = \{<A, p>\}$ .

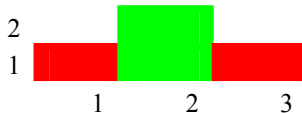


Рис. 4.24 – Разбиение штрихкод для долготы в задаче «Ураганы»

Примечание: легко убедиться, что разбиение по широте мало информативно для задачи классификации.

Построим конфигуратор для долготы, используя любое разбиение  $A$  из таблицы 4.15:

$$\begin{aligned}
 & \text{Долгота} \{ \\
 & 3 \{1 \wedge 1 \ 3; 2 \wedge 2\} \\
 & 2 \{1 \_ 1 \wedge 1 [59; q1]; 2 \wedge 2 [q1; q2]; 1 \_ 2 \wedge 3 [q2; 71,5]\} \\
 & 1 \{[59; 71,5]\}.
 \end{aligned}$$

Параметры  $q1$  и  $q2$  зависят от типа разбиения, например, для левого штрихкода:  $q1 = 63$ ,  $q2 = 68$ .

Легко убедиться, что конфигуратор теста «Долгота» позволяет однозначно классифицировать все 37 прецедентов из БД. Правило классификации следующее:

Если  $\text{Долгота}/3 \geq 1$ , то  $z = 1$  (*baro*), иначе  $z = 2$  (*trop*).

Поскольку метод штрихкода строит, как минимум, три различных разбиения (левое, правое, срединное), то в один конфигуратор можно

объединить все три разбиения. Каноническое представление объединенного конфигуриатора:

*Долгота (объединенная)* {3 #0 {7; 8; 9} 2 #0 {4; 5; 6} 1 {1; 2; 3} 0 {}}.

На рис. 4.25 показан структурно-завершенный орграф объединенного конфигуриатора теста «Долгота».

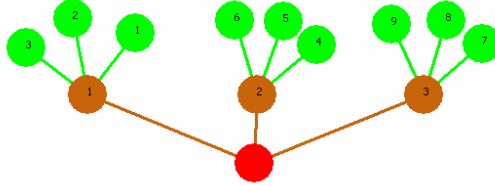


Рис. 4.25 – Орграф доменов объединенного орграфа теста «Долгота»

#### 4.3.3 Пример «Поиск зависимостей энергетических возможностей и стоимости носителей от их технических характеристик»

Основу современных транспортных космических систем (ТКС) составляют ракеты-носители (РН). Основными «потребительскими» характеристиками РН являются масса выводимого на орбиту полезного груза (энергетические возможности) и стоимость пуска. В работе [169] с помощью методов прикладного статистического анализа решается задача нахождения зависимостей между указанными "потребительскими" характеристиками и некоторыми техническими характеристиками РН.

Особенностью подобных задач является большое число влияющих и во многом зависимых факторов, что, безусловно, усложняет решение перечисленных задач статистическими методами. Это проявляется, в частности, в отсутствии точных априорных гипотез об исследуемых объектах, что характеризуется следующими видами структурной неопределенности [2, 169]: 1) по числу однородных групп наблюдений, объективно существующих в исходной выборке данных (базе прецедентов); 2) по количеству и составу входных переменных в регрессионных моделях зависимостей выходных переменных от входных переменных; 3) по степени статистической зависимости между аддитивными случайными составляющими в выходных переменных в системе регрессионных моделей; 4) по количеству и составу признаков в задачах статистической классификации. Первопричиной данных проблем в рамках статистического подхода является стремление найти одну финальную статическую закономерность, а не алгоритм перманентной самоорганизации модели.

Предлагаемый нами подход к решению данного класса задач принципиально отличается от статистического подхода. Главное отличие состоит в констатации факта, что разные ситуации действительности в базе прецедентов (в данном примере – пуски РН) могут иметь разные параметры порядка и, следовательно, разные определяющие фундаментальные закономерности, а не одну закономерность, как постулирует регрессионный анализ. Предлагаемое нами решение заключается в поиске всех конкурирующих наборов параметров порядка, определяющих тенденции и закономерности в рассматриваемой ПрО. Каждый набор параметров порядка – это один предельный формальный синдром. Все наборы образуют предельную синдромную модель знаний. Другое отличие состоит в перманентной самоорганизации моделей знаний в процессе поступления новых прецедентов (применяется *механизм фальсификации*, отсутствующий в статистическом подходе). Такой подход позволяет использовать все имеющиеся априорные данные (как правило, весьма ограниченные), а не делить их на обучающую и контрольную выборки.

Все что необходимо для построения предельных синдромных моделей знаний – это банк тестов и база прецедентов, в которой каждый прецедент отнесен к определенному классу из  $Z$ . Например, если нас интересуют закономерности, определяющие стоимость пусков, то весь интервал возможных цен разбивается на конечное число подинтервалов, а множество заключений  $Z$  представляет собой список номеров интервалов. Если необходимо повысить точность искомой зависимости, то либо расширяется само  $Z$  (исходный интервал разбивается на большее число подинтервалов), либо указанная процедура (формирование  $Z$  и построение синдромных моделей знаний) повторяется для найденного подинтервала. Таким образом, получаем иерархическую процедуру определения требуемой закономерности заданного уровня точности. Правило построения  $Z$  распространяется и на массу полезного груза, а также всевозможные комбинации характеристик.

Приведем пример описания базы прецедентов  $\Omega$  и банка тестов для рассматриваемой ПрО, а заодно продемонстрируем полезность метода штрихкода. Множество потребительских характеристик обозначим следующим образом:

Y1 – масса полезного груза, выводимого на низкую круговую орбиту (кг);

Y2 – стоимость пуска (заявленная, млн. долл.).

Пусть множество факторов образовано такими характеристиками [169]: X1 – время с даты пуска до фиксированного момента (в годах); X2 – стартовая масса РН (т); X3 – широта местности расположения космодрома (град.); X4

– стартовая тяга – тяга у земли первой ступени (кгс);  $X_5$  – удельный импульс у земли первой ступени (с);  $X_6$  – удельный импульс в пустоте верхней ступени (с). В скобках приведены единицы измерения в рамках базовых доменов.

Примечание: информация о РН мирового парка получена из базы данных, разработанной и пополняемой по открытым источникам в Институте технической механики НАНУ и НКАУ.

Ясно, что базовые домены всех характеристик – тестов являются непрерывными или конструктивно-непрерывными как тест «Время с даты пуска до фиксированного момента». Пример описания множества прецедентов на уровне базовых доменов приведен в таблице 4.16.

Таблица 4.16 – Фрагмент базы прецедентов

Входные характеристики – тесты						Выходные тесты*		
X1	X2	X3	X4	X5	X6	Y1	Y2	Z
18	420	5,23	67800	248,5	444,2	8250	110	3
15	358	5,23	67800	248,5	444,2	6800	100	3
4	780	5,23	122700	392	446	22000	155	6
10	749,1	5,23	900000	391	321	19450	155	6
...	...	...	...	...	...	...	...	...

\* Z зависит от Y1 и/или Y2.

Качество найденных закономерностей (синдромов), т.е. степень их истинности (инвариантности) целиком определяется конфигураторами тестов и базой прецедентов. Причем наиболее критичным является переход от базового – непрерывного домена к первому дискретному домену. Рассматривалось несколько вариантов разбиения базового домена на подинтервалы.

В первом варианте интервал [Min; Max] базового домена теста  $X_i$  разбивался на  $N_i$  одинаковых подинтервалов. Во втором варианте тот же интервал разбивался на  $N_i$  подинтервалов таким образом, чтобы в каждый подинтервал попало примерное одинаковое число прецедентов. В третьем варианте эксперт «вручную» разбивал каждый базовый домен (число подинтервалов для каждого теста получалось разным).

В таблице 4.17 показан пример автоматического построения и описания конфигулятора теста X1 по первому варианту с помощью электронной таблицы. Элементы на 3-м и 4-м уровне (доменах) объединялись таким образом, чтобы в них было примерно равное число элементов. Видно, что домен «4» является одним из листьев домена «3».

Таблица 4.17 – Тест X1: вариант 1

ТЕСТ X1: разбиение «одинаковые интервалы»						
домен 1		домен 2			домен 3	домен 4
от	до	от	до	значение	значение	значение
0	46	0	9,2	1	1	1
		9,2	18,4	2	2	2
		18,4	27,6	3	3	2
		27,6	36,8	4	3	2
		36,8	46	5	3	2

В таблице 4.18 показан пример автоматического построения и описания конфигуратора теста X1 по второму варианту с помощью электронной таблицы. Число интервалов разбиения выбрано четным, чтобы в дальнейшем можно было попарно объединить соседние интервалы. Объединение продолжается до тех пор, пока не останется два или три интервала.

Таблица 4.18 – Тест X1: вариант 2

ТЕСТ X1: разбиение «одинаковое число точек»					
домен 1		домен 2			домен 3
от	до	от	до	значение	значение
0	46	0	4	1	1
		4	7	2	1
		7	11	3	3
		11	15	4	3
		15	19	5	5
		19	46	6	5

Если объединить все три разбиения в один конфигуратор, то для всех тестов получается схожая структура орграфа доменов. Ниже приведен канонический орграф теста X1:

X1 {6 {12 13; 14 15 16} 5 #0 {12; 13; 14; 15; 16} 4 {6 7; 8 9; 10 11} 3 #0 {6; 7; 8; 9; 10; 11} 2 {1; 2; 3 4 5} 1 {1; 2; 3; 4; 5} 0 {}.

Листья любого орграфа в описании, как правило, не входят, так как они порождаются автоматически в момент использования или отображения орграфа. Исключения составляют те случаи, когда пользователь не запускает функцию автоматического порождения листьев, но при этом хочет включить в анализ те или иные листья. Примером такого исключения служит домен «4» в первом варианте разбиения теста X1 (таблица 4.17).

Пример предельного структурно-завершенного орграфа доменов для X1 показан на рис. 4.26. Каждая «ветка» орграфа, выходящая из базовой

вершины, определяет один из трех первоначальных типов разбиения (левая «ветка» – первый вариант разбиения, средняя – второй вариант разбиения, правая – третий вариант разбиения).

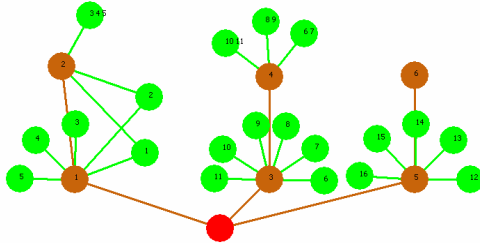


Рис. 4.26 – Орграф доменов  $G^{++}$ , объединяющий три разбиения

В таблице 4.19 показан пример автоматического построения и описания конфигуратора теста X3 по методу штрихкода (вариант 4) с помощью электронной таблицы. Принцип обобщения такой же, как и для теста X1.

Таблица 4.19 – Тест X3: метод штрихкода (левый)

Тест X3: разбиение по методу штрихкода								
домен 1		домен 2			дом. 3	дом. 4	Класс	ТИП
от	до	от	до	значение	знач.	знач.	A	Z
0,01	63,25	0,01	2,3	1	1	1	1	4
		2,3	5,23	2	2	2	2	1
		5,23	9,0	3	3	3	3	2;3;4;6
		9,0	13,78	4	2	2	2	1
		13,78	28,2	5	4	1	4	1;2;3
		28,2	28,45	6	5	3	5	2;3;4
		28,45	30,4	7	6	3	6	2;3;4;5;6
		30,4	31,25	8	7	1	7	1;2;4
		31,25	34,62	9	2	2	2	1
		34,62	37,83	10	8	2	8	1;2;5
		37,83	38,5	11	2	2	2	1
		38,5	45,98	12	9	3	9	2
		45,98	63,25	13	10	1	10	1;2;3;4;6
		63,25	63,25	14	11	2	11	1;2
Итого элементов:				14	11	3		

Необходимость разбиения по методу штрихкода базовых доменов всех параметров определялась тем, что предыдущие три варианта разбиения оказались неудовлетворительными: любые модели знаний содержали синдромы, которые опирались на базовые домены. На рис. 4.27 показан

график разбиения штрихкод для теста ХЗ.

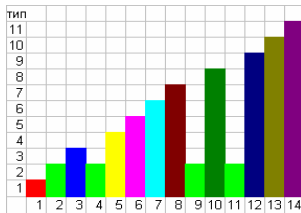


Рис. 4.27 – Разбиение штрихкод для теста ХЗ

Пример оптимального по энтропии структурно – заверщенного орграфа доменов для теста ХЗ показан на рис. 4.28. Каждая «ветка» орграфа, выходящая из базовой вершины, определяет один из четырех типов разбиения.

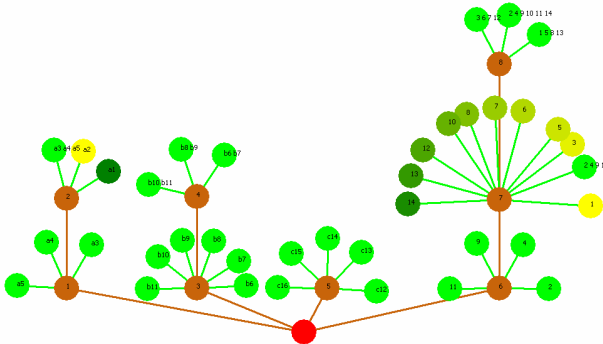


Рис. 4.28 – Орграф доменов  $G^+$ , объединяющий четыре разбиения

Следует отметить принципиальное положение о том, что разбиение штрихкод гарантирует отсутствие артефактов на уровне первых дискретных доменов, что означает получение удовлетворительных синдромных моделей знаний.

#### 4.4 Реализация базы прецедентов и конфигураторов с помощью электронной таблицы

Ставится задача найти относительно простой способ реализации базы прецедентов и конфигураторов тестов. Способ должен позволить непрограммистам самостоятельно формировать синдромные модели знаний с использованием метода предельных обобщений.

Рассмотрим пример. В спортивном клубе ведется база данных



спортсменов спринтеров с текущей оценкой их перспективности. Анализ перспективности спортсменов осуществляется экспертами на основании двух тестов «Индекс массы тела (ИМТ)» и «Возраст». Пусть возможны три заключения о перспективности, т.е.  $Z = \{1; 2; 3\}$ . Пример базы прецедентов в формате электронной таблицы приведен в таблице 4.20.

Таблица 4.20 – БД «Спортсмены спринтеры»

Возраст	ИМТ	Z
16	23	1
17	24	2
18	25	1
19	26	3
20	27	2
21	28	1
22	23	2
30	25	2
32	26	3

Для выявления предпочтений экспертов и последующего создания интеллектуальных приложений необходимо сформировать синдромные модели знаний. Для построения синдромных моделей знаний с помощью метода предельных обобщений необходимо задать конфигураторы всех тестов. Зададим конфигураторы тестов «Индекс массы тела» и «Возраст» следующим образом:

*Индекс массы тела* ^ИМТ {  
 3 {удовлетворительный ^b c d ; неудовлетворительный ^a e}  
 2 { значительно выше нормы ^a [27; 28];  
     выше нормы ^b [26; 27];  
     норма ^c [25; 26];  
     ниже нормы ^d [24; 25];  
     значительно ниже нормы ^e [23; 24]}  
 1 { [23; 28] } }

*Возраст* {  
 3 {Подходящий ^b c; Предельный ^a}  
 2 {старый атлет ^a [29; 32];  
     атлет средних лет ^b [23; 28];  
     молодой атлет ^c [16; 22]}  
 1 { [16; 32] } }.

Листья в орграфе доменов теста  $G(z)$  обозначим следующим образом:

$Z_1 = \{1; -1\} = \{1, 4\}$ ;  $Z_2 = \{2; -2\} = \{2, 5\}$ ;  $Z_3 = \{3; -3\} = \{3, 6\}$ .

С помощью конфигураторов тестов и средств электронной таблицы EXCEL исходная база прецедентов преобразуется в расширенную таблицу, включающую все домены всех тестов. Пример расширенной базы прецедентов «Спортсмены спринтеры» приведен в таблице 4.21.

Таблица 4.21 – Расширенная БД «Спортсмены спринтеры»

Возраст			ИМТ			Заключение			
1	2	3	1	2	3	Z	Z1	Z2	Z3
16	c	b	23	e	a	1	1	5	6
17	c	b	24	d	b	2	4	2	6
18	c	b	25	c	b	1	1	5	6
19	c	b	26	b	b	3	4	5	3
25	b	b	24	d	b	1	1	5	6
26	b	b	27	a	a	3	4	5	3
22	c	b	23	e	a	2	4	2	6
30	a	a	25	c	b	2	4	2	6
32	a	a	26	b	b	3	4	5	3

Данных расширенной таблицы достаточно для реализации метода предельных обобщений, в частности, нахождения критических описаний, синдромных и вероятностных моделей знаний (см. главу 5).

С прикладной точки зрения важно отметить, что расширенную базу прецедентов достаточно легко реализовать с помощью встроенных функций EXCEL. Действительно, значения в рамках 2-го домена теста «Возраст» вычисляются, например, с помощью аналогов следующей формулы:

=ЕСЛИ(И(A2>15;A2<23);"c";ЕСЛИ(И(A2>22;A2<29);"b";ЕСЛИ(И(A2>28;A2<33);"a"))),

где в ячейке A2 хранится значение теста из 1-го домена.

Значения в рамках 3-го домена теста «Возраст» вычисляются с помощью аналогов следующей формулы:

=ЕСЛИ(A24="a";"a"; ЕСЛИ(A24="b";"b"; ЕСЛИ(A24="c";"b";))),

где в ячейке A24 хранится значение теста из 2-го домена. Аналогичным образом вычисляются значения других тестов, включая результат. Таким образом, с помощью встроенных функций реализуются конфигураторы.

## ГЛАВА 5. МЕТОД ПРЕДЕЛЬНЫХ ОБОБЩЕНИЙ

Фундаментальной логической активностью разума является гносеологический цикл движения от частного к общему, который состоит из следующих этапов [2, 10, 105]:

1. Возникают некоторые частные знания без общего.
2. Выдвигается гипотеза о природе того единства, которое лежит за частными знаниями.
3. Для обоснования указанной гипотезы, частные знания выводятся тем или иным способом из общего знания. Если это удастся сделать, то общее знание получает определенное обоснование.
4. Возможно, что из гипотезы общего выводятся дополнительные частные знания, которые ранее не были получены в опыте. Теперь они ищутся в опыте, и, в случае подтверждения, общее получает дополнительное подкрепление.
5. Рано или поздно расширение частных знаний превышает ресурсы первоначального общего, и возникает потребность в гипотезе более глубокого общего, которое смогло бы охватить расширенное многообразие частного и породить новое подтверждаемое в будущем многообразие частного. Далее цикл повторяется.

Метод предельных обобщений (МПО) позволяет сформировать гипотезы относительно общего знания, описывающего частные ситуации действительности с известным исходом. Какие-то из гипотез могут оказаться корректными, т.е. охватывать все возможные в данной ПрО ситуации действительности. Особенностью выдвигаемых гипотез (моделей знаний) является то, что все они обладают свойством предельной фальсифицируемости на заданном множестве ситуаций действительности в рамках заданной онтологии (Банка тестов).

### **5.1 Вероятностные закономерности, формальные синдромы, синдромные и вероятностные модели знаний**

Дано: банк тестов  $\{G(\tau)\}$  и множество ситуаций действительности (прецедентов) с известными исходами  $\Omega = \{\alpha(\{x/T\}, \underline{z}/Z)\}$ , где  $Z = \{1, \dots, N\}$ . В качестве исходов  $Z$  могут выступать диагнозы, прогнозы, варианты управления. Без потери общности будем считать, что каждый тест входит в описание ситуации (прецедента) один раз. Множество заключений или исходов  $Z$  является полным и альтернативным. Относительно количества и качества прецедентов потребуем минимальной представительности

выборки по каждому заключению  $z/Z$ , а также наличия всех необходимых данных (рассматриваются прецеденты с полной информацией).

Обозначим базовые домены всех тестов через  $T_0$ . Соответственно, первичное описание базы прецедентов будет иметь вид  $\{\tau T_0\} \equiv \{\tau T\}_0$ . Для фиксации конкретного описания (наброска) базы прецедентов будем использовать нотацию  $\Omega(\{\tau T\})$  или  $\Omega(\{\tau T\}, Z)$ , например,  $\Omega(\{\tau T_0\})$ . Данная нотация означает, что все прецеденты в базе прецедентов описаны на одном и том же фиксированном уровне общности  $\{\tau T\}$ . Для фиксации конкретного описания (наброска) произвольной ситуации/прецедента будем использовать нотацию  $\alpha(\{\tau T\})$  или  $\alpha(\{\tau T\}, z/Z)$ . В общем случае  $\Omega$  может содержать прецеденты разного уровня общности описания. Термины «прецедент», «ситуация» и «ситуация действительности» считаются синонимами.

Будем говорить, что база прецедентов  $\Omega$  не содержит *конфликтов* на уровне общности  $\{\tau T\}$  (в рамках описания  $\{\tau T\}$ ), если нет двух ситуаций с разными исходами (заклечениями), но совпадающими значениями тестов. Любые две ситуации, у которых совпадают значения тестов, но разные исходы назовем *артефактом первого рода* (genus). Любые  $M$  ситуаций, у которых совпадают значения тестов, но все исходы разные назовем *артефактом (M-1)-рода*. Предполагается, что первоначально все прецеденты  $\alpha$  описаны с использованием максимально точных доменов (базовых доменов) всех тестов и на этом первичном уровне база прецедентов не содержит конфликтов. Другими словами, база прецедентов  $\Omega(\{\tau T_0\})$  конфликтов не содержит.

*Вероятностной закономерностью* появления заключения  $z/Z$  назовем правило вида

$$R = (\{\tau T\} \rightarrow J_z z/Z), \quad p(R) \geq p^*, \quad \nu_R = \nu(R), \quad (5.1)$$

где  $\{\tau T\}$  – избыточная совокупность значений тестов;  $J_z$  и  $p(R)$  – ранг или вероятность получения заключения  $z/Z$  при условии  $\{\tau T\}$ ;  $p^*$  – порог (например, 0.9);  $\nu_R$  – вес правила, пропорциональный количеству прецедентов с заключением  $z/Z$ , отвечающих правилу  $R$ . Вероятностную закономерность нельзя редуцировать с сохранением ранга. Сослаться на конкретную закономерность можно следующим образом:  $R(\{\tau T\}, z/Z)$ . Для веса  $\nu_R$ , как и для любого теста, существует оргграф  $G(\nu_R)$ .

Синонимами вероятностной закономерности можно считать *ассоциативное правило* (обобщение импликаций, когда допускается некоторая доля исключений) и *предвестник*.

Под *формальным синдромом* будем понимать избыточную совокупность значений тестов, позволяющую однозначно установить заключение  $z/Z$  (в такой трактовке формальные синдромы отличаются от

синдромов в медицине). Ранее мы договорились формальный синдром называть просто синдромом. Произвольный синдром представим в виде:

$$S = (\{\underline{t}/T\} \rightarrow \underline{z}/Z), \quad \nu_S = \nu(S), \quad (5.2)$$

где  $\{\underline{t}/T\}$  – избыточная совокупность значений тестов;  $\underline{z}/Z$  – заключение;  $\nu_S$  – вес синдрома пропорциональный количеству прецедентов, для которых выполняется данный синдром. Сослаться на конкретный синдром можно следующим образом:  $S(\{\underline{t}/T\}, \underline{z}/Z)$ . Любой тест в описание синдрома входит только один раз. Для веса  $\nu_S$  существует орграф  $G(\nu_S)$ .

Легко видеть, что формальный синдром – это вероятностная закономерность ранга 1, так как заключение определяется однозначно. Синдром нельзя усилить, т.е. повысить ранг и нельзя редуцировать. Следует отметить, что на основе формального синдрома может быть сформировано продукционное правило вида: *Если  $\{\underline{t}/T\}$ , то  $\underline{z}/Z$* . Однако формальный синдром обладает свойствами, которых нет у продукционного правила, в частности, *обобщением* и *конкретизацией*.

*Схемой вероятностной закономерности* назовем список тестов, входящих в закономерность, и заключение:  $R(\{\tau\}, \underline{z}/Z)$ . *Детальной схемой вероятностной закономерности* назовем список тестов с указанием доменов и заключение:  $R(\{\tau/T\}, \underline{z}/Z)$ . Аналогично, *схемой формального синдрома* назовем список тестов, входящих в синдром, и заключение:  $S(\{\tau\}, \underline{z}/Z)$ . *Детальной схемой формального синдрома* назовем список тестов с указанием доменов и заключение:  $S(\{\tau/T\}, \underline{z}/Z)$ . Таким образом, у синдрома есть, как минимум, три уровня детальности описания:  $S(\{\underline{t}/T\}, \underline{z}/Z)$ ,  $S(\{\tau/T\}, \underline{z}/Z)$  и  $S(\{\tau\}, \underline{z}/Z)$ .

Пусть имеется детальная схема синдрома  $S(\{\tau/T\}_S, \underline{z}/Z)$ . *Дополнением* к  $\{\tau/T\}_S$  назовем произвольное множество  $\{\tau/T\}_S^\perp$ , которое в совокупности с  $\{\tau/T\}_S$  образует полное описание базы прецедентов  $\{\tau/T\}$ . Ясно, что

$$\{\tau\}_S \cap \{\tau\}_S^\perp = \emptyset; \quad \{\tau\}_S \cup \{\tau\}_S^\perp = \{\tau\}.$$

**Предложение 5.1.** Синдром  $S(\{\underline{t}/T\}_S, \underline{z}/Z)$  является синдромом во всех описаниях базы прецедентов  $\Omega(\{\tau/T\}_S \cup \{\tau/T\}_S^\perp)$ .

**Предложение 5.2.** Если существует синдром  $S(\{\underline{t}/T\}_S, \underline{z}/Z)$ , то для любого  $\{\underline{t}/T\}'$  такого, что  $\{\underline{t}/T\}' \rightarrow \{\underline{t}/T\}_S$  верно:  $\{\underline{t}/T\}' \rightarrow \underline{z}/Z$ . При этом синдрома  $S(\{\underline{t}/T\}', \underline{z}/Z)$  может не существовать.

Суть предложения 5.2 заключается в следующем: понижение уровня общности значений тестов, входящих в синдром, также гарантирует однозначную классификацию с тем же результатом. Поскольку выполнимость операции  $\{\underline{t}/T\}' \rightarrow \{\underline{t}/T\}_S$  проверяется исключительно на основе банка тестов  $\{G(\tau)\}$ , то какой-либо совокупности данных  $\{\underline{t}/T\}'$

может не оказаться в базе прецедентов  $\Omega(\{\tau/T\}' \cup (\{\tau/T\} \setminus \{\tau/T\}'))$ , но даже если такая совокупность данных есть, то она не обязательно является синдромом.

*Предельной вероятностной закономерностью* назовем вероятностную закономерность, которую нельзя обобщить ни по одному тесту с сохранением ранга. Предельную вероятностную закономерность ранга 1 назовем *предельным синдромом*. Предельный синдром является предельным в трех смыслах: его нельзя усилить, т.е. повысить ранг; его нельзя редуцировать и его нельзя обобщить ни по одному входящему тесту. Другими словами, любой предельный синдром претендует на роль *инварианта* контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ . Предельную вероятностную закономерность обозначим  $R^*$ , предельный синдром –  $S^*$ .

С каждой вероятностной закономерностью  $R$  связано «облако» предельных вероятностных закономерностей  $\{R^*\}$ , которое получается путем всех допустимых обобщений в рамках банка тестов  $\{G(\tau)\}$ . Ранги всех закономерностей из  $\{R^*\}$  одинаковы и равняются  $p(R)$ . С каждым синдромом  $S$  связано, соответственно, «облако» предельных синдромов  $\{S^*\}$ , которое также получается путем всех допустимых обобщений. «Облака» для каждого синдрома и каждой вероятностной закономерности определяются единственным образом на  $\langle \Omega(Z), \{G(\tau)\} \rangle$ . Можно записать:

$$\forall R, r/\mu_r: R \rightarrow \{R^*\}_R; \quad \forall S, s/\mu_s: S \rightarrow \{S^*\}_S, \quad (5.3)$$

где  $r, s$  – отображения, а  $\mu_r, \mu_s$  – механизмы реализации отображений. Очевидно, справедливо следующее:

$$\forall R^*, r/\mu_r: R^* \rightarrow R^*; \quad \forall S^*, s/\mu_s: S^* \rightarrow S^*. \quad (5.4)$$

Отображения  $r, s$  существенно опираются на контекст  $\langle \Omega(Z), \{G(\tau)\} \rangle$ .

Множество  $\{R^*\}_R$  будем называть множеством *сопряженных предельных закономерностей* для  $R$ . Множество  $\{S^*\}_S$  будем называть множеством *сопряженных предельных синдромов* для  $S$ .

Фактически можно говорить о том, что для любой вероятностной закономерности  $R$  и любого синдрома  $S$  строятся орграфы набросков  $G_S(R)$  и  $G_S(S)$  соответственно (в качестве множества независимых операций выступают операции обобщения в рамках орграфов доменов тестов). В таком случае сопряженные множества  $\{R^*\}_R$  и  $\{S^*\}_S$  являются ничем иным, как экстремальными пограничными слоями набросков (см. главы 3 и 6).

Полезность введения информационных объектов «предельная вероятностная закономерность» и «предельный синдром» заключается в том, что в силу обобщения сопряженные объекты охватывают более широкое множество априорных данных, чем первичные объекты.

**Предложение 5.3.** Для любой вероятностной закономерности  $R$

существует единственное непустое множество сопряженных предельных вероятностных закономерностей  $\{R^*\}_R$ . Соответственно, для любого синдрома  $S$  существует единственное непустое множество сопряженных предельных синдромов  $\{S^*\}_S$ .

Опишем одну из возможных процедур построения множества сопряженных предельных синдромов для произвольного синдрома  $S$ , т.е. реализации отображения  $s/\mu_s$ . Множество сопряженных предельных вероятностных закономерностей строится аналогичным образом.

Если в рамках орграфа  $G(\tau)$  домен  $T'$  имеет более высокий уровень общности, чем домен  $T$ , т.е. элементы домена  $T'$  выводятся однозначно из элементов домена  $T$ , то будем говорить, что домен  $T'$  *строго доминирует* домен  $T$  и использовать для этого нотацию:  $T' > T$ . В случае *нестромого доминирования* (домены могут совпадать) будем писать:  $T' \geq T$ . Если между двумя доменами  $T$  и  $T'$  не выполняется отношение доминирования, то данный факт будем отражать нотацией  $T' \ngtr T$ .

Для любого теста  $\tau$  и любого домена  $T \in G(\tau)$  определено множество доминирующих вершин (в нестрогом смысле)

$$\{T^*\}_{\tau, T} = \{T' \mid T' \in G(\tau) \ \& \ T' \geq T\}. \quad (5.5)$$

Пример. Пусть  $G(\tau) = \{1 \rightarrow 2 \rightarrow 3 \rightarrow 4, 2 \rightarrow 5 \rightarrow 6\}$ , тогда

$$\{T^*\}_{\tau, 2} = \{2, 3, 4, 5, 6\}; \quad \{T^*\}_{\tau, 3} = \{3, 4\}; \quad \{T^*\}_{\tau, 6} = \{6\}.$$

Если задан синдром  $S(\{\underline{\tau}/T\}, \underline{z}/Z)$ , то каждый тест  $\tau/T$ , входящий в описание синдрома, порождает множество доминирующих вершин  $\{T^*\}_{\tau, T}$ . Произвольным образом пронумеруем все тесты, входящие в описание синдрома. Получим множество  $\{(\tau/T)_1, \dots, (\tau/T)_n\}_S$ . Нумерация тестов позволяет ввести нумерацию множеств доминирующих вершин:  $\{\{T^*\}_1, \dots, \{T^*\}_n\}$ . Определим декартово произведение всех множеств доминирующих вершин для синдрома  $S$

$$\{T^*\}_1 \times \{T^*\}_2 \times \dots \times \{T^*\}_n = \{\theta = \langle \dots, \rangle\}_S = \{\theta\}_S. \quad (5.6)$$

Таким образом, любой кортеж  $\theta$  представляет собой упорядоченный список доменов для всех тестов, входящих в описание синдрома  $S$ . Исходный список доменов синдрома  $S$  обозначим  $\theta_0$ . Кортеж  $\theta$  однозначно определяет уровень общности  $\{\tau/T_\theta\}$ , при этом  $\{\tau/T_\theta\}$  однозначно выводится из  $\{\underline{\tau}/T\}$ .

Следует отметить, что  $\{\theta\}_S$  является, по сути, реляционным отношением и легко может быть синтезировано с помощью реляционных баз данных.

В рамках контекста  $\langle \Omega(\{\tau/T_0\}, Z), \{G(\tau)\} \rangle$  определим функцию *Sindrom* ( $\{\underline{\tau}/T\}, \underline{z}/Z$ ) следующим образом:

$$Sindrom(\{\underline{t}/T\}, \underline{z}/Z) = true \Leftrightarrow \exists S(\{\underline{t}/T\}, \underline{z}/Z).$$

Для существования синдрома  $S$  необходимо и достаточно выполнение двух требований:

(i) не существует двух прецедентов  $\alpha$  и  $\beta$  таких, что  $z_\alpha = \underline{z}/Z$ ,  $\{\underline{t}/T\}_\alpha = \{\underline{t}/T\}_\beta$ , но  $z_\alpha \neq z_\beta$ ;

(ii) набор данных  $\{\underline{t}/T\}$  избыточен, т.е. у него нет подмножества, удовлетворяющего требованию (i).

На множестве кортежей доменов  $\{\theta\}_S$  с помощью функции  $Sindrom$  можно определить все синдромы, а именно:

$$\{S_\theta\}_S = \{S(\{\underline{t}/T_\theta\}, \underline{z}/Z) \mid Sindrom(\{\underline{t}/T_\theta\}, \underline{z}/Z) = true \ \& \ \theta \in \{\theta\}_S\}. \quad (5.7)$$

Множество  $\{S_\theta\}_S$  не пусто, так как  $S \in \{S_\theta\}_S$ . Функция  $Sindrom$  опирается на однозначное преобразование:  $\forall \theta, \{\underline{t}/T\}_S \rightarrow \{\underline{t}/T_\theta\}$ .

Будем говорить, что описание  $\{\tau/T\}'$  по уровню общности доминирует описание  $\{\tau/T\}$ , если для любого теста  $\tau$  элементы домена  $T'$  выводятся из элементов домена  $T$  в рамках орграфа  $G(\tau)$ . Для отношения *нестрогого доминирования* (домены  $T$  и  $T'$  могут совпадать, кроме одного) будем использовать нотацию  $\{\tau/T\}' \geq \{\tau/T\}$ , а для отношения *строгого доминирования* (все домены разные) – нотацию  $\{\tau/T\}' > \{\tau/T\}$ .

Если описание  $\{\tau/T\}'$  по уровню общности доминирует описание  $\{\tau/T\}$ , то это означает, что для любых данных  $\{\underline{t}/T\}$  существует однозначное преобразование:  $\{\underline{t}/T\} \rightarrow \{\underline{t}/T'\}$ .

Если между двумя описаниями  $\{\tau/T\}$  и  $\{\tau/T'\}$  не выполняется отношение доминирования и они не совпадают, то данный факт будем отражать нотацией  $\{\tau/T'\} \gg \{\tau/T\}$ .

Отношение «доминирование» будем применять к любым одинаковым наборам тестов, причем каждый тест входит в набор только один раз.

Будем говорить, что синдром  $S(\{\underline{t}/T'\}, \underline{z}/Z)$  доминирует синдром  $S(\{\underline{t}/T\}, \underline{z}/Z)$ , если  $\{\tau\}' = \{\tau\}$ ,  $\{\tau/T'\}' \geq \{\tau/T\}$  и  $\{\underline{t}/T\} \rightarrow \{\underline{t}/T'\}$ . Для отношения доминирования между синдромами будем использовать нотацию  $S' \geq S$ . Если отсутствует обратное отображение  $\{\underline{t}/T'\}' \rightarrow \{\underline{t}/T\}$ , то будем говорить о строгом доминировании и использовать нотацию  $S' > S$ .

Заметим, что обратное отображение  $\underline{t}/T' \rightarrow \underline{t}/T$  существует в единственном случае, когда  $T.a \rightarrow T'.a$ , т.е. в схеме обобщения элементов «1 в 1». Строгое доминирование между наборами тестов ( $\{\tau/T'\}' > \{\tau/T\}$ ) не означает автоматически строгое доминирование между синдромами ( $S' > S$ ).

Результирующее множество предельных синдромов для синдрома  $S$  определяется следующим образом:

$$\{S^*\}_S = \{S^* \mid (S^* \in \{S_\theta\}_S) \ \& \ (\neg \exists S \in \{S_\theta\}_S: S \geq S^*)\}. \quad (5.8)$$



Сведем в один алгоритм все этапы построения множества сопряженных предельных синдромов для синдрома  $S$  (под  $\Omega$  понимается полное множество всех описаний базы прецедентов, а не одно конкретное описание).

### Алгоритм 5.1

Вход: Синдром  $S(\{\underline{z}/T\}, \underline{z}/Z)$ , контекст  $K = \langle \Omega, \{G(\tau)\} \rangle$ .

Выход:  $\{S^*\}_S$ .

$LimS(S, K)$

**begin**

**for**  $\tau \in \{\tau\}_S$  **do begin**

$\{T^*\}_{\tau T} = \emptyset$ ;

**for**  $T' \in G(\tau)$  **do if**  $T' \geq T$  **then**  $\{T^*\}_{\tau T'} := \{T^*\}_{\tau T} \cup T'$  **end**;

$n := |\{\tau\}_S|$ ;

$\{\theta\}_S := \{T^*\}_1 \times \{T^*\}_2 \times \dots \times \{T^*\}_n$ ;

$\{S_\theta\}_S := S$ ;

**for**  $\theta \in \{\theta\}_S$  **do if**  $Sindrom(\{\underline{z}/T_\theta\}, \underline{z}/Z) = true$

**then**  $\{S_\theta\}_S := \{S_\theta\}_S \cup S(\{\underline{z}/T_\theta\}, \underline{z}/Z)$ ;

$\{S^*\}_S := \emptyset$ ;

**for**  $S' \in \{S_\theta\}_S$  **do begin**  $flag := 1$ ;

**for**  $(S'' \in \{S_\theta\}_S) \ \& \ (S'' \neq S')$  **do if**  $S'' \geq S'$  **then**  $flag := 0$ ;

**if**  $flag = 1$  **then**  $\{S^*\}_S := \{S^*\}_S \cup S'$  **end**;

**return**  $(\{S^*\}_S)$

**end.**

Алгоритм 5.1 очевидно избыточен, но, тем не менее, он решает поставленную задачу.

**Предложение 5.4.** Между любыми двумя предельными синдромами не существует отношения доминирования.

Предельные синдромы, претендующие на статус инвариантов (с большими весами и высоким уровнем обобщения), можно рассматривать в качестве параметров порядка развития сложных ситуаций. Действительно, смена одного предельного синдрома другим может привести к качественному изменению поведения системы или ситуации, которое вызвано сменой заключения  $\underline{z}/Z$ . Что касается синдромов в целом, то их смена также может привести к качественному изменению развития ситуации (поведения системы). Можно считать, что такие синдромы формируют предпараметры порядка. Наличие разных синдромов предопределяет конкуренцию параметров порядка [180].

Факт возможности смены качественного поведения ситуации/системы путем целенаправленного изменения текущих синдромов лежит в основе

синдромного принципа управления (глава 7).

Для иллюстрации понятий и положений данного раздела и главы в целом введем в рассмотрение объекты примера «Диагностика». Речь идет о диагностике двух состояний ( $z = 1$ ) и ( $z = 2$ ) по двум тестам  $\tau_1$  – Температура и  $\tau_2$  – Возраст. Конфигураторы тестов приведены ниже.

$\tau_1$  – Температура {

T3 {Нормальная ^Н; Ненормальная ^ПН ПВ В};

T2 {Пониженная ^ПН [35, 36.5]; Нормальная ^Н [36.5, 36.9]; Повышенная ^ПВ (36.9, 37.5]; Высокая ^В (37.5, 40)};

T1 {[35; 40]}.

G(Температура) = {T1 → T2 → T3}.

$\tau_2$  – Возраст {

B3 {Допенсионный ^М СЛ; Пенсионный ^П СТ};

B2 {Молодой ^М [0; 33]; Средних лет ^СЛ [34; 59]; Пожилой ^П [60; 75]; Старческий ^СТ [76; 100]};

B1 {[0; 100]}.

G(Возраст) = {B1 → B2 → B3}.

Исходная база прецедентов  $\Omega(\{\tau/T_0\})$  представлена в описании «T1 – B1».

База содержит четыре прецедента:  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ .

Описание «T1 – B1»			
$\alpha$	$\tau_1/T1$	$\tau_2/B1$	$z$
1	36.0	12	1
2	36.7	50	
3	37.2	32	2
4	36.4	87	

С помощью конфигураторов тестов порождаются еще 8 описаний  $\Omega(\{\tau/T\})$ , которые приведены ниже (номера прецедентов опущены, но они сохраняются). Всего в примере 9 описаний, включая  $\Omega(\{\tau/T_0\}) \equiv \Omega(\{\tau_1/T1, \tau_2/B1\})$ .

Таблица 5.1 – Все описания базы прецедентов (кроме базового)

Описание «T1 – B2»		
$\tau_1/T1$	$\tau_2/B2$	$z$
36.0	Молодой	1
36.7	Сред. лет	
37.2	Молодой	2
36.4	Старческ.	

Описание «T2 – B1»		
$\tau_1/T2$	$\tau_2/B1$	$z$
Пониженная	12	1
Нормальная	50	
Повышенная	32	2
Пониженная	87	

Описание «Т1 – В3»		
$\tau_1/T1$	$\tau_2/B3$	$z$
36.0	Допенс.	1
36.7	Допенс.	
37.2	Допенс.	2
36.4	Пенсион.	

Описание «Т3 – В1»		
$\tau_1/T3$	$\tau_2/B1$	$z$
Ненорм.	12	1
Нормальная	50	
Ненорм.	32	2
Ненорм.	87	

Описание «Т2 – В2»		
$\tau_1/T2$	$\tau_2/B2$	$z$
Пониженная	Молодой	1
Нормальная	Сред. лет	
Повышенная	Молодой	2
Пониженная	Старческ.	

Описание «Т3 – В2»		
$\tau_1/T3$	$\tau_2/B2$	$z$
Ненорм.	Молодой	1
Нормальная	Сред. лет	
Ненорм.	Молодой	2
Ненорм.	Старческ.	

Описание «Т2 – В3»		
$\tau_1/T2$	$\tau_2/B3$	$z$
Пониженная	Допенс.	1
Нормальная	Допенс.	
Повышенная	Допенс.	2
Пониженная	Пенсион.	

Описание «Т3 – В3»		
$\tau_1/T3$	$\tau_2/B3$	$z$
Ненорм.	Допенс.	1
Нормальная	Допенс.	
Ненорм.	Допенс.	2
Ненорм.	Пенсион.	

Приведем примеры множеств  $\{S^*\}_S$ .

$$S = (\tau_1/T1? 37.2 \rightarrow (z=2)); \quad \{S^*\}_S = \{S^* = (\tau_1/T2? \text{Повышенная} \rightarrow (z=2))\}.$$

$$S = (\tau_1/T1? 36.7 \rightarrow (z=1)); \quad \{S^*\}_S = \{S^* = (\tau_1/T3? \text{Нормальная} \rightarrow (z=1))\}.$$

$$S = (\tau_2/B1? 87 \rightarrow (z=2)); \quad \{S^*\}_S = \{S^* = (\tau_2/B3? \text{Пенсионный} \rightarrow (z=2))\}.$$

$$S = (\tau_2/B1? 12 \rightarrow (z=1)); \quad \{S^*\}_S = \{S^* = (\tau_2/B1? 12 \rightarrow (z=1))\}, \text{ т.е. } S = S^*.$$

$$S = (\tau_2/B2? \text{Сред. лет} \rightarrow (z=1));$$

$$\{S^*\}_S = \{S^* = (\tau_2/B2? \text{Сред. лет} \rightarrow (z=1))\}, \text{ т.е. } S = S^*.$$

$$S = (\tau_1/T2? \text{Пониженная}, \tau_2/B2? \text{Молодой} \rightarrow (z=1));$$

$$\{S^*\}_S = \{S^* = (\tau_1/T2? \text{Пониженная}, \tau_2/B3? \text{Допенсионный} \rightarrow (z=1))\}.$$

Совокупность синдромов образует *синдромную модель знаний*, если она позволяет определить заключение, как минимум, для любого прецедента из  $\Omega(\{\tau/T_0\})$ . Модель знаний является *полной* в рамках фиксированного описания базы прецедентов  $\Omega(\{\tau/T\})$ , если она содержит все возможные синдромы в рамках данного описания. Для полной синдромной модели знаний будем использовать нотацию:  $\{S\}_{Full}$  на  $\Omega(\{\tau/T\})$  или  $\{S\}_{Full, \{\tau/T\}}$ .

Для оценки применимости той или иной модели знаний к прецедентам

из  $\Omega(\{\tau/T_0\})$  активно используется механизм построения замыканий множества априорных данных  $\{\underline{\tau}/T\}^+$  (см. главу 2).

Пример  $\{S\}_{Full}$  на  $\Omega(\{\tau_1/T2, \tau_2/B3\})$ :

$S_1 = (\tau_1/T2? \text{Пониженная}, \tau_2/B3? \text{Допенсионный} \rightarrow (z=1));$   
 $S_2 = (\tau_1/T2? \text{Нормальная} \rightarrow (z=1)); \quad S_3 = (\tau_1/T2? \text{Повышенная} \rightarrow (z=2));$   
 $S_4 = (\tau_2/B3? \text{Пенсионный} \rightarrow (z=2)).$

Пример  $\{S\}_{Full}$  на  $\Omega(\{\tau_1/T3, \tau_2/B1\})$ :

$S_1 = (\tau_1/T3? \text{Нормальная} \rightarrow (z=1)); \quad S_2 = (\tau_2/B1? 12 \rightarrow (z=1));$   
 $S_3 = (\tau_2/B1? 50 \rightarrow (z=1)); \quad S_4 = (\tau_2/B1? 32 \rightarrow (z=2)); \quad S_5 = (\tau_2/B1? 87 \rightarrow (z=2)).$

Легко убедиться, что в рамках описания  $\Omega(\{\tau_1/T1, \tau_2/B1\})$  модель  $\{S\}_{Full}$  включает восемь синдромов. В рамках описания «ТЗ – ВЗ» моделей знаний нет (из-за конфликтности данных), однако имеются два синдрома:

$S_1 = (\tau_1/T3? \text{Нормальная} \rightarrow (z=1)); \quad S_2 = (\tau_2/B3? \text{Пенсионный} \rightarrow (z=2)).$

Приведем общий алгоритм построения полной синдромной модели знаний в рамках фиксированного описания базы прецедентов  $\Omega(\{\tau/T\})$ , в которой отсутствуют артефакты.

### Алгоритм 5.2

Вход: База прецедентов  $\Omega(\{\tau/T\})$ .

Выход:  $\{S\}_{Full}$  на  $\Omega(\{\tau/T\})$ .

$KB(\Omega(\{\tau/T\}))$

**begin**

$\{S\}_{Full} := \emptyset;$

**for**  $\alpha(\{\underline{\tau}/T\}, \underline{z}/Z) \in \Omega$  **do**

**for**  $\{\tau/T\}' \subseteq \{\tau/T\}$  **do**

**if**  $Sindrom(\{\tau/T\}', \alpha, z_\alpha) = true$

**then**  $\{S\}_{Full} := \{S\}_{Full} \cup S(\{\tau/T\}', \alpha, z_\alpha);$

**return**  $(\{S\}_{Full})$

**end.**

Функция *Sindrom* выполняет одно единственное действие: она проверяет существование синдрома  $S(\{\tau/T\}', \alpha, z_\alpha)$ , данные для которого берутся из прецедента  $\alpha(\{\underline{\tau}/T\}, \underline{z}/Z)$ . Отсутствие артефактов в базе прецедентов  $\Omega(\{\tau/T\})$  гарантирует, что алгоритм 5.2 формирует непустое множество  $\{S\}_{Full}$ .

**Предложение 5.5.** Если база прецедентов  $\Omega(\{\tau/T_0\}, Z)$  не содержит конфликтов, то синдромные модели знаний на  $\langle \Omega, \{G(\tau)\} \rangle$  существуют.

Так как по условию  $\Omega(\{\tau/T_0\})$  не содержит конфликтов, то в рамках

описания  $\{\tau/T_0\}$  может быть построена, в частности, полная синдромная модель знаний с помощью алгоритма 5.2, что и доказывает предложение.

Пусть  $\{S\}$  – произвольная синдромная модель знаний, тогда, используя отображение  $s/\mu_s$ , соответствующую предельную модель знаний можно определить так

$$\forall \{S\}, \{S^*\}_{\{S\}} = \cup_{S \in \{S\}} \{S^*\}_S. \quad (5.9)$$

Модель знаний  $\{S^*\}_{\{S\}}$  назовем *сопряженной предельной* моделью для  $\{S\}$ . На процедурном уровне операция (5.9) реализуется с помощью алгоритма 5.1. Справедливо следующее предложение.

**Предложение 5.6.** Для любой синдромной модели знаний существует единственная сопряженная предельная модель знаний на  $\langle \Omega, \{G(\tau)\} \rangle$ .

Предельная синдромная модель знаний, образуется путем объединения сопряженных множеств («облаков») предельных синдромов, соответствующих синдромам исходной модели. Так как сопряженные множества отдельных синдромов определяются единственным образом, то их объединение также единственно.

Приведем алгоритм построения сопряженной предельной синдромной модели знаний  $(\{S\}_{Full, \{\tau/T\}})^*$ , который основан на предварительном построении полной синдромной модели знаний  $\{S\}_{Full}$  в рамках фиксированного описания базы прецедентов  $\Omega(\{\tau/T\})$ .

### Алгоритм 5.3

Вход: Описание  $\{\tau/T\}$ , контекст  $K = \langle \Omega, \{G(\tau)\} \rangle$ .

Выход:  $(\{S\}_{Full, \{\tau/T\}})^*$ .

$LKB(\{\tau/T\}, K)$

**begin**

$(\{S\}_{Full})^* := \emptyset;$

$\{S\}_{Full} := KB(\Omega(\{\tau/T\}));$

**for**  $S \in \{S\}_{Full}$  **do begin**

$\{S^*\}_S := LimS(S, K);$

$(\{S\}_{Full})^* := (\{S\}_{Full})^* \cup \{S^*\}_S$  **end;**

**return**  $((\{S\}_{Full})^*)$

**end.**

Синдромная модель знаний *минимальна*, если из нее нельзя удалить ни один синдром без потери полноты охвата прецедентов из  $\Omega(\{\tau/T_0\})$ . В рамках описания базы прецедентов  $\Omega(\{\tau/T\})$  может существовать множество минимальных синдромных моделей знаний, содержащих разное количество синдромов. Следовательно, в рамках описания  $\Omega(\{\tau/T\})$  существуют *абсолютно минимальные синдромные модели знаний*

(возможно одна). Для минимальной синдромной модели знаний будем использовать нотацию:  $\{S\}_{Min}$  на  $\Omega(\{\tau/T\})$  или  $\{S\}_{Min, \{\tau/T\}}$ . Для абсолютно минимальной синдромной модели знаний будем использовать нотацию:  $\{S\}^*_{Min}$  на  $\Omega(\{\tau/T\})$  или  $\{S\}^*_{Min, \{\tau/T\}}$ .

В рамках описания  $\Omega(\{\tau_1/T3, \tau_2/B1\})$  существуют две абсолютно минимальные синдромные модели знаний:

$$(\{S\}^*_{Min})_1 = \{S_1 = (\tau_1/T3? \text{Нормальная} \rightarrow (z=1)); S_2 = (\tau_2/B1? 12 \rightarrow (z=1)); S_4 = (\tau_2/B1? 32 \rightarrow (z=2)); S_5 = (\tau_2/B1? 87 \rightarrow (z=2))\}.$$

$$(\{S\}^*_{Min})_2 = \{S_2 = (\tau_2/B1? 12 \rightarrow (z=1)); S_3 = (\tau_2/B1? 50 \rightarrow (z=1)); S_4 = (\tau_2/B1? 32 \rightarrow (z=2)); S_5 = (\tau_2/B1? 87 \rightarrow (z=2))\}.$$

В рамках описания  $\Omega(\{\tau_1/T2, \tau_2/B3\})$  полная модель знаний является одновременно и абсолютно минимальной.

Пусть имеется произвольная синдромная модель знаний  $\{S\}$  в рамках описания  $\{\tau/T\}$ . Объединив все  $\{\tau/T\}_S$ , где  $S \in \{S\}$ , получим множество  $\{\tau/T\}_{\{S\}}$ . Ниже приведен общий алгоритм формирования множества  $\{\tau/T\}_{\{S\}}$  для произвольной модели знаний  $\{S\}$ .

#### Алгоритм 5.4

Вход: Модель знаний  $\{S\}$ .

Выход:  $\{\tau/T\}_{\{S\}}$ .

*Tests* ( $\{S\}$ )

**begin**

$\{\tau/T\}_{\{S\}} := \emptyset;$

**for**  $S(\{\tau/T, z/Z\} \in \{S\}$  **do**  $\{\tau/T\}_{\{S\}} := \{\tau/T\}_{\{S\}} \cup \{\tau/T\}_S;$

**return** ( $\{\tau/T\}_{\{S\}}$ )

**end.**

*Дополнением* к  $\{\tau/T\}_{\{S\}}$  назовем произвольное множество  $\{\tau/T\}^\perp_{\{S\}}$ , которое в совокупности с  $\{\tau/T\}_{\{S\}}$  реализует полное описание базы прецедентов  $\{S\}$ . Ясно, что

$$\{\tau\}_{\{S\}} \cap \{\tau\}^\perp_{\{S\}} = \emptyset; \quad \{\tau\}_{\{S\}} \cup \{\tau\}^\perp_{\{S\}} = \{\tau\}.$$

**Предложение 5.7.** Синдромная модель знаний  $\{S\}$ , построенная в рамках некоторого описания  $\{\tau/T\}$ , является синдромной моделью знаний во всех описаниях базы прецедентов  $\Omega(\{\tau/T\}_{\{S\}} \cup \{\tau/T\}^\perp_{\{S\}})$ .

Приведем пример. В рамках описания «Г1 – В1» существует синдромная модель:

$$\{S\} = \{S_2(\tau_2/B1? 12; z?1); S_3(\tau_2/B1? 50; z?1); S_4(\tau_2/B1? 32; z?2); S_5(\tau_2/B1? 87; z?2)\}.$$

Видно, что  $\{\tau/T\}_{\{S\}} = \tau_2/B1$ . Дополнением для  $\{\tau/T\}_{\{S\}}$  могут быть множества  $\{\tau_1/T1\}$ ,  $\{\tau_1/T2\}$  и  $\{\tau_1/T3\}$ . Согласно предложению 5.1 модель  $\{S\}$  является также моделью знаний в рамках описаний «Т2 – В1» и «Т3 – В1». Таким образом, модель  $\{S\}$  является моделью знаний одновременно в трех описаниях (не путать с применимостью модели).

В рамках описания «Т3 – В1» существует синдромная модель:

$$\{S\} = \{S_1 = (\tau_1/T3? \text{ Нормальная} \rightarrow (z=1)); S_2 = (\tau_2/B1? 12 \rightarrow (z=1)); S_4 = (\tau_2/B1? 32 \rightarrow (z=2)); S_5 = (\tau_2/B1? 87 \rightarrow (z=2))\}.$$

Видно, что  $\{\tau/T\}_{\{S\}} = \{\tau_1/T3; \tau_2/B1\}$ , следовательно,  $\{\tau\}_{\{S\}}^\perp = \emptyset$ . Таким образом, модель  $\{S\}$  является моделью знаний только в одном описании.

Следствием предложения 5.1 является тот факт, что модель знаний  $\{S\}_{Min}$  на  $\Omega(\{\tau/T\})$  правильнее было бы именовать «модель знаний  $\{S\}_{Min}$  на  $\Omega(\{\tau/T\}_{\{S\}} \cup \{\tau/T\}_{\{S\}}^\perp)$ , где  $\{S\} = \{S\}_{Min}$ », т.е. минимальная модель знаний на множестве или кластере описаний. Точно также модель знаний  $\{S\}_{Min}^*$  на  $\Omega(\{\tau/T\})$  правильнее было бы именовать «модель знаний  $\{S\}_{Min}^*$  на  $\Omega(\{\tau/T\}_{\{S\}} \cup \{\tau/T\}_{\{S\}}^\perp)$ ». То же справедливо и для  $\{S\}_{Full}$ .

Общим для моделей знаний на  $\Omega(\{\tau/T\})$  является то, что все они классифицируют все прецеденты в рамках описания  $\Omega(\{\tau/T\})$ .

Из предложения 5.7 вытекает следствие.

**Следствие 5.1.** Сопряженная предельная модель знаний  $\{S^*\}$  для синдромной модели знаний  $\{S\}$ , построенной в рамках некоторого описания  $\{\tau/T\}$ , классифицирует все прецеденты во всех описаниях базы прецедентов  $\Omega(\{\tau/T\}_{\{S\}} \cup \{\tau/T\}_{\{S\}}^\perp)$ .

**Предложение 5.8.** Если для классификации выбрана синдромная модель знаний  $\{S\}$ , то описание всех прецедентов в базе прецедентов может быть редуцировано до множества тестов  $\{\tau\}_{\{S\}}$ .

Можно найти все предельные синдромы на всех уровнях общности для каждой ситуации  $\alpha \in \Omega$ . Их объединение представляет собой *полную предельную синдромную модель знаний* на  $\langle \Omega, \{G(\tau) \rangle$ . Разные синдромы предельной синдромной модели знаний могут опираться на разные описания  $\{\tau/T\}$ . Для полной предельной синдромной модели знаний будем использовать нотацию:  $\{S^*\}_{Full}$  на  $\langle \Omega, \{G(\tau) \rangle$  или просто –  $\{S^*\}_{Full}$ .

Пусть  $g$  отображение такое, что

$$\forall \alpha \in \Omega, g/\mu_g: \alpha \rightarrow \{S^*\}_\alpha \quad (5.10)$$

где  $\{S^*\}_\alpha$  – все предельные синдромы на всех уровнях общности для ситуации  $\alpha$ ,  $\mu_g$  – механизм реализации отображения  $g$ , опирающийся на  $\langle \Omega, \{G(\tau) \rangle$ .

Обозначим через  $\{S\}_{\alpha, \{\tau/T\}}$  множество всех синдромов для ситуации  $\alpha$  на

уровне описания базы прецедентов  $\Omega(\{T/T\})$ . На множестве  $\{S\}_{\alpha, \{T/T\}}$  можно определить сопряженные предельные синдромы  $(\{S\}_{\alpha, \{T/T\}})^*$ , объединение которых по всем описаниям и даст  $\{S^*\}_{\alpha}$ , а именно:

$$\{S^*\}_{\alpha} = \cup_{\{T/T\}} (\{S\}_{\alpha, \{T/T\}})^* \quad (5.11)$$

В базе прецедентов примера «Диагностика» имеются четыре прецедента:  $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ . Построим для них множества  $\{S^*\}_{\alpha}$  (таблица 5.2).

Таблица 5.2 – Нахождение предельных синдромов для прецедентов

<b>Прецедент <math>\alpha_1 = (\tau_1/T1? 36.0; \tau_2/B1? 12; z?1)</math></b>		
Описание	Синдромы	Предельные синдромы
«Т1 – В1»	$S = (\tau_1/T1? 36.0);$ $S = (\tau_2/B1? 12)$	$S^* = (\tau_1/T1? 36.0);$ $S^* = (\tau_2/B1? 12)$
«Т1 – В2»	$S = (\tau_1/T1? 36.0)$	$S^* = (\tau_1/T1? 36.0)$
«Т1 – В3»	$S = (\tau_1/T1? 36.0)$	$S^* = (\tau_1/T1? 36.0)$
«Т2 – В1»	$S = (\tau_2/B1? 12)$	$S^* = (\tau_2/B1? 12)$
«Т2 – В2»	$S = (\tau_1/T2? \text{Пониженная},$ $\tau_2/B2? \text{Молодой})$	$S^* = (\tau_1/T2? \text{Пониженная}, \tau_2/B3?$ $\text{Допенсионный})$
«Т2 – В3»	$S = (\tau_1/T2? \text{Пониженная},$ $\tau_2/B3? \text{Допенсионный})$	$S^* = (\tau_1/T2? \text{Пониженная}, \tau_2/B3?$ $\text{Допенсионный})$
«Т3 – В1»	$S = (\tau_2/B1? 12)$	$S^* = (\tau_2/B1? 12)$
«Т3 – В2»	-	-
«Т3 – В3»	-	-
$\{S^*\}_{\alpha_1}$	$S^* = (\tau_1/T1? 36.0); S^* = (\tau_2/B1? 12); S^* = (\tau_1/T2? \text{Пониженная},$ $\tau_2/B3? \text{Допенсионный})$	

Примечание: в синдромах опущены заключения, так как они совпадают.

<b>Прецедент <math>\alpha_2 = (\tau_1/T1? 36.7; \tau_2/B1? 50; z?1)</math></b>		
Описание	Синдромы	Предельные синдромы
«Т1 – В1»	$S = (\tau_1/T1? 36.7);$ $S = (\tau_2/B1? 50)$	$S^* = (\tau_1/T3? \text{Нормальная});$ $S^* = (\tau_2/B2? \text{Средних лет})$
«Т1 – В2»	$S = (\tau_1/T1? 36.7);$ $S = (\tau_2/B2? \text{Средних лет})$	$S^* = (\tau_1/T3? \text{Нормальная});$ $S^* = (\tau_2/B2? \text{Средних лет})$
«Т1 – В3»	$S = (\tau_1/T1? 36.7)$	$S^* = (\tau_1/T3? \text{Нормальная})$
«Т2 – В1»	$S = (\tau_2/B1? 50);$ $S = (\tau_1/T2? \text{Нормальная})$	$S^* = (\tau_2/B2? \text{Средних лет});$ $S^* = (\tau_1/T3? \text{Нормальная})$
«Т2 – В2»	$S = (\tau_2/B2? \text{Средних лет});$ $S = (\tau_1/T2? \text{Нормальная})$	$S^* = (\tau_2/B2? \text{Средних лет});$ $S^* = (\tau_1/T3? \text{Нормальная})$
«Т2 – В3»	$S = (\tau_1/T2? \text{Нормальная})$	$S^* = (\tau_1/T3? \text{Нормальная})$
«Т3 – В1»	$S = (\tau_1/T3? \text{Нормальная});$ $S = (\tau_2/B1? 50)$	$S^* = (\tau_1/T3? \text{Нормальная});$ $S^* = (\tau_2/B2? \text{Средних лет})$



«Т3 – В2»	$S = (\tau_1/T3? \text{Нормальная});$ $S = (\tau_2/B2? \text{Средних лет})$	$S^* = (\tau_1/T3? \text{Нормальная});$ $S^* = (\tau_2/B2? \text{Средних лет})$
«Т3 – В3»	$S = (\tau_1/T3? \text{Нормальная})$	$S^* = (\tau_1/T3? \text{Нормальная})$
$\{S^*\}_{\alpha 1}$	$S^* = (\tau_2/B2? \text{Средних лет}); S^* = (\tau_1/T3? \text{Нормальная})$	

<b>Прецедент <math>\alpha_3 = (\tau_1/T1? 37.2; \tau_2/B1? 32; z?2)</math></b>		
Описание	Синдромы	Предельные синдромы
«Т1 – В1»	$S = (\tau_1/T1? 37.2);$ $S = (\tau_2/B1? 32)$	$S^* = ((\tau_1/T2? \text{Повышенная});$ $S^* = (\tau_2/B1? 32)$
«Т1 – В2»	$S = (\tau_1/T1? 37.2)$	$S^* = ((\tau_1/T2? \text{Повышенная})$
«Т1 – В3»	$S = (\tau_1/T1? 37.2)$	$S^* = ((\tau_1/T2? \text{Повышенная})$
«Т2 – В1»	$S = ((\tau_1/T2? \text{Повышенная});$ $S = (\tau_2/B1? 32)$	$S^* = ((\tau_1/T2? \text{Повышенная});$ $S^* = (\tau_2/B1? 32)$
«Т2 – В2»	$S = ((\tau_1/T2? \text{Повышенная})$	$S^* = ((\tau_1/T2? \text{Повышенная})$
«Т2 – В3»	$S = ((\tau_1/T2? \text{Повышенная})$	$S^* = ((\tau_1/T2? \text{Повышенная})$
«Т3 – В1»	$S = (\tau_2/B1? 32)$	$S^* = (\tau_2/B1? 32)$
«Т3 – В2»	-	-
«Т3 – В3»	-	-
$\{S^*\}_{\alpha 1}$	$S^* = (\tau_2/B1? 32); S^* = ((\tau_1/T2? \text{Повышенная})$	

<b>Прецедент <math>\alpha_4 = (\tau_1/T1? 36.4; \tau_2/B1? 87; z?2)</math></b>		
Описание	Синдромы	Предельные синдромы
«Т1 – В1»	$S = (\tau_1/T1? 36.4);$ $S = (\tau_2/B1? 87)$	$S^* = (\tau_1/T1? 36.4);$ $S^* = (\tau_2/B3? \text{Пенсионный})$
«Т1 – В2»	$S = (\tau_1/T1? 36.4);$ $S = (\tau_2/B2? \text{Старческий})$	$S^* = (\tau_1/T1? 36.4);$ $S^* = (\tau_2/B3? \text{Пенсионный})$
«Т1 – В3»	$S = (\tau_1/T1? 36.4);$ $S = (\tau_2/B3? \text{Пенсионный})$	$S^* = (\tau_1/T1? 36.4);$ $S^* = (\tau_2/B3? \text{Пенсионный})$
«Т2 – В1»	$S = (\tau_2/B1? 87)$	$S^* = (\tau_2/B3? \text{Пенсионный})$
«Т2 – В2»	$S = (\tau_2/B2? \text{Старческий})$	$S^* = (\tau_2/B3? \text{Пенсионный})$
«Т2 – В3»	$S = (\tau_2/B3? \text{Пенсионный})$	$S^* = (\tau_2/B3? \text{Пенсионный})$
«Т3 – В1»	$S = (\tau_2/B1? 87)$	$S^* = (\tau_2/B3? \text{Пенсионный})$
«Т3 – В2»	$S = (\tau_2/B2? \text{Старческий})$	$S^* = (\tau_2/B3? \text{Пенсионный})$
«Т3 – В3»	$S = (\tau_2/B3? \text{Пенсионный})$	$S^* = (\tau_2/B3? \text{Пенсионный})$
$\{S^*\}_{\alpha 4}$	$S^* = (\tau_1/T1? 36.4); S^* = (\tau_2/B3? \text{Пенсионный})$	

**Предложение 5.9.** Для любого прецедента  $\alpha \in \Omega$  множество предельных синдромов  $\{S^*\}_\alpha$  определяется единственным образом. Приведем алгоритм построения  $\{S^*\}_\alpha$  для любого прецедента  $\alpha$ .

### Алгоритм 5.5

Вход: Прецедент  $\alpha$ , контекст  $K = \langle \Omega, \{G(\tau)\} \rangle$ .

Выход:  $\{S^*\}_\alpha$

$LimSP(\alpha, K)$

**begin**

$\{S^*\}_\alpha := \emptyset;$

**for** каждого описания  $\{\tau/T\}$  **do**

**for**  $\{\tau/T\}' \subseteq \{\tau/T\}$  **do**

**if**  $Sindrom(\{\tau/T\}', \alpha, z_\alpha) = true$

**then**  $\{S^*\}_\alpha := \{S^*\}_\alpha \cup LimS(S(\{\tau/T\}', \alpha, z_\alpha), K);$

**return**  $(\{S^*\}_\alpha)$

**end.**

Одной из задач метода предельных обобщений является построение полной предельной синдромной модели знаний в виде

$$\{S^*\}_{Full} = \bigcup_{\alpha \in \Omega} \{S^*\}_\alpha \quad (5.12)$$

Для примера «Диагностика»

$\{S^*\}_{Full} = \{S^*\}_{\alpha1} \cup \{S^*\}_{\alpha2} \cup \{S^*\}_{\alpha3} \cup \{S^*\}_{\alpha4} = \{S^*(\tau_1/T1? 36.0; z?1); S^*(\tau_2/B1? 12; z?1); S^*(\tau_1/T2? \text{Пониженная, } \tau_2/B3? \text{Допенсионный; } z?1)\}_{\alpha1} \cup \{S^*(\tau_2/B2? \text{Средних лет; } z?1); S^*(\tau_1/T3? \text{Нормальная; } z?1)\}_{\alpha2} \cup \{S^*(\tau_2/B1? 32; z?2); S^*(\tau_1/T2? \text{Повышенная; } z?2)\}_{\alpha3} \cup \{S^*(\tau_1/T1? 36.4; z?2); S^*(\tau_2/B3? \text{Пенсионный; } z?2)\}_{\alpha4}.$

Ясно, что такие синдромы, как  $S^*(\tau_1/T1? 36.0; z?1)$ , не могут претендовать на статус инвариантов (параметров порядка) из-за низкого уровня обобщения, но такие синдромы, как  $S^*(\tau_1/T2? \text{Пониженная, } \tau_2/B3? \text{Допенсионный; } z?1)$ ,  $S^*(\tau_2/B2? \text{Средних лет; } z?1)$ ,  $S^*(\tau_1/T3? \text{Нормальная; } z?1)$ ,  $S^*(\tau_1/T2? \text{Повышенная; } z?2)$  и  $S^*(\tau_2/B3? \text{Пенсионный; } z?2)$ , имеют достаточно высокий уровень обобщения и поэтому вполне могут претендовать на статус параметров порядка.

**Предложение 5.10.** Если база прецедентов  $\Omega(\{\tau/T_0\}, Z)$  не содержит конфликтов, то полная предельная синдромная модель знаний  $\{S^*\}_{Full}$  на  $\langle \Omega, \{G(\tau)\} \rangle$  существует и единственна.

Приведем алгоритм построения полной предельной синдромной модели знаний  $\{S^*\}_{Full}$  на  $\langle \Omega, \{G(\tau)\} \rangle$ , который основан на выражении (5.12).

### Алгоритм 5.6

Вход: Контекст  $K = \langle \Omega, \{G(\tau)\} \rangle$ .

Выход:  $\{S^*\}_{Full}$ .

$LimSFull(K)$

**begin**

$\{S^*\}_{Full} := \emptyset;$

**for**  $\alpha \in \Omega$  **do**  $\{S^*\}_{Full} := \{S^*\}_{Full} \cup LimSP(\alpha, K)$ ;  
**return**  $(\{S^*\}_{Full})$   
**end.**

Строго говоря, полное множество предельных синдромов можно найти, даже если описание  $\Omega(\{\tau/T_0\}, Z)$  конфликтно. Другое дело, что данное множество не будет моделью знаний, позволяющей однозначно классифицировать все прецеденты.

На основе полной предельной синдромной модели знаний могут быть построены *минимальные предельные синдромные модели знаний* и, соответственно, *абсолютно минимальные предельные синдромные модели знаний*. Последние модели представляют особенный интерес. Для минимальных предельных синдромных моделей знаний будем использовать нотацию:  $\{S^*\}_{Min}$  на  $\langle \Omega, \{G(\tau)\} \rangle$  или просто  $\{S^*\}_{Min}$ . Для абсолютно минимальных предельных синдромных моделей знаний будем использовать нотацию:  $\{S^*\}^*_{Min}$  на  $\langle \Omega, \{G(\tau)\} \rangle$  или просто  $\{S^*\}^*_{Min}$ .

Для нахождения всех  $\{S^*\}^*_{Min}$  достаточно вначале построить все  $\{S^*\}_{Min}$ , а затем выбрать из них все  $\{S^*\}^*_{Min}$ . Все модели  $\{S^*\}^*_{Min}$  применимы к базовому описанию базы прецедентов  $\Omega(\{\tau/T_0\})$ , но в целом сфера их применения (охват разных описаний прецедентов) может существенно различаться, что предоставляет возможности для поиска максимальных по уровню общности моделей.

**Предложение 5.11.** Пусть  $\Omega = \{\alpha_1, \dots, \alpha_n\}$ . Тогда, если  $\{S^*\}_{\alpha_i} \cap \{S^*\}_{\alpha_j} = \emptyset$  для  $i, j = 1, \dots, n$  ( $i \neq j$ ), то справедливо следующее

$$\{\{S^*\}^*_{Min}\} = \{\{S^*\}_{Min}\} = \{S^*\}_{\alpha_1} \times \{S^*\}_{\alpha_2} \times \dots \times \{S^*\}_{\alpha_n}. \quad (5.13)$$

$$\forall \{S^*\}_{Min}, \forall \{S^*\}^*_{Min}, |\{S^*\}_{Min}| = |\{S^*\}^*_{Min}| = n, \quad (5.14)$$

$$|\{\{S^*\}^*_{Min}\}| = |\{\{S^*\}_{Min}\}| = \prod_{j=1, \dots, n} |\{S^*\}_{\alpha_j}|. \quad (5.15)$$

Соотношение (5.14) дает верхнюю границу мощности любой  $\{S^*\}_{Min}$  и любой  $\{S^*\}^*_{Min}$ .

Для примера «Диагностика» выполняется условие предложения 5.11, поэтому верно представление (5.13). Приведем примеры

$$(\{S^*\}^*_{Min})_1 = \{S^*(\tau_1/T1? 36.0; z?1); S^*(\tau_2/B2? Средних лет; z?1); S^*(\tau_2/B1? 32; z?2); S^*(\tau_1/T1? 36.4; z?2)\};$$

$$(\{S^*\}^*_{Min})_2 = \{S^*(\tau_1/T2? Пониженная, \tau_2/B3?Допенсионный; z?1); S^*(\tau_1/T3? Нормальная; z?1); S^*(\tau_1/T2? Повышенная; z?2); S^*(\tau_2/B3? Пенсионный; z?2)\};$$

$$(\{S^*\}^*_{Min})_3 = \{S^*(\tau_1/T2? Пониженная, \tau_2/B3?Допенсионный; z?1); S^*(\tau_2/B2? Средних лет; z?1); S^*(\tau_1/T2? Повышенная; z?2); S^*(\tau_2/B3? Пенсионный;$$

$z?2)\}$ .

Легко убедиться, что  $|\{\{S^*\}_{Min}\}| = 3 \cdot 2 \cdot 2 \cdot 2 = 24$ . Можно также убедиться в том, что из 24 минимальных моделей только две модели ( $\{S^*\}_{Min}\}_2$  и ( $\{S^*\}_{Min}\}_3$ ) обладают максимально возможным уровнем обобщения и, следовательно, только они претендуют на статус инвариантных моделей.

Пусть  $\Omega_j = \{\alpha(\{T_j\}, z_j)\}$ ,  $j \in Z$ . Ясно, что  $\Omega = \cup_{j \in Z} \Omega_j$ . Определим *ядро синдромов* для каждого  $j \in Z$  (директивная зона)

$$\{S^*\}_{\Omega_j} = \cap_{\alpha \in \Omega_j} \{S^*\}_{\alpha} \quad j \in Z. \quad (5.16)$$

**Предложение 5.12.** Если для любого  $j \in Z$ ,  $\{S^*\}_{\Omega_j} \neq \emptyset$ , то справедливо следующее

$$\{\{S^*\}_{Min}\} = \{S^*\}_{\Omega_1} \times \{S^*\}_{\Omega_2} \times \dots \times \{S^*\}_{\Omega_N}. \quad (5.17)$$

$$\forall \{S^*\}_{Min} \quad |\{S^*\}_{Min}| = N, \quad (5.18)$$

$$|\{\{S^*\}_{Min}\}| = \prod_{j=1, \dots, N} |\{S^*\}_{\Omega_j}|. \quad (5.19)$$

Соотношение (5.18) дает оценку мощности любой  $\{S^*\}_{Min}$ .

Важной задачей метода предельных обобщений является построение директивных зон (ядер)  $\{S^*\}_{\Omega_j}$ ,  $j \in Z$ . Дело в том, что ядра содержат, как правило, предельные синдромы максимального уровня общности и не содержат информационный «шум», т.е. редкие и малозначащие синдромы. Ядра могут существовать не для каждого  $j \in Z$ .

Важными задачами метода предельных обобщений является также построение  $\{S^*\}_{Min}$  и  $\{S^*\}_{Min}$  по разным критериям (например, с максимальными весами). Конечная цель исследований будет достигнута, когда хотя бы одна из минимальных моделей знаний не будет изменяться с неограниченным увеличением базы прецедентов  $\Omega$ . Данный факт означает *инвариантность* или *корректность* соответствующей модели знаний.

Отметим одно важное обстоятельство, касающееся представительности базы прецедентов  $\Omega$ . Корректные модели знаний могут быть построены на  $\langle \Omega, \{G(\tau)\} \rangle$ , даже если не будет выполняться условие

$$\forall \tau \in \{G(\tau)\}, \cup_{\alpha \in \Omega} (\{T_0\}_{\alpha}) = T_0.$$

Другими словами, если по каждому тесту не будет покрытия области возможных значений (на уровне базовых доменов). Обобщение в рамках банка тестов  $\{G(\tau)\}$  позволяет преодолеть такую неполноту информации. Кроме того, минимальные синдромные модели знаний используют, как правило, далеко не все тесты, что еще более ослабляет требования к полноте априорной информации.

Будем говорить, что синдромная модель знаний  $\{S^*\}$  *доминирует*

модель знаний  $\{S\}$ , если она применима к большему числу описаний прецедентов  $\alpha(\{\tau/T\}, \underline{z}/Z)$ , включая и те описания прецедентов, к которым применима  $\{S\}$ . Факт доминирования будем отражать нотацией  $\{S'\} > \{S\}$ . Будем говорить, что синдромные модели знаний  $\{S'\}$  и  $\{S\}$  эквивалентны в плане доминирования, если они применимы к одному и тому же множеству описаний прецедентов. Факт эквивалентности в плане доминирования будем отражать нотацией  $\{S'\} \sim \{S\}$ .

Подобная трактовка доминирования на моделях знаний дает возможность сравнивать предельные модели знаний, синдромы которых имеют разные уровни общности, а также применять модели знаний к описаниям прецедентов, принадлежащих конфликтным описаниям базы прецедентов. Например, синдром  $S^*(\tau_2/V3? \text{ Пенсионный}; z?2)$  модели  $(\{S^*\}_{Min})_3$  применим к прецедентам конфликтного описания базы прецедентов «ТЗ – ВЗ», для которого не существует синдромной модели знаний.

Если между двумя моделями знаний не выполняется отношение доминирования и они не эквивалентны в плане доминирования, то будем говорить, что такие модели знаний *несравнимы между собой в плане доминирования*. Данный факт будем отражать нотацией  $\{S'\} \times \{S\}$ .

Отношение доминирования на моделях знаний, также как и отношение эквивалентности, является транзитивным.

Для установления отношения доминирования, эквивалентности или несравнимости между произвольными синдромными моделями знаний  $\{S\}$  и  $\{S'\}$  достаточно построить таблицу вида 5.3, где  $M = |\{\Omega(\{\tau/T\})\}|$  – общее количество всех описаний,  $n = |\Omega|$  – общее количество прецедентов.

Таблица 5.3 – Покрытие прецедентов двумя моделями знаний

	$\{S\}$				$\{S'\}$			
$\{\tau/T\} \setminus \alpha$	$\alpha_1$	$\alpha_2$	...	$\alpha_n$	$\alpha_1$	$\alpha_2$	...	$\alpha_n$
$\{\tau/T\}_1$	+	+	...	+	+	+	...	+
$\{\tau/T\}_2$	+/-	+/-	...	+/-	+/-	+/-	...	+/-
...	...	...	...	...	...	...	...	...
$\{\tau/T\}_M$	+/-	+/-	...	+/-	+/-	+/-	...	+/-

Символ ‘+’ в пересечении строки  $\{\tau/T\}_i$  и столбца  $\alpha_j$  означает, что соответствующая модель знаний на уровне описания  $\{\tau/T\}_i$  классифицирует прецедент  $\alpha_j$ . Символ ‘-’ означает, что классификация невозможна. Если области покрытия совпадают, то модели эквивалентны по доминированию. Если одно покрытие включает другое покрытие, то соответствующая модель знаний доминирует другую модель. В иных случаях имеет место несравнимость моделей знаний по доминированию.

Если через  $|G|$  обозначить общее количество вершин в произвольном орграфе доменов  $G$ , то общее количество всех описаний определяется выражением:

$$M = \prod_{\tau \in \{G(\tau)\}} |G(\tau)|. \quad (5.20)$$

*Долей покрытия модели знаний* назовем отношение площади области покрытия (клетки с «+» в таблице 5.3) к площади всей области (суммарному количеству клеток равному  $M \cdot n$ ). Для модели знаний  $\{S\}$  долю покрытия обозначим  $\lambda(\{S\})$  или  $\lambda_{\{S\}}$ .

Если  $\lambda(\{S\}) = 1$ , то это означает, что модель знаний  $\{S\}$  покрывает все прецеденты во всех описаниях, т.е.  $\{S\}$  имеет максимальную зону покрытия. Ясно, что такая модель является недоминируемой.

**Предложение 5.13.** Если модель знаний  $\{S\}$  доминирует модель знаний  $\{S'\}$ , то  $\lambda(\{S\}) > \lambda(\{S'\})$ . Если модели знаний  $\{S\}$  и  $\{S'\}$  эквивалентны по доминированию, то  $\lambda(\{S\}) = \lambda(\{S'\})$ .

Обозначим через  $D_{\{\tau T\}}$  множество всех описаний, которые доминируются описанием  $\{\tau T\}$ , т.е.

$$D_{\{\tau T\}} = \{\{\tau T\}' \mid \{\tau T\} \geq \{\tau T\}'\}. \quad (5.21)$$

**Предложение 5.14.** Если модель  $\{S\}$  является моделью знаний в рамках описания  $\{\tau T\}$ , то справедлива оценка

$$\lambda(\{S\}) \geq (1 + |D_{\{\tau T\}}|) / M, \quad (5.22)$$

где  $M$  – общее количество всех описаний, определяемое формулой (5.20).

**Предложение 5.15.** Если  $\{\tau T\}_{\{S'\}} \geq \{\tau T\}_{\{S\}}$ , то синдромная модель знаний  $\{S'\}$  доминирует синдромную модель знаний  $\{S\}$  или эквивалентна ей по доминированию.

Тот факт, что  $\{\tau T\}_{\{S'\}} \geq \{\tau T\}_{\{S\}}$ , означает, в частности, следующее:

$$\{\tau\}_{\{S'\}} = \{\tau\}_{\{S\}}; \quad \{\tau\}^{\perp}_{\{S'\}} = \{\tau\}^{\perp}_{\{S\}},$$

$$\forall \alpha \in \Omega, \forall \{\tau T\}^{\perp}_{\{S\}}, \alpha(\{\tau T\}_{\{S\}} \cup \{\tau T\}^{\perp}_{\{S\}}) \rightarrow \alpha(\{\tau T\}_{\{S'\}} \cup \{\tau T\}^{\perp}_{\{S\}}),$$

следовательно, к  $\alpha(\{\tau T\}_{\{S\}} \cup \{\tau T\}^{\perp}_{\{S\}})$  и  $\alpha(\{\tau T\}_{\{S'\}} \cup \{\tau T\}^{\perp}_{\{S\}})$  применима модель  $\{S'\}$ . Таким образом, может иметь место либо доминирование  $\{S'\}$  над  $\{S\}$ , либо эквивалентность  $\{S'\}$  и  $\{S\}$ .

Предложение 5.15 обосновывает целесообразность построения моделей знаний максимального уровня общности.

**Предложение 5.16.** Для любой синдромной модели знаний  $\{S\}$  сопряженная модель  $\{S^*\}$  доминирует ее или эквивалентна ей по доминированию.

Действительно, сопряженная модель классифицирует все ситуации,

которые классифицирует исходная модель знаний, но при этом сопряженная модель может классифицировать и другие ситуации (более высокого уровня общности). Таким образом, имеется полное покрытие со стороны сопряженной модели, что означает либо доминирование, либо эквивалентность.

**Предложение 5.17.** Полная предельная синдромная модель знаний  $\{S^*\}_{Full}$  доминирует любую синдромную модель знаний или эквивалентна ей по доминированию.

Доказательство данного утверждения базируется на том, что сопряженная предельная модель знаний  $\{S^*\}_{\{S\}}$  эквивалентна или доминирует произвольную модель знаний  $\{S\}$ , но любая  $\{S^*\}_{\{S\}}$  является частью  $\{S^*\}_{Full}$ , следовательно,  $\{S^*\}_{Full}$  эквивалентна или доминирует  $\{S^*\}_{\{S\}}$ , а значит и  $\{S\}$ . В частности, модель  $\{S^*\}_{Full}$  эквивалентна в плане доминирования или доминирует любую  $\{S^*\}_{Min}$ .

Большой практический интерес представляет поиск моделей знаний  $\{S^*\}_{Min}$  эквивалентных по доминированию  $\{S^*\}_{Full}$ .

Установим, существует ли отношение доминирования между моделями  $(\{S^*\}_{Min})_2$  и  $(\{S^*\}_{Min})_3$ , а также эквивалентность какой-либо из них модели  $\{S^*\}_{Full}$ . Представим данные по покрытию прецедентов в рамках всех описаний в виде таблицы 5.4.

Таблица 5.4 – Покрытие прецедентов двумя минимальными моделями знаний и полной моделью

	$(\{S^*\}_{Min})_2$				$(\{S^*\}_{Min})_3$				$\{S^*\}_{Full}$			
$\{\pi/T\} \setminus \alpha$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$
«T1 – B1»	+	+	+	+	+	+	+	+	+	+	+	+
«T1 – B2»	+	+	+	+	+	+	+	+	+	+	+	+
«T1 – B3»	+	+	+	+	+	-	+	+	+	+	+	+
«T2 – B1»	+	+	+	+	+	+	+	+	+	+	+	+
«T2 – B2»	+	+	+	+	+	+	+	+	+	+	+	+
«T2 – B3»	+	+	+	+	+	-	+	+	+	+	+	+
«T3 – B1»	-	+	-	+	-	+	-	+	+	+	+	+
«T3 – B2»	-	+	-	+	-	+	-	+	-	+	-	+
«T3 – B3»	-	+	-	+	-	-	-	+	-	+	-	+

Как видно из таблицы модель  $(\{S^*\}_{Min})_2$  доминирует модель  $(\{S^*\}_{Min})_3$ , а модель  $\{S^*\}_{Full}$  доминирует  $(\{S^*\}_{Min})_2$ . Сравнение  $(\{S^*\}_{Min})_2$  с другими моделями из  $\{\{S^*\}_{Min}\}$  показывает, что она является недоминируемой моделью знаний (на данном множестве моделей), при этом площадь ее области покрытия максимальна из всех моделей  $\{S^*\}_{Min}$ . Это говорит о том, что на множестве всех недоминируемых моделей знаний следует искать модели с максимальной площадью покрытия (долей покрытия), так как они

наиболее близки по покрытию к модели  $\{S^*\}_{Full}$ .

Для моделей знаний из таблицы 5.4 рассчитаем долю покрытия  $\lambda$ :

$$\lambda(\{S^*\}_{Min})_2 = 30/36 = 5/6; \quad \lambda(\{S^*\}_{Min})_3 = 27/36 = 3/4;$$

$$\lambda(\{S^*\}_{Full}) = 32/36 = 8/9 = \max.$$

Легко убедиться, что модель

$$\{S^*\}_{Min})_4 = \{S^*(\tau_2/B1? 12; z?1); S^*(\tau_2/B1? 50; z?1); S^*(\tau_2/B1? 32; z?2); S^*(\tau_2/B1? 87; z?2)\}$$

является несравнимой по доминированию ни с моделью  $\{S^*\}_{Min})_2$ , ни с моделью  $\{S^*\}_{Min})_3$ . Действительно, при всей ограниченности модель  $\{S^*\}_{Min})_4$  позволяет классифицировать прецедент  $\alpha_1(\langle T3 - B1 \rangle) = (\tau_1/T3? \text{Ненорм.}; \tau_2/B1? 12; z?1)$ , а две другие модели нет.

В таблице 5.5 приведены области покрытия модели  $\{S^*\}_{Min})_4$  и модели  $\{S^*\}_{Min})_5$  следующего вида:

$$\{S^*\}_{Min})_5 = \{S^*(\tau_1/T1? 36.0; z?1); S^*(\tau_1/T1? 36,7; z?1); S^*(\tau_1/T1? 37,2; z?2); S^*(\tau_1/T1? 36.4; z?2)\}.$$

Таблица 5.5 – Покрытие прецедентов двумя минимальными моделями

$\{\tau/T\} \setminus \alpha$	$\{S^*\}_{Min})_4$				$\{S^*\}_{Min})_5$			
	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$
«T1 – B1»	+	+	+	+	+	+	+	+
«T1 – B2»	-	-	-	-	+	+	+	+
«T1 – B3»	-	-	-	-	+	+	+	+
«T2 – B1»	+	+	+	+	-	-	-	-
«T2 – B2»	-	-	-	-	-	-	-	-
«T2 – B3»	-	-	-	-	-	-	-	-
«T3 – B1»	+	+	+	+	-	-	-	-
«T3 – B2»	-	-	-	-	-	-	-	-
«T3 – B3»	-	-	-	-	-	-	-	-

Как видно из таблицы модели  $\{S^*\}_{Min})_4$  и  $\{S^*\}_{Min})_5$  являются несравнимыми по доминированию, хотя в процентном отношении они покрывают одинаковые площади (доли покрытия одинаковы и равны 1/3).

Можно также убедиться в том, что в примере «Диагностика» ни одна из моделей  $\{S^*\}_{Min})$  не эквивалентна по доминированию модели  $\{S^*\}_{Full}$ . Можно говорить лишь о минимальном *обогащении* моделей  $\{S^*\}_{Min})$  до полной эквивалентности с моделью  $\{S^*\}_{Full}$ . Так, если к модели  $\{S^*\}_{Min})_2$  добавить два синдрома  $S^*(\tau_2/B1? 12; z?1)$  и  $S^*(\tau_2/B1? 32; z?2)$ , то результирующая обогащенная модель будет эквивалентна  $\{S^*\}_{Full}$ .

Таким образом, одной из центральных задач в рамках метода



предельных обобщений является поиск всех *абсолютно минимальных синдромных моделей знаний, которые эквивалентны*  $\{S^*\}_{Full}$ . Такие модели всегда существуют, если описание  $\Omega(\{\tau/T_0\})$  бесконфликтно (они могут совпадать с  $\{S^*\}_{Full}$ ). Введем для них обозначение  $\{S^*\}_{Min-Full}$ . Пример:

$$\{S^*\}_{Min-Full} = (\{S^*\}_{Min})_2 \cup \{S^*(\tau_2/V1? 12; z?1); S^*(\tau_2/V1? 32; z?2)\}.$$

Общий алгоритм построения  $\{S^*\}_{Min-Full}$  таков: вначале находятся все недоминируемые модели  $\{S^*\}_{Min}$  с максимальными площадями покрытия, затем осуществляется их обогащение до уровня  $\{S^*\}_{Min-Full}$ . Процедура обогащения следующая: находятся прецеденты  $\alpha(\{\underline{\tau}/T\})$ , которые покрываются  $\{S^*\}_{Full}$ , но не покрываются выбранной для обогащения моделью  $\{S^*\}_{Min}$ , затем любой предельный синдром для  $\alpha(\{\underline{\tau}/T\})$  добавляется к  $\{S^*\}_{Min}$ .

Рассмотрим один частный, но практически важный случай, когда гарантированно существуют модели  $\{S^*\}_{Min-Full}$ .

Терминальные вершины орграфа доменов обозначим  $T^*$ . Если для теста  $\tau$  орграф  $G(\tau)$  содержит только одну терминальную вершину, то в большинстве случаев такой орграф можно представить линейной структурой:  $G(\tau) = \{T_0 \rightarrow \dots \rightarrow T^*\}$ . Орграфы доменов примера «Диагностика» являются линейными.

**Предложение 5.18.** Пусть орграфы всех тестов содержат лишь одну терминальную вершину и описание базы прецедентов  $\Omega(\{\tau/T^*\})$  не содержит конфликтов, тогда модели знаний  $\{S^*\}_{Min-Full}$  принадлежат описанию  $\{\tau/T^*\}$ . Покрытие моделями  $\{S^*\}_{Min-Full}$  является полным.

Действительно, так как описание  $\Omega(\{\tau/T^*\})$  не содержит конфликтов, то в рамках этого описания существует модель знаний  $\{S\}_{Full}$ , которая совпадает с сопряженной моделью знаний  $(\{S\}_{Full})^*$  из-за невозможности дальнейшего обобщения. Модель  $(\{S\}_{Full})^*$  на  $\Omega(\{\tau/T^*\})$  является частью полной модели  $\{S^*\}_{Full}$  и при этом она позволяет классифицировать все прецеденты во всех описаниях, т.е. она эквивалентна  $\{S^*\}_{Full}$  по доминированию. Любая абсолютно минимальная модель знаний  $\{S^*\}_{Min}$ , построенная на основе  $(\{S\}_{Full})^*$  в рамках описания  $\{\tau/T^*\}$ , является эквивалентной самой  $(\{S\}_{Full})^*$ , а значит и  $\{S^*\}_{Full}$  (по свойству транзитивности).

Таблица вида 5.3, построенная для двух моделей –  $\{S^*\}_{Full}$  и  $\{S^*\}_{Min-Full}$  на  $\{\tau/T^*\}$ , заполнена исключительно символами ‘+’, что также доказывает их эквивалентность (области покрытия полные).

Из предложения 5.11 вытекает следствие.

**Следствие 5.2** Справедлива следующая оценка количественного состава любой абсолютно минимальной, предельной синдромной модели знаний  $\{S^*\}_{Min}$  на  $\langle \Omega, \{G(\tau)\} \rangle$ :

$$|Z| \leq |\{S^*\}_{Min}| \leq |\Omega|. \quad (5.23)$$

Вместе с тем, любая даже абсолютно минимальная модель знаний может быть еще редуцирована (в ряде случаев значительно). Такая возможность вытекает из того, что  $Z$  образует полную и несовместную группу заключений. Следовательно, если установлено отсутствие любого из  $z_1, \dots, z_{N-1}$  заключений (с точностью до нумерации), то прецедент может иметь только заключение  $z_N$ , но в таком случае можно обойтись без синдромов, которые позволяют установить  $z_N$ . Рассмотрим более детально данную возможность.

Пусть в произвольной синдромной базе знаний  $\{S\}$  заключению  $z_j$  отвечает подмножество синдромов  $\{S\}_j$  ( $j = 1, \dots, N$ ). Определим мощность (индекс) каждого из подмножеств синдромов:

$$L_j = |\{S\}_j|, \quad j = 1, \dots, N. \quad (5.24)$$

По определению синдромной модели знаний  $L_j > 0$  ( $j = 1, \dots, N$ ). Мощность всей модели знаний  $\{S\}$  определяется выражением:

$$L_{\{S\}} = \sum_{j \in Z} L_j. \quad (5.25)$$

Таким образом, для произвольной синдромной модели знаний  $\{S\}$  определен кортеж индексов:  $\langle L_1, \dots, L_N \rangle$ . Найдем такое подмножество номеров  $I = \{i_1, \dots, i_{N-1}\} \subset Z$ , которое обеспечивает минимум сумме индексов, а именно:

$$L_I = \sum_{j \in I} L_j \rightarrow \min. \quad (5.26)$$

Ясно, что задача (5.26) может решаться не единственным образом. В таком случае можно выбрать либо любое решение, либо продолжить оптимизацию по другим критериям, например, дополнительно минимизировать общую суммарную стоимость тестов или минимизировать число используемых тестов.

Для произвольной синдромной модели знаний  $\{S\} = \cup_{j \in Z} \{S\}_j$  определим *редуцированную модель знаний*  $\{S\}^{\sim}$  следующим образом:

$$\{S\}^{\sim} = \cup_{j \in I} \{S\}_j \quad (5.27)$$

$$z_k, \text{ где } k = Z \setminus I \Leftrightarrow \&_{j \in I} \neg z_j \equiv \neg (\vee_{j \in I} z_j) = \text{true}, \quad (5.28)$$

где множество  $I$  определяется путем решения задачи (5.26). Правило (5.28) является неотъемлемой частью модели  $\{S\}^{\sim}$ . Оно позволяет установить заключение  $z_k$  в том случае, когда доказано отсутствие заключений из  $I$ .

Принципиальное отличие синдромной модели знаний от редуцированной модели заключается в том, что в рамках синдромной модели для любого прецедента из базы прецедентов существует хотя бы

один синдром. В редуцированной модели это правило не выполняется: для прецедентов  $\alpha(\{\underline{z}T\}, z_k)$  синдромы отсутствуют.

Таким образом, редуцированная модель знаний является более сложной структурой, чем просто синдромная модель знаний, но это усложнение позволяет сократить (иногда значительно) как общее число синдромов, так и, возможно, число используемых тестов.

Приведем пример. Пусть  $Z = \{1, 2, 3\}$ ;  $\{S\}_1 = \{S_1; S_2; S_3\}$ ;  $\{S\}_2 = \{S_4; S_5; S_6\}$ ;  $\{S\}_3 = \{S_7; S_8\}$ . Получаем две эквивалентные (по числу синдромов) редуцированные модели знаний:

$$\begin{aligned} \{S\}^{\sim} &= \{S\}_1 \cup \{S\}_3 \quad \& (z_2 \Leftrightarrow \neg z_1 \ \& \ \neg z_3 = true); \quad I_1 = \{1, 3\}; \\ \{S\}^{\sim} &= \{S\}_2 \cup \{S\}_3 \quad \& (z_1 \Leftrightarrow \neg z_2 \ \& \ \neg z_3 = true); \quad I_2 = \{2, 3\}. \end{aligned}$$

Выбор первой или второй редуцированной модели может зависеть от числа используемых тестов в каждой модели, от суммарного веса синдромов, от суммарной стоимости тестов или от других свойств тестов (например, комфортности, опасности и т.д.).

Редуцированные модели можно также минимизировать, как и синдромные модели  $\{S\}_{Min}$ . Более того, редуцированные модели целесообразно строить на основе  $\{S\}_{Min}$ . Уточним данное положение.

Решим задачу (5.26) для множества моделей  $\{S^*\}_{Min}$  на  $\langle \Omega, \{G(\tau)\} \rangle$ . Получим величину  $L^* = L_I$  и множество  $\{I\}$ , где каждое  $I \in \{I\}$  обеспечивает достижение  $L^*$ .

**Предложение 5.19.** Во всех моделях  $\{S^*\}_{Min}$  на  $\langle \Omega, \{G(\tau)\} \rangle$  мощность каждого сегмента  $\{S\}_j$  ( $j \in Z$ ) одинакова. Константа  $L^*$  и множества  $\{I\}$  для всех моделей  $\{S^*\}_{Min}$  также одинаковы.

Действительно, каждый сегмент  $\{S\}_j$  ( $j \in Z$ ) оптимизируется (минимизируется) независимо от других, следовательно, глобальный оптимум у них один. Совпадение глобальных оптимумов по каждому сегменту  $\{S\}_j$  ( $j \in Z$ ) означает универсальность константы  $L^*$  и множества  $\{I\}$  для всех моделей  $\{S^*\}_{Min}$ .

**Предложение 5.20.** Константа  $L^*$  определяет абсолютно минимальное число синдромов, которое может быть использовано в редуцированной модели знаний для решения задачи классификации на  $\Omega$ . Оптимальные редуцированные модели знаний строятся на основе любой  $\{S^*\}_{Min}$ .

Пусть  $K^*$  – общее число моделей знаний  $\{S^*\}_{Min}$ , а  $K$  – общее число оптимальных редуцированных моделей знаний  $(\{S^*\}_{Min})^{\sim}$ . Справедлива оценка

$$K \leq K^* \cdot |\{I\}|. \quad (5.29)$$

Неравенство в выражении (5.29) означает, что некоторые оптимальные модели знаний могут совпадать.

Уточним величину  $K^*$ . Мощность каждого сегмента  $\{S\}_j$  ( $j \in Z$ ) в

моделях знаний  $\{S^*\}_{Min}$  фиксирована и равна  $L_j$ . Однако состав сегментов может быть разным в разных моделях. Разные наборы синдромов  $j$ -го сегмента обозначим  $\{S\}_{j,m}$  ( $j \in Z$ ,  $m = 1, \dots, M_j$ ). Ясно, что  $\forall m, |\{S\}_{j,m}| = L_j$ .

Определим множества  $X_j$  ( $j \in Z$ ) следующим образом:

$$X_j = \{\{S\}_{j,1}; \{S\}_{j,2}; \dots; \{S\}_{j,M_j}\}, j \in Z. \quad (5.30)$$

**Предложение 5.21.** Полное множество предельных моделей знаний  $\{S^*\}_{Min}$  определяется декартовым произведением следующего вида

$$\{\{S^*\}_{Min}\} = X_1 \times X_2 \times \dots \times X_N, \quad (5.31)$$

а их общее количество определяется выражением

$$K^* = |X_1 \times X_2 \times \dots \times X_N| = M_1 \cdot \dots \cdot M_N, \text{ где } M_j = |X_j|. \quad (5.32)$$

На основе выражения (5.32) может быть получена точная оценка для  $K^*$ .

**Предложение 5.22.** Полное множество оптимальных редуцированных моделей знаний  $(\{S^*\}_{Min})^\sim$  определяется выражением

$$\{(\{S^*\}_{Min})^\sim\} = \bigcup_{I \in \{I\}} (X_{i_1} \times \dots \times X_{i_{N-1}} \& (z_{zI} \Leftrightarrow \&_{j \in I} \neg z_j = true)), \quad (5.33)$$

а их общее количество определяется выражением

$$K^\sim = \sum_{I \in \{I\}} M_{i_1} \cdot \dots \cdot M_{i_{N-1}}, \text{ где } I = \{i_1, \dots, i_{N-1}\}, M_j = |X_j|. \quad (5.34)$$

Рассмотрим вопрос о построении минимальных синдромных моделей знаний с максимальными весами.

Пусть  $\{S\}_j$  – это  $j$ -й фрагмент ( $j \in Z$ ) модели знаний  $\{S\}_{Full}$ . Для каждого синдрома  $S$  определен вес  $v_S$ . Упорядочим по убыванию веса все синдромы из  $\{S\}_j$ . Упорядоченное множество обозначим  $\langle S \rangle_j$  ( $j \in Z$ ). Поскольку веса синдромов могут совпадать, операция упорядочивания выполняется не единственным образом. Правило формирования компоненты  $(\{S\}_{Min})_j$  следующее: синдромы из  $\langle S \rangle_j$  добавляются в  $(\{S\}_{Min})_j$  в порядке убывания веса, при этом очередной синдром добавляется только в том случае, если он покрывает новые прецеденты из  $\Omega_j$ . Правило применяется для всех  $j \in Z$ .

Построение *минимальных синдромных моделей знаний с максимальными весами*  $\{S\}_{Min}^v$  на основе  $\{S\}_{Full}$  повышает вероятность нахождения инвариантных моделей знаний, так как исключаются редкие и, возможно, ошибочные синдромы (последние могут появиться в результате «шума данных»). Алгоритм построения  $\{S\}_{Min}^v$  приведен ниже. Нотация  $\langle S(\underline{T}, j) \in \langle S \rangle_j \downarrow \rangle$  означает перебор синдромов в порядке убывания весов.

### Алгоритм 5.7

Вход: Модель знаний  $\{S\}_{Full}$ .

Выход: Модель знаний  $\{S\}^v_{Min}$ .  
 $KB\_MaxWeight(\{S\}_{Full})$

**begin**

$\{S\}^v_{Min} := \emptyset;$

$\{S\}_{Full} := \cup_{j \in Z} \langle S \rangle_j;$

**for**  $j \in Z$  **do begin**

$\{\alpha\} := \emptyset;$

**for**  $S(\{\underline{T}\}, j) \in \langle S \rangle_j \downarrow$  **do**

**if**  $\Omega_j \setminus \{\alpha\} \neq \emptyset$  **then begin**

$\{\beta\} := \emptyset;$

**for**  $\alpha \in \Omega_j \setminus \{\alpha\}$  **do**

**if**  $\{\underline{T}\} = \{\underline{T}\}_\alpha$  **then**  $\{\beta\} := \{\beta\} \cup \alpha;$

$\{\alpha\} := \{\alpha\} \cup \{\beta\};$

**if**  $\{\beta\} \neq \emptyset$  **then**  $\{S\}^v_{Min} = \{S\}^v_{Min} \cup S(\{\underline{T}\}, j)$  **end**

**end;**

**return**  $(\{S\}^v_{Min})$

**end.**

Следует отметить, что алгоритм 5.7 не гарантирует решение следующей целевой задачи:

$$\sum_{S \in \{S\}} \nu_S \rightarrow \max_{\{S\}}, \text{ где } \{S\} \in \{\{S\}_{Min} \mid \{S\}_{Min} \in \{S\}_{Full}\}.$$

Вместе с тем, данный алгоритм дает вполне удовлетворительные для практики решения, которые совпадают с оптимальным решением или близки к нему.

Рассмотрим вопрос о том, сколько тестов и какого уровня общности минимально достаточно, чтобы установить заключение  $z/Z$  для прецедентов из  $\Omega$ .

Пусть  $\{\tau/T\}^*_{Min}$  – полный набор тестов с указанием доменов, которые фигурируют в какой-либо абсолютно минимальной, предельной модели знаний  $\{S^*\}^*_{Min}$ . Может оказаться так, что один и тот же тест присутствует в наборе  $\{\tau/T\}^*_{Min}$  несколько раз, но с разными доменами.

Для модели знаний  $\{S\} = (\{S^*\}^*_{Min})_2$  получим:

$$\{\tau/T\}_{\{S\}} = \{\tau_1/T2; \tau_1/T3; \tau_2/B3\}.$$

Для модели знаний  $\{S\} = (\{S^*\}^*_{Min})_3$  получим:

$$\{\tau/T\}_{\{S\}} = \{\tau_1/T2; \tau_2/B2; \tau_2/B3\}.$$

Если на  $\langle \Omega, \{G(\tau)\} \rangle$  моделей знаний  $\{S^*\}^*_{Min}$  несколько, то присвоим им произвольным образом индексы:  $(\{S^*\}^*_{Min})_1, \dots, (\{S^*\}^*_{Min})_q$ , где  $q = K^*$ . Соответственно, получим  $q$  наборов тестов:  $(\{\tau/T\}^*_{Min})_1, \dots, (\{\tau/T\}^*_{Min})_q$  или

$(\{\tau\}^*_{Min})_1, \dots, (\{\tau\}^*_{Min})_q$ . Не все наборы могут различаться. Решим следующую задачу: найдем множество индексов  $J$  такое, что

$$\forall j \in J, |(\{\tau\}^*_{Min})_j| = \min_{i=1, \dots, q} |(\{\tau\}^*_{Min})_i|. \quad (5.35)$$

Ясно, что для проведения классификации на  $\langle \Omega, \{G(\tau)\} \rangle$  с помощью предельных синдромных моделей знаний  $\{S^*\}^*_{Min}$  и с учетом минимизации числа тестов следует выбирать модели  $(\{S^*\}^*_{Min})_j$ , где  $j \in J$ .

Приведем две абсолютно минимальные предельные модели знаний для примера «Диагностика»:

$(\{S^*\}^*_{Min})_6 = \{S^*(\tau_1/T1? 36.0; z?1); S^*(\tau_1/T3? \text{Нормальная}; z?1); S^*((\tau_1/T2? \text{Повышенная}; z?2); S^*(\tau_1/T1? 36.4; z?2))\};$

$(\{T/T\}^*_{Min})_6 = \{\tau_1/T1; \tau_1/T2; \tau_1/T3\}; (\{\tau\}^*_{Min})_4 = \{\tau_1\} = \min;$

$(\{S^*\}^*_{Min})_7 = \{S^*(\tau_2/B1? 12; z?1); S^*(\tau_2/B2? \text{Средних лет}; z?1); S^*(\tau_2/B1? 32; z?2); \tau_2/B3? \text{Пенсионный}; z?2)\};$

$(\{T/T\}^*_{Min})_7 = \{\tau_1/B1; \tau_1/B2; \tau_1/B3\}; (\{\tau\}^*_{Min})_5 = \{\tau_2\} = \min.$

Для сравнения приведем тесты наилучшей по покрытию модели знаний  $(\{T/T\}^*_{Min})_2$ :

$(\{T/T\}^*_{Min})_2 = \{\tau_1/T2; \tau_1/T3; \tau_2/B3\}; (\{\tau\}^*_{Min})_2 = \{\tau_1; \tau_2\} = \max.$

Представим данные по покрытию прецедентов моделями знаний  $(\{T/T\}^*_{Min})_2, (\{T/T\}^*_{Min})_6$  и  $(\{T/T\}^*_{Min})_7$  в виде таблицы 5.6.

Таблица 5.6 – Покрытие прецедентов тремя минимальными моделями знаний

$\{T/T\} \setminus \alpha$	$(\{S^*\}^*_{Min})_2$				$(\{S^*\}^*_{Min})_6$				$(\{S^*\}^*_{Min})_7$			
	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$
«T1 – B1»	+	+	+	+	+	+	+	+	+	+	+	+
«T1 – B2»	+	+	+	+	+	+	+	+	-	+	-	+
«T1 – B3»	+	+	+	+	+	+	+	+	-	-	-	+
«T2 – B1»	+	+	+	+	-	+	+	-	+	+	+	+
«T2 – B2»	+	+	+	+	-	+	+	-	-	+	-	+
«T2 – B3»	+	+	+	+	-	+	+	-	-	-	-	+
«T3 – B1»	-	+	-	+	-	+	-	-	+	+	+	+
«T3 – B2»	-	+	-	+	-	+	-	-	-	+	-	+
«T3 – B3»	-	+	-	+	-	+	-	-	-	-	-	+

Как видим, модели знаний  $(\{S^*\}^*_{Min})_6$  и  $(\{S^*\}^*_{Min})_7$  содержат абсолютно

минимальное число тестов (по одному), но при этом значительно уступают модели знаний  $(\{S^*\}_{Min})_2$  по площади области покрытия (доле покрытия). Видно также, что модель  $(\{S^*\}_{Min})_2$  доминирует модель  $(\{S^*\}_{Min})_6$ , но не доминирует модель  $(\{S^*\}_{Min})_7$ .

В целом можно высказать следующее общее соображение: чем меньше число тестов используется в модели знаний, тем ниже уровень общности по каждому тесту и наоборот, чем большее число тестов используется, тем возможен более высокий уровень общности по каждому тесту.

В общем случае модели  $(\{S^*\}_{Min})^{\sim}$  не гарантируют использование минимального числа тестов.

Пусть  $\{\tau\}_S$  – множество тестов, которые фигурируют в синдроме  $S$ . Эти тесты содержатся в схеме синдрома  $S(\{\tau\}, \underline{z}/Z)$ . Множество всех тестов, которые фигурируют в произвольной синдромной модели знаний  $\{S\}$  определяется следующим образом:

$$\{\tau\}_{\{S\}} = \cup_{S \in \{S\}} \{\tau\}_S. \quad (5.36)$$

Для выполнения операции (5.36) можно использовать функцию  $Tests(\{S\})$ .

**Предложение 5.23.** Для разных бесконфликтных описаний базы прецедентов минимальные наборы тестов, достаточные для классификации на  $\Omega$ , могут отличаться.

В качестве доказательства приведем минимальные наборы тестов для следующих описаний (пример «Диагностика»):

Описание «Т1 – В1»: существуют два набора -  $\{\tau_1\}$  и  $\{\tau_2\}$ .

Описание «Т2 – В3»: существует один набор -  $\{\tau_1; \tau_2\}$ .

**Предложение 5.24.** Множества тестов модели знаний  $\{S\}$  и сопряженной предельной модели  $\{S^*\}$  совпадают.

Действительно,  $\{S^*\} = \cup_{S \in \{S\}} \{S^*\}_S$ , где  $\forall S, \{S^*\}_S \subseteq \{S_\theta\}_S, \theta \in \{\theta\}_S$ . По построению  $\{\theta\}_S$  любой кортеж доменов  $\theta$  не выходит за рамки  $\{\tau\}_S$ , что и доказывает предложение.

Пусть задана синдромная модель знаний  $\{S\}$ . Определим *минимальные по тестам* синдромные модели знаний  $\{S\}\# \subseteq \{S\}$  следующим образом

$$|\{\tau\}_{\{S\}\#}| = \min_{\{S'\} \subseteq \{S\}} |\{\tau\}_{\{S'\}}|. \quad (5.37)$$

Так как задача (5.37) может иметь не единственное решение, то важной задачей является поиск всех моделей  $\{S\}\# \subseteq \{S\}$ .

**Предложение 5.25.** Пусть  $\{S\}\#$  минимальна по тестам в рамках  $\{S\}$ , тогда  $\{S^*\}\#$  минимальна по тестам в рамках  $\{S^*\}$ .

Кроме синдромных моделей знаний выше были определены редуцированные модели знаний, которые строятся на базе синдромных моделей. Редуцированные модели, возможно, еще больше сокращают число используемых тестов. Это относится и к абсолютно минимальным,

предельным синдромным моделям знаний.

Построим *редуцированные по тестам* модели знаний. Пусть в произвольной синдромной базе знаний  $\{S\}$  заключению  $z_j$  отвечает подмножество синдромов  $\{S\}_j$  ( $j = 1, \dots, N$ ). Каждому  $\{S\}_j$  отвечает множество тестов

$$\{\tau\}_{\{S\}_j} = \cup_{S \in \{S\}_j} \{\tau\}_S. \quad (5.38)$$

Для каждого подмножества номеров  $I = \{i_1, \dots, i_{N-1}\} \subset Z$  определено объединение тестов:

$$\{\tau\}_I = \cup_{j \in I} \{\tau\}_{\{S\}_j}. \quad (5.39)$$

Найдем такое подмножество номеров  $I^\# \subset Z$ , которое обеспечивает решение следующей задачи:

$$|\{\tau\}_{\{S\}}^\#| = \min_{I \subset Z} |\{\tau\}_I|, \text{ при условии } |\{\tau\}_{\{S\}}^\#| < |\{\tau\}_{\{S\}}|. \quad (5.40)$$

Из-за ограничения решение задачи (5.40) может не существовать. Отсутствие решения означает, что за счет редуцирования не удастся уменьшить число тестов. Отметим, что редуцированная синдромная модель знаний существует всегда.

Для произвольной синдромной модели знаний  $\{S\} = \cup_{j \in Z} \{S\}_j$  определим *редуцированную по тестам модель знаний*  $\{S\}^\#$  следующим образом:

$$\{S\}^\# = \cup_{j \in I^\#} \{S\}_j, \quad (5.41)$$

$$z_k, \text{ где } k = Z \setminus I^\# \Leftrightarrow \&_{j \in I^\#} \neg z_j = \text{true}, \quad (5.42)$$

где множество  $I^\#$  определяются путем решения задачи (5.40). Правило (5.42) позволяет установить заключение  $z_k$  в том случае, когда доказано отсутствие всех заключений из  $I^\#$ . Модель  $\{S\}^\#$  существует только тогда, когда существует решение задачи (5.40).

Строя редуцированные модели для всех моделей  $\{S\}' \subseteq \{S\}$ , можно найти решение следующей задачи:

$$|\{\tau\}_{\{S\}}^\#| = \min_{\{S\}' \subseteq \{S\}} |\{\tau\}_{\{S\}' }^\#|, \quad (5.43)$$

$$\{\{S\}^\#\} = \operatorname{argmin}_{\{S\}' \subseteq \{S\}} |\{\tau\}_{\{S\}' }^\#|. \quad (5.44)$$

В качестве  $\{S\}$  может быть выбрана, например,  $\{S\}_{Full}$  в рамках любого описания  $\Omega(\{T\})$ .

Приведем пример построения максимально редуцированных по тестам моделей знаний  $\{\{S\}^\#\}$ . В таблице 5.7 представлена простая схема синдромной модели знаний.



Таблица 5.7 – Пример схемы синдромной модели знаний

$S \setminus \tau$	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$z$	$\Omega_1$	$\Omega_2$
$S_1$	+	+				1	1	1
$S_2$		+		+		2	2	2
$S_3$			+			3	3	3
$S_4$				+		3	4	4, 5
$S_5$			+		+	3	5	5

Дано:  $\{S\} = \{S_1; S_2; S_3; S_4; S_5\}$  и две базы прецедентов:  $\Omega_1$  и  $\Omega_2$ . Необходимо найти все максимально редуцированные по тестам модели знаний  $\{\{S\}^{\#\#\}$  для каждой из баз прецедентов.

Легко убедиться, что для базы прецедентов  $\Omega_1$  существует единственная максимально редуцированная по тестам модель  $\{S\}^{\#\#}$ .

$$\begin{aligned} \{\tau\}_{\{S\}1} &= \{\tau_1; \tau_2\}; \{\tau\}_{\{S\}2} = \{\tau_2; \tau_4\}; \{\tau\}_{\{S\}3} = \{\tau_3; \tau_4; \tau_5\}; \\ I^{\#} &= \{1, 2\}; \{\tau^{\#}\}_{\{S\}} = \{\tau_1; \tau_2; \tau_4\}; |\{\tau^{\#}\}_{\{S\}}| = |\{\tau^{\#}\}_{\{S\}}| = 3; \\ \{S\}^{\#} &= \{S\}^{\#\#} = \{S_1; S_2\} \ \& \ (z_3 \leftrightarrow \neg z_1 \ \& \ \neg z_2 = true). \end{aligned}$$

Таким образом, всего трех тестов из пяти достаточно, чтобы провести полную классификацию на  $\Omega_1$ .

Для базы прецедентов  $\Omega_2$  существуют уже две модели  $\{S\}^{\#\#}$ . Первая модель описана выше, вторая модель имеет вид:

$$\begin{aligned} \{S\}^{\#\#} &= \{S_2; S_3; S_4\} \ \& \ (z_1 \leftrightarrow \neg z_2 \ \& \ \neg z_3 = true); \\ I^{\#} &= \{2, 3\}; \{\tau^{\#}\}_{\{S\}} = \{\tau_2; \tau_3; \tau_4\}; |\{\tau^{\#}\}_{\{S\}}| = 3. \end{aligned}$$

Отметим, что число синдромов в обеих моделях  $\{S\}^{\#\#}$  разное. Что касается сути оптимизации, то первая и вторая модели отличаются только тестами  $\tau_1$  и  $\tau_2$ . Если необходимо выбрать одну из двух моделей, то следует привлекать для анализа дополнительные характеристики тестов  $\tau_1$  и  $\tau_2$ , например, стоимость или сроки выполнения.

**Предложение 5.26.** Все абсолютно минимальные наборы тестов, которые достаточны для решения задачи классификации на  $\Omega$ , определяются из решения задач (5.40) – (5.43) для  $\{S\} = \{S^*\}_{Full}$ . Минимальное количество тестов определяется величиной  $|\{\tau^{\#}\}_{\{S\}}|$ .

Величина  $L^{\#} = |\{\tau^{\#}\}_{\{S\}}|$ , где  $\{S\} = \{S^*\}_{Full}$ , является важной константой на  $\langle \Omega, \{G(\tau)\} \rangle$ . Она показывает, каково минимальное число тестов, которое необходимо и достаточно для полной классификации на базовом уровне всех прецедентов из  $\Omega$ .

Поиск всех минимальных наборов тестов  $\{\tau^{\#}\}_{\{S\}}$  и соответствующих редуцированных моделей знаний  $\{\{S\}^{\#\#\}$  является важной задачей в рамках метода предельных обобщений.

Рассмотрим вопрос о передаче моделей знаний. Пусть  $\{\pi T\}_{\{S\}}$  наборы тестов какой-либо модели знаний  $\{S\}$ , например,  $\{S^*\}_{Min-Full}$ . Сократим число тестов и редуцируем орграфы банка тестов  $\{G(\tau)\}$  таким образом, чтобы в новом банке тестов  $\{G'(\tau)\}$  присутствовали только орграфы доменов для тестов  $\{\tau\}_{\{S\}}$  и домены из набора  $\{\pi T_0\} \cup \{\pi T\}_{\{S\}}$ . Такую редукцию всегда можно выполнить, например, из орграфа  $G(\tau) = \{T_0 \rightarrow \dots \rightarrow T^*\}$  получить орграф  $G'(\tau) = \{T_0 \rightarrow T^*\}$ . При сокращении числа тестов с  $\{\tau\}$  до  $\{\tau\}_{\{S\}}$  соответственно упрощается и описание прецедентов. В результате получаем новое – более простое описание банка прецедентов  $\Omega'$ . Обучаемому агенту передается когнитивная структура  $\langle \{G'(\tau)\}, \{S\} \rangle$ . Достаточность такой информации подтверждается следующим предложением.

**Предложение 5.27.** На  $\langle \Omega, \{G(\tau)\} \rangle$  и  $\langle \Omega', \{G'(\tau)\} \rangle$  существуют совпадающие модели знаний  $\{S\}$ .

Таким образом, после того как определены (инвариантные) синдромные модели знаний на  $\langle \Omega, \{G(\tau)\} \rangle$ , орграфы банка тестов  $\{G(\tau)\}$  и описание банка прецедентов могут быть значительно редуцированы с сохранением найденных моделей знаний. Данный факт наглядно подтверждает известную истину: «Все гениальное просто!». Действительно, для нахождения инвариантных моделей знаний требовался развернутый банк тестов и, зачастую, избыточное (по тестам) описание базы прецедентов, но после того как модели установлены, их использование может опираться на значительно более простые когнитивные структуры. Данный факт играет ключевую роль при обучении и передаче знаний.

Отметим, что передать одну базу знаний  $\{S\}$  без соответствующего банка тестов нельзя, так как исходные данные о новом прецеденте могут иметь уровень общности, отличающийся от синдромов модели  $\{S\}$ , и необходимо использовать банк тестов для соответствующих преобразований. Для примера рассмотрим один частный, но важный случай.

**Предложение 5.28.** Пусть орграфы всех тестов из  $\{G(\tau)\}$  содержат лишь одну терминальную вершину и описание базы прецедентов  $\Omega(\{\pi T^*\})$  не содержит конфликтов, тогда все орграфы доменов банка тестов  $\{G'(\tau)\}$  имеют вид  $G'(\tau) = \{T_0 \rightarrow T^*\}$ , а модель знаний  $\{S\}$  принадлежит описанию  $\{\pi T^*\}_{\{S\}}$ .

## 5.2 Применение синдромных и вероятностных моделей знаний

Опишем алгоритм применения моделей знаний к новым ситуациям действительности (новым прецедентам).

Пусть  $\alpha(\{\underline{t}/T\}, z/Z?)$  – новая ситуация действительности, для которой нужно установить заключение. Строим замыкание  $\{\underline{t}/T\}^+$  с использованием банка тестов  $\{G(\tau)\}$ . Далее берем любую модель знаний в рамках описания  $\Omega(\{\tau/T\}')$  такого, что  $\{\tau/T\}' \geq \{\underline{t}/T\}$ , или полную предельную модель знаний  $\{S^*\}_{Full}$  в рамках контекста  $\langle \Omega, \{G(\tau)\} \rangle$  и проверяем для  $\alpha$  наличие соответствующих синдромов. Если все установленные синдромы говорят об одном и том же заключении  $\underline{z}/Z$ , то следует принять именно это заключение. Если же найдутся два синдрома, которые выдают разные заключения, то это означает, что  $\alpha$  фальсифицирует выбранную модель знаний. Такую модель следует удалить, а еще лучше построить заново все модели знаний после независимой верификации заключения для  $\alpha$  и добавления  $\alpha$  в  $\Omega$ . Возможна также ситуация, при которой ни один синдром выбранной модели знаний не дал результата. В этом случае прецедент  $\alpha$  также фальсифицирует выбранную модель знаний.

Если появление значительного числа новых прецедентов не фальсифицирует выбранную синдромную модель знаний, то такая модель претендует на статус *инвариантной модели* или *истинной модели*. Статус инвариантной модели коррелирует с понятием *устойчивости модели*.

Простейший алгоритм классификации нового прецедента следующий (предполагается, что  $Z$  не содержит кода '0').

### Алгоритм 5.8

Вход: Прецедент  $\alpha(\{\underline{t}/T\}, z/Z?)$ , модель знаний  $\{S\}$ .

Выход:  $z_\alpha$

Онтология: Банк тестов  $\{G(\tau)\}$ .

*Classif*( $\alpha, \{S\}$ )

**begin**

$z_\alpha := 0; \{\underline{t}/T\}^+;$

**for**  $S(\{\underline{t}/T\}', \underline{z}/Z) \in \{S\}$  **do**

**if**  $(\{\underline{t}/T\}'_\alpha \subseteq \{\underline{t}/T\}^+) \& (\{\underline{t}/T\}'_\alpha = \{\underline{t}/T\}')$  **then**

**begin**  $z_\alpha := \underline{z}/Z$ ; **return** ( $z_\alpha$ ) **end**;

**return** ( $z_\alpha$ )

**end.**

Проверка условия  $(\{\underline{t}/T\}'_\alpha \subseteq \{\underline{t}/T\}^+)$  означает следующее: синдром  $S(\{\underline{t}/T\}', \underline{z}/Z)$  задает уровень общности  $\{\tau/T\}'$ ; если в замыкании  $\{\underline{t}/T\}^+$  имеются данные такого уровня, то условие истинно, иначе – ложно.

Если функция *Classif* выдает код '0', то это означает, что новый прецедент  $\alpha$  фальсифицирует модель знаний  $\{S\}$ , так как ни один синдром не применим к ситуации  $\alpha$ . Приведенный выше алгоритм не позволяет установить наличие конфликта, при котором имеет место несколько

конкурирующих заключений.

Если функция *Classif* выдает код '0', то проблема устраняется достаточно просто: необходимо установить все синдромы  $\{S\}_\alpha$  и выполнить следующие две операции:

$$\Omega := \Omega \cup \alpha; \quad \{S\} := \{S\} \cup \{S\}_\alpha.$$

Приведенный ниже алгоритм позволяет исключить оба вида фальсификаций: как отсутствие заключений, так и появление нескольких конкурирующих заключений.

### Алгоритм 5.9

Вход: Прецедент  $\alpha(\{z/T\}, z/Z?)$ , модель знаний  $\{S\}$ .

Выход:  $z_\alpha$ .

Онтология: Банк тестов  $\{G(\tau)\}$ .

*Classif\_2*( $\alpha, \{S\}$ )

**begin**

$Z_\alpha := \emptyset; \quad \{z/T\}^+;$

**for**  $S(\{z/T\}', z/Z) \in \{S\}$  **do**

**if**  $(\{z/T\}'_\alpha \subseteq \{z/T\}^+) \ \& \ (\{z/T\}'_\alpha = \{z/T\}')$  **then**  $Z_\alpha := Z_\alpha \cup z/Z;$

**if**  $|Z_\alpha| = 1$  **then**  $z_\alpha := z/Z$  **else**  $z_\alpha := 0;$

**return** ( $z_\alpha$ )

**end.**

Недостатком алгоритма 5.9 является необходимость проведения всех тестов, которые фигурируют в синдромах модели знаний  $\{S\}$ .

Приведем рациональный вариант алгоритма классификации нового прецедента, в котором проверяются не все возможные синдромы, а только те из оставшихся, которые совместимы с уже найденными.

Совместимость синдромов  $S$  и  $S'$  означает, что у некоторого прецедента  $\alpha$  могут быть одновременно синдромы  $S$  и  $S'$ . Несовместимость синдромов  $S$  и  $S'$  означает, что у любого прецедента  $\alpha$  не могут иметь место одновременно синдромы  $S$  и  $S'$ . Однако, несмотря на кажущуюся простоту понятия «совместимость синдромов», дать формальное определение совместимости достаточно непросто, так как в текущей базе прецедентов могут отсутствовать прецеденты, имеющие одновременно синдромы  $S$  и  $S'$  даже если они совместимы.

Множество синдромов назовем *совместимым*, если в базе прецедентов  $\Omega$  имеется прецедент, содержащий все синдромы из множества.

Соответственно, произвольное множество синдромов  $\{S\}$  назовем *условно совместимым*, если для него не выполняется условие совместимости, но для любого теста  $\tau$ , принадлежащего одновременно  $\{\tau\}_S$

и  $\{\tau\}_S$  ( $\{S, S^*\} \subseteq \{S\}$ ), выполняется хотя бы одно из преобразований:  $\underline{d}/T \rightarrow \underline{d}/T'$  или  $\underline{d}/T' \rightarrow \underline{d}/T$ . Ясно, что если домены  $T'$  и  $T$  совпадают, то и значения должны совпадать. Вместе с тем, между доменами  $T'$  и  $T$  может не быть даже отношения доминирования. Если  $\{\tau\}_S \cap \{\tau\}_{S'} = \emptyset$ , то такие синдромы также будем считать условно совместимыми. Очевидно, из условной совместимости не всегда следует совместимость.

Произвольное множество синдромов назовем *несовместимым*, если оно не является совместимым и не является условно совместимым.

В рамках  $\{S\}_{Full}$  на  $\Omega(\{\tau_1/T2, \tau_2/B3\})$  имеются два синдрома:

$$S_3 = (\tau_1/T2? \text{Повышенная} \rightarrow (z=2)); S_4 = (\tau_2/B3? \text{Пенсионный} \rightarrow (z=2)).$$

Так как в базе прецедентов  $\Omega(\{\tau_1/T2, \tau_2/B3\})$  отсутствует прецедент  $\alpha(\tau_1/T2? \text{Повышенная}; \tau_2/B3? \text{Пенсионный}; z=2)$ , то синдромы  $S_3$  и  $S_4$  являются условно совместимыми.

В рамках  $\{S\}_{Full}$  на  $\Omega(\{\tau_1/T3, \tau_2/B1\})$  имеются три синдрома:

$$S_1 = (\tau_1/T3? \text{Нормальная} \rightarrow (z=1)); S_2 = (\tau_2/B1? 12 \rightarrow (z=1));$$

$$S_3 = (\tau_2/B1? 50 \rightarrow (z=1)).$$

Легко убедиться, что  $S_1$  совместим с  $S_3$  и условно совместим с  $S_2$ . Синдромы  $S_2$  и  $S_3$  несовместимы.

Если база прецедентов отсутствует, а имеется только модель знаний, то к синдромам последней будем применять исключительно понятие «условной совместимости». Приведем пример. Пусть множество синдромов для заключения  $\underline{z}$  представлено таблицей 5.8.

Таблица 5.8 – Пример таблицы синдромов

$S \setminus \tau$	$a/A$	$b/B$	$c/C$	$z/Z$
$S_1$	1	6		$\underline{z}$
$S_2$	1		8	$\underline{z}$
$S_3$		9	8	$\underline{z}$
$S_4$	2	10		$\underline{z}$

Условно совместимыми являются только пары синдромов  $\{S_1; S_2\}$  и  $\{S_2; S_3\}$ . Синдром  $S_4$  несовместим ни с одним из других синдромов. Например, синдромы  $S_4$  и  $S_1$  несовместимы потому, что у любого прецедента  $\alpha$  не могут быть одновременно  $a/A?1$  и  $a/A?2$  (по предположению каждый тест входит в описание прецедента только один раз).

Отношения совместимости и условной совместимости между синдромами не являются транзитивными. Действительно,  $S_1$  условно совместим с  $S_2$ ,  $S_2$  условно совместим с  $S_3$ , но  $S_1$  несовместим с  $S_3$ , так как у прецедента не могут быть одновременно  $b/B?6$  и  $b/B?9$ .

Пусть, как и ранее  $\{S\}_j$  – множество всех синдромов, отвечающих заключению  $z_j$  ( $j \in Z$ ). Множество синдромов  $\{S\} \subseteq \{S\}_j$  назовем *предельной группой (условно) совместимых синдромов*, если все синдромы из  $\{S\}$  совместимы или условно совместимы и в  $\{S\}$  нельзя добавить ни один новый синдром из  $\{S\}_j$ . Предельная группа  $\{S\}$  может совпадать с  $\{S\}_j$ . Ясно, что различных предельных групп (условно) совместимых синдромов для каждого  $j \in Z$  конечное число. Разные предельные группы могут пересекаться между собой. Проиндексируем все предельные группы произвольным образом:  $\{S\}_j^p$ , где  $p = 1, \dots, P_j$ . Можно записать:

$$\forall j \in Z, \quad \{S\}_j = \cup_{p=1, \dots, P_j} \{S\}_j^p. \quad (5.45)$$

Таким образом, предельные группы (условно) совместимых синдромов образуют своеобразное покрытие множества  $\{S\}_j$  ( $j \in Z$ ).

Для примера из таблицы 5.2 имеются всего три предельные группы (условно) совместимых синдромов:

$$\{S\}_{\bar{z}} = \{S_1; S_2\} \cup \{S_2; S_3\} \cup \{S_4\}.$$

Видно, что каждый из синдромов  $S_1$ ,  $S_3$  и  $S_4$  входит в единственную группу, а синдром  $S_2$  входит в две группы.

В рамках  $\{S\}_{Full}$  на  $\Omega(\{\tau_1/T3, \tau_2/B1\})$  множество синдромов  $\{S\}_{z=1}$  представимо в виде объединения двух предельных групп:

$$\{S\}_{z=1} = \{S_1; S_2\} \cup \{S_1; S_3\}.$$

Для построения рационального алгоритма важно знать, что у нового прецедента  $\alpha$  могут быть одновременно все синдромы из любой группы  $\{S\}_j^p$ , а также любое их подмножество и не могут быть одновременно синдромы из разных предельных групп (кроме пересечения).

Пусть к определенному моменту времени для прецедента  $\alpha$  установлены синдромы  $\{S\}_{\alpha}$ . Область поиска новых возможных синдромов ограничена множеством

$$\{S(z_j)\}_{\alpha}^{\perp} = \cup_p (\{S\}_j^p \setminus \{S\}_{\alpha} \subseteq \{S\}_j^p) \setminus \{S\}_{\alpha}. \quad (5.46)$$

К примеру, если обнаружен синдром  $S_1$  для модели из таблицы 5.8, то проверять следует только синдром  $S_2$ .

Конечной целью разрабатываемого алгоритма является поиск всех имеющихся у прецедента синдромов  $\{S\}_{\alpha}$ . Целесообразность поиска всех имеющихся синдромов можно объяснить стремлением повысить достоверность заключения. Алгоритм 5.10 решает поставленную задачу в предположении, что используемая синдромная база знаний истинна (это означает, что прецедент не может ее фальсифицировать). Будем также

предполагать для базы знаний известным представление (5.45).

### Алгоритм 5.10

Вход: Прецедент  $\alpha(\{\underline{z}/T\}, z/Z?)$ , модель знаний  $\{S\}$ .

Выход:  $\{S\}_\alpha$ .

Онтология: Банк тестов  $\{G(\tau)\}$ .

$Classif\_3(\alpha, \{S\})$

**begin**

$\{S\}_\alpha := \emptyset; \{S\}_i := \emptyset; \{\underline{z}/T\}^+;$

// Находим  $z_\alpha$

**for**  $S(\{\underline{z}/T\}^?, z/Z) \in \{S\}$  **do begin**

$\{S\}_i := \{S\}_i \cup S(\{\underline{z}/T\}^?, z/Z);$

**if**  $(\{\underline{z}/T\}^? \subseteq \{\underline{z}/T\}^+) \& (\{\underline{z}/T\}^? \subseteq \{\underline{z}/T\}^?)$  **then**

**begin**  $z_\alpha := z/Z; \{S\}_\alpha := \{S\}_\alpha \cup S(\{\underline{z}/T\}^?, z/Z);$  **exit end end;**

// Находим все синдромы  $\{S\}_\alpha$

$j := z_\alpha;$

$\{S\}_\alpha^\perp := \cup_p (\{S\}_j^p \mid \{S\}_\alpha \subseteq \{S\}_j^p) \setminus \{S\}_i;$

**while**  $\{S\}_\alpha^\perp \neq \emptyset$  **do**

**for**  $S(\{\underline{z}/T\}^?, z_\alpha) \in \{S\}_\alpha^\perp$  **do begin**

$\{S\}_i := \{S\}_i \cup S(\{\underline{z}/T\}^?, z_\alpha);$

**if**  $(\{\underline{z}/T\}^? \subseteq \{\underline{z}/T\}^+) \& (\{\underline{z}/T\}^? \subseteq \{\underline{z}/T\}^?)$  **then begin**

$\{S\}_\alpha := \{S\}_\alpha \cup S(\{\underline{z}/T\}^?, z_\alpha);$

$\{S\}_\alpha^\perp := \cup_p (\{S\}_j^p \mid \{S\}_\alpha \subseteq \{S\}_j^p) \setminus \{S\}_i;$  **exit end end;**

**return**  $(\{S\}_\alpha)$

**end.**

В главе 7 будет введена операция ультраоснащения синдромных моделей знаний. Данная операция позволит существенно оптимизировать алгоритм установления заключения для произвольного прецедента  $\alpha$ .

Вероятностные модели знаний  $\{R\}$  и особенно  $\{R^*\}$  можно рассматривать как предвестники событий. Во многих случаях вероятностные закономерности можно применить тогда, когда еще нельзя установить ни один синдром (из-за нехватки данных). Особенно важную роль играют ранние предвестники событий, так как они позволяют заблаговременно установить риск неблагоприятного развития ситуации (например, инсульта, инфаркта или отказа техники).

Изменения банка тестов, непрерывное появление новых прецедентов и возможная периодическая фальсификация моделей знаний приводят к явлению когнитивной самоорганизации.

Сократим число тестов и редуцируем орграфы банка тестов  $\{G(\tau)\}$  таким образом, чтобы в новом банке тестов  $\{G^*(\tau)\}^*$  присутствовали только

орграфы доменов для тестов  $\{\tau\}_{\{S^*\}}$  и  $\{\tau\}_{\{R^*\}}$ , а также домены из набора  $\{\tau/T_0\} \cup \{\tau/T\}_{\{S^*\}} \cup \{\tau/T\}_{\{R^*\}}$ . Для каждой  $Z$ -задачи, определяемой  $\Omega(Z)$ , имеет место свой критический банк тестов  $\{G^*(\tau)\}^*$ .

Под *когнитивной самоорганизацией* в рамках  $Z$ -задачи будем понимать, в частности, последовательность скачкообразных переходов

$$\{S^*\}_1, \{R^*\}_1, \{G^*(\tau)\}_1^* \dots \rightarrow \{S^*\}_n, \{R^*\}_n, \{G^*(\tau)\}_n^* \quad (5.47)$$

которая возникает вследствие изменения контекста  $K = \langle \Omega, \{G(\tau)\} \rangle$ , в частности, фальсификации текущих моделей знаний. Модели  $\{S^*\}_j, \{R^*\}_j, \{G^*(\tau)\}_j^*$  являются примерами самоорганизующихся критических структур в рамках любой  $Z$ -задачи.

К результатам когнитивной самоорганизации относятся также экстремальные пограничные слои набросков (см. главу 6) и критические пути в вычислительных средах (см. глава 7).

### 5.3 Критические описания базы прецедентов. Синдромные модели знаний для критических описаний

Пусть имеются база прецедентов  $\Omega_0 = \{\alpha(\{\tau/T_0\}, \underline{z}/Z)\}$ , где  $Z = \{1, \dots, N\}$  и банк тестов  $\{G(\tau)\}$ . В рамках любого описания  $\{\tau/T\}$ , допускаемого банком тестов  $\{G(\tau)\}$ , база прецедентов будет иметь вид

$$\Omega(\{\tau/T\}, Z) = \{\alpha(\{\tau/T\}, \underline{z}/Z)\}.$$

Другими словами, если задано описание  $\{\tau/T\}$ , то все прецеденты имеют одинаковый уровень общности, задаваемый  $\{\tau/T\}$ . В общем случае, каждый прецедент в базе прецедентов может иметь собственное описание, что при необходимости будем отражать нотацией  $\Omega = \{\alpha(\{\tau/T\}_\alpha, \underline{z}/Z)\}$ .

Так как любое  $\{\tau/T\} \geq \{\tau/T_0\}$ , то всегда существует однозначное преобразование  $\Omega(\{\tau/T_0\}, Z) \rightarrow \Omega(\{\tau/T\}, Z)$ . В общем случае, если  $\{\tau/T'\} \geq \{\tau/T\}$ , то всегда существует однозначное преобразование

$$\Omega(\{\tau/T\}, Z) \rightarrow \Omega(\{\tau/T'\}, Z).$$

Поставим следующую задачу: найти все предельные описания базы прецедентов, которые не содержат конфликтов, но любое доминирующее описание содержит конфликты. Такие описания назовем критическими. Дадим точные определения.

Описание  $\{\tau/T\}$  называется *критическим*, если база прецедентов  $\Omega(\{\tau/T\}, Z)$  не содержит конфликтов, но любое описание базы прецедентов  $\Omega(\{\tau/T'\}, Z)$  такое, что  $\{\tau/T'\} \geq \{\tau/T\}$  содержит конфликты. Если доминирующих описаний  $\{\tau/T'\}$  нет, то описание  $\{\tau/T\}$  также является критическим. Критические описания будем обозначать нотацией  $\{\tau/T\}^*$ .



Описание  $\{\tau/T\}$  называется *докритическим*, если база прецедентов  $\Omega(\{\tau/T\}, Z)$  не содержит конфликтов и найдется описание базы прецедентов  $\Omega(\{\tau/T\}', Z)$  такое, что  $\{\tau/T\}' \geq \{\tau/T\}$ , которое также не содержит конфликтов.

Описание  $\{\tau/T\}$  называется *закритическим* (или *надкритическим*), если база прецедентов  $\Omega(\{\tau/T\}, Z)$  содержит конфликты.

В определенном смысле критические описания являются границей между порядком (однозначностью интерпретаций) и хаосом (многозначностью интерпретаций, т.е. наличием артефактов). Граница эта неустойчива и может изменяться с изменением базы прецедентов или онтологии (банка тестов).

В примере «Диагностика» из 9 описаний только два описания являются критическими. Это описания «Г2 – В3» и «Г3 – В1». Закритических описаний также два: «Г3 – В2» и «Г3 – В3». Оставшиеся 5 описаний являются докритическими.

**Предложение 5.29.** Если описание  $\{\tau/T\}$  является закритическим, то любое описание  $\{\tau/T\}'$  такое, что  $\{\tau/T\}' \geq \{\tau/T\}$ , также является закритическим. Если описание  $\{\tau/T\}$  является критическим или докритическим, то любое описание  $\{\tau/T\}'$  такое, что  $\{\tau/T\} \geq \{\tau/T\}'$ , является докритическим.

Суть предложения 5.29 состоит в том, что если некоторое описание действительности (базы прецедентов) содержит артефакты, то дальнейшее обобщение по любому из тестов не устраняет конфликтность, хотя сами артефакты, естественно, изменяются. И наоборот, если некоторое описание не содержит артефакты, то повышение точности описания (понижение уровня общности) по любому из тестов не приведет к появлению артефактов, а значит и конфликтности описания.

**Предложение 5.30.** Если базовое описание  $\Omega(\{\tau/T_0\}, Z)$  неконфликтно, то критические описания на  $\langle \Omega, \{G(\tau)\} \rangle$  существуют.

Пусть  $\{\tau/T\}$  – закритическое описание. Множество всех артефактов данного описания обозначим через  $\{\beta\}_{\{\tau/T\}}$ , где

$$\beta = (\{\underline{\tau}/T\}_{\beta}, \{\underline{z}/Z(c_z)\}_{\beta}), \text{ genus } \beta = |\{\underline{z}/Z\}_{\beta}| - 1. \quad (5.48)$$

Для любого артефакта  $|\{\underline{z}/Z\}_{\beta}| \geq 2$ . Нотация ' $\{\underline{z}/Z(c_z)\}_{\beta}$ ' означает мультимножество, в котором  $c_z$  – количество прецедентов с заключением  $\underline{z}/Z$ , входящих в артефакт  $\beta$ . На любой артефакт можно сослаться следующим образом:  $\beta(\{\underline{\tau}/T\}, \{\underline{z}/Z(c_z)\})$  или  $\beta(\{\underline{\tau}/T\}, \{\underline{z}/Z\})$ .

В рамках закритического описания «Г3 – В2» имеем один артефакт:

$$\beta = (\{\tau_1/\text{Г3? Ненорм.}, \tau_2/\text{В2? Молодой}\}, \{z?1(1); z?2(1)\}), \text{ genus } \beta = 1.$$

В рамках закритического описания «Г3 – В3» имеем также один артефакт:

$\beta = (\{\tau_1/T_3? \text{ Ненорм.}, \tau_2/B_3? \text{ Допенс.}\}, \{z?1 (1); z?2 (1)\})$ ,  $\text{genus } \beta = 1$ .

**Предложение 5.31.** Пусть  $\{\tau/T\}$  и  $\{\tau'/T'\}$  – два закритических описания, причем  $\{\tau'/T'\} \geq \{\tau/T\}$ , соответственно  $\{\beta\}_{\{\tau/T\}}$  и  $\{\beta'\}_{\{\tau'/T'\}}$  – множества артефактов, тогда справедливо следующее:

- (i)  $\forall \beta \exists ! \beta': \{\tau'/T'\}_{\beta} \rightarrow \{\tau/T\}'_{\beta'}$ ,  $\{\underline{z}/Z\}_{\beta} \subseteq \{\underline{z}/Z\}_{\beta'}$ , т.е.  $\text{genus } \beta' \geq \text{genus } \beta$ ;
- (ii) Если  $\beta_1 \rightarrow \beta'$  и  $\beta_2 \rightarrow \beta'$ , тогда

$$(\{\underline{z}/Z (c_z)\}_{\beta_1} \cup \{\underline{z}/Z (c_z)\}_{\beta_2}) \subseteq \{\underline{z}/Z (c_z)\}_{\beta'}; \tag{5.49}$$

- (iii) Могут существовать  $\alpha \in \Omega(\{\tau/T\}) \setminus \{\beta\}_{\{\tau/T\}}$ :  $\alpha \rightarrow \beta'$ .

Таким образом, с повышением уровня обобщения закритических описаний имеющиеся артефакты не могут исчезнуть, при этом род артефактов также не может уменьшиться. Кроме того, некоторые прецеденты могут перейти в разряд артефактов. В выражении (5.49) объединение выполняется по правилу объединения мультимножеств. Пример применения правила:

$$\{a (2), b (3), c\} \cup \{a (3), b\} = \{a (5), b (4), c\}.$$

Следует отметить, что в ряде случаев с повышением уровня обобщения общее количество артефактов может уменьшаться. Это связано с тем, что артефакты при обобщении могут объединяться. Приведем пример.

В таблице 5.9 приведена база прецедентов в рамках закритического описания  $\{a/A, b/B\}$ .

Таблица 5.9 – Пример закритического описания  $\Omega(\{a/A, b/B\})$

$\alpha$	$a/A$	$b/B$	$z$
1	a1	b1	1
2	a1	b1	2
3	a2	b2	1
4	a2	b2	3

В рамках описания  $\Omega(\{a/A, b/B\})$  имеем два артефакта первого рода:

$$\beta_1 = (\{a/A? a1, b/B? b1\}_1, \{z?1 (1); z?2 (1)\}_1), \text{genus } \beta_1 = 1;$$

$$\beta_2 = (\{a/A? a2, b/B? b2\}_2, \{z?1 (1); z?3 (1)\}_2), \text{genus } \beta_2 = 1.$$

Представим фрагменты конфигураторов тестов  $a$  и  $b$ :

$$a \{A' \{a \wedge a1 a2; \dots\} A \{a1; a2; a3; \dots\} A_0 \{\dots\}\};$$

$$b \{B' \{b \wedge b1 b2; \dots\} B \{b1; b2; b3; \dots\} B_0 \{\dots\}\}.$$

$$G(a) = \{A_0 \rightarrow A \rightarrow A'\}, \quad G(b) = \{B_0 \rightarrow B \rightarrow B'\}.$$

Из конфигураторов следует:  $A.\{a1; a2\} \rightarrow A'.a$ ;  $B.\{b1; b2\} \rightarrow B'.b$ . Ясно также, что  $\{a/A', b/B'\} > \{a/A, b/B\}$ , т.е. имеет место строгое

доминирование. При переходе от описания  $\Omega(\{a/A, b/B\})$  к описанию  $\Omega(\{a/A', b/B'\})$  оба артефакта  $\beta_1$  и  $\beta_2$  переходят в один артефакт второго рода  $\beta$  следующего вида:

$$\beta = (\{a/A' \text{ ? } a, b/B' \text{ ? } b\}, \{z?1 (2); z?2 (1); z?3 (1)\}), \text{genus } \beta = 2.$$

При переходе от описания  $\Omega(\{a/A, b/B\})$  к доминирующему описанию  $\Omega(\{a/A', b/B'\})$  артефакт  $\beta_1$  переходит в артефакт  $\beta'_1$ , а артефакт  $\beta_2$  переходит в артефакт  $\beta'_2$ :

$$\beta'_1 = (\{a/A' \text{ ? } a, b/B' \text{ ? } b1\}_1, \{z?1 (1); z?2 (1)\}_1), \text{genus } \beta'_1 = 1;$$

$$\beta'_2 = (\{a/A' \text{ ? } a, b/B' \text{ ? } b2\}_2, \{z?1 (1); z?3 (1)\}_2), \text{genus } \beta'_2 = 1.$$

При переходе от описания  $\Omega(\{a/A', b/B'\})$  к доминирующему описанию  $\Omega(\{a/A', b/B'\})$  оба артефакта  $\beta'_1$  и  $\beta'_2$  переходят в один артефакт второго рода  $\beta$ . В серии обобщений  $\Omega(\{a/A, b/B\}) \rightarrow \Omega(\{a/A', b/B'\}) \rightarrow \Omega(\{a/A', b/B'\})$  множества артефактов будут преобразовываться следующим образом:  $\{\beta_1, \beta_2\} \rightarrow \{\beta'_1, \beta'_2\} \rightarrow \{\beta\}$ .

Из предложения 5.31 могут быть сделаны два важных следствия.

Определим множество прецедентов участвующих в артефактах в рамках описания  $\Omega(\{\tau/T\}, Z)$  следующим образом:

$$\{\alpha\}_\beta = \{\alpha \in \Omega(\{\tau/T\}, Z) \mid \exists \beta \in \{\beta\}_{\{\tau/T\}}: \{\tau/T\}_\alpha = \{\tau/T\}_\beta\}. \quad (5.50)$$

В ряде случаев может быть полезно числовое выражение доли прецедентов, участвующих в артефактах (в %):  $100 \cdot |\{\alpha\}_\beta|/|\Omega|$ .

**Следствие 5.3.** Доля прецедентов участвующих в артефактах с повышением уровня общности не уменьшается.

Данное следствие напрямую вытекает из заключения (iii) предложения 5.31.

Пусть  $\{T^*\}_\tau$  – множество терминальных вершин в орграфе  $G(\tau)$ , тогда любое описание базы прецедентов  $\Omega(\{\tau/T^*\})$  назовем *финальным*. Так как каждый тест входит в описание один раз, то общее количество финальных описаний базы прецедентов определяется выражением:

$$|\{\Omega(\{\tau/T^*\})\}| = \prod_{\tau \in G(\tau)} |\{T^*\}_\tau|. \quad (5.51)$$

**Следствие 5.4.** Пусть фиксировано финальное описание  $\{\tau/T^*\}$ . Максимальная доля прецедентов участвующих в артефактах среди всех описаний  $\{\tau/T\}$  таких, что  $\{\tau/T^*\} \geq \{\tau/T\}$ , имеет место в финальном описании базы прецедентов  $\Omega(\{\tau/T^*\}, Z)$ .

По определению, если финальное описание базы прецедентов  $\Omega(\{\tau/T^*\}, Z)$  не содержит конфликтов, то данное описание является критическим.

Для маркировки критичности описания будем использовать литеры ‘А’,

‘В’, ‘С’, где А – докритическое описание, В – критическое описание, С – закритическое описание. На рис. 5.1 приведена иллюстрация того, что критические описания являются границей между порядком и хаосом (граница изображена красной – темной – линией).

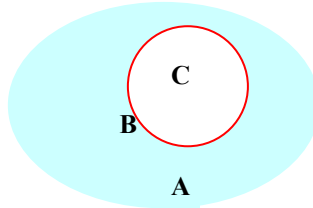


Рис. 5.1 – Соотношение описаний разных типов

Приведем общий алгоритм определения критичности произвольного описания  $\{\tau T\}$ . Для начала построим функцию  $PostCritical(\Omega)$ , которая для любого описания базы прецедентов определяет, является ли оно закритическим или нет.

#### Алгоритм 5.11

Вход: Описание базы прецедентов  $\Omega(\{\tau T\}, Z)$ .

Выход: *true*, если описание закритическое, иначе – *false*.

$PostCritical(\Omega)$

**begin**

**for**  $\alpha(\{\tau T\}_{\alpha}, z_{\alpha}) \in \Omega$  **do**

**for**  $\chi(\{\tau T\}_{\chi}, z_{\chi}) \in \Omega \setminus \alpha$  **do**

**if**  $(\{\tau T\}_{\alpha} = \{\tau T\}_{\chi}) \ \& \ (z_{\alpha} \neq z_{\chi})$  **then return** (*true*);

**return** (*false*)

**end.**

Функцию  $PostCritical(\Omega)$  используем в алгоритме определения критичности произвольного описания  $\{\tau T\}$ .

Для любой вершины  $T$  орграфа доменов  $G(\tau)$  определим множество вершин, которые являются прямыми потомками  $T$ :

$$\{T^*\}_{\tau T} = \{T^* \mid (T \rightarrow T^*) \in G(\tau)\}.$$

Ясно, что для любой терминальной вершины  $T^*$  имеет место  $\{T^*\}_{\tau T^*} = \emptyset$ .

Искомый алгоритм определения критичности произвольного описания базы прецедентов приведен ниже.

#### Алгоритм 5.12

Вход: Описание базы прецедентов  $\Omega(\{\tau T\}, Z)$ .

Выход: Маркер критичности – А, В или С.

Онтология:  $\{G(\tau)\}$ .

$Critical(\Omega)$

**begin**

**if**  $PostCritical(\Omega)$  **then return** ('C');

**for**  $\tau \in \{\tau\}$  **do**

**if**  $\{T^*\}_{\neq T} \neq \emptyset$  **then**

**for**  $T \in \{T^*\}_{\neq T}$  **do begin**

$\{\neq T\}' := (\{\neq T\} \setminus \tau) \cup \neq T'$ ;

$\Omega(\{\neq T\}, Z) \rightarrow \Omega(\{\neq T\}', Z)$ ;

**if**  $PostCritical(\Omega(\{\neq T\}', Z)) = false$  **then return** ('A') **end;**

**return** ('B')

**end.**

Суть алгоритма 5.12 в следующем: если описание неконфликтно, то поочередно для каждого теста осуществляется элементарное обобщение во всех направлениях, и если какое-либо обобщение также неконфликтно, то описание докритичное, иначе – критичное.

В главе 3 было введено понятие «орграф набросков». Орграф набросков порождается множеством операций  $\{F\}$ , которые можно рассматривать в качестве автоматизмов вычислительной среды. В данном разделе множество  $\{F\}$  порождается элементарными операциями обобщения ( $T \rightarrow T'$ ) в рамках орграфов доменов из  $\{G(\tau)\}$ .

Пусть  $G_S(\Omega(\{\neq T_0\}, Z)) = G_S(\Omega_0)$  – орграф набросков базы прецедентов  $\Omega(\{\neq T_0\}, Z)$ , которая выступает в качестве исходного образа. Множество заключений  $Z$  фиксировано, поэтому каждый набросок можно идентифицировать описанием  $\{\neq T\}$ , включая первичный набросок  $\{\neq T_0\}$ .

Если через  $|G|$  обозначить общее количество вершин в произвольном орграфе  $G$ , то общее количество набросков в орграфе  $G_S(\Omega_0)$  определяется выражением:

$$|G_S(\Omega_0)| = \prod_{\tau \in \{G(\tau)\}} |G(\tau)|. \quad (5.52)$$

Пример. Пусть  $|\{\tau\}| = 10$ ;  $\forall \tau, |G(\tau)| = 3$ . Получим:  $|G_S(\Omega_0)| = 3^{10} = 59049$ .

Количество конечных набросков в орграфе  $G_S(\Omega_0)$ , т.е. набросков, к которым не применимы операции из  $\{F\}$ , определяется выражением:

$$|Finite(\Omega_0)| = \prod_{\tau \in \{G(\tau)\}} |\{T^*\}_{\neq T}|. \quad (5.53)$$

В частности, если все орграфы из  $\{G(\tau)\}$  линейны (нет ветвлений), то это значит, что в каждом орграфе доменов всего одна терминальная вершина. В таком случае, в орграфе набросков  $G_S(\Omega_0)$  всего один конечный

набросок.

Выражение (5.53) полностью совпадает с выражением (5.51). Это говорит о совпадении финальных описаний базы прецедентов с финитными набросками орграфа набросков.

Состав экстремального пограничного слоя набросков в раках орграфа  $G_S(\Omega_0)$  определяет следующее утверждение.

**Предложение 5.32.** В орграфе набросков  $G_S(\Omega_0)$  по внутреннему критерию «конфликтность» экстремальный пограничный слой набросков состоит из критических описаний.

На рис. 5.2 приведен орграф  $G_S(\Omega_0)$  для примера «Диагностика». В качестве  $\Omega_0$  выступает базовое описание «Т1 – В1». Серым цветом выделены критические описания.

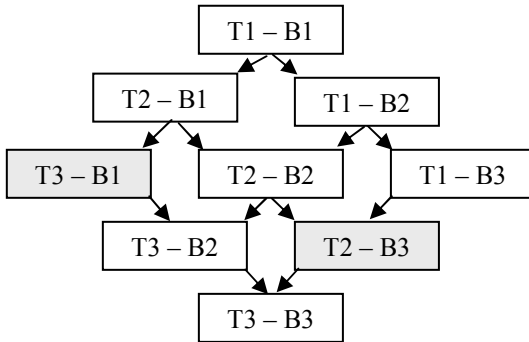


Рис. 5.2 – Орграф  $G_S(\Omega_0)$  для примера «Диагностика»

Общее количество критических описаний непросто рассчитать, но в ряде случаев можно сделать определенные оценки. Так, например, если финитный набросок не содержит конфликтов, то такой набросок принадлежит к числу критических описаний (по определению), а значит и к экстремальному пограничному слою набросков.

**Следствие 5.5.** Если все финальные описания не содержат конфликтов, то все они принадлежат экстремальному пограничному слою набросков, а мощность экстремального пограничного слоя набросков определяется выражением (5.53).

Пусть  $\{\sigma/T\}^*$  – критическое описание. Из определения критического описания следует, что не существует моделей знаний для описаний базы прецедентов  $\Omega(\{\sigma/T\}', Z)$ , где  $\{\sigma/T\}' \geq \{\sigma/T\}^*$ , т.е. для закритических описаний. В рамках любого закритического описания существуют артефакты. Для всех уникальных наборов заключений  $\{\underline{z}/Z\}_\beta$ , где  $\beta$  –

произвольный артефакт, введем числовые коды (отличные от кодов  $Z$ ). В рамках примера «Диагностика» возможна только одна комбинация –  $\{z?1; z?2\}$ . Присвоим данной комбинации код заключения «3». Ниже приведены два описания, в которых произведены замены старых заключений на новый код.

Описание «Т3 – В2»		
$\tau_1/T3$	$\tau_2/V2$	$z$
Ненорм.	Молодой	3
Нормальная	Сред. лет	1
Ненорм.	Молодой	3
Пониженная	Пенсион.	2

Описание «Т3 – В3»		
$\tau_1/T3$	$\tau_2/V3$	$z$
Ненорм.	Допенс.	3
Нормальная	Допенс.	1
Ненорм.	Допенс.	3
Ненорм.	Пенсион.	2

В результате такой замены артефакты пропадают во всех описаниях и, следовательно, для них можно строить синдромные модели знаний. Другими словами, новые коды трактуются как обычные заключения, при этом «старые» прецеденты переходят в «новые» прецеденты (с новыми заключениями). Для новых прецедентов можно находить свои синдромы и свои предельные синдромы.

#### 5.4 Дуальные синдромные модели знаний

Если  $Z$  содержит более двух классов (заключений), то автоматизмами среды порождаются еще  $N$  множеств заключений:

$$Z_1 = \{1; \neg 1\} = \{1; 2 \dots N\}, Z_2 = \{2; \neg 2\} = \{2; 1 \ 3 \dots N\}, \dots, Z_N = \{N; \neg N\}.$$

В совокупности множества  $Z, Z_1, Z_2, \dots, Z_N$  образуют структурно-завершенный орграф доменов заключения  $z$ :

$$G^+(z) = \{Z \rightarrow Z_1; Z \rightarrow Z_2; \dots; Z \rightarrow Z_N\}$$

или в виде конфигуратора

$$G^+(z) = \{Z_N \# Z \{N; 1 \dots N-1\}; \dots; Z_1 \{1; 2 \dots N\}; Z \{1; \dots; N\}\}.$$

Орграф  $G^+(z)$  имеет типовую структуру, изображенную на рис. 5.3 (для  $Z = \{1, 2, 3, 4, 5\}$ ).

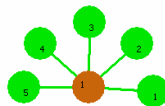


Рис.5.3 – Структурно-завершенный орграф  $G^+(z)$

Листья орграфа  $G^+(z)$  имеют одинаковый вид:  $(z_i; -z_i)$ . Если заключению  $(-z_i)$  присвоить номер  $(N + i)$  и ввести в рассмотрение множество заключений  $Z' = \{1, \dots, N, N+1, \dots, 2N\}$ , то множество вершин-листьев примет вид:  $\{(z_1; z_{N+1}); (z_2; z_{N+2}), \dots, (z_N; z_{2N})\}$ . Каждая пара  $Z_i = (z_i; z_{N+i})$  образует полную группу несовместных заключений. Для каждой такой пары заключений (дуады) автоматически формируются описания базы прецедентов  $\Omega(\{\tau T\}, Z_i)$  и свои синдромные модели знаний, которые будем называть *дуальными синдромными моделями знаний*.

Главной особенностью дуальных моделей знаний является то, что их можно строить для некоторых конфликтных описаний базы прецедентов  $\Omega(\{\tau T\}, Z)$ . Другими словами, уровень общности дуальных моделей знаний, как правило, значительно выше, чем моделей знаний для исходного множества  $Z$ . Справедливо предложение.

**Предложение 5.33.** Множество закритических описаний любой базы прецедентов  $\Omega(Z_i)$ , где  $i \in Z$  ( $|Z| > 2$ ), принадлежит множеству закритических описаний базы прецедентов  $\Omega(Z)$ .

Таким образом, для каждой из  $(N+1)$ -й групп заключений  $Z, Z_1, \dots, Z_N$  строятся свои синдромные модели знаний, включая предельные модели. Причем только модели знаний для  $Z$  позволяют полностью решить целевую задачу. Дуальные модели для  $Z_1, \dots, Z_N$  имеют, как правило, более высокий уровень общности, но решают целевую задачу для каждого  $z_i$  в отдельности.

Рассмотрим пример. Объект – Клиника. Решается вопрос об угрозе некоторого заболевания. Возможны три заключения  $Z = \{1; 2; 3\}$  и, соответственно, три производных заключения:

$$Z1 = \{1; -1\} = \{1, 4\}; \quad Z2 = \{2; -2\} = \{2, 5\}; \quad Z3 = \{3; -3\} = \{3, 6\}.$$

Вопрос об угрозе заболевания решается на основании двух тестов:

- курение;
- количество калорий в сутки.

Требуется найти критические описания для текущей базы прецедентов.

Тесты задаются орграфами доменов:

*Курение*  $\wedge t1$  {

3 {Не курит  $\wedge N$ ; Курит  $\wedge \mathbf{b}$  d}

2 {Не курит  $\wedge N$ ; Меньше пачки сигарет в день  $\wedge \mathbf{b}$  c; Свыше пачки сигарет в день  $\wedge \mathbf{d}$ }

1 {Не курит  $\wedge N$ ; 10-15 сигарет в день  $\wedge \mathbf{b}$ ; 16-20 сигарет в день  $\wedge \mathbf{c}$ ; Свыше пачки сигарет в день  $\wedge \mathbf{d}$ }

*Количество калорий в сутки*  $\wedge t2$  {

4 {Норма  $\wedge N$ ; Отклонение от нормы  $\wedge \mathbf{a}$  c}



- 3 {Сниженное  $\hat{a}$ ; Норма  $\hat{N}$ ; Повышенное  $\hat{c}$   $\hat{d}$  }  
 2 {Ниже нормы  $\hat{a}$  [1800; 2500]; Норма  $\hat{N}$  (2500; 3500]; Избыточное  $\hat{c}$  (3500; 4200]; Очень избыточное  $\hat{d}$  (4200; 5000] }  
 1 {[1800; 5000]}.

Пусть множество прецедентов  $\Omega$  содержит 12 случаев, которые представлены в таблице 5.10 в формате (t1/1, t2/1), т.е. в базовых доменах.

Таблица 5.10 – Исходная база прецедентов  $\Omega$

Курение	Кол-во калорий в сутки	Z
c	4300	2
d	4200	1
b	2200	2
b	2500	2
N	2200	3
N	2100	3
b	2700	1
c	3400	1
b	3500	2
b	3200	2
d	4300	3
d	5000	3

В таблице 5.11 приведены сведения о прецедентах на всех уровнях обобщения.

Таблица 5.11 – Расширенное описание базы прецедентов

t1			t2				Z	Z1	Z2	Z3
1	2	3	1	2	3	4				
c	b	b	4300	d	c	a	2	4	2	6
d	d	b	4200	c	c	a	1	1	5	6
b	b	b	2200	a	a	a	2	4	2	6
b	b	b	2500	a	a	a	2	4	2	6
N	N	N	2200	a	a	a	3	4	5	3
N	N	N	2100	a	a	a	3	4	5	3
b	b	b	2700	N	N	N	1	1	5	6
c	b	b	3400	N	N	N	1	1	5	6
b	b	b	3500	N	N	N	2	4	2	6
b	b	b	3200	N	N	N	2	4	2	6
d	d	b	4300	d	c	a	3	4	5	3
d	d	b	5000	d	c	a	3	4	5	3

В таблице 5.12 приведены все комбинации доменов для двух тестов и для каждой комбинации указана конфликтность базы прецедентов (0/1). Напомним, что конфликт будет иметь место тогда и только тогда, когда найдутся два прецедента с одинаковыми значениями тестов, но разными заключениями. Все конфликты помечены символом «С». Символом «В» помечены критические описания. Докритические описания помечены символом «А». С процедурной точки зрения таблицы 5.11 и 5.12 легко формируются с помощью инструментов электронных таблиц, например, MS EXCEL.

Таблица 5.12 – Результирующая таблица набросков

t1	t2	Z		Z1		Z2		Z3	
		0/1	A,B,C	0/1	A,B,C	0/1	A,B,C	0/1	A,B,C
1	1	0	A	0	A	0	A	0	A
1	2	1	C	1	C	1	C	0	A
1	3	1	C	1	C	1	C	1	C
1	4	1	C	1	C	1	C	1	C
2	1	0	B	0	A	0	B	0	A
2	2	1	C	1	C	1	C	0	B
2	3	1	C	1	C	1	C	1	C
2	4	1	C	1	C	1	C	1	C
3	1	1	C	0	B	1	C	1	C
3	2	1	C	1	C	1	C	1	C
3	3	1	C	1	C	1	C	1	C
3	4	1	C	1	C	1	C	1	C

Данный пример иллюстрирует тот факт, что среди закритических описаний базы прецедентов  $\Omega(Z)$  могут быть докритические и критические описания баз прецедентов  $\Omega(Z_i)$ . Например, описание  $\Omega(t1/3; t2/1; Z)$  является закритическим, в то время как описание  $\Omega(t1/3; t2/1; Z1)$  является критическим.

Построение дуальных моделей знаний значительно расширяет спектр синдромных моделей знаний за счет введения обобщающих заключений. Вместе с тем, справедливо утверждение.

**Предложение 5.34.** Синдромная модель знаний  $\{S^*\}_{Full}$  на  $\Omega(Z)$  включает все предельные синдромы баз прецедентов  $\Omega(Z_i)$ , относящихся к заключению  $z_i$ , где  $i \in Z$ .

Отметим, что в предыдущем разделе также вводились обобщающие заключения, которые соответствовали артефактам. Однако подобные заключения выводят за рамки  $\{S^*\}_{Full}$  на  $\Omega(Z)$ .

Главный результат данного параграфа заключается в понимании того факта, что вместо решения задачи классификации на  $\Omega(Z)$ , возможно, следует решить серию подобных задач на  $\Omega(Z_i)$ , которые порождаются автоматизмами вычислительной среды. Другими словами, предварительно нужно решить задачу «выбора задачи». В итоге получаем иерархию задач. В седьмой главе будет рассмотрена общая схема функционирования подобной иерархии задач.

### **5.5 Анализ формальных понятий, бикластеризация и когнитивный подход**

Трудно найти что-то более важное для нормального функционирования мышления, восприятия, деятельности и речи, чем процессы категоризации, в результате которых формируются понятия, концепты и категории (по большей части, неосознаваемые). Как отмечает Лакофф [85], вне способности к категоризации мы не смогли бы функционировать ни в материальном мире, ни в социальной и интеллектуальной жизни. Изучение процессов категоризации особенно важно для разработки формальных моделей различных аспектов интеллектуальной деятельности. Рассмотрим возможность построения некоторых моделей, опираясь на методологию бикластеризации, анализа формальных понятий и принцип предельных обобщений.

Традиционные методы кластеризации предлагают разнообразный набор средств для решения задач группировки объектов с учетом их сходства в самых разных предметных областях. Однако эти методы имеют ряд недостатков. Как правило, методы кластеризации разбивают объекты на группы, но не отвечают на вопрос, в чем заключается сходство сгруппированных в один кластер объектов. Если исходные данные представлены в виде объектно-признаковой таблицы, то под кластером понимается множество строк (столбцов) такой таблицы, при этом сходство рассчитывается по всем значениям, записанным в такой строке или столбце, хотя реально у таких строк (столбцов) могут быть похожими только некоторые подмножества признаков (объектов). Зачастую требуется выявлять не столько числовое сходство объектов, сколько их общие свойства, т.е. то, благодаря чему они были признаны сходными. Подход к разрешению данной проблемы получил название «бикластеризация». Бикластеризация позволяет отыскивать «бикластеры», включающие, помимо множества объектов, множество их общих признаков [69, 84].

Прикладное алгебраическое направление Анализ Формальных Понятий (АФП) предоставляет решеточные модели бикластеризации особого вида, позволяющие сохранять объектно-признаковое описание сходства группы

объектов внутри кластера и, кроме того, строить иерархии таких кластеров по отношению «быть более общим, чем» [197, 197]. В рамках этой области сформулировано математическое определение формального понятия (ФП) и описано построение иерархий ФП. Исходно ФП является парой вида (объем, содержание), где под объемом понимается некоторое множество объектов, а под содержанием — множество их общих признаков. Как видим, это определение напоминает описание бикластера. Исходные данные в АФП представляются в виде объектно-признаковой матрицы, состоящей из нулей и единиц, а ФП является максимальный прямоугольник (с точностью до перестановок столбцов и строк) такой матрицы, заполненный единицами. Это означает, что данное подмножество объектов обладает всеми признаками некоторого подмножества признаков. Множество всех ФП упорядоченно и образует полную решетку, называемую «решеткой формальных понятий».

В ряде работ исследуются возможности «ослабления» требований к определению ФП, например, рассматриваются шумоустойчивые понятия и нечеткие решетки понятий [69]. Необходимость такого рода ослаблений вызвана излишне жёсткой структурой ФП, требующей наличия всех признаков из содержания понятия у всех объектов его объема. Однако в случае наличия шума возможны «выпадения» некоторых признаков из содержания понятия.

Сообщество FIMI (Frequent Itemset Mining Implementation) изучает проблемы поиска частых (замкнутых) множеств признаков в больших базах данных. Задача поиска частых множеств признаков является одной из центральных тем в Data Mining. Замкнутые множества признаков являются в точности содержаниями ФП. Поэтому методы FIMI можно отнести к моделям бикластеризации [69, 198].

Под термином бикластеризация в настоящее время понимается широкий круг методов, а потому для него в научной литературе существует целый ряд синонимов [69]: совместная кластеризация (simultaneous clustering), кокластеризация (co-clustering), двуходовая кластеризация (two-way clustering), кластеризация подпространства (subspace clustering), двумерная кластеризация (bidimensional) и бокс-кластеризация (box-clustering). Соответственно, требования, предъявляемые к понятию бикластера, различаются, а потому формальные определения даются только для конкретных случаев.

Несмотря на такое разнообразие подходов, многие вопросы остаются открытыми, в частности, данные подходы не учитывают современных представлений о понятии (категории), как о сложно устроенном когнитивном феномене, не имеющем однозначной семантической репрезентации. Стало ясно [85], что процесс категоризации (выработки

понятия) устроен куда сложнее, чем определение сущности, объединяющей элементы с общими признаками. Свойственная человеку эмпирическая категоризация основывается на принципах, далеко отстоящих от тех закономерностей, которые воплощены в отмеченных выше подходах к бикластеризации. Когнитивный подход допускает ситуации, когда одни из объектов (прецедентов) в большей степени соответствуют представлению о понятии (категории), чем другие (понятиям, категориям свойственно иметь наилучших представителей). В отличие от обычной таксономии, категории, занимающие «серединное» положение в когнитивной иерархии, являются *базовыми* (примеры – экстремальный пограничный слой набросков, срединные домены). Основная часть знания структурируется именно на этом уровне. Таким образом, понятия, категории зависят от специфически человеческих способностей к восприятию, к созданию ментальных образов, к обучению и запоминанию, к эффективной коммуникации. Классические теории АФП и бикластеризации в этом отношении идеальны, поскольку в них понятия определяются исключительно в терминах общих характеристик их членов, а не в терминах специфических свойств человеческого понимания [85].

Принцип предельных обобщений предлагает когнитивный подход к бикластеризации, который включает в себя две группы моделей:

- модели на основе замкнутых множеств (с учетом многоуровневости описаний);
- модели на основе синдромного подхода, включая модели на основе орграфов набросков (главы 3, 6).

Один из основных тезисов когнитивного подхода к АФП и бикластеризации заключается в том, что ключевое понятие «сходство» (задается операторами соответствия) имеет смысл только в рамках фиксированного уровня общности описания. Различные уровни общности описания прецедентов формируются системой координат феноменологического пространства наблюдателя в виде Банка тестов. Другой важный тезис заключается в априорной «многоликости» понятия (категории, образа), в суперпозиции смыслов, вкладываемых в понятие, в существовании ядра и периферии системных признаков понятия. Как следствие, предлагаемые модели обладают высокой устойчивостью к шумам за счет предельного обобщения и восстановления данных на основе моделей знаний.

Первоначально рассмотрим модели на основе замкнутых множеств (отнесем их к классу *идеализированных формальных понятий*).

*Контекстом* в АФП называют тройку  $K = (G, M, I)$ , где  $G$  – множество объектов, например, множество ситуаций действительности  $\Omega$  (множество прецедентов);  $M$  – множество признаков (в нашем случае – банк тестов), а

бинарное отношение  $I \subseteq G \times M$  говорит о том, какие объекты какими признаками обладают [69, 197].

*Многозначный формальный контекст* в АФП есть четвёрка  $(G, M, W, J)$ , где  $G, M, W$  – множества (объектов, признаков и значений признаков, соответственно), а  $J$  – тернарное отношение  $J \subseteq G \times M \times W$ , задающее значение  $w$  признака  $m$ , причём:

$$(g, m, w) \in J \text{ и } (g, m, v) \in J \text{ влечёт } w = v.$$

Процедура сведения многозначных контекстов к однозначным называется шкалированием (scaling). Для шкалирования каждый признак многозначного контекста представляется формальным контекстом, называемым шкалой [69].

При анализе прецедентов  $\Omega = \{\alpha(\{\underline{g}T\})\}$  под контекстом  $K$  в общем случае можно понимать тройку  $\langle \Omega, \{G(\tau)\}, \{\tau T\} \rangle$ , а в традиционной нотации  $K_{\{\tau T\}} = (\Omega, \{G(\tau)\}, I_{\{\tau T\}})$ , что означает фиксирование определенного уровня описания (доменов) по всем тестам. Без потери общности будем предполагать, что каждый тест входит в описание прецедента один раз, принимает одно значение и что для каждого прецедента известны значения всех тестов на всех уровнях (база прецедентов с полной информацией). Последнее требование фактически означает, что прецеденты должны быть описаны с помощью базовых доменов всех тестов:  $\Omega = \{\alpha(\{\underline{g}T_0\})\}$ .

Главной особенностью контекста  $(\Omega, \{G(\tau)\}, I_{\{\tau T\}})$  является зависимость результатов всех операций АФП от рассматриваемого уровня общности. Бинарное отношение  $I_{\{\tau T\}}$  в контексте  $K_{\{\tau T\}}$  можно представить следующим образом (все элементы всех доменов считаются различными):

$$I_{\{\tau T\}} \subseteq \Omega \times M_{\{\tau T\}}, \text{ где } M_{\{\tau T\}} = T_{t1} \cup \dots \cup T_{tN}, N = |\{\tau\}|.$$

Прецеденты могут описываться с помощью событий:  $\Omega = \{\alpha(\{\langle \underline{g}T, \underline{t}/A \rangle\})\}$ , где  $\langle \underline{g}T, \underline{t}/A \rangle$  – событие. Под контекстом  $K$  в этом случае можно понимать пару  $\langle \Omega, \{G(\tau)\} \rangle$ . Для сведения к классическому варианту АФП необходимо синхронизировать и согласовать по уровню общности события во всех прецедентах. Если на каком-то уровне общности  $\{\tau T\}$ ,  $t/A$  такое согласование удастся сделать, то фактически это означает, что переопределением названий тестов на данном уровне общности можно уйти от событий, т.е. получить описание  $\Omega = \{\alpha(\{\underline{g}T\})\}$ . В таком случае приходим к контексту  $K_{\{\tau T\}}$ .

Если уровень общности не фиксирован, то такой контекст будем называть свободным и обозначать  $K_{\cup\{\tau T\}} = \langle \Omega, \{G(\tau)\} \rangle$ .

Прецеденты могут быть описаны на уровне синдромов той или иной

синдромной модели знаний  $\{S\}$ , т.е.  $\Omega(Z) = \{\alpha(\{S\})\} \equiv \{\{S\}_\alpha\}$ . Под контекстом  $K_{\{S\}}$  в этом случае можно понимать пару  $\langle \Omega(\{\{S\}_\alpha\}), \{S\} \rangle$ . В традиционной интерпретации можно записать:  $K_{\{S\}} = (\Omega(Z), \{S\}, I_{\{S\}})$ , где  $I_{\{S\}} \subseteq \Omega \times \{S\}$ . Ясно, что  $K_{\cup\{\tau T\}}, Z \rightarrow K_{\{S\}} \rightarrow K_{\{S^*\}}$ . В рамках данной цепочки происходит радикальное изменение объектно-признакового описания.

Контекст  $K_{\{S\}}$  имеет значительно более высокий системный уровень, чем контексты  $K_{\{\tau T\}}$  и  $K_{\cup\{\tau T\}}$ , особенно, если речь идет о контекстах  $K_{\{S^*\}}$ . В рамках контекста  $K_{\{S\}}$  категоризация осуществляется в два этапа: на первом этапе формируется синдромная модель знаний (формируются новые системные признаки), а на втором этапе формируются понятия уже на основе системных признаков. Контекст  $K_{\{S^*\}}$  с полным основанием можно отнести к *базовому уровню* концептуальной организации. Следовательно, и понятия контекста  $K_{\{S^*\}}$  также можно отнести к базовому уровню. К этому уровню относится и редуцированный Банк тестов  $\{G^*(\tau)\}^*$ , с помощью которого можно формировать контексты  $K_{\{\tau T\}}$  базового уровня («*пучок признаков*» является предшественником синдрома). Поскольку предельные синдромы являются гештальтами, то и идеализированные понятия базового уровня несут многие черты гештальтов.

База прецедентов может представлять собой множество образов  $\Omega(\{w\})$ . Для каждого образа  $w$  автоматизмами среды строятся орграфы набросков  $G_s(w)$ . Орграф набросков содержит слои набросков  $\{p\}/G_s(w)$ , где  $p$  – набросок, а также экстремальный пограничный слой набросков  $\{p^*\}/G_s(w)$ . Соответственно можно рассматривать множество разных контекстов, например:  $K_{\{G_s\}} = \langle \Omega(\{w\}), \{G_s(w)\}, \{G(\tau)\} \rangle$  – свободный контекст,  $K_{\{p/G_s\}}$  – контекст фиксированного уровня общности,  $K_{\{p^*/G_s\}}$  – контекст базового уровня и т.д. На основе орграфов набросков  $\{G_s(w)\}$  могут быть построены (предельные) синдромные модели знаний  $\{S_w\}$ , следовательно, контекстом базового уровня является контекст  $K_{\{S/G_s\}} = \langle \Omega(\{w\}), \{G_s(w)\}, \{S_w\} \rangle$ . Более детально вопросы построения таких контекстов будут рассмотрены в главе 6.

Для произвольных  $A \subseteq G$  и  $B \subseteq M$  в АФП определены операторы соответствия Галуа [84, 197]:

$$A' = \{m \in M \mid \forall g \in A (g I m)\}; \quad B' = \{g \in G \mid \forall m \in B (g I m)\}. \quad (5.54)$$

Оператор " (двукратное применение оператора ') является *оператором замыкания*: он идемпотентен ( $A''' = A'$ ), монотонен ( $A \subseteq B$  влечет  $A' \subseteq B'$ ) и экстенсивен ( $A \subseteq A'$ ). Множество объектов  $A \subseteq G$  такое, что  $A' = A$ , называется *замкнутым*. Аналогично для замкнутых множеств признаков – подмножеств множества  $M$ .

Пара множеств  $(A, B)$ , таких, что  $A \subseteq G$ ,  $B \subseteq M$ ,  $A' = B$  и  $B' = A$ , называется *формальным понятием контекста*  $K$  [84, 197]. Множества  $A$  и  $B$  замкнуты и называются *объемом* и *содержанием* формального понятия  $(A, B)$  соответственно. Для множества объектов  $A$  множество их общих признаков  $A'$  служит описанием сходства объектов из множества  $A$ , а замкнутое множество  $A''$  является кластером сходных объектов (с множеством общих признаков  $A'$ ).

Понятия, упорядоченные отношением  $(A_1, B_1) \geq (A_2, B_2) \Leftrightarrow A_1 \supseteq A_2$ , образуют полную решетку, называемую *решеткой понятий*.

Для произвольного  $B \subseteq M$  величина  $|B'|$  называется *поддержкой* (support)  $B$  и обозначается  $\text{sup}(B)$ . Нетрудно видеть, что множество  $B$  замкнуто тогда и только тогда, когда для любого  $D \supset B$  имеет место  $\text{sup}(D) < \text{sup}(B)$ . Именно это свойство используется для определения замкнутости в методах Data Mining. Множество  $B \subseteq M$  называется *k-частым*, если  $|B'| > k$  (то есть множество признаков  $B$  встречается в более чем  $k$  объектах), где  $k$  – параметр. Вычисление частых замкнутых множеств признаков (содержаний) приобрело важность в Data Mining благодаря тому, что по этим множествам эффективно вычисляются множества всех ассоциативных правил [69].

Для произвольных  $\{\alpha\} \subseteq \Omega$  и  $\{\underline{b}/B\} \subseteq M_{\{\tau T\}}$  в рамках контекста  $K_{\{\tau T\}}$  запишем *операторы соответствия* следующим образом:

$$\{\alpha\}'_{\{\tau T\}} = \{\underline{b}/B \in M_{\{\tau T\}} \mid \forall \alpha \in \{\alpha\} (\alpha \text{ I}_{\{\tau T\}} \underline{b}/B)\}; \quad (5.55)$$

$$\{\underline{b}/B'\}_{\{\tau T\}} = \{\alpha \in \Omega \mid \forall \underline{b}/B \in \{\underline{b}/B\} (\alpha \text{ I}_{\{\tau T\}} \underline{b}/B)\}. \quad (5.56)$$

Каждый тест во множество  $\{\underline{b}/B\}$  входит только один раз. Справедливо предложение.

**Предложение 5.35.** Результат операции  $\{\underline{b}/B'\}_{\{\tau T\}}$  будет одинаковым во всех контекстах  $K_{\{b/B\} \cup \{a/A\}}$ , где  $\{a/A\}$  – произвольное дополнение  $\{b/B\}$  до полного описания  $\{\tau T\}$ .

Можно говорить о *кластере контекстов* с общим ядром  $\{b/B\}$ :  $\{K\}_{\{b/B\}} = \cup_{\{a/A\}} \{K_{\{b/B\} \cup \{a/A\}}\}$ . Другими словами, результат операции  $\{\underline{b}/B'\}_{\{\tau T\}}$  будет одинаковым в рамках кластера контекстов  $\{K\}_{\{b/B\}}$ .

Операторы соответствия в рамках контекста  $K_{\{S\}}$  имеют стандартный вид:

$$\{\alpha\}'_{\{S\}} = \{S \in \{S\} \mid \forall \alpha \in \{\alpha\} (\alpha \text{ I}_{\{S\}} S)\}; \quad (5.57)$$

$$\{Si\}'_{\{S\}} = \{\alpha \in \Omega \mid \forall S \in \{Si\} (\alpha \text{ I}_{\{S\}} S)\}. \quad (5.58)$$

Как известно, любую синдромную модель знаний  $\{S\}$  в рамках  $Z$ -задачи можно представить в виде  $\{S\} = \cup_j \{S\}_j$ , где  $\{S\}_j$  отвечает  $z_j \in Z$ . Соответственно,  $\Omega(Z) = \cup_j \Omega(z_j)$ .

**Предложение 5.36.** Пусть дано множество синдромов  $\{Si\} \subseteq \{S\}$ . Если



существуют  $i, j$  такие, что  $i \neq j$ ,  $\{Si\} \cap \{S\}_j \neq \emptyset$ ,  $\{Si\} \cap \{S\}_i \neq \emptyset$ , то  $\{Si\}'|_{\{S\}} = \emptyset$ .

**Следствие 5.6.** Если  $\{Si\}'|_{\{S\}} \neq \emptyset$ , то  $\exists j: \{Si\} \subseteq \{S\}_j$ , а  $\{\alpha\} \subseteq \Omega(z_j)$ .

Дадим определения *обобщенных операторов соответствия в рамках свободного контекста*  $K_{\cup\{\tau T\}}$ :

$$\{\alpha\}'|_{\cup\{\tau T\}} = \{\underline{b}/B\}_+ = \cup_{b \in \{\tau\}} (\cap_{\alpha \in \{\alpha\}} (\underline{b}/B)_\alpha), \quad (5.59)$$

$$\{\underline{b}/B\}'|_{\cup\{\tau T\}} = \{\beta\}, \text{ таких что } \{\underline{b}/B\} \subseteq \cup_{c \in \{\tau\}} (\cap_{\beta \in \{\beta\}} (\underline{c}/C)_\beta), \quad (5.60)$$

где  $(\ )^+$  – замыкание множества значений тестов (см. главу 2); нижний индекс «+» является маркером операции. Каждый тест во множество  $\{\underline{b}/B\}$  может входить несколько раз, но с разными доменами (при этом не должно быть противоречивых данных). Ясно, что операторы (5.59) – (5.60) представить в традиционном виде, например:

$$\{\underline{b}/B\}'|_{\cup\{\tau T\}} = \{\alpha \in \Omega \mid \forall \underline{b}/B \in \{\underline{b}/B\} (\alpha \text{ I}_{\cup\{\tau T\}} \underline{b}/B)\},$$

где  $I_{\cup\{\tau T\}} \subseteq \Omega \times M_{\cup\{\tau T\}}$ , а  $M_{\cup\{\tau T\}}$  объединяет все домены всех орграфов Банка тестов. Однако подобное представление нивелирует эффект обобщения, связанный с замыканием (эффект движения информации).

Приведем пример применения операторов соответствия (5.59) – (5.60).

Пусть  $\{G(\tau)\} = \{G^+(a), G(b)\}$ , где конфигураторы  $G(a)$  и  $G(b)$  имеют вид:

$$a \{A \{1; 2; 3; 4\}\}; \quad b \{B \{1; 2; 3\}\}; \quad G(b) = B; \quad G(a) = A.$$

Соответственно структурно-завершенный орграф  $G^+(a)$  будет иметь вид:

$$a \{A4 \#A \{4 \wedge 4; -4 \wedge 1 \ 2 \ 3\} A3 \#A \{3 \wedge 3; -3 \wedge 1 \ 2 \ 4\} A2 \#A \{2 \wedge 2; -2 \wedge 1 \ 3 \ 4\} \\ A1 \{1 \wedge 1; -1 \wedge 2 \ 3 \ 4\} A \{1; 2; 3; 4\}\};$$

$$G^+(a) = \{A \rightarrow A1; A \rightarrow A2; A \rightarrow A3; A \rightarrow A4\}.$$

Домены  $A$  и  $B$  являются базовыми. С их помощью описываются прецеденты. Пусть  $\Omega = \{\alpha, \beta, \gamma\}$ , где

$$\alpha = (a/A?1; b/B?1), \quad \beta = (a/A?2; b/B?2), \quad \gamma = (a/A?1; b/B?3).$$

Построим замыкания по всем тестам каждого прецедента:

$$(\underline{a}/A)_\alpha^+ = (\underline{a}/A)_\gamma^+ = \{a/A?1; a/A1?1; a/A2?-2; a/A3?-3; a/A4?-4\};$$

$$(\underline{a}/A)_\beta^+ = \{a/A?2; a/A1?-1; a/A2?2; a/A3?-3; a/A4?-4\};$$

$$(\underline{b}/B)_\alpha^+ = \{b/B?1\}; \quad (\underline{b}/B)_\beta^+ = \{b/B?2\}; \quad (\underline{b}/B)_\gamma^+ = \{b/B?3\}.$$

Применим обобщенные операторы соответствия в рамках контекста  $\langle \Omega, \{G^+(a), G(b)\} \rangle$ :

$$\{\alpha, \beta\}'|_{\cup\{\tau T\}} = \cup_{\tau \in \{a, b\}} (\cap_{\gamma \in \{\alpha, \beta\}} (\underline{\tau}/T)_\gamma^+) = \{a/A3?-3; a/A4?-4\};$$

$$\{a/A3?-3; a/A4?-4\}'|_{\cup\{\tau T\}} = \{\alpha, \beta, \gamma\}, \text{ так как}$$

$$\{a/A3?-3; a/A4?-4\} = \{\alpha, \beta, \gamma\}'|_{\cup\{\tau T\}}.$$

Видно, что пара множеств  $(\{\alpha, \beta, \gamma\}, \{a/A3?-3; a/A4?-4\})$  образует

формальное понятие контекста  $K_{\cup\{\tau T\}} = \langle \Omega, \{G^+(a), G(b)\} \rangle$ .

Изменим контекст. Рассмотрим модернизированный Банк тестов  $\{G(\tau)\} = \{G^+(a), G^+(b)\}$ , где  $G^+(b)$  – структурно-завершенный орграф для орграфа  $G(b)$ . По умолчанию он имеет такую же структуру, как и орграф  $G^+(a)$ , а именно:

$$b \{B3 \#B \{3 \wedge 3; \neg 3 \wedge 1 \ 2\} B2 \#B \{2 \wedge 2; \neg 2 \wedge 1 \ 3\} B1 \{1 \wedge 1; \neg 1 \wedge 2 \ 3\} B \{1; 2; 3\}\}; \quad G^+(b) = \{B \rightarrow B1; B \rightarrow B2; B \rightarrow B3\}.$$

База прецедентов остается прежней. В рамках контекста  $K_{\cup\{\tau T\}} = \langle \Omega, \{G^+(a), G^+(b)\} \rangle$  легко получить следующий результат

$$\{\alpha, \beta\}'_{\cup\{\tau T\}} = \cup_{\tau \in \{a, b\}} (\cap_{\chi \in \{\alpha, \beta\}} (\underline{\tau/T})_{\chi}^+) = \{a/A3? \neg 3; a/A4? \neg 4; b/B3? \neg 3\}; \\ \{a/A3? \neg 3; a/A4? \neg 4; b/B3? \neg 3\}'_{\cup\{\tau T\}} = \{\alpha, \beta\}.$$

Как видим, изменив в контексте лишь один орграф Банка тестов и не меняя базу прецедентов, мы получили совсем другой результат, в частности, формальное понятие образует иная пара множеств  $(\{\alpha, \beta\}, \{a/A3? \neg 3; a/A4? \neg 4; b/B3? \neg 3\})$ .

Возьмем в качестве контекста данные примера «Диагностика»  $K_{\cup\{\tau T\}} = \langle \Omega = \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}, \{G(\tau_1), G(\tau_2)\} \rangle$  ( $\tau_1$  – Температура;  $\tau_2$  – Возраст). Построим замыкания по всем тестам и всем прецедентам:

$$(\underline{\tau_1/T})_{\alpha_1}^+ = \{\tau_1/T1? 36.0; \tau_1/T2? \text{Пониж.}; \tau_1/T3? \text{Ненорм.}\}; \\ (\underline{\tau_1/T})_{\alpha_2}^+ = \{\tau_1/T1? 36.7; \tau_1/T2? \text{Норм.}; \tau_1/T3? \text{Норм.}\}; \\ (\underline{\tau_1/T})_{\alpha_3}^+ = \{\tau_1/T1? 37.2; \tau_1/T2? \text{Повыш.}; \tau_1/T3? \text{Ненорм.}\}; \\ (\underline{\tau_1/T})_{\alpha_4}^+ = \{\tau_1/T1? 36.4; \tau_1/T2? \text{Пониж.}; \tau_1/T3? \text{Ненорм.}\}; \\ (\underline{\tau_2/T})_{\alpha_1}^+ = \{\tau_2/B1? 12; \tau_2/B2? \text{Молод.}; \tau_2/B3? \text{Допенс.}\}; \\ (\underline{\tau_2/T})_{\alpha_2}^+ = \{\tau_2/B1? 50; \tau_2/B2? \text{Ср.лет.}; \tau_2/B3? \text{Допенс.}\}; \\ (\underline{\tau_2/T})_{\alpha_3}^+ = \{\tau_2/B1? 32; \tau_2/B2? \text{Молод.}; \tau_2/B3? \text{Допенс.}\}; \\ (\underline{\tau_2/T})_{\alpha_4}^+ = \{\tau_2/B1? 87; \tau_2/B2? \text{Старч.}; \tau_2/B3? \text{Пенс.}\}.$$

Сформируем результат операции  $\{\alpha\}'$  в разных контекстах:

$$\{\alpha_1, \alpha_3\}'_{\langle \tau_2 - B1 \rangle} = \emptyset; \quad \{\alpha_1, \alpha_3\}'_{\langle \tau_1 - B2 \rangle} = \{\tau_2/B2? \text{Молодой}\}; \\ \{\alpha_1, \alpha_3\}'_{\cup\{\tau T\}} = \{\tau_1/T3? \text{Ненорм.}; \tau_2/B2? \text{Молодой}; \tau_2/B3? \text{Допенс.}\}; \\ \{\alpha_1, \alpha_2, \alpha_3\}'_{\cup\{\tau T\}} = \{\tau_2/B3? \text{Допенс.}\}; \quad \{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}'_{\cup\{\tau T\}} = \emptyset; \\ \{\alpha_1, \alpha_4\}'_{\cup\{\tau T\}} = \{\tau_1/T2? \text{Пониж.}; \tau_1/T3? \text{Ненорм.}\}.$$

Пусть  $F$  – некоторый оператор вида

$$F: \{\alpha\}, \{G(\tau)\} \rightarrow \{\underline{b}/B\}_F. \quad (5.61)$$

С использованием оператора  $F$  определим *обобщенные операторы соответствия контекста*  $K_F$ :

$$\{\alpha\}'|_F = \{\underline{b}/B\}'_F = F(\{\alpha\}, \{G(\tau)\}), \quad (5.62)$$

$$(\{\underline{b}/B\})'|_F = \{\beta\}, \text{ таких что } \{\underline{b}/B\} \subseteq F^+(\{\beta\}, \{G(\tau)\}), \quad (5.63)$$

где  $F^+$  – замыкание множества значений тестов. Видно, что обобщенные операторы соответствия контекста  $K_{\cup\{\tau T\}}$  являются частным случаем обобщенных операторов соответствия контекста  $K_F$ . Приведем для примера еще одну интерпретацию оператора  $F$ :

$$\{\alpha\}'|_F = \{\underline{b}/B\}'_F = \cup_{b \in \{\tau\}} (\cap_{\alpha \in \{\alpha\}} (\underline{b}/B)^+_{\alpha})^-. \quad (5.64)$$

где  $\{\}^-$  – нижний предел множества значений тестов (см. главу 2). С функциональной точки зрения операторы (5.59) и (5.64) схожи, только оператор (5.64) не содержит избыточных данных. Сравните с приведенными выше результаты следующих операций

$$\{\alpha_1, \alpha_3\}'|_F = \{\tau_1/T3? \text{ Ненорм.}; \tau_2/B2? \text{ Молодой}\};$$

$$\{\alpha_1, \alpha_4\}'|_F = \{\tau_1/T2? \text{ Пониж.}\}.$$

Очевидно, не стало избыточности данных. Действительно, поскольку есть значение « $\tau_2/B2?$  Молодой», то значение « $\tau_2/B3?$  Допенс.» является избыточным (оно выводится из первого путем замыкания). То же касается и теста «Температура». Учитывая, что число равноуровневых описаний может составлять многие миллиарды, сокращение объема результирующих данных может быть очень существенным.

Приведем нотации операторов замыкания в разных контекстах:

$\{\alpha\}'|_{\{\tau T\}} = \{\beta\}$  – замыкание по прецедентам, т.е. определение кластера сходных прецедентов по заданному исходному множеству прецедентов. Уровень общности фиксирован;

$\{\alpha\}'|_{\cup\{\tau T\}} = \{\beta\}$  – замыкание по прецедентам, т.е. определение кластера сходных прецедентов по заданному исходному множеству прецедентов в рамках свободного контекста  $K_{\cup\{\tau T\}}$  (уровень общности не фиксирован);

$\{\alpha\}'|_{\{S\}} = \{\beta\}$  – замыкание по прецедентам, т.е. определение кластера сходных прецедентов в рамках контекста  $K_{\{S\}}$ ;

$\{\underline{b}/B\}'|_{\{\tau T\}} = \{\underline{a}/A\}$  – замыкание по тестам, т.е. определение всех совпадающих результатов тестов у прецедентов, у которых совпадают некоторые априорно заданные результаты тестов;

$\{\underline{b}/B\}'|_{\cup\{\tau T\}} = \{\underline{a}/A\}$  – замыкание по тестам, т.е. определение всех совпадающих результатов тестов у прецедентов, у которых совпадают некоторые априорно заданные результаты тестов. Уровень описания не фиксирован;

$\{Si\}'|_{\{S\}} = \{S'\}$  – замыкание по синдромам в рамках контекста  $K_{\{S\}}$ .

Уточним формулировки идеализированных формальных понятий в

рамках контекстов  $K_{\{S\}}$ ,  $K_{\{\sigma T\}}$  и  $K_{\{S \cup \sigma T\}}$ .

Пара множеств  $(\{\alpha\}, \{S\}_{\{\alpha\}})$  таких, что  $\{\alpha\} \subseteq \Omega$ ,  $\{S\}_{\{\alpha\}} \subseteq \{S\}$ ,  $\{\alpha\}' = \{S\}_{\{\alpha\}}$  и  $(\{S\}_{\{\alpha\}})' = \{\alpha\}$ , называется *формальным понятием контекста*  $K_{\{S\}}$ . В качестве модели знаний  $\{S\}$  рекомендуется выбирать предельную синдромную модель  $\{S^*\}$ . Понятия контекста  $K_{\{S^*\}}$  относятся к базовому уровню концептуализации. Справедливо утверждение.

**Предложение 5.37.** Все формальные понятия контекста  $K_{\{S\}}$  разбиваются на  $N$  непересекающихся классов, где  $N = |Z|$ . Каждый  $j$ -ый класс определяется подконтекстом  $\langle \Omega(z_j), \{S\}_j \rangle$ , где  $\{S\}_j$  отвечает  $z_j \in Z$ .

Доказательство опирается на предложение 5.36.

Пару множеств  $(\{\alpha\}, \{\underline{a}/A\})$  таких, что  $\{\alpha\} \subseteq \Omega$ ,  $\{\alpha\}'_{\{\sigma T\}} = \{\underline{a}/A\}$  и  $\{\underline{a}/A\}'_{\{\sigma T\}} = \{\alpha\}$ , назовем *формальным понятием контекста*  $K_{\{\sigma T\}}$  (при условии, что каждый тест входит в описание один раз и база прецедентов с полной информацией).

Пару множеств  $(\{\alpha\}, \{\underline{b}/B\})$  таких, что  $\{\alpha\} \subseteq \Omega$ ,  $\{\alpha\}'_{\{ \cup \sigma T \}} = \{\underline{b}/B\}$  и  $\{\underline{b}/B\}'_{\{ \cup \sigma T \}} = \{\alpha\}$ , назовем *формальным понятием контекста*  $K_{\{ \cup \sigma T \}}$ . В данном варианте уровень общности заранее не фиксирован (свободный контекст).

Приведем несколько понятий, сформированных на базе примера «Диагностика», но в разных контекстах ( $\tau_1$  – Температура;  $\tau_2$  – Возраст).

В рамках контекста  $K(\langle T2 - B2 \rangle)$  имеются такие формальные понятия:

$(\{\alpha_1, \alpha_4\}, \{\tau_1/T2? \text{Пониженная}\})$  и  $(\{\alpha_1, \alpha_3\}, \{\tau_2/B2? \text{Молодой}\})$ .

В рамках контекста  $K(\langle T3 - B2 \rangle)$  имеется такое формальное понятие:

$(\{\alpha_1, \alpha_3\}, \{\tau_1/T3? \text{Ненорм.}, \tau_2/B2? \text{Молодой}\})$ .

В рамках максимально обобщенного контекста  $K(\langle T3 - B3 \rangle)$  имеется такое формальное понятие:

$(\{\alpha_1, \alpha_3\}, \{\tau_1/T3? \text{Ненорм.}, \tau_2/B3? \text{Допенсионный}\})$ .

Обобщение формального понятия фиксированного контекста может уже не быть формальным понятием. В качестве доказательства приведем пример. Пара множеств  $(\{\alpha_1, \alpha_4\}, \{\tau_1/T2? \text{Пониженная}\})$  является формальным понятием контекстов  $K(\langle T2 - B1 \rangle)$ ,  $K(\langle T2 - B2 \rangle)$  и  $K(\langle T2 - B3 \rangle)$ . Однако пара множеств  $(\{\alpha_1, \alpha_4\}, \{\tau_1/T3? \text{Ненорм.}\})$  не является формальным понятием, в частности, контекста  $K(\langle T3 - B3 \rangle)$ . Действительно:  $\{\tau_1/T3? \text{Ненорм.}\}' = \{\alpha_1, \alpha_3, \alpha_4\}$ .

Справедливы следующие утверждения.

**Предложение 5.38.** Если пара множеств  $(\{\alpha\}, \{\underline{a}/A\})$  является формальным понятием контекста  $K_{\{a/A\} \cup \{b/B\}}$ , то она будет формальным понятием в рамках любого контекста  $K_{\{a/A\} \cup \{b/B'\}}$ , где  $\{b/B\} \geq \{b/B'\}$ .

Пример. Понятие  $\{\{\alpha_1, \alpha_3, \alpha_4\}, \{\tau_1/T3? \text{Ненорм.}\}\}$  контекста  $K(\langle T3 -$

В3») является также понятием в рамках контекстов  $K(\langle T3 - B2 \rangle)$  и  $K(\langle T3 - B1 \rangle)$ .

**Предложение 5.39.** Если пара множеств  $(\{\alpha\}, \{\underline{a}/A\})$  является формальным понятием свободного контекста  $K_{\cup\{\tau/T\}}$  и каждый тест входит в  $\{\underline{a}/A\}$  только один раз, то эта пара будет также формальным понятием фиксированных контекстов  $K_{\{\underline{a}/A\} \cup \{b/B\}}$ , где  $\{b/B\}$  – произвольное дополнение  $\{a/A\}$  до полного описания  $\{\tau/T\}$ .

Действительно, если нашлось бы какое-то значение  $\underline{b}/B$  принадлежащее всем  $\{\alpha\}$ , то оно обязательно вошло бы в замыкание  $\{\underline{a}/A\}^+$ , что противоречит условию. Следовательно, никакие значения  $\underline{b}/B$  не могут быть добавлены к  $\{\underline{a}/A\}$ . При фиксированном контексте каждый тест входит в  $\{\underline{a}/A\}$  только один раз (при свободном контексте это условие, как правило, не выполняется). Таким образом, выполнены все требования понятия фиксированного (по уровню общности) контекста.

Рассмотрим модели АФП и бикластеризации на основе синдромного подхода. Зададим контекст  $K_Z = \langle \Omega(\{\alpha(z_\alpha)\}), Z, I_Z \rangle$  следующим образом:

$$\alpha \quad I_Z \quad \underline{z}/Z \Leftrightarrow z_\alpha = \underline{z}/Z.$$

Справедливо предложение.

**Предложение 5.40.** Полное множество формальных понятий контекста  $K_Z$  образуют пары  $(\{\alpha\}_j, z_j)$ , где  $z_j \in Z$ , а  $\{\alpha\}_j = \Omega(z_j)$ . Всего имеется  $N$  понятий, где  $N = |Z|$ .

Каждому понятию  $(\{\alpha\}_j, z_j)$  контекста  $K_Z$  однозначно соответствует сопряженная пара множеств  $(\{\alpha\}_j, \{S\}_j)$  контекста  $K_{\{S\}}$ , где  $\{S\}_j$  отвечает  $z_j \in Z$ , следовательно, между объектами  $j$ -го понятия (бикластера) имеется лишь *семейное сходство*: они связаны общим заключением  $z_j$ , при этом у них может не быть общих свойств (синдромов). Действительно, в рамках любого  $\{S\}_j$  может существовать *директивная зона* (общее ядро синдромов)

$$\{S\}_{\Omega_j} = \bigcap_{\alpha \in \Omega_j} \{S\}_\alpha, \quad j \in Z$$

и *зона возможности*:  $\{S\}_j^\perp = \{S\}_j \setminus \{S\}_{\Omega_j}$ . Если  $\{S\}_{\Omega_j} = \emptyset$ , то это означает, что у объектов  $j$ -го бикластера нет общих свойств (кроме семейного сходства).

Большая директивная зона в некоторых мыслительных процессах может замещать всю категорию в целом (переносное или «метонимическое» рассуждение). Директивная зона и зона возможности формируют структуру понятия «центр – периферия».

Благодаря наличию зоны возможности процедура распознавания семейного класса  $\underline{z}/Z$  (диагностика, классификация) для нового или знакомого прецедента обладает максимальной *пластичностью* и *устойчивостью* к шумам. Устойчивость к пропускам данных повышает

*процедура восстановления данных на основе моделей знаний:* для каждого дискретного домена теста, входящего в описание прецедентов, по требованию строится синдромная модель знаний (домен теста выступает в качестве  $Z$ ); с использованием моделей знаний производится восстановление значения требуемого уровня общности.

Учитывая также предвестники  $\{R\}_j$ , сопряженность пар множеств ( $\{\alpha\}_j$ ,  $z_j$ ) и ( $\{\alpha\}_j$ ,  $\{S\}_j$ ,  $\{R\}_j$ ) будем обозначать следующим образом ( $j \in Z$ ):

$$(\{\alpha\}_j, z_j) \leftrightarrow (\{\alpha\}_j, \{S\}_j, \{R\}_j), \quad \{\alpha\}_j = \Omega(z_j), \quad \{S\}_j \subset \{S\}, \quad \{R\}_j \subset \{R\} \quad (5.65)$$

или в сокращенной нотации: ( $\{\alpha\}_j$ ,  $z_j$ ,  $\{S\}_j$ ,  $\{R\}_j$ ),  $j \in Z$ . Структура (5.65) означает, что любое эмпирическое понятие должно быть понимаемо как *микротеория* со своей внутренней логической структурой. Только обращение к внутренней структуре понятий или образов позволяет иерархически сопрягать их с мыслительными образами и теоретическими положениями более высоких структурных уровней (с суждениями, умозаключениями, предложениями). Именно эта тождественность внутреннего строения любых информационных массивов любого масштаба дает основание говорить, что любое эмпирическое понятие есть своеобразная микротеория [85].

Важно отметить, что пары ( $\{\alpha\}_j$ ,  $z_j$ ) относятся к априорной информации, в то время как системные свойства  $\{S\}_j$  и  $\{R\}_j$  вычисляются, а точнее говоря, формируются в результате естественной самоорганизации. Системные свойства базового уровня (признаки) представляют собой отражение сущности (природных) способностей наблюдателя и опыта его функционирования в материальном и социальном окружении. Категории базового уровня с эпистемологической и функциональной точки зрения приоритетны по отношению к следующим факторам [85]: восприятию гештальтов (предельные синдромы и закономерности являются гештальтами), формированию образов, моторике, организации знаний, легкости протекания когнитивных процессов (обучения, распознавания, запоминания и т. д.) и к легкости вербализации языковых выражений. Подобные понятия, категории, по-видимому, не осознаются, они используются автоматически, бессознательно и без заметных усилий, воспринимаясь просто как часть нормальной жизнедеятельности наблюдателя. Используемые таким образом понятия обладают иным, более важным, психологическим статусом по сравнению с теми понятиями, которые обязательно осознаются [85].

В таблице 5.13 приведен фрагмент базы прецедентов для формирования эмпирического понятия «Фрукт» (данные взяты из [84]).

Таблица 5.13 – Обучающая выборка

$\Omega \setminus \{T\}$	цвет	жесткий	гладкий	форма	фрукт
Яблоко	желтое	нет	да	круглое	+
Грейпфрут	желтый	нет	нет	круглый	+
Киви	зеленое	нет	нет	овальное	+
Слива	синяя	нет	да	овальная	+
Кубик	зеленая	да	да	кубический	-
Яйцо	белое	да	да	овальное	-
Теннис.мяч	белый	нет	нет	круглый	-

# Обращаем внимание, что данные представлены не в формате  $\{T_0\}$

В данном примере  $Z = \{1, 2\} \equiv \{\text{фрукт+}; \text{фрукт-}\} \equiv \{\text{образ+}; \text{образ-}\}$ . При такой оппозиции можно говорить о *понятии положительного контекста*. В общем случае речь будет идти о *понятии Z-контекста*. На основе контекста  $\langle \Omega(Z), \{G(\text{Цвет}), G(\text{Жесткость}), G(\text{Гладкость}), G(\text{Форма})\} \rangle$  формируется предельная синдромная модель знаний  $\{S^*\}$  и бикластеры  $(\{\alpha\}_j, z_j, \{S^*\}_j)$ . Банк тестов  $\{G(\tau)\}$  выполняет роль опытного знания существующего достаточно автономно от текущей Z-задачи. Он формируется в процессе решения большого числа разнообразных Z-задач и обучения. Данный пример показывает, что развиваемый подход кардинально отличается от известных методов классификации, в частности, ДСМ-метода [2, 12, 179].

Эмпирические понятия положительного и Z-контекстов можно сформировать, например, на основе «Базы мужских и женских лиц», «Базы химических элементов (активны/неактивны)» и фармакологических препаратов, «Базы генов/белков», «Базы голосов дикторов» и т.д. При анализе видео и аудио баз прецедентов целесообразно строить орграфы набросков (глава 6).

Среди членов бикластера  $(\{\alpha\}_j, z_j, \{S^*\}_j, \{R\}_j)$  могут оказаться и более, и менее удачные примеры. Другими словами, *степень центральности* членов бикластера (понятия, концепта, категории), определенно лежащих в его границах, может быть различной (градуированной). Действительно, все синдромы имеют веса (например, % охвата прецедентов или поддержку), следовательно, чем больше суммарный вес синдромов у прецедента, тем выше степень его центральности. Прецеденты с крайне низкой степенью центральности в ряде приложений можно отнести к «информационному шуму».

Таким образом, понятие сформировано, если удалось построить решающее правило, которое для любого примера из обучающей выборки указывает, принадлежит ли этот пример понятию, или нет. Решающее правило является *корректным*, если оно в дальнейшем успешно распознаёт

объекты, не вошедшие первоначально в обучающую выборку. Развиваемый когнитивный подход предлагает универсальную мультисенсорную технологию индуктивного формирования эмпирических понятий и категорий базового уровня.

Резюме. Предлагаемый подход показывает, как понятия и категории связаны с *доконцептуальной структурой* опытного знания (орграфами доменов тестов, орграфами набросков, экстремальными слоями набросков, синдромами, ассоциативными связями и т.д.). Другими словами, когнитивные модели приобретают фундаментальную значимость благодаря своей способности органично вписываться в рамки доконцептуальной структуры опытного знания [85].

### 5.6 Пример «Оценка и прогноз течения заболевания»

*Постановка задачи.* В одном медицинском НИИ была сформирована база клинических наблюдений пациентов, перенесших инфаркт миокарда (ИМ) [159]. Наблюдения фиксируют состояния пациентов в разные сроки после ИМ. Методом экспертных оценок для каждого пациента (159 случаев) был определен характер течения заболевания: прогрессирующее – 26 случаев, стабильное – 63 случая, регрессирующее – 70 случаев, т.е.  $Z = \{1, 2, 3\}$ . Требовалось с помощью метода предельных обобщений (МПО) построить синдромную модель знаний, описывающую эмпирическую выборку наблюдений. Модель знаний должна не только обладать 100% чувствительностью и специфичностью на исходной выборке, но и быть, по возможности, корректной, т.е. пригодной для оценки новых клинических наблюдений.

*Решение задачи.* Для описания базы прецедентов врачами были выбраны 24 медико-социальных теста, включая клинические и функциональные тесты: Возраст во время ИМ; Пол; Q, не Q инфаркт; Локализация инфаркта; Функциональный класс (ФК); Группа инвалидности; Сахарный диабет (СД); Сердечная недостаточность (СН); Гипертоническая болезнь (ГБ); Желудочковая экстрасистолия (ЖЭ); Мерцательная аритмия (МА); Давность ИМ; Работа в момент осмотра; Вид труда; Тяжесть/напряженность труда; Образование; Индекс массы тела (ИМТ); Индекс массы миокарда левого желудочка (ИММЛЖ); Индекс левого предсердия (Инд. ЛП); Индекс конечно-диастолического объема (Инд. КДО); ФВ\_Т; ФВ\_С; Тип ремоделирования ЛЖ; Рубцовые изменения в миокарде.

Согласно МПО первоначально были построены конфигураторы всех тестов. Примеры конфигураторов тестов:

*Индекс левого предсердия*  $\wedge 19$  S ИндЛП {3 {Норма  $\wedge 0$ ; Увеличение иЛП  $\wedge 1$



2} 2 {Норма  $\wedge^0$  [1,40; 2,10]; Умеренное увеличение иЛП  $\wedge^1$  [2,11; 2,50]; Значительное увеличение иЛП  $\wedge^2$  [2,51; 5,0] } 1 ([1,40; 5,0]);

*Сердечная недостаточность*  $\wedge^8$  Н СН {

4 #1 {ПБ ст.  $\wedge^3$ ; I ст. или ПА ст  $\wedge^1$  2}

3 #1 {ПА ст.  $\wedge^2$ ; I ст. или ПБ ст  $\wedge^1$  3}

2 {I ст.  $\wedge^1$ ; ПА - ПБ ст  $\wedge^2$  3}

1 {I ст.  $\wedge^1$ ; ПА ст  $\wedge^2$ ; ПБ ст.  $\wedge^3$ }}

*Индекс массы миокарда ЛЖ*  $\wedge^{18}$  R ИММЛЖ {

3 {Норма  $\wedge^0$ ; Увеличение ИММЛЖ  $\wedge^1$  2 3}

2 {Норма  $\wedge^0$  M[30; 124] Ж[30; 109];  
Умеренное увеличение ИММЛЖ  $\wedge^1$  M[125; 149] Ж[110; 134];  
Значительное увеличение ИММЛЖ  $\wedge^2$  M[150; 199] Ж[135; 184];  
Резкое увеличение ИММЛЖ  $\wedge^3$  M[200; 400] Ж[185; 400]}

1 ([30; 400])}

*Рубцовые изменения в миокарде*  $\wedge^{24}$  X РубецАн {

2 {Признаков рубцовых изменений нет  $\wedge^0$ ;  
Признаки рубцовых изменений есть  $\wedge^1$  2}

1 {Признаков рубцовых изменений нет  $\wedge^0$ ;  
Признаки рубцовых изменений без аневризмы  $\wedge^1$ ;  
Признаки рубцовых изменений с аневризмой  $\wedge^2$ }}

*Группа инвалидности*  $\wedge^6$  F ГрИнв {

3 {не инвалид  $\wedge^1$ ; инвалид  $\wedge^2$  3}

2 {не инвалид  $\wedge^1$  4 5; II группа  $\wedge^2$ ; III группа  $\wedge^3$ }

1 {не инвалид  $\wedge^1$ ; II группа  $\wedge^2$ ; III группа  $\wedge^3$ ; первично не признаны инвалидами  $\wedge^4$ ; не освидетельствовались  $\wedge^5$ }}.

Номера и латинские буквы после названий тестов (алиасы) означают порядковый номер теста и колонки в таблице Excel. Приведенные конфигураторы задают вычислительные схемы для обобщений первичных результатов тестов. База прецедентов  $\Omega(\{\pi/1\}, Z)$  была реализована с помощью электронной таблицы. Общее количество описаний определялось произведением числа доменов для всех тестов и превысило 143 млн.

На следующем этапе было построено расширенное описание базы прецедентов по типу таблицы 5.11. Далее программно определялись критические описания. Всего критических описаний оказалось 16. Для построения полной синдромной модели знаний  $\{S\}_{Full, \{\pi/T\}}$  и минимальной избыточной модели знаний  $\{S\}_{Min}^v$  было выбрано одно из критических описаний  $\{\pi/T\}^*$  с максимальной долей покрытия  $\lambda$ .

Общее количество найденных синдромов в рамках модели  $\{S\}_{Full}$  на  $\{\pi/T\}^*$  составило: для прогрессирующего течения – 505, для стабильного течения – 1470, для регрессирующего течения – 1780. Каждый синдром

включал результаты от 2 до 5 тестов (было задано ограничение на количество тестов в синдромах – не более 5). Поиск минимального количества синдромов при условии максимального процента покрытия соответствующей группы пациентов (модель знаний  $\{S\}^v_{Min}$  на  $\{T/T\}^*$ ) дал следующие результаты: для прогрессирующего течения – 11, для стабильного течения – 16, для регрессирующего течения – 12. Фрагмент найденных множеств синдромов приведен на рис. 5.4.

52,86	Гр. Инв: не инв	МА: Нет	Инд. КДО: Норма	
50,00	ФК: II ФК	Гр. Инв: не инв	МА: Нет	
50,00	ФК: II ФК	ИММЛЖ: Норма	ФВ_Т : ФВ в норме	
41,43	Вид тр: Физ./Умств.	ИММЛЖ: Норма	ФВ_Т : ФВ в норме	
38,57	ФК: II ФК	Вид тр: Физ./Умств.	Инд ЛП: Норма	ФВ_С : ФВ в норме
37,14	Q, не Q: Q	ФК: II ФК	Инд ЛП: Норма	ФВ_С : ФВ в норме
35,71	Q, не Q: Q	ФК: II ФК	ЖЭ: Нет	Инд ЛП: Норма
35,71	Лок: Передний	Вид тр: Физ./Умств.	Образов: Не высш	ФВ_С : ФВ в норме
30,00	Вид тр: Физ./Умств.	Тяж/напр тр: 3	Инд ЛП: Норма	
28,57	Q, не Q: Q	Лок: Передний	ФВ_С : ФВ в норме	
25,71	ГБ: III ст.	ИММЛЖ: Норма	ФВ_Т : ФВ в норме	РубецАн: Призн. Нет
24,29	Давн ИМ: > 5	Тяж/напр тр: 1 или 2	Инд ЛП: Норма	

Рис. 5.4 – Фрагмент списка синдромов с указанием % покрытия для регрессирующего течения

Список синдромов приведен не в полном объеме, так как предварительный анализ найденных синдромов экспертами показал необходимость фальсификации некоторых из них путем корректировки конфигураторов, корректировки некоторых заключений, а также добавления новых прецедентов, опровергающих соответствующий синдром. После этапа корректировки банка тестов и фальсификации всех вызывающих сомнение синдромов можно говорить о стабилизации модели знаний. После стабилизации модель  $\{S\}^v_{Min}$  может быть применена для прогноза течения заболевания в разные периоды после инфаркта миокарда.

В главе 7 будет произведено ультраоснащение модели знаний  $\{S\}^v_{Min}$  в виде среды радикалов. Каждый радикал является элементарной функциональной системой, стремящейся установить наличие определенного синдрома. Среда радикалов позволяет реализовать разнообразные, в том числе самоорганизующиеся схемы работы когнитивной динамической системы.

Данный пример показывает, что метод предельных обобщений может быть эффективно применен для решения широкого класса клинических задач, включая диагностику, прогнозирование и оценку эффективности лечения.

## ГЛАВА 6. МНОГОУРОВНЕВЫЕ НАБРОСКИ ОБРАЗОВ, СИТУАЦИЙ, МОДЕЛЕЙ

Для иллюстрации сути модели многоуровневых набросков приведу арабскую притчу о том, как один шейх решил переехать жить с одного края огромной пустыни на другой [178]. Построив новый дворец и перевезя в него своих жен и сокровища, шейх повелел перевезти, наконец, свою огромную библиотеку. Караван верблюдов с книгами отправился через пустыню в долгий путь. В дороге верблюды и люди валились от усталости. Тогда живые еще и грамотные погонщики останавливались и конспектировали содержание тех книг, которые уже не могли нести верблюды дальше. По мере движения каравана книг становилось все меньше и меньше, а конспектов – все больше. Наконец, понадобилось конспектировать сами конспекты, конспекты конспектов и так далее. К концу путешествия от каравана остался только один погонщик, который нес в руке листочек с последним конспектом конспектов всей громадной библиотеки. Войдя во дворец, он упал у ног шейха. «Прочти, что было в моих книгах!» – повелел шейх. Погонщик развернул на ладони листок и прочел: «Велик Аллах и Магомет – пророк его!».

В качестве исходной ситуации (исходного образа) в притче выступала вся библиотека. В процессе движения погонщики создавали наброски исходного образа, затем наброски набросков и так до тех пор, пока не остался один – финитный набросок с предельно сжатым смыслом.

### **6.1 Краткий обзор методических подходов к построению набросков**

На этапе индукции исследователь детально выписывает все имеющиеся у него результаты, все, что представляет интерес. И начинает группировать, «стягивать» по определенным основаниям классификаций в первичные обобщения, затем – в обобщения второго порядка (опять же по определенным основаниям), и так далее, – происходит индуктивный процесс – абстрагирование – восхождение от конкретного к абстрактному – пока все результаты не сведутся в авторскую концепцию – короткую, буквально в несколько строк емкую формулировку, отражающую в самом общем сжатом виде всю совокупность результатов, всю суть работы. Подобная сжатая формулировка является примером финитного наброска, а упорядоченная совокупность обобщений разного порядка представляет собой оргграф набросков. Отметим, что для успешного восприятия и понимания требуется множественность точек зрения. Множественность точек зрения, их сталкивание между собой является ключевым положением

при построении орграфа набросков любого образа, любой ситуации и практической реализацией Принципа полимодельной дополнителности, конкурентности и отбора.

Для исследования критических систем часто применяется прием, называемый *масштабным преобразованием*, или перенормировкой. Перенормировка заключается в огрублении пространственного разрешения, в результате чего несколько элементов исходной системы объединяются в единый блок, который на новом масштабе уже рассматривается как самостоятельный элемент [96]. Перенормировка является примером построения набросков все более высокого уровня общности.

Как отмечается в работе [99], основной целью исследований при решении сложных задач является построение иерархии упрощенных моделей. Идея построения иерархии моделей достаточно проста. Вначале выделяются основные, ключевые процессы, играющие главную роль в изучаемом явлении на данных пространственных и временных масштабах. Затем строятся еще более простые модели явления с меньшей областью применимости и учитывающие еще меньшее количество факторов. И так происходит до тех пор, пока не возникают простейшие модели, поведение которых уже понятно. Только после того как модели нижнего уровня изучены и поняты, удастся перейти на следующий, более высокий уровень. Перечислим некоторые процедуры «генерации» простых моделей или малоразмерных моделей (набросков) из сложных [13, 99, 187].

1. Метод замены переменных. Этот метод состоит в построении замен для искомым функций и независимых переменных, преобразующих уравнения исходной модели в новую систему уравнений, которая поддается детальному исследованию известными численно-аналитическими методами. Так, в преобразованной системе уравнений может обнаружиться малый параметр  $\varepsilon$ , тем самым расчищается путь для применения асимптотических методов. В некоторых случаях при  $\varepsilon = 0$  оказывается возможной декомпозиция преобразованной системы на более простые подсистемы меньшей размерности.

2. Введение агрегированных переменных, или макропеременных. Так как макропеременных существенно меньше по сравнению с исходными, то относительно макропеременных получается (точно или приближенно) замкнутая система уравнений, которая и представляет собой существенно более простую малоразмерную модель, доступную для детального численно-аналитического исследования.

3. Применение методов теории сингулярных возмущений. Многие математические модели представляют собой системы обыкновенных дифференциальных уравнений первого порядка, которые после

обезразмеривания содержат в некоторых уравнениях малые параметры при производных. В этом случае существенное упрощение модели (понижение её размерности) достигается методами теории сингулярных возмущений.

4. Применение фундаментального принципа теории бифуркаций – принципа сведения. Суть его состоит в том, что вблизи некоторых бифуркационных многообразий в пространстве параметров многомерной математической модели её исследование сводится к исследованию другой более простой модели существенно меньшей размерности. Практическое значение рассматриваемого механизма упрощения сложных математических моделей связано с тем, что при эксплуатации некоторых технологических систем максимум эффективности (например, продуктивности) достигается вблизи опасных границ области устойчивости в пространстве управляемых параметров. Возникающая там возможность более простого математического описания поведения сложной системы позволяет сформулировать эффективные критерии приближения к этим опасным границам.

5. Применение теории самоорганизации сложных систем. Концепция самоорганизации связана с представлением об эволюции сложных систем к финальным динамическим режимам, допускающим описание математическими моделями невысокой размерности.

Итеративное агрегирование основано на двух взаимообратных операциях [72]: агрегировании и дезагрегировании. *Агрегирование* – преобразование исходной математической модели в модель с меньшим числом переменных или ограничений (агрегированную модель, набросок). Такая модель дает приближенное описание исследуемого объекта или процесса в сравнении с исходной моделью. Под *дезагрегированием* понимается отображение из пространства переменных агрегированной модели в пространство переменных исходной модели, позволяющее решению агрегированной модели сопоставить решение исходной модели.

Важно отметить, что *операторов агрегирования* (как и *операторов дезагрегирования*) может быть сколь угодно много, поэтому можно рассматривать выделенное множество конкурирующих и взаимодополняющих операторов (с различным количеством агрегатов). К каждому наброску (агрегированной модели) можно применять все или некоторое подмножество операторов агрегирования. Крупноагрегированные динамические модели являются примерами финитных набросков. Таким способом строится оргграф набросков моделей, а решение сложной и громоздкой задачи вычисления точных воздействий (детализированного плана) подменяется значительно более простой и гибкой стратегией адаптивного «подруливания» при сохранении требуемой точности результата. На этом же принципе основана иерархия синдромных

моделей знаний: первоначально строятся синдромные модели для укрупненных заключений, затем для более детализированных и т.д.

Примером математической технологии построения орграфа набросков может служить *символическая динамика* [122]. При данном подходе фазовое пространство динамической системы разбивается на области таким образом, что любая ее траектория однозначно определяется последовательностью прохождения этих областей (символической последовательностью). По заданному разбиению фазового пространства строится ориентированный граф, называемый символическим образом.

Математическим аппаратом моделирования иерархического (пирамидального) способа размещения информации в памяти, а также вытекающего из данного размещения процесса иерархической обработки знаний, могут служить *Растущие пирамидальные сети* (РПС) [44, 45, 162]. РПС представляет собой семантическую сеть, или направленный граф, состоящий из узлов и связей между ними. Количество входных связей превышает количество выходных, что обуславливает пирамидальную архитектуру сети. В процессе обучения на вход РПС последовательно поступают объекты обучающей выборки, а количество конечных узлов сети равно длине выборки. Формулы верхнего уровня, обобщающие внутренние логические комбинации, задают модель той ситуации, которая отражена в выборке. Сотрудники Института математики СО РАН под руководством Загоруйко Н. Г. разработали на основе РПС способ семантического представления текстов в виде Q-сетей для автоматического построения онтологий [62].

Моделью построения некоторых орграфов набросков в методическом плане может служить применение *системы итерирующих функций* (СИФ) [83]. Функции из СИФ  $\{T_i\}$  являются сжимающими преобразованиями. Рассматривая  $x_0$  в качестве исходного образа, множество  $\{T_i\}$  в качестве полного множества взаимодополнительных агрегирующих операторов, а любую операцию  $x' = T_i(x)$  в качестве одной из операций построения наброска  $x'$  на основе наброска  $x$ , получим качественную аналогию с процессом построения орграфа набросков (к любому наброску применяются все операторы из  $\{T_i\}$ ). Ключевое отличие (в основной схеме) состоит в конечности итераций. Процесс завершается построением финитных набросков, а не фрактального множества  $F$ . Применение любого  $T_i$  к финитному наброску либо невозможно, либо приводит к пустому наброску.

Даниэл Деннет (Dennett, Daniel, американский философ) предложил «модель многократных набросков», согласно которой ментальная активность в мозге осуществляется в виде параллельных и перекрещивающихся процессов выбора, ревизии и интерпретации

получаемой чувственной информации. Каждая такая интерпретация гипотетична и может мгновенно сменяться другой, более соответствующей реальной ситуации [53].

Величковский Б. М. отмечает [38], что «...мозговые механизмы, лежащие в основе когнитивных модулей, отнюдь не похожи на случайную мозаику, образуя скорее упорядоченные иерархические структуры градиентного типа». Эти градиенты, по его представлениям, служат доказательством вертикального, многоуровневого измерения когнитивной организации. Обсуждаемые теоретические представления созвучны с развиваемыми в [70] положениями об иерархическом принципе интеграции *скоростных, медленных и сверхмедленных* информационно-управляющих систем головного мозга, формируемых для обеспечения психических состояний и познавательной деятельности человека, включая вербальную ассоциативно-мыслительную деятельность.

Орграфы набросков относятся к иерархическим когнитивным структурам градиентного типа (градиент задает обобщение). Очевидна также разница в скорости обработки набросков разного уровня общности (ввиду разного объема информации). Орграф набросков порождает огромное разнообразие качественных трансформаций, "фазовых переходов", сопряженных с процессами самоорганизации информации и возникновением новых параметров порядка (динамических аттракторов), в результате чего возникают новые знаки и символы (происходит категоризация). Осуществление когнитивной самоорганизации в рамках интеллектуальных приложений является одной из важнейших задач.

Как мы уже убедились, различных схем построения набросков очень много, поэтому ограничимся рассмотрением лишь нескольких схем, которые позволяют проиллюстрировать ключевые идеи общего подхода. В частности, это следующие схемы:

- орграфы доменов тестов как элементарные орграфы набросков базовых доменов;
- орграфы набросков на основе банка тестов;
- орграфы набросков на основе моделей знаний и редукции описания;
- орграфы набросков на основе грануляции и сжатия;
- орграфы набросков системопаттернов;
- орграфы набросков формируемых в процессе диагностического познания;
- орграфы набросков на основе индикаторов (экономика, критические состояния в медицине).

Важной особенностью первых пяти схем является то, что они могут быть реализованы с помощью автоматизмов среды.

Что касается орграфа доменов теста, то с помощью автоматизмов среды

строятся, в частности, структурно-завершенные орграфы (см. главу 2). Ствол орграфа доменов создается чаще всего с помощью экспертов ПрО.

Построению орграфов набросков на основе моделей знаний посвящена глава 5. Редукция описаний рассматривалась в главе 3 при анализе сходства. Базовый формализм построения орграфов набросков рассмотрен в разделе 3.4. В этом же разделе рассмотрены орграфы набросков на основе банка тестов. В частности можно отметить следующее.

Модель многоуровневых набросков образа является нелинейной математической моделью восприятия образа (ситуации, текста), описывающей предельные структуры разного уровня общности, суперпозиция которых дает полную характеристику образа. Можно считать, что некоторый образ понят (осмыслен) тогда и только тогда, когда построен орграф набросков данного образа (ситуации, текста, модели), а также построены синдромные и вероятностные модели знаний, экстремальные пограничные слои набросков. Суперпозиция набросков олицетворяет своеобразный *принцип неопределенности* в репрезентации образа.

Ряд набросков разного уровня общности для произвольного образа формируется на основе индуктивной рекуррентной схемы обобщений-абстракций, которая базируется на (полной) системе итерирующих сжимающих отображений. Пределом обобщения являются финитные наброски. К финитным наброскам либо не удастся применить ни одно из отображений, либо они исчезают в результате применения любого из итерирующих отображений. Существует экстремальный пограничный слой набросков, каждый из которых обладает предельными свойствами: к ним нельзя применить ни одно отображение без потери однозначности трактовки наброска. Каждый из экстремальных набросков по-своему минимален. Наброски, уровень обобщения которых выше уровня экстремальных набросков, имеют многозначную интерпретацию в рамках имеющейся базы прецедентов, поэтому они служат основой для ассоциаций и метафор. Многозначную интерпретацию имеют, как правило, и финитные наброски. Экстремальный пограничный слой набросков образуется в результате когнитивной самоорганизации и относится к базовому уровню. Оперирование критическими набросками обеспечивает максимальную скорость когнитивным процессам. В целом слои набросков орграфа отражают иерархический принцип интеграции скоростных, медленных и сверхмедленных информационно-управляющих подсистем когнитивной системы.

Сравнение разных образов начинается с максимального уровня общности, т.е. с финитных набросков. Если финитные наброски сравниваемых образов совпадают, то продолжается сравнение на



предпоследнем слое и так до тех пор, пока не будут найдены различающиеся наброски.

Орграфы набросков системопаттернов будут рассмотрены в разделе «Квантовая семантика» и в главе 7.

## **6.2 Многоуровневая морфология феномена действительности. Качественная прерывность**

Так как когнитивная динамическая система (КДС) является активной системой, т.е. активно взаимодействует с внешней средой, изучает внешнюю среду, то в ряде случаев будем ассоциировать КДС с наблюдателем. С наблюдателем связан мир качественных структур – неустранимое присутствие проявлений и свойств наблюдателя в наблюдаемом. Концентрация внимания наблюдателя на тех или иных аспектах ситуации, а также эмоциональный окрас существенно искажают пространство субъективной (феноменологической) реальности.

Примем, что локальное проявление качества в качественной действительности может быть представлено как степень соответствующих интенсивностей (цвет посредством частот, состояние холода температурой и т.д.). Интенсивности, как и любой тест, задаются орграфами доменов. Действительность (физическая), воспринимаемая наблюдателем, не однородна: в любой локальной области действительности в зависимости от контекста возникает свой уровень общности интенсивностей качеств действительности (в соответствии с конфигураторами интенсивностей). Изменения величин интенсивности качеств на разных уровнях могут быть как прерывными, так и непрерывными (квазинепрерывными).

Примем за основу следующую гипотезу [126]: «чувственные» феномены наблюдателя фиксируются относительно других феноменов в точности там, где имеется качественная прерывность. Качественная прерывность связана с масштабной грануляцией по пространству – времени и уровнями обобщения интенсивностей качеств.

Пусть  $W$  представляет собой пространственно-временной феномен. Различные наполняющие  $W$  качества  $\{q/Q: G(q)\}$  (прочность, цвет, текстура, температура, отражающая способность и т.д.) выражаются посредством степеней интенсивностей. Качества внутренне присущи самим объектам и связаны с определенными возможностями измерения. Примеры конфигураторов многих физических качеств приведены в главе 4.

Следуя [126], точку  $w \in W$  назовем *регулярной на заданном уровне общности*, если все интенсивности качеств заданного уровня общности непрерывны (квазинепрерывны) в окрестности  $w$ . Точку  $w \in W$  назовем *сингулярной на заданном уровне общности* (мерцающей), если и только

если имеется как минимум одно качество, являющееся прерывным на заданном уровне общности в  $w$ .

*Примечание:* если домен дискретный, то непрерывность (точнее – квазинепрерывность) означает либо тождественность значений в окрестности  $w$ , либо «соседние» значения (кроме оппозиций). Предполагается, что сенсоры наблюдателя могут настраиваться на заданный уровень общности  $\{q/Q\}$ .

Пусть  $M_{\{q/Q\}}$  – множество всех сингулярных точек  $W$  на уровне общности  $\{q/Q\}$ .  $M_{\{q/Q\}}$  назовем *морфологией* наполняющего  $W$  феномена на уровне общности  $\{q/Q\}$ . Базовый автоматизм наблюдателя должен выявлять качественную прерывность действительности на заданном уровне общности  $\{q/Q\}$ , т.е. морфологию  $M_{\{q/Q\}}$ .

*Примечание:* построение морфологии на разных уровнях общности является принципиальным для настоящего исследования, так как на одних уровнях качественная прерывность может иметь место, а на других нет.

*Набросок  $W$*  назовем результатом конечного числа итераций «обобщения – выделения морфологий» исходного феномена, который на сенсорном уровне описан с использованием базовых доменов  $\{q/Q_0\}$ . Наброски предельного уровня общности называются *финитными*. Все слои набросков феномена  $W$  образуют *орграф набросков*  $G_{S\Sigma}(W)$ , где  $\Sigma$  – схема сжимающих отображений. На рис.6.1 показаны примеры набросков для образа «Кораблик».

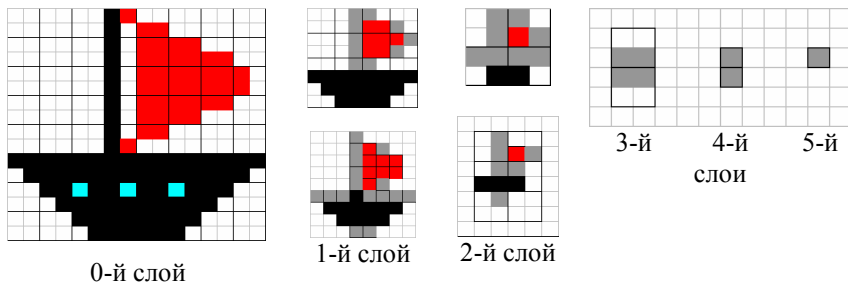


Рис. 6.1 – Фрагменты слоев орграфа набросков

Морфология во всех набросках обозначена серым цветом, характеризующим неопределенный – «мерцающий», многозначный цвет. Единственный финитный набросок на 5-ом слое состоит из одной сингулярной точки. Подобный финитный набросок имеют большинство компактных визуальных образов (так выглядят визуальные феномены либо на очень большом расстоянии, либо при очень слабой фокусировке – слабом внимании). Столь примитивные наброски могут играть роль

«заколок» в оперативной памяти когнитивной системы (как известно оперативная память у живых организмов очень маленькая). При приближении к объекту или при повышении внимания объект будет разворачиваться во все более детальный образ. В памяти КДС оргграф набросков представлен суперпозицией всех набросков (ни один набросок не может полностью представлять весь образ). Более детально один из вариантов данной схемы будет описан в следующем разделе.

Оргграф набросков  $G_{S\Sigma}(W)$  обладает предсказательной способностью при усилении внимания на  $W$  (например, при приближении), т.е. реализует модель «память – предсказание» и соответствующий макроавтоматизм [39, 199].

Если действительность представлена несколькими феноменами – объектами  $\{W\}$ , то возможно автоматическое выделение уникальных набросков максимального уровня общности для каждого объекта из  $\{W\}$ . Уникальные наброски образуют экстремальный пограничный слой набросков.

Фундаментальным следствием явления качественной прерывности является банк тестов  $\{G(\tau)\}$ , который служит системой координат феноменологического пространства (ментального многообразия).

Формируемая теория многоуровневой и многомодальной качественной прерывности может предложить полезную исходную точку для понимания природных процессов категоризации.

### 6.3 Оргграфы набросков на основе грануляции и сжатия

Рассмотрим для примера простейшие (модельные) формы обобщений, непосредственно осуществляемых в образном материале.

Пусть  $U$  – бесконечный  $N$ -мерный универсум, состоящий из однородных дискретных элементов  $u$ , содержащих коды  $\underline{u}$ . Все или почти все элементы  $U$  относятся к фону. Конечное число элементов  $U$  может относиться к какому-либо образу  $W$  или его наброску. Следовательно, каждый элемент универсума содержит либо код фона, либо код, относящийся к образу  $W$ . Коды образа отвечают определенной характеристике образа и принадлежат алфавиту  $L$ , совпадающему с базовым доменом конфигуратора характеристики. Алфавит  $L$  содержит выделенный элемент, отвечающий «неопределенному» значению (код морфологии). Количество элементов образа –  $|W|$ .

Введем наблюдателя с помощью локальной системы координат, связанной с образом  $W$ . Если существуют разные варианты задания локальной системы координат, то выбирается один из вариантов. Разные образы будем считать идентичными, если они совпадают в локальной

системе координат.

Введем множество сжимающих отображений  $\{T: U \rightarrow U\}$ . Отображения  $T$  могут изменять коды элементов универсума  $U$ , относящихся к образу или его наброскам (в локальных системах координат до и после преобразования). Любой набросок  $P$  получается в результате применения серии отображений из  $\{T\}$  к исходному образу  $W$ . Исходный образ  $W$  также является наброском. Наброски могут содержать морфологию.

Обозначим через  $T(P)$  множество элементов  $U$ , отличных от фона, полученных в результате  $T$ -преобразования наброска  $P$ . Может оказаться, что  $T(P) = \emptyset$ . Потребуем от  $T$  выполнения четырех условий: 1) если набросок  $P$  отсутствует, то  $T(U) \equiv U$ ; 2) если набросок  $P$  присутствует, то  $T$  связано с локальной системой координат наброска; 3)  $\forall u \in T(P), \underline{u} \in L$ ; 4)  $|T(P)| < |P|$ .

Отображения  $T \in \{T\}$  могут отличаться между собой как коэффициентом сжатия, так и привязкой к локальной системе координат. Некоторые отображения  $T$  можно отнести к числу операций «грануляция + сжатие». Выделим из множества  $\{T\}$  подмножество отображений «грануляция + сжатие»  $\{T\}'$  с одинаковой схемой сжатия (одинаковым коэффициентом сжатия), отличающихся только привязкой к локальной системе координат наброска. Примем, что число различных привязок конечно, следовательно,  $\{T\}'$  может быть полным множеством сжимающих однотипных отображений «грануляция + сжатие». В дальнейшем, без потери общности,  $\{T\}'$  будем отождествлять с  $\{T\}$ .

Пусть дан первичный образ  $W$  и фиксировано  $\{T\}$ . Первичный образ будем считать наброском 0-го слоя  $F_0$ . Определим по индукции наброски следующих слоев:

наброски 1-го слоя:  $F_1 = \cup_{T \in \{T\}} T(F_0)$ ;

наброски  $(k+1)$  слоя:  $F_{k+1} = \cup_{T \in \{T\}} T(F_k)$ ,

где  $F_j$  означает множество всех набросков  $j$ -ого слоя, а  $T(F_j)$  – применение  $T$ -отображения к каждому из набросков  $j$ -ого слоя.

Несложно убедиться, что для любого образа  $W$  существует  $K$ :

$$F_K \neq \emptyset, \text{ но } F_{K+1} = \emptyset.$$

Доказательство вытекает из конечности образа и условия (4), предъявляемого к отображениям.

Число  $K(W)$  является константой образа  $W$  (при фиксированном  $\{T\}$ ). Последний слой  $F_K$  состоит исключительно из финитных набросков, т.е. таких набросков, которые «исчезают» в результате применения любого  $T \in \{T\}$ , т.е. если  $P$  – финитный набросок, то  $\forall T \in \{T\}, |T(P)| = 0$ . Финитные наброски могут быть и в других слоях. Множество всех отличающихся финитных набросков обозначим через  $FS(W)$ . Если  $F_1 = \emptyset$ ,

то  $K(W) = 0$  и  $FS(W) \equiv W$ , т.е. множество  $FS(W)$  не бывает пустым.

Финитные наброски являются аттракторами итерационного процесса «грануляция + сжатие» любого образа. Произвольные наброски могут играть роль внутренних кодов разных уровней общности-абстракции для первичного образа  $W$ . Наброски высоких уровней общности инвариантны к множеству деталей. Кроме константы  $K(W)$ , константами образа являются также общее число набросков во всех слоях  $Kp(W)$ , число уникальных набросков, число набросков в каждом слое.

Ряд слоев набросков можно представить в виде ориентированного графа набросков  $G_s(W)$ . Ребра графа соединяют два наброска, если один переходит в другой в результате применения одного из отображений множества  $\{T\}$ . Отметим, что для любого наброска  $P \in G_s(W)$  оргграф  $G_s(P)$  является частью оргграфа  $G_s(W)$ . Следовательно, оргграф  $G_s(W)$  может быть точным фрагментом (подграфом) оргграфа набросков, построенного для более детального образа.

Движение по оргграфу  $G_s(W)$  от более общих набросков к более детальным позволяет реализовать *предсказание* и *замещение* образов в процессе узнавания действительности. При замещении интеллект «восполняет» отсутствующую информацию. Предсказания – это первичная функция и фундамент КДС [199].

Конкретизируем *индуктивную рекуррентную схему построения оргграфа набросков (визуального) образа* с учетом морфологии.

Пусть имеется бесконечное поле, состоящее из одинаковых клеток. Каждая клетка имеет определенный цвет (кодом цвета может быть число или слово, состоящее из набора символов). Имеются два выделенных цвета: цвет фона и неопределенный цвет. При отсутствии образа все клетки поля имеют цвет фона. Произвольный графический образ  $W$  представляет собой конечное множество клеток, цвет которых отличен от фона.

Без потери общности клетки будем считать прямоугольниками, образующими регулярную сеть. Определим сетку, состоящую из макро ячеек размером  $M_1 \times M_2$  клеток поля. Общее число клеток в ячейке сетки:  $N = M_1 \times M_2$ . Примеры образа и его набросков показаны на рис. 6.1 (сетка с макро ячейками размером  $2 \times 2$ ).

Пусть каким либо образом сетка первый раз размещена на поле. Легко убедиться, что существует ровно  $(N-1)$  вариантов размещения сетки, отличных от первого – базового (всего  $N$  вариантов). Если левый нижний угол любой ячейки базовой сетки принять за начало координат, то все  $(N-1)$  вариантов размещения сетки получаются из базового переносом начала координат в остальные клетки ячейки.

Зададим *схему итерационного обобщения-абстрагирования*  $\Sigma[N:J]$ , состоящую из двух пунктов:

1. Пусть положение сетки фиксировано. При обобщении множество клеток каждой ячейки преобразуются в одну клетку наброска следующего уровня, при этом цвет клетки определяется следующим образом: если  $J$  клеток ячейки имеют одинаковый цвет (возможно, цвет фона), то цвет клетки нового наброска принимает тот же цвет, в противном случае, неопределенный цвет (клетка является сингулярной, т.е. принадлежит морфологии).

2. Пункт 1 выполняется для первичного образа  $N$  раз с различными положениями сетки. В результате образуются  $N$  набросков 1-го слоя. При этом часть набросков могут совпадать. Оставляем в 1-ом слое только различные наброски. Их число обозначим  $K_1(W)$ . Подобную операцию проделываем с каждым из набросков 1-го слоя. Получаем наброски 2-го слоя. Оставляем во втором слое только уникальные наброски. Их число обозначим  $K_2(W)$ . И так продолжаем до тех пор, пока на каком-то шаге итерации не окажется ни одного наброска, т.е. при обобщении любого из набросков последнего слоя получается поле с цветом фона.

Общее число слоев набросков обозначим через  $K_\Sigma(W)$ . На  $j$ -ом слое число уникальных набросков  $K_j(W)$ , где  $j = 1, \dots, K_\Sigma(W)$ . Общее число набросков во всех слоях, включая 0-й слой, составляет  $K_p(W)$ . Число уникальных набросков во всех слоях обозначим  $K'(W)$ . Очевидно:  $K'(W) \leq K_p(W)$ .

На практике  $J$  выбирают близким к  $N$ . Примеры схем обобщения:  $\Sigma[3:2]$ ,  $\Sigma[4:3]$ ,  $\Sigma[16:12]$ ,  $\Sigma[9:6]$  и т.д. Наброски рис. 6.1 для  $W$  – «Кораблик» получены по схеме  $\Sigma[4:3]$ . Константы следующие:  $K_2(W) = 5$ ,  $K_p(W) = 31$ ,  $K'(W) = 28$ . Для сравнения на рис. 6.2 показаны некоторые наброски для образа  $W'$  – «Зонтик», которые также получены по схеме  $\Sigma[4:3]$ . Константы:  $K_2(W') = 6$ ,  $K_p(W') = 62$ . Морфология обозначена серым цветом.

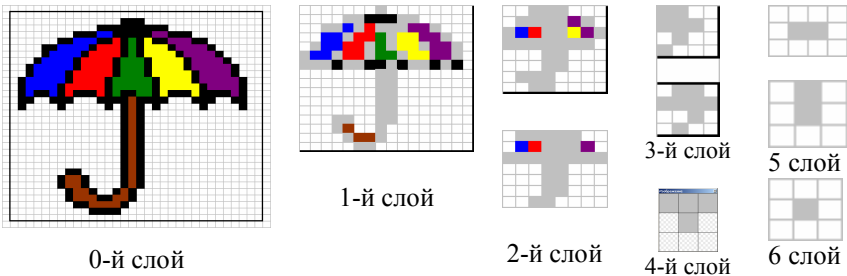


Рис. 6.2 – Фрагменты слоев орграфа набросков для «Зонтик»

На рис. 6.3 показан полный орграф набросков для образа «Зонтик» (вершина черного цвета вверху – исходный образ; единственный финитный набросок обозначен красным цветом).

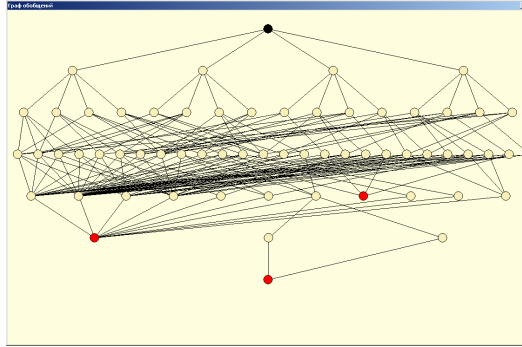


Рис. 6.3 – Орграф набросков для образа «Зонтик»

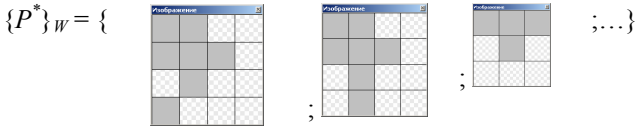
Как видно из рисунка 6.3, существует множество путей разворачивания одних и тех же набросков, в частности, финитных набросков. Это необходимо учитывать при реализации механизма «память–предсказание».

Сравнивая орграфы образов «Кораблик» и «Зонтик» можно отметить следующее. Увеличение числа слоев в орграфе «Зонтика» можно отнести на счет более детального исходного образа. Финитные наброски одинаковы. Более того, совпадают наброски предпоследнего слоя. Это значит, что на большом расстоянии или слабой фокусировке оба образа нельзя отличить. Отличающиеся наброски, т.е. наброски экстремального пограничного слоя для обоих образов появляются только на 3-м слое от конца. Рассмотрим этот вопрос подробнее.

Пусть  $\Omega = \{W\}$  – множество различающихся образов (ни один образ не совпадает с наброском другого). Множество  $\{T\}$  фиксировано. Через  $\{P\}_W$  обозначим множество всех различающихся набросков в орграфе  $Gs(W)$ . Через  $\{P^*\}_W$  обозначим множество набросков экстремального пограничного слоя. Наброски из  $\{P^*\}_W$  обладают двумя важными свойствами:

- они уникальны для  $W$ ;
- любое их обобщение операторами из  $\{T\}$  приводит к многозначным наброскам, т.е. наброскам, принадлежащим нескольким  $\{P\}_W$ .

Приведем примеры критичных набросков для образа  $W = \text{«Зонтик»}$  при условии, что  $\Omega = \{\text{«Зонтик»}; \text{«Кораблик»}\}$ , а  $\{T\} = \{T_1, \dots, T_4\}$  определяется схемой грануляции – сжатия  $\Sigma[4:3]$ :



Из примера видно, что критические наброски используют разный объем памяти. Справедливо следующее: чем меньше занимаемый объем памяти, тем выше уровень обобщения – абстрагирования наброска.

Таким образом, в рамках фиксированного контекста  $\langle \Omega, \{G_s(W)\}, \{G(\tau)\}, \Sigma \rangle$  множество всех набросков  $\{P\}_W$  для любого  $W$  делится на три непересекающихся подмножества: докритических, критических и закритических набросков. Каждый набросок является множеством значений некоторых тестов, т.е.  $P = \{\underline{z}/T\}$ . Ясно, что на множествах критических набросков можно построить модели знаний:

$$k = \{P^*_W \rightarrow z_W \mid W \in \Omega, P^*_W \in \{P^*\}_W\}. \quad (6.1)$$

Модель (6.1) является протосиндромной моделью знаний, но не синдромной и не предельной моделью. Дело в том, что даже критические наброски могут содержать избыточную информацию. Обратите внимание на последний набросок в приведенном выше ряду набросков  $\{P^*\}_W$ . Так вот, следующая комбинация значений тестов была бы минимально достаточной для распознавания образа «Зонтик»:



Однако эта комбинация не является наброском. На языке синдромов можно записать:  $S^* = (\tau_{2,1}?m; \tau_{1,2}?m; \tau_{3,2}?m; \rightarrow z/Z?1)$ , где  $m$  – код цвета морфологии;  $z?1 = \langle \text{Зонтик} \rangle$ ;  $Z = \{1, 2\}$ .

Выше уже отмечалось, что в сфере восприятия синдромы являются *перцептивными категориями* – целостными перцептивными гипотезами, свернутыми до некоего единичного перцептивного эталона, определяющего построение и распознавание образа. Следовательно, в рамках фиксированного контекста  $\langle \Omega(\{W\}), \{G_s(W)\}, \{G(\tau)\}, \Sigma \rangle$  может быть построена полная предельная синдромная модель знаний  $\{S^*\}_{Full}$ , которую можно представить следующим образом:

$$k = \{S^*_W \rightarrow z_W \mid W \in \Omega, S^*_W \in \{S^*\}_W\}. \quad (6.2)$$

Могут быть также построены модели (предельных) вероятностных закономерностей  $\{R^*\}_W$ .

Для любого множества набросков  $V \in \{P\}_W$  определим множество набросков-прообразов следующим образом:



$$(V/Gs(W))^{-1} = \{P' \mid \exists T \in \{T\}: T(P') \in V\}. \quad (6.3)$$

Могут быть определены прообразы набросков из  $(V/Gs(W))^{-1}$ , а именно:

$$((V/Gs(W))^{-1})^{-1} = (V/Gs(W))^{-2} \quad (6.4)$$

и т.д. по индукции. В результате получаем распространяющуюся волну от набросков  $V$ , которая играет фундаментальную роль в задачах предсказания и классификации. Если различные финитные наброски генерируют собственные волны, то орграф набросков образа является результатом интерференции таких волн. Поскольку финитные (закритические) наброски имеют многозначный смысл, то возникает интерференция волн разных образов, что порождает суперпозицию смыслов.

Относительно сенсориума КДС сделаем важное предположение, что он способен воспроизводить все многообразие набросков реальных образов при изменении фокусировки (внимания) или при изменении расстояния до объекта. Другими словами, перцептивные структуры сенсориума и механизм построения орграфа набросков должны быть синхронизированы.

Раскроем суть решения задачи динамического распознавания нового образа.

Пусть необходимо распознать новый разворачивающийся во времени образ, который в текущий момент представлен совокупностью набросков  $V_t$ . Наброски  $V_t$  проверяются в каждом множестве  $\{P\}_W$ , где  $W \in \Omega$ . Если все наброски входят в единственное  $\{P\}_W$ , то заключением становится образ  $W$ . Пусть наброски  $V_t$  входят в  $\{P\}_W$  и  $\{P\}_{W'}$ . В этой ситуации у сенсориума КДС запрашиваются уточненные данные относительно нового образа – следующий слой набросков  $V_{t+1}$  (если это оказывается возможным). Параллельно из памяти КДС извлекаются прообразы  $V_t$ , а именно:

$$X = (V_t/Gs(W))^{-1}; \quad Y = (V_t/Gs(W'))^{-1}. \quad (6.5)$$

Другими словами, выполняется операция предсказания, являющаяся автоматизмом КДС. Далее вычисляются множества  $X' = X \setminus Y$  и  $Y' = Y \setminus X$ . Наброски  $V_{t+1}$  сравниваются с  $X'$  и  $Y'$ , и если есть пересечение с одним из них, то распознавание завершается. В противном случае, у сенсориума запрашивается следующий слой набросков  $V_{t+2}$ , а операция (6.5) выполняется для  $V_{t+1}$ . И так до тех пор, пока не завершится распознавание. Приведенная схема легко обобщается на произвольное число возможных образов. Суть алгоритма не меняется и в том случае, если входной образ предъявлен сенсориуму КДС сразу с максимальным разрешением. КДС строит орграф набросков входного образа, и распознавание начинается с финитных набросков, т.е. с предельного огрубления.

Алгоритм динамического распознавания существенным образом опирается на природную категоризацию (перцептивные категории), что обеспечивает наивысшую степень абстракции и помехозащищенности. Отметим также, что описанный алгоритм учитывает принципиальную многозначность проявлений образа, что выражается в суперпозиции набросков (более подробно об этом пойдет речь в разделе «Квантовая семантика»).

Рассмотрим пример. Пусть  $\Omega = \{W = \text{«Зонтик»}; W' = \text{«Кораблик»}\}$ , а предъявляемым объектом является какой-то зонтик (может отличаться от эталонного зонтика в БД), но КДС об этом не знает. Первоначально предъявляется  $V_t$ , содержащее всего один финитный набросок в виде сингулярной точки (клетки серого цвета – 6-й слой на рис. 6.2). Ясно, что  $V_t$  принадлежит и  $\{P\}_W$ , и  $\{P\}_{W'}$ , т.е. данных для распознавания недостаточно. КДС запрашивает у сенсориума  $V_{t+1}$ , а также извлекает из памяти прообразы  $X$  и  $Y$ . Поскольку  $X'$  и  $V_{t+1}$  содержат как минимум один общий набросок, напоминающий букву «Т» (см. выше примеры критичных набросков для «Зонтика»), то за один шаг уточнения (т.е. при очень слабом разрешении) удастся провести распознавание.

Приведенный выше алгоритм динамического распознавания реализован нами в ряде учебных и медицинских приложений (рис. 6.4). В каждом приложении применялись свои схемы построения орграфов набросков.



а) Распознавание символов

б) Вер. диагноз – гайморит

Рис. 6.4 – Примеры динамического распознавания

Приведем примеры критичных набросков для образа  $W = \text{«9»}$  при условии, что  $\Omega = \{\text{«0»}; \text{«1»}; \dots; \text{«9»}\}$ , а  $\{T\}$  задается схемой  $\Sigma[4:3]$ :

$$\{P^*_{W'}\} = \{ \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \end{array} ; \begin{array}{|c|c|c|c|} \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \square & \square & \square & \square \\ \hline \end{array} ; \dots \}$$

Орграф набросков для образа  $W = \langle 9 \rangle$  с отображением экстремального пограничного слоя набросков (синие вершины в среднем слое) показан на рис. 6.5 (сравните с орграфом набросков «Зонтика»).

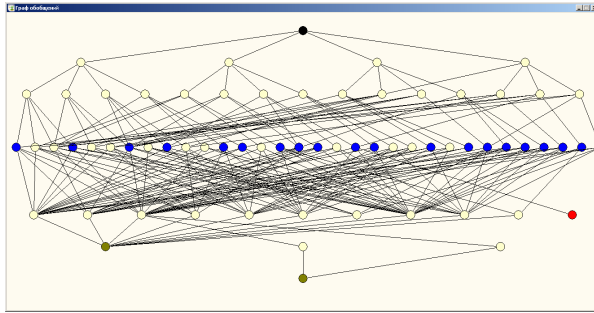


Рис. 6.5 – Орграф набросков для образа «9»

Единственным минимальным по объему критическим наброском для образа  $W = \langle 0 \rangle$  является следующий набросок:



На рис. 6.6 показаны диалоговые окна приложения, с помощью которого создаются различные базы прецедентов  $\Omega$ , строятся орграфы набросков визуальных образов и проводится динамическое распознавание.

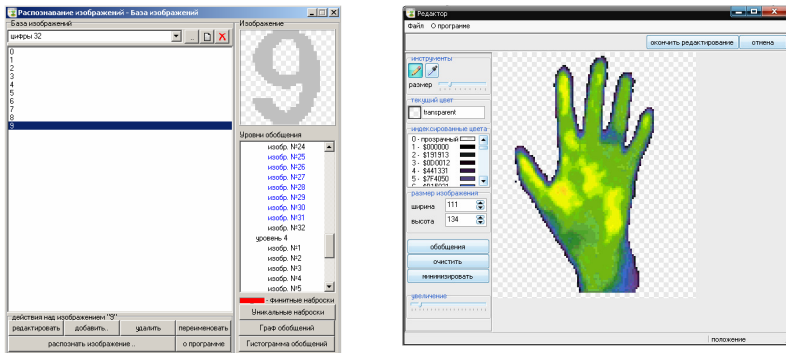


Рис. 6.6 – Окна приложения для работы с визуальными образами

Рассмотрим внутреннюю структуру финитных набросков, образующихся в рамках схемы  $\Sigma[N:J]$ .

*Предельной точкой (предельным многообразием)* назовем совокупность клеток образа/наброска, которая помещается в одной ячейке сетки и при однократной итерации по схеме  $\Sigma[N:J]$  сливается с фоном («исчезает»).

**Предложение 6.1.** По схеме обобщения-абстрагирования  $\Sigma[N:J]$  предельное многообразие (точка) содержит не менее одной клетки, но не более  $(N-J)$  клеток образа/наброска.

Действительно, при любом (из  $N$ ) наложении ячейки на изолированное скопление из  $(N - J)$  точек образа она (ячейка) будет содержать не менее  $J$  точек с цветом фона. Согласно схеме однократного обобщения произойдет сжатие в одну клетку с цветом фона. Это означает, что рассматриваемая совокупность клеток являлась предельной точкой. По схеме обобщения  $\Sigma[N:(N-1)]$  предельная точка содержит строго одну клетку образа/наброска.

**Предложение 6.2.** Пусть задана итерационная схема обобщения-абстрагирования  $\Sigma[N:J]$  и произвольный образ  $W$ , тогда финитные наброски образа  $W$  состоят из конечного множества предельных точек, возможно, одной точки.

Как правило, предельные многообразия являются морфологией (пример – рис. 6.1 и 6.2), но бывают исключения, например, у «Черного квадрата» по схеме обобщения  $\Sigma[4:3]$  имеется два различных финитных наброска с предельными точками: один набросок содержит черную клетку (не является сингулярной), другой сингулярную (состоит из морфологии).

С образом «Черный квадрат» связано следующее утверждение. Пусть  $W_0$  – «Черный квадрат», тогда если  $W \subset W_0$ , то  $K_\Sigma(W) \leq K_\Sigma(W_0)$ . Данный факт позволяет оценить сверху константу  $K_\Sigma(W)$  для любого образа  $W$ .

Таким образом, когнитивная динамика восприятия любого образа описывается двумя разнонаправленными процессами: (1) сворачивание исходного образа до предельного уровня обобщения – финитных набросков; (2) развертывание образа от финитных набросков до максимума – исходного образа. При взаимодействии с действительностью на основе  $\{Gs(W)\}$  реализуется громадное количество предсказаний, возникающих за пределами понимания. Но когда поступает образ, который КДС не знает, предсказание нарушается. И ошибка привлекает внимание.

Выше в качестве примеров использовались двумерные образы. Однако все рассуждения остаются в силе для пространств любой размерности, в частности, одномерного. Приведем простой пример.

Пусть имеется произвольный дискретный сигнал конечной длины. Закодируем цветом значения сигнала. Дополним сигнал слева и справа бесконечным фоновым сигналом. Для построения орграфа набросков применим схему  $\Sigma[3:2]$ . На рис. 6.7 – 6.8 показан пример задания сигнала и

построения орграфа набросков сигнала с использованием MS Excel.

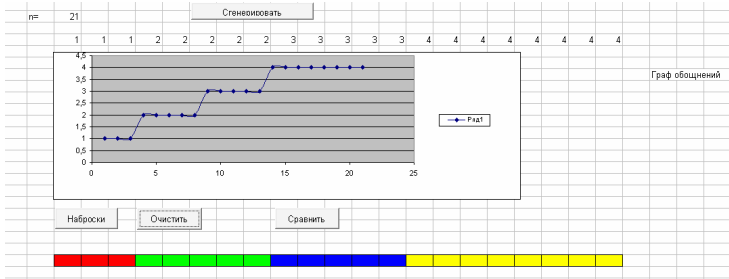
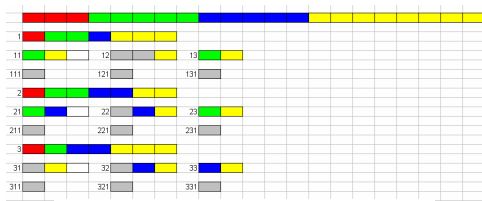
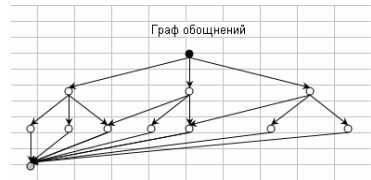


Рис. 6.7 – Задание и кодирование сигнала



а) Построение набросков сигнала



б) Орграф набросков

Рис. 6.8 – Автоматическое построение орграфа набросков сигнала

Как видим финитный набросок сигнала, как и большинства визуальных образов, состоит из одной сингулярной точки. Как и образы, сигналы сравниваются между собой, начиная с финитных набросков (динамическое распознавание).

Схема обобщения – абстрагирования  $\Sigma[N:L]$  может применяться не только к визуальным образам или сигналам. Приведем пример схемы для анализа надежности. Рассмотрим иерархическую систему, фрагмент которой изображен на рис. 6.9. Система разбита на уровни, которые можно интерпретировать как степени детализации описания (чем ниже уровень, тем детальнее) или как различные стадии развития системы.



Рис. 6.9 – Фрагмент иерархической системы (модель Подлазова А.В.)

Каждый элемент уровня  $i > 0$  состоит из трех элементов предыдущего

уровня  $i-1$ . Элементы системы могут быть исправны или дефектны. Элементы нижележащих уровней, тройками объединяющиеся в элемент следующего уровня, передают ему свое состояние в соответствии с его восприимчивостью к дефектам, под которой понимается минимальное число дефектных составляющих, необходимых чтобы и он стал дефектным. Так, элемент с восприимчивостью  $k$  будет дефектным, если дефектны не менее  $k$  его составляющих. Таким образом, возможны три типа элементов: с  $k = 1; 2; 3$ , т.е. становящиеся дефектными, соответственно, если дефектна хотя бы одна составляющая, если дефектны не менее двух составляющих и если дефектны все три составляющие. Соответственно, получили схему  $\Sigma[N:L]$  применительно к оценке надежности.

#### 6.4 Орграф набросков формируемый в процессе диагностического познания (на примере медицины)

Эффективные методы сжатия информации и построения набросков сложных образов выработаны в клинической медицине. Проанализируем логическую структуру диагностического познания с целью выявления формальных техник и процедур, которые помогут реализовать подобный процесс в рамках интеллектуальных технологий (в том числе и в других сферах).

Представленная в таблице 6.1 система логических операций, осуществляемых в процессе диагностического познания, включает в себя восемь этапов, объединенных четырьмя уровнями (описываемая структура диагностического познания соответствует [172]).

Таблица 6.1 – Этапы диагностического познания (построения набросков)

Этапы познания	Название этапа	Уровни познания
I этап	Установление наличия проблемной ситуации (состояния «нездоровья»). Сбор эмпирических данных.	Симптомологический
II этап	Определение локализации проблемы (определение анатомической локализации болезни)	
III этап	Формирование (клинических) синдромов	Синдромологический
IV этап	Выделение ведущего синдрома	
V этап	Определение характера общепатологического процесса	

VI этап	Формирование дифференциально - диагностического ряда проблем (нозологий)	Проблемный (нозологический)
VII этап	Выделение одной проблемы (нозологической единицы)	
VIII этап	Установление развернутого заклучения о состоянии сложной системы (клинического диагноза)	Заклучительный (клинический)

На практике, в зависимости от конкретной ситуации, определяющее значение приобретает тот или иной этап диагностического познания.

Первый этап представляет собой этап сбора эмпирических данных  $\{\underline{z}/T\}$ . Задачей первого этапа диагностического процесса является оценка состояния системы в целом на основании анализа наиболее общих неспецифических признаков и заключение о наличии или отсутствии проблемной ситуации.

Задачей второго этапа является выделение подсистемы сложной системы, которая максимально поражена патологическим процессом. Этим решается вопрос о «месте локализации проблемы». Однако в ряде случаев, несмотря на тщательно проведенный анализ, так и не удастся определить локализацию проблемы. Причиной такого положения может быть то, что проблема еще не проявилась на уровне того или иного элемента системы, либо диагностируемая проблема вообще не имеет четкой локализации. Например, при наличии интоксикации инфекционной природы (сепсис) источник патологического процесса может быть определен позже с помощью лабораторно-инструментальных методов на этапе дифференциальной диагностики.

Методологическая общность I и II этапов диагностического процесса состоит в анализе симптоматики проблемы (внешних проявлений проблемы), т.е. множества значений первичных тестов  $\{\underline{z}/T\}$ . Это позволяет условно рассматривать начальные этапы диагностики в рамках единого *симптомологического* уровня познания.

Получаемая на симптомологическом уровне диагностическая информация предоставляет возможность формирования «абстрактных целостных образов» – синдромов (предметных), характеризующих определенные стороны сущности проблемы (третий этап диагностического процесса). *Синдромы* (предметные, например, клинические) – это принципиально новые укрупненные и специфические по своей природе диагностические ориентиры, а не просто совокупность первично выявленных симптомов и признаков (тестов). Свойством синдрома как целостного образа является неделимость. Предметные синдромы привязаны к определенной предметной области (отсюда и характеристика –

предметные). Синдромы могут появляться, изменяться, исчезать или сочетаться друг с другом в различных комбинациях, представляя собой весьма динамичный образ. Простейший оргграф доменов (конфигуратор) теста-синдрома приведен ниже:

$$\text{Синдром } \wedge Si \{1 \{ \text{Имеется } (f/\mu); \text{Отсутствует} \} \}, \quad (6.6)$$

где  $f/\mu: \{\underline{z}/T\} \rightarrow Si$  – известное отображение ( $\mu$  – механизм реализации отображения), позволяющее установить наличие синдрома по имеющимся эмпирическим данным  $\{\underline{z}/T\}$ . Отображение  $f/\mu$  можно отнести к онтологическим соглашениям. Для иллюстрации связи между первичными тестами  $\{\underline{z}/T\}$  и синдромом  $Si$  будем применять нотацию  $Si(\{\underline{z}/T\})$ .

*Примечание:* некоторые формальные синдромы (особенно предельные) можно рассматривать как частный случай *предметных синдромов*. Последние будем обозначать  $Si$ . Отличительной особенностью формальных синдромов является то, что формальные синдромы всегда указывают на определенное заключение –  $S(\{\underline{z}/T\}, \underline{z}/Z)$ , а предметные синдромы  $Si(\{\underline{z}/T\})$  далеко не всегда можно связать с единственной проблемой (заключением). К предметным синдромам можно отнести также некоторые (предельные) вероятностные закономерности из  $\{R\}$ . Обнаружение нескольких вероятностных закономерностей может служить примером сложного синдрома, в том числе патогномичного.

Автоматически формируемые синдромные модели знаний значительно расширяют круг известных предметных синдромов. В этом их главное предназначение. Выявление любого формального синдрома на любой стадии диагностического познания позволяет сразу же предположить (или обосновать) определенное заключение. В отличие от синдрома *симптомокомплекс*  $\{\underline{z}/T\}$  представляет собой всего лишь набор отдельных симптомов и признаков (тестов), не обладающий неделимой целостностью.

Выявление множества предметных синдромов позволяет перейти от первичного наброска (образа), представленного эмпирическими данными  $\{\underline{z}/T\}$ , к обобщенному образу – наброску  $\{Si\}$ . При этом все последующие процедуры диагностического процесса осуществляются с использованием гораздо меньшего количества «единиц памяти». Выявление больших (сложных) синдромов позволяет построить еще более сжатые наброски.

Задачей четвертого этапа является выделение ведущего синдрома  $Si^*$ . *Ведущий синдром* – это один или, реже, два синдрома, проявления которых наиболее выражены и преобладают в картине развития патологического процесса. Основной смысл выделения ведущего синдрома заключается в том, что этим достигается исключение из первоочередного рассмотрения значительного количества проблем, проявления которых не могут быть выражены данным синдромом. Таким образом, суть четвертого этапа



заключается в проведении межсиндромной дифференциальной диагностики и выделении множества ведущих синдромов  $\{Si\}^*$ .

Выделение ведущего синдрома позволяет выбрать конкретный алгоритм дифференциальной диагностики (АДД) из множества алгоритмов [111]. Если выделено несколько ведущих синдромов (как правило, не больше двух), то необходимо пройти АДД для каждого синдрома и сравнить результаты. Все АДД для каждого синдрома содержатся в Многоцелевом Банке Знаний ПрО (см. главу 7). Пример АДД был приведен в главе 3.

Задачей пятого этапа является установление общепатологического процесса, характеризующего синдромы, в том числе и ведущий. Возможность выделения того или иного общепатологического процесса определяется тем, что большинство проблем условно можно разделить на сравнительно небольшое число групп, в каждую из которых входят проблемы, имеющие сходную сущность. Например, в медицине выделяют следующие общепатологические процессы [172]: воспаление (как реакция на повреждение ткани); дистрофия и некроз (как результат повреждения клеток и тканей); опухолевый процесс (как неудержимый клонированный рост клеток, приводящий к разрушению окружающих тканей и гибели организма); иммуно-патологические процессы (как реакция на инвазию чужеродных клеток и веществ); нарушение общего и местного кровообращения; врожденные аномалии развития; невротические расстройства (функциональные).

Задачей шестого этапа является формирование дифференциально-диагностического ряда проблем (например, нозологий)  $\{Pr\}$ , обладающих патогенетической общностью в рамках единого ведущего синдрома. Основной смысл данного этапа состоит, таким образом, в исключении из первоочередного рассмотрения групп проблем, которые не могут быть обусловлены предварительно выявленным ведущим синдромом, рассматриваемым в рамках определенного общепатологического процесса. Необходимо также, чтобы проблемы, включенные в дифференциально-диагностический ряд, обладали общими доминирующими симптомами, выявленными на первом уровне.

Задачей седьмого этапа является выделение одной проблемы  $Pr^*$  из дифференциально-диагностического ряда  $\{Pr\}$ , сформированного на предыдущем этапе. Выделение проблемы осуществляется методом дифференциальной диагностики. Инструментом проведения такой диагностики являются, в частности, алгоритмы частной диагностики (АЧД) [111, 134]. Подобные алгоритмы являются неотъемлемой частью Многоцелевого Банка Знаний интеллектуального приложения. Если проблемы описаны как концепты в рамках онтологии ПрО (примеры

концептов приведены в главе 3), то для проведения дифференциальной диагностики можно применять методы обобщенного формального анализа понятий. В результате выполнения этапа выделяется та проблема, степень вероятности которой оказывается наибольшей исходя из выявляемых условно-специфических признаков. В случае обнаружения специфических признаков (например, формальных синдромов), суждения о вероятности той или иной проблемы достигают достоверного уровня. Выделение проблемы  $Pr^*$  является финитным наброском в орграфе набросков.

На восьмом этапе формируется развернутое заключение о состоянии сложной системы. Общая схема развернутого заключения содержится в отраслевом стандарте описания проблемы  $Pr^*$ .

Результирующий орграф набросков, формируемых в процессе диагностического познания, можно представить следующим образом:

$$\{\underline{T}\} \Rightarrow \{Si\} \Rightarrow \{Si\}^* \Rightarrow \{Pr\} \Rightarrow \{Pr\}^*. \quad (6.7)$$

Орграф набросков вида (6.7) можно рассматривать в качестве концептуальной схемы, которая универсальна по отношению к ряду предметных областей.

Построим формальную модель процесса диагностического познания, которая объединяла бы в себе синдромный и алгоритмический принцип диагностики, а также механизм обоснования диагноза (частная диагностика) [134]. Все компоненты модели должны принадлежать Многоцелевому банку знаний. Сама модель (логическая схема) также помещается в банк знаний (БкЗ). Введем следующие обозначения:

$s$  – симптом (факт, тест  $\underline{T}$ ),  $SM$  – все симптомы из БкЗ;

$Ps(s)$  – предикат, устанавливающий факт наличия или отсутствия симптома  $s$  у изучаемой системы ( $\forall s, Ps(s) = f \vee t$ );

$Sp$  – все симптомы (тесты  $\underline{T}$ ), обнаруженные у системы;

$Si$  – (предметный) синдром,  $SN$  – все синдромы из БкЗ;

$P(Si, Sp)$  – предикат, устанавливающий факт наличия или отсутствия синдрома  $Si$  ( $\forall Si, P(Si, Sp) = f \vee t$ ). Для реализации предиката  $P$  используется онтология (6.2);

$\{Pr\}$  – дифференциальный комплекс заключений, на выходе алгоритма – результирующий комплекс заключений;

$\{ADD(Si) \mid Si \in SN\}$  – древовидные алгоритмы дифференциальной диагностики из БкЗ (с алгоритмом работает интеллектуальный агент, например, врач), на выходе алгоритма формируется  $\{Pr\}$ ;

$\{CD(Pr) \mid Pr \in БкЗ\}$  – предикаты, отражающие результаты работы частных диагностических алгоритмов из БкЗ ( $\forall Pr, CD(Pr) = f \vee t$ ).

Формальная модель процесса диагностического познания:

Шаг 1. Выявляем наличие у системы симптомов  $s$  из  $SM$ :

$Sp := \emptyset$ ; **for** всех  $s \in SM$  **do if**  $P_s(s) = t$  **then**  $Sp := Sp \cup s$ ;

Шаг 2. Выявляем наличие у системы синдромов  $Si$  из  $SN$  (на основе  $Sp$ ):

$\{Si\} := \emptyset$ ; **for** всех  $Si \in SN$  **do if**  $P(Si, Sp) = t$  **then**  $\{Si\} := \{Si\} \cup Si$ ;

Шаг 3. Для каждого  $Si$  из  $\{Si\}$  запускаем алгоритм диагностики АДД( $Si$ ), а результат добавляем к  $\{Pr\}$ :

$\{Pr\} := \emptyset$ ; **for** всех  $Si \in \{Si\}$  **do**  $\{Pr\} := \{Pr\} \cup АДД(Si)$ ;

Шаг 4. Для каждой  $Pr$  из  $\{Pr\}$  проводим частную диагностику, которая либо подтверждает, либо отвергает  $Pr$ :

**for** всех  $Pr \in \{Pr\}$  **do if**  $ЧД(Pr) = f$  **then**  $\{Pr\} := \{Pr\} \setminus Pr$ ;

Шаг 5. Выход -  $\{Pr\}$

**return** ( $\{Pr\}$ ).

Построенная (упрощенная) модель процесса диагностического познания далеко не исчерпывает всех сценариев диагностики. Приведем примеры орграфов набросков, формируемых в процессе диагностического познания:

$\{\underline{g}T\} = \{Кашель? \text{ имеется}; Частота дыхания? \text{ увеличена}; Одышка? \text{ имеется}; \{Характер одышки? \text{ смешанная}\}; Температура тела? \text{ повышенная}; Аппетит? \text{ снижен}; Перкуторный звук? \text{ притупленный}; Распространенность перкуторных изменений? \text{ локальные}; Характер дыхания при аускультации? \text{ ослабленное}; Хрипы? \text{ влажные мелкопузырчатые}; Распространенность аускультативных изменений? \text{ локальные}; Рентгенограмма: легочной рисунок? \text{ усилен}; Прозрачность легочных полей? \text{ снижена}; Распространенность изменений на рентгенограмме? \text{ локальные}; Количество лейкоцитов в крови? \text{ повышено}\} \Rightarrow \{Si\} = \{Синдром интоксикации? \text{ имеется}; Синдром уплотнения легочной ткани? \text{ имеется}; Синдром дыхательной недостаточности? \text{ имеется}; Катаральный синдром? \text{ имеется}\} \Rightarrow Pr^* = \text{“Острая пневмония”};$

$\{\underline{g}T\} = \{Кашель? \text{ отсутствует}; Частота дыхания? \text{ норма}; Одышка? \text{ отсутствует}; Температура тела? \text{ норма}; Аппетит? \text{ норма}; Перкуторный звук? \text{ ясный легочный}; Распространенность перкуторных изменений? \text{ диффузные}; Характер дыхания при аускультации? \text{ жесткое}; Хрипы? \text{ влажные среднепузырчатые}; Распространенность аускультативных изменений? \text{ диффузные}; Рентгенограмма: легочной рисунок? \text{ усилен}; Прозрачность легочных полей? \text{ норма}; Распространенность изменений на рентгенограмме? \text{ диффузные}; Количество лейкоцитов в крови? \text{ норма}\} \Rightarrow \{Si\} = \{Синдром интоксикации? \text{ отсутствует}; Синдром уплотнения легочной ткани? \text{ имеется}; Синдром дыхательной недостаточности? \text{ отсутствует}\}.$

отсутствует; *Катаральный синдром?* имеется}  $\Rightarrow Pr^*$  = “Острый (простой) бронхит”.

В данных примерах синдромы и заключения устанавливались на основе онтологических соглашений (см. главу 3).

### 6.5 Орграфы набросков на основе индикаторов

Эффективным методом построения орграфов набросков является расчет всевозможных индикаторов (показателей, индексов), характеризующих течение сложного процесса. Индикаторы применяются практически во всех областях знаний: экономике, медицине, экологии, политике, социальных исследованиях. Каждый индикатор  $p$  является тестом и, следовательно, задается орграфом доменов  $G(p)$ . Среди экономических индикаторов первостепенное значение имеют показатели состояния и результатов функционирования экономики в целом, которые часто называют *агрегированными показателями*  $\{p/P\}^*$ .

В контексте построения орграфа набросков первичным этапом является определение *компонентов* (агрегированных) экономических показателей  $\{c/C\}$  [102]. Например, ВВП может быть первоначально разделен на три составные части: продажи конечным потребителям внутри страны, чистый экспорт и изменение товарно-материальных запасов и т.д. Другими словами, уровней компонентов может быть несколько. Обозначим промежуточные уровни через  $\{c/C\}$ .

Пусть  $\{\underline{z}/T\}$  – значения первичных социально-экономических, политических, экологических или медицинских показателей (тестов),  $UB$  – уровень воздействия (например, уровень помощи [134]). Конструкт орграфа набросков на основе индикаторов и индексов можно изобразить следующим образом:

$$\{\underline{z}/T\} \Rightarrow \{c/C\}' \Rightarrow \{c/C\} \Rightarrow \{p/P\} \Rightarrow \{p/P\}^* \Rightarrow \{Риск/ UB/ Прогноз\}. \quad (6.8)$$

Финитный набросок  $\{Риск/ UB/ Прогноз\}$  задает интерпретацию исходного образа  $\{\underline{z}/T\}$ .

Орграф набросков может использоваться в рамках метода предельных обобщений с целью построения синдромных моделей знаний. Покажем это.

Пусть изучаются факторы, влияющие на наступление некоторого события, например, банкротства предприятия. Положим  $Z = \{\text{событие}^+; \text{событие}^-\} = \{+; -\} = \{1, 2\}$ . Сформируем первичную базу прецедентов следующего вида:  $\Omega = \{\alpha\{\underline{z}/T\}, \underline{z}/Z\}$ , где  $Z = \{1, 2\}$ . Если речь идет о банкротстве предприятий, то  $\Omega^+ = \{\alpha\{\underline{z}/T\}, \underline{z}=1\}$  – множество обанкротившихся предприятий, соответственно –  $\Omega^- = \{\alpha\{\underline{z}/T\}, \underline{z}=2\}$  –

множество стабильных (успешных) предприятий.

Орграф набросков (6.8) позволяет перейти от первичного описания базы прецедентов к расширенному описанию следующего вида:

$$\Omega' = \{\alpha\{\tau/T\}, \{\underline{c}/C\}', \{\underline{c}/C\}, \{\underline{p}/P\}, \{\underline{p}/P\}^*, \underline{z}/Z\}, \text{ где } Z = \{1, 2\}.$$

Переход от описания  $\Omega$  к описанию  $\Omega'$  позволяет задействовать все виды обобщений, как по горизонтали (в рамках одного прецедента), так и по вертикали (пример комбинации нескольких видов орграфов набросков). Далее, в рамках контекста  $K = \langle \Omega', \{G(\tau)\} \rangle \equiv \langle \Omega', \{G(\tau)\}, \{G(c)\}, \{G(p)\} \rangle$  строятся синдромные модели знаний  $\{S\}$ , в частности, предельные синдромные модели знаний  $\{S^*\} = \{S^*\}^+ \cup \{S^*\}^-$ , а также модели (предельных) вероятностных закономерностей  $\{R^*\} = \{R^*\}^+ \cup \{R^*\}^-$ .

Множество синдромов  $\{S^*\}^+$  определяет все возможные наборы параметров порядка (в рамках контекста  $K$ ), которые приводят к наступлению «события+» (например, банкротству), а закономерности из  $\{R^*\}^+$  являются предвестниками события. Напротив, множество синдромов  $\{S^*\}^-$  определяет все возможные наборы параметров порядка, которые обеспечивают благоприятный режим функционирования. На основе знания указанных множеств синдромов может быть реализован *синдромный принцип управления*, целью которого является выход на благоприятный режим функционирования и его стабилизация (см. главу 7).

Конкретизируем приведенную выше схему применения орграфа набросков в рамках МПО для изучения факторов банкротства предприятия.

Текущий риск банкротства предприятия будем оценивать с помощью пяти тестов:  $X1$  – общая оборачиваемость активов предприятия;  $X2$  – отношение чистой прибыли к активам предприятия;  $X3$  – отношение краткосрочных и долгосрочных заемных средств к активам предприятия;  $X4$  – отношение долгосрочных обязательств к активам предприятия;  $X5$  – натуральный логарифм выручки предприятия. Собственно риск  $P$  задается некоторой функцией от параметров  $X1, \dots, X5$ , т.е.  $P = P(X1, \dots, X5)$ . Банк тестов  $\{G(\tau)\} = \{G(X1), \dots, G(X5), G(P)\}$ . Без потери общности базовый домен во всех орграфах обозначим «1». Текущий риск банкротства является примером интегрального показателя (индикатора).

В базе прецедентов по каждому предприятию будем хранить данные за несколько лет предшествовавших банкротству, т.е. множество кортежей вида  $\{\langle X1, \dots, X5 \rangle_{1, \dots}, \langle X1, \dots, X5 \rangle_N\}$ , где  $N$  – количество лет, включая год банкротства. Для успешных предприятий также берутся данные за  $N$  лет. Следовательно, первичная база прецедентов имеет вид:

$$\Omega(\{\tau/T_0\}) = \{\alpha\{\langle X1/1, \dots, X5/1 \rangle_{1, \dots}, \langle X1/1, \dots, X5/1 \rangle_N\}, \underline{z}/Z\},$$

где  $Z = \{1, 2\}$ . Расширенная база прецедентов, включающая индикаторы

(риски), имеет вид:

$$\Omega^*(\{T_0\}) = \{\alpha(\{<X_1/1, \dots, X_5/1, P/1>_1, \dots, <X_1/1, \dots, X_5/1, P/1>_N\}, \underline{z}/Z)\}.$$

На основе контекста  $K = \langle \Omega^*, \{G(\tau)\} \rangle$  строятся предельные синдромные модели знаний  $\{S^*\}$  и множества предельных вероятностных закономерностей  $\{R^*\}$ . Каждый предельный синдром из  $\{S^*\}^+$  описывает, по сути, набор параметров порядка (сценарий) банкротства предприятия и, соответственно, характер изменения финансовых показателей в ретроспективной динамике. Каждая предельная вероятностная закономерность из  $\{R^*\}^+$  является предвестником события. На основе синдромных моделей  $\{S^*\}$  и моделей  $\{R^*\}$  может быть выполнено построение комплекса динамических моделей прогнозирования банкротства путем выделения синдромов и вероятностных закономерностей разной степени удаленности от интересующего события, в частности, ранних предвестников события.

По аналогии строятся синдромные модели (наборы параметров порядка) и модели вероятностных закономерностей для произвольных экономических, социальных, политических и иных событий.

## 6.6 Орграфы набросков на основе оценочных и прогностических шкал

Объем данных, полученных с помощью клинических, лабораторных и инструментальных методов исследования, экспоненциально растет в течение нескольких последних десятилетий. Возросшие объемы информации потенцируют большие трудности в интеграции этих данных для получения достоверных оценочных и прогностических решений. Необходимость одновременно использовать большие объемы информации может привести к неэффективности процесса принятия решений, неоправданным различиям в терапевтических подходах и к ошибкам. Неправильный прогноз у пациента отделения интенсивной терапии чреват или неоправданной эскалацией терапии, которая зачастую сама является небезопасной или же наоборот отказом от терапии в пользу мероприятий по элементарному поддержанию жизни [3, 43, 111].

Для того чтобы принять правильное решение в отношении оценки тяжести и прогноза, касающегося пациента отделения интенсивной терапии, необходимо использование специальных инструментов, например, шкал оценки тяжести и прогнозирования исхода. С помощью данных шкал возможно эффективное поэтапное сжатие информации, другими словами – построение орграфа набросков ситуации действительности (в данном случае – клинической ситуации). Пример – шкала Апгар для оценки

состояния новорожденного (таблица 6.3).

Таблица 6.3 – Шкала Апгар

Признак	Баллы		
	0	1	2
ЧСС	Отсутствуют	< 100 уд/мин	Больше 100 уд/мин
Дыхание	Отсутствует	Слабый крик, гиповентиляция	Хорошее, сильный крик
Мышечный тонус	Вялый	Отд. движения	Активные движения
Рефлексы	Не определ.	Гримаса	Крик или актив. движ.
Цвет	Синий или белый	Выраженный акроцианоз	Полностью розовый

В настоящее время, шкала Апгар используется для диагностики асфиксии и оценки степени ее тяжести [3]. У большинства новорожденных состояние через одну минуту после рождения по шкале Апгар оценивается в 7-8 баллов. Через 5 минут состояние этих новорожденных соответствует 8-10 баллам. Оценка по шкале Апгар на первой минуте равная 4-6 баллам соответствует асфиксии средней степени тяжести, при условии, если к пятой минуте она стала равна 8-10 баллам. Диагноз асфиксии тяжелой степени является правомочным в случае, если оценка по шкале Апгар равна 0-3 баллам на первой минуте и не превышает 7 баллов спустя пять минут.

Рассмотрим специфику построения орграфа набросков на примере шкалы Апгар. Зададим банк тестов  $\{G(\tau)\}$  в следующем виде:

*Число сердечных сокращений*  $\wedge$  ЧСС  $\{1 \{Отсутствуют \wedge 0; \text{Меньше } 100 \text{ ударов в минуту} \wedge 1; \text{Больше } 100 \text{ ударов в минуту} \wedge 2\}\}$

*Дыхание*  $\{1 \{Отсутствует \wedge 0; \text{Слабый крик, гиповентиляция} \wedge 1; \text{Хорошее, сильный крик} \wedge 2\}\}$

*Мышечный тонус*  $\{1 \{\text{Вялый} \wedge 0; \text{Отдельные движения} \wedge 1; \text{Активные движения} \wedge 2\}\}$

*Рефлексы*  $\{1 \{\text{Не определяются} \wedge 0; \text{Гримаса} \wedge 1; \text{Крик или активные движения} \wedge 2\}\}$

*Цвет*  $\{1 \{\text{Синий или белый} \wedge 0; \text{Выраженный акроцианоз} \wedge 1; \text{Полностью розовый} \wedge 2\}\}$

*Сумма баллов*  $\wedge \Sigma b \{1 \{[0; 10]\}\}$

*Ст. тяжести асфиксии*  $\{1 (f/\mu) \{Отсутствует; \text{Легкая}; \text{Средняя}; \text{Тяжелая}\}\}$

*Время (мин)*  $\wedge t \{1 \{[1; 1000]\}\}$ .

Договоримся метки значений тестов в рамках базового домена трактовать как баллы (это возможно только в том случае, если баллы уникальны для каждого значения теста). Функция  $(f/\mu)$  определена на

множестве результатов тестов  $\{\underline{t}/1\}$  и является онтологическим соглашением. С ее помощью вычисляется значение теста «Ст. тяжести асфиксии/1». Тест «Сумма баллов» является распространенным вариантом свертки (суммирования) значений разнородных тестов.

Напомним, что произвольным событием называется пара  $\langle \underline{t}/T, \underline{t}/\Lambda \rangle$ . Если время совпадает у нескольких событий, то для упрощения фиксации событий можно использовать нотации:  $\{\{\underline{t}/T\}, \underline{t}/\Lambda\}$  или  $\langle \{\underline{t}/T\}, \underline{t}/\Lambda \rangle$ .

Логiku оценки по шкале Аппар изобразим следующим оргграфом набросков:

$$\begin{aligned} & \{\{\underline{t}/1\}, \underline{t}/1?1\} \cup \{\{\underline{t}/1\}, \underline{t}/1?5\} \Rightarrow \{\{\underline{b}_\tau\}, \underline{t}/1?1\} \cup \{\{\underline{b}_\tau\}, \underline{t}/1?5\} \Rightarrow \\ & \Rightarrow \{\Sigma \underline{b}_\tau, \underline{t}/1?1\} \cup \{\Sigma \underline{b}_\tau, \underline{t}/1?5\} \Rightarrow \text{Ст. тяжести асфиксии/1}, \end{aligned}$$

где  $\{\underline{b}_\tau\}$  – множество баллов, соответствующих тестам  $\{\underline{t}/1\}$  (баллы в отдельный тест не выделяются, так как являются неотъемлемой частью исходных тестов);  $\Sigma \underline{b}_\tau$  – значение теста «Сумма баллов». Приведем пример (в качестве значений тестов могут использоваться алиасы):

$$\begin{aligned} & \{\{\text{ЧСС}/1?1; \text{Дыхание}/1?1; \text{Мышечный тонус}/1?1; \text{Рефлексы}/1?1; \text{Цвет}/1?1\}, \\ & \underline{t}/1?1\} \cup \{\{\text{ЧСС}/1?1; \text{Дыхание}/1?2; \text{Мышечный тонус}/1?1; \text{Рефлексы}/1?2; \\ & \text{Цвет}/1?2\}, \underline{t}/1?5\} \Rightarrow \{\{\underline{b}_\tau\}, \underline{t}/1?1\} \cup \{\{\underline{b}_\tau\}, \underline{t}/1?5\} \Rightarrow \\ & \Rightarrow \{\Sigma \underline{b}_\tau/1? 5, \underline{t}/1?1\} \cup \{\Sigma \underline{b}_\tau/1? 8, \underline{t}/1?5\} \Rightarrow \text{Ст. тяжести асфиксии}/1? \text{Средняя}. \end{aligned}$$

Рассмотрим особенности построения оргграфа набросков на примере шкалы АРАСНЕ (шкала оценки острых физиологических расстройств и хронических нарушений состояния) [3]. Шкала АРАСНЕ II состоит из трех блоков: (1) Оценка острых физиологических изменений (acute physiology score -APS); (2) Оценка возраста; (3) Оценка хронических заболеваний.

Данные по блоку «Оценка острых физиологических изменений» собираются в течение первых 24 часов поступления пациента в ОРИТ (отделение реанимации и интенсивной терапии). Принцип бальной оценки у всех результатов тестов одинаков. Он приведен в таблице 6.4.

Таблица 6.4 – Оценка острых физиологических изменений (фрагмент)

Признак	Значение	Баллы
Среднее артериальное давление (мм рт.ст.)	$\geq 160$	+4
	130 – 159	+3
	110 – 129	+2
	70 – 109	0
	50 – 69	+2
	$\leq 49$	+4

Зададим конфигурактор теста «Среднее АД» следующим образом:



Среднее АД ^САД {

3 {Высокое ^d e f; Нормальное ^c; Низкое ^a b}

2 {≥ 160 ^f [160; 220]; 130 - 159 ^e [130; 159]; 110 - 129 ^d [110; 129]; 70 - 109 ^c [70; 109]; 50 - 69 ^b [50; 69]; ≤ 49 ^a [30; 49]}

1 {[30; 220]}.

Для каждого первичного теста  $\tau \in \{G(\tau)\}$  создадим сопряженный тест  $b_\tau$ , результатом которого являются соответствующие баллы. Для теста «САД» сопряженный тест будет иметь вид:

$b_{САД} \{1 \{4 (САД/2? a \vee f); 3 (САД/2? e) ; 2 (САД/2? b \vee d); 0 (САД/2? c)\}\}$ .

Дополнительные тесты пришлось создавать потому, что баллы не уникальны для каждого значения теста  $\tau$  и это не позволяет использовать алиасы, как в случае со шкалой Апгар. Значения тестов  $b_\tau$  рассчитываются автоматически на основе значений тестов  $\tau$ . Таким образом, банк тестов удваивается:  $\{G(\tau), G(b_\tau)\}$ . В рамках каждого блока шкалы АРАСНЕ баллы  $\{b_\tau/B\}$  всех тестов  $\{\tau/T\}$  суммируются. В результате получаются индикаторы, оценки  $\{p/P\}$ .

Оценка по АРАСНЕ II = (Баллы по шкале острых физиологических изменений) + (баллы за возраст) + (баллы за хронические заболевания).

Высокие оценки по шкале АРАСНЕ II связаны с высоким риском летальности в ОРИТ, что позволяет сформировать финальную интерпретацию данных:  $\{Риск/ УП/ Прогноз\}$ , где УП – уровень помощи.

Интегральные показатели  $\{p/P\}^*$  по шкале АРАСНЕ III расширены и значительно усовершенствованы по сравнению со шкалой АРАСНЕ II [3].

Приведенные примеры отражают общий конструкт орграфа набросков на основе оценочных и прогностических шкал, который можно изобразить следующим образом:

$$\{\underline{\tau}/T\} \Rightarrow \{b/B\} \Rightarrow \{p/P\} \Rightarrow \{p/P\}^* \Rightarrow \{Риск/ УП/ Прогноз\}. \quad (6.9)$$

Финитный набросок  $\{Риск/ УП/ Прогноз\}$  задает максимально сжатую интерпретацию исходного образа  $\{\underline{\tau}/T\}$ .

Орграф набросков (6.9) может использоваться в рамках метода предельных обобщений с целью построения (предельных) синдромных и вероятностных моделей знаний. Покажем это.

Пусть изучаются факторы, влияющие на наступление некоторого события, например, летального исхода, инвалидизации, неблагоприятного течения заболевания (после инсульта, инфаркта и т.д.) или неэффективности лечения. Положим  $Z = \{\text{событие}^+; \text{событие}^-\} = \{+; -\} = \{1, 2\}$ . Сформируем первичную базу прецедентов следующего вида:  $\Omega = \{\alpha(\{\underline{\tau}/T\}, \underline{z}/Z)\}$ , где  $Z = \{1, 2\}$ . Если речь идет о летальном исходе, то

$\Omega^+ = \{\alpha(\{\underline{t}/T\}, \underline{z}=1)\}$  – множество умерших пациентов, соответственно –  
 $\Omega^- = \{\alpha(\{\underline{t}/T\}, \underline{z}=2)\}$  – множество выживших пациентов.

Орграф набросков (6.9) позволяет перейти от первичного описания базы прецедентов к расширенному описанию следующего вида:

$$\Omega' = \{\alpha(\{\underline{t}/T\}, \{\underline{b}/B\}, \{\underline{p}/P\}, \{\underline{p}/P\}^*, \underline{z}/Z)\}, \text{ где } Z = \{1, 2\}.$$

Переход от описания  $\Omega$  к описанию  $\Omega'$  позволяет задействовать все виды обобщений (пример комбинации нескольких видов орграфов набросков). Далее, в рамках контекста  $K = \langle \Omega', \{G(\tau)\} \rangle \equiv \langle \Omega', \{G(\tau)\}, \{G(b)\}, \{G(p)\} \rangle$  строятся предельные синдромные модели знаний  $\{S^*\} = \{S^*\}^+ \cup \{S^*\}^-$  и предельные вероятностные модели  $\{R^*\} = \{R^*\}^+ \cup \{R^*\}^-$ .

Множество синдромов  $\{S^*\}^+$  определяет все возможные наборы параметров порядка (в рамках контекста  $K$ ) приводящие к наступлению «события+», а закономерности из  $\{R^*\}^+$  являются предвестниками события. Напротив, множество синдромов  $\{S^*\}^-$  определяет все возможные наборы параметров порядка, которые обеспечивают благоприятный режим функционирования. На основе знания множеств синдромов и предвестников может быть реализован *синдромный принцип управления*, целью которого является выход на благоприятный режим функционирования и его стабилизация (см. главу 7).

Главный вывод состоит в следующем: интеллектуальные приложения на основе синдромных и вероятностных моделей знаний способны в перспективе полностью заменить оценочные и прогностические шкалы как основы принятия решений. Вместе с тем, орграфы набросков с использованием шкал могут служить средством расчета дополнительных индикаторов, которые наравне с другими тестами участвуют в построении синдромных и вероятностных моделей.

## 6.7 Квантовая семантика: пути реализации

В последние десятилетия в научных публикациях и на научных конференциях активно обсуждаются вопросы о возможности алгоритмизации интеллекта (мышления), о влиянии квантовых эффектов и т.д. Так Пенроуз Р. решительно выступает против точки зрения, согласно которой нашу сознательную мыслительную деятельность во всех ее разнообразных проявлениях можно, в принципе, адекватно описать в рамках тех или иных вычислительных моделей [125]. Авторы учебника «Теория интеллекта» [26] считают, что машинный интеллект способен воспроизводить лишь детерминированные, дискретные и конечные информационные процессы. Данная дискуссия имеет прямое отношение к поиску и развитию фундаментальных основ интеллектуальных

информационных технологий. В этом плане с Роджером Пенроузом можно согласиться: необходимо продолжить поиск базовых принципов и инструментов, позволяющих мозгу (добавим – когнитивной динамической системе) осуществлять свою деятельность.

Традиционно разграничивают три вида реальности: классическая, квантовая и феноменологическая (субъективная реальность).

*Примечание:* квантовая физика (механика) имеет достаточно много интерпретаций, но феноменологическая физика и квантовая семантика в частности разработаны гораздо слабее, поэтому и возможных интерпретаций у них на данном этапе значительно больше.

Классическая физика описывает реальность как объективную, находящуюся «вне нас», существующую независимо от нас и эволюционирующую согласно тем или иным детерминистским законам. Простые объекты, сцепляясь друг с другом, образуют более сложные. Координатная система -  $x, y, z, t$ .

Кроме традиционного дуализма «волна–частица» (вернее сказать – дуализма «локальность–нелокальность»), который существует для всех тел, всех частиц вне зависимости от их размера), современная квантовая теория рассматривает связи между частью и целым, обмен энергией и информацией, взаимные переходы между непроявленной квантовой реальностью и наблюдаемым классическим миром. Главными в современной квантовой теории являются понятия «состояние» и «нелокальность» (особенность запутанных состояний, которым невозможно поставить в соответствие локальные элементы реальности); для всего, что может быть охарактеризовано этими понятиями, можно надеяться применить квантовые методы. Тем самым квантовая теория неизмеримо расширила сферу своего возможного применения [29, 104].

Основой для понимания квантовой картины мира является наличие двух типов состояний — смеси и суперпозиции. Состояния системы, когда реализуется только один из множества вариантов, в квантовой механике называют *смешанными*, или *смесью*. Смешанные состояния являются по сути классическими — система может быть с определенной вероятностью обнаружена в одном из состояний, но никак не в нескольких состояниях сразу. Однако известно, что в природе имеет место и совершенно другая ситуация, когда объект находится в нескольких состояниях одновременно. В таком случае говорят о *суперпозиции состояний*, или о чисто-квантовом состоянии. Суперпозиционные состояния могут существовать лишь при отсутствии взаимодействия рассматриваемой системы с окружением. Они описываются посредством так называемой *волновой функции*, которую также называют *вектором состояния*. Координатная система– гильбертово пространство (задается набором собственных базисных или основных

состояний). Процесс перехода суперпозиции в смесь (из нелокализованного в пространстве квантового состояния в наблюдаемое) называется *декогеренцией* [65, 103]. Обратный процесс называется *рекогеренцией*.

Одним из наиболее известных проявлений суперпозиции состояний является *интерференция*. Интерференция имеет место только для когерентных состояний. Декогеренцию (нарушение когерентности) в этом случае можно рассматривать как подавление интерференции. Важную роль в квантовой теории играют *нелокальные корреляции* (квантовые корреляции) — специфический эффект несепарабельности (квантовой запутанности), который заключается в согласованном поведении отдельных частей системы (независимо от расстояния между ними).

В поиске квантово-семантических методологических подходов и методов нужно вспомнить химическую теорию резонанса, описывающую объект исследования (молекулу) как результат суперпозиции (резонансного взаимодействия) множества соответствующих этой молекуле частных несамодостаточных предельных структур [55].

Можно согласиться со многими исследователями, которые считают, что развитие фундамента интеллектуальных технологий непременно приведет к освоению пластов реальности, на которых имеет место квантовая нелокальность, то есть обычные пространственно-временные ограничения теряют своё значение. В этой связи стоит отметить появление квантовых компьютеров, устройств, имеющих в качестве рабочего ресурса нелокальные состояния, то есть состояния, которые нельзя соотнести с привычным нам пространством-временем [28, 33].

Квантовый, а точнее – квантово-семантический подход к интерпретации когнитивных феноменов предполагает выделение двух противоположных процессов (по аналогии с физической декогеренцией и рекогеренцией): квантово-семантическую декогеренцию (QS-декогеренцию) и квантово-семантическую рекогеренцию (QS-рекогеренцию).

*QS-декогеренция* – это процесс перехода пластов реальности из семантически многозначного состояния (суперпозиции смыслов, «мерцающих» состояний) в однозначно проявленное состояние, значение (смесь состояний) в рамках феноменологического пространства наблюдателя.

В качестве координатной системы феноменологического пространства наблюдателя может выступать банк тестов  $\{G(\tau)\}$ , который обеспечивает преобразование информации от *сенсориума* (первичной информации сенсоров) к знаковой системе, в частности, языковой системе. Функционирование феноменологического пространства в значительной степени осуществляется на основе «врожденных» и приобретенных автоматизмов (разновидность системопаттернов), что делает большую

часть этого пространства «невидимой».

Для разных наблюдателей в одной и той же ситуации действительности может проявиться разный смысл. Это связано с тем, что каждый наблюдатель обладает собственным феноменологическим пространством, которое непрерывно развивается и модифицируется.

*Примечание:* наблюдатель может быть как естественный, включая человека и социум, так и искусственный, включая автономный и распределенный ИИ, а также гибридные системы. Термины «наблюдатель» и «когнитивная система» будем считать синонимами.

В ходе *QS*-декогеренции пласты реальности, ситуации действительности, подсистемы, оргграфы приобретают однозначно проявленные свойства и формы. Математическая формализация с ее стремлением к однозначной интерпретации является примером *QS*-декогеренции. *QS*-декогеренция осуществляется только в процессе взаимодействия наблюдателя с внешним и внутренним миром. Речь является примером такого взаимодействия. Весь материал глав 1 – 5 посвящен изучению *QS*-декогерированных структур.

Примером неразделяемой суперпозиции слов (*интерференции смыслов*) может служить парадокс лжеца – некто говорит «Я лгу» (два неразделяемых смысловых состояния – «истина» и «ложь»). В первой главе были приведен пример визуального образа с неразделяемым «мерцающим» (многозначным) смыслом.

На разрушение суперпозиции смыслов, т.е. на *QS*-декогеренцию, большое влияние оказывает контекст взаимодействия, в частности, перцептивная установка. Физическая декогеренция также может оказывать существенное влияние на *QS*-декогеренцию. Одним из специфических эффектов *QS*-декогеренции можно считать *мерцание (смысловых) состояний* – скачкообразные неконтролируемые переходы между разными состояниями (примерами служат бистабильные образы). Прекращению мерцаний может способствовать изменение контекста. Важнейшим специфическим эффектом *QS*-декогеренции является возникновение *качественной прерывности* в полотне реальности. Частным случаем такого эффекта являются *аттракционы*, возникающие, например, в художественной реальности [60].

Обратный процесс обретения определенными структурами квантово-семантических свойств назовем *QS-рекогеренцией*. В ходе него информационные оболочки "расплываются", а границы между объектами начинают исчезать, происходит объединение образов ситуаций действительности, функциональных подсистем, оргграфов набросков, других когнитивных структур в единую нелокальную квантово-семантическую систему (виртуальную сплошную среду).

Построение орграфа набросков любого феномена является примером порождения многозначности смыслов (каждый набросок представляет отдельный смысл). Кроме того, закритические наброски порождают метафоры и ассоциации, что генерирует дополнительные смыслы. Процесс построения орграфов набросков реализует по-сути *квантово-семантическое запутывание* (по аналогии с квантовым запутыванием), что делает невозможным разделение образной системы на отдельные самостоятельные и полностью независимые составные части.

Одним из специфических проявлений QS-рекогеренции является *гештальт-восприятие*, когда восприятие целого доминирует над восприятием частей. Очевидные для восприятия объекты макроуровня конституируются посредством некоторого рода граничных (категориальных) шаблонов (набросков, синдромных моделей знаний и т.д.). С этим эффектом тесно связан эффект *замещения*.

Важным специфическим эффектом QS-рекогеренции является *восприятие стабильного окружения*, которое не может быть связано с существованием сколько-нибудь детального, удерживаемого в памяти образа. Данный эффект видимо можно объяснить существованием орграфов набросков, что позволяет оперировать критическими или финитными набросками, которые не содержат деталей характеристик и пространственно-временной локализации феномена действительности.

Специфическим эффектом QS-рекогеренции является *дефокусировка*, т.е. увеличение размаха *концептуального внимания*. Считается, что расширение фокуса внимания способствует творчеству, развитию креативности. Это одна из важнейших закономерностей дивергентного мышления. Специфическим эффектом QS-рекогеренции является *предсказание* когнитивной системой определенных изменений действительности, в частности, имеет место *эффект предупреждения сигнала*.

В результате QS-рекогеренции качественная картина когнитивной репрезентации описывается суперпозицией предельных структур (чистых состояний) и не может быть правильно изображена ни одним из отдельно взятых набросков, а только их набором.

Таким образом, «состояние» в рамках квантовой физики (и гильбертова пространства) заменяется на «смысловое состояние (значение)» в рамках квантовой семантики (и феноменологического пространства). Квантовая суперпозиция состояний переходит в суперпозицию смыслов в квантовой семантике. Квантовая смесь состояний переходит в *смесь смыслов* (значений) в квантовой семантике (смесь разделяема на альтернативные смыслы, а суперпозиция нет). Можно заключить, что анализ и синтез когнитивного процесса невозможен без сочетания процессов QS-декогеренции и QS-рекогеренции.

Большую роль в квантовой теории играет принцип дополнителности. Выше уже упоминался дуализм «локальное – нелокальное». Дуализм выступает как неотъемлемое свойство нашего мира. Рекогерируя (развивая, формируя) противоположности, система или структура приобретает цельность. Выражением такой цельности в восточной философии является двойственное отношение инь – ян. В структуре орграфов доменов и набросков проявляется, в частности, дуализм дискретного и непрерывного: непрерывный базовый домен (образ, модель) обязательно дополняется множеством дискретных доменов (набросков, моделей).

В квантовой семантике знаки, символы, слова рассматриваются в различных когнитивных (познавательных) конструкциях. Так слова можно представить в виде квантующихся явлений, состоящих из наблюдаемых «частиц» и невидимых «волн» [164]. Тогда за каждым словом предполагаются теневые объекты, «теневые» слова, латентные значения слов. Следовательно, смыслы слов (фраз) нужно интерпретировать с учетом «тени». Другими словами, реальный текст характеризуется словесными пропусками того, что «само собой» известно реципиенту, того, что образует «имплицитные смысловые аксиомы» – «знания о мире» (источник сообщения должен обладать рефлексией относительно «знаний о мире» реципиента). В этом также проявляется принцип дополнителности – в смысл каждого слова (знака) может вкладываться все то, что дополняет общепринятое значение слова или высказывания.

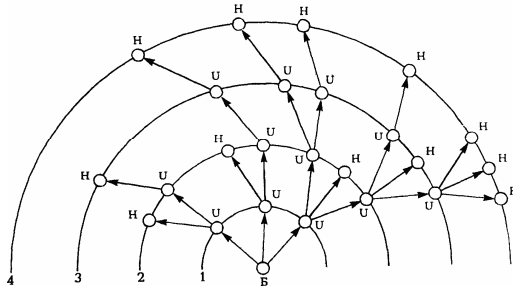
Поскольку любое слово в тексте или код в сообщении являются значением какого-либо теста, то общее количество равноуровневых представлений текста  $\langle a, b, c, \dots \rangle$  определяется произведением числа вершин в соответствующих орграфах доменов. Более того, для каждого масштабного представления  $W_\gamma = \langle a/A, b/B, c/C, \dots \rangle_\gamma$  автоматами среды строятся орграфы набросков  $Gs(W_\gamma)$ . На множестве всех текстов или сообщений (базы прецедентов  $\Omega(W)$ ) в рамках каждого орграфа набросков  $Gs(W_\gamma)$  формируются докритические, критические и закритические наброски. Критические наброски требуют минимум памяти для отображения специфики (уникальности) текста. Закритические наброски служат для реализации ассоциативных связей и построения метафор. Далее, в рамках контекста  $\langle \{G(\tau)\}, \Omega(W) \rangle$ , отражающего опыт и знания наблюдателя, для любого текста (фразы)  $W$  автоматами среды строится синдромная модель знаний  $\{S\}_W$ , а также множество вероятностных закономерностей  $\{R\}_W$ . Таков предварительный результат скрытой квантово-волновой динамики для текста  $\langle a, b, c, \dots \rangle$ .

Квантовая семантика подтверждает вывод многих исследователей о том, что мышление – это существенно нелокальный процесс (не путать с распараллеливанием). При взаимодействии семантических «волн»

происходит генерирование колоссального числа разноуровневых описаний (разных уровней реальности), заполняется семиотическое пространство, пространство смыслов и, соответственно, переходных слоев. В переходных слоях происходит перестройка картины мира (текста, сообщения, ситуации) с одного масштаба на другой. Здесь встречаются разные законы, действуют смешанные языки, рождаются новые смыслы. Отсюда актуальность методов связывания, сращивания, соединения образов, набросков решений, моделей в переходных слоях.

Развиваемая нами парадигма когнитивного моделирования полностью отвечает духу квантовой семантики. Покажем это на примерах.

Локализацией теста можно считать задание конкретного значения теста  $\underline{a}/T$  в рамках орграфа  $G(\tau)$ . Смысловым дополнением являются все сопряженные значения в рамках других доменов орграфа  $G(\tau)$ . В результате делокализации отдельного значения теста – распространения семантической волны – происходит переход от  $\underline{a}/T$  к множеству значений в рамках орграфа  $G(\tau)$ . На рис. 6.10 показан пример такой волны с компактным носителем, распространяющейся от базового домена.



Б – базовый домен, H – терминальные вершины, U – промежуточные вершины.

Рис. 6.10 – Пример семантической волны, распространяющейся от базового домена орграфа  $G(\tau)$

Обозначив, как и ранее, замыкание через  $\{\underline{a}/T\}^+$ , а множество всех значений, образованных в результате нисходящей-восходящей детерминации через  $\{\underline{a}/T\}^\infty$ , распространение семантической волны в рамках банка тестов можно схематически изобразить следующим образом:

$$\underline{a}/T \rightarrow \{\underline{a}/T\}^+, \{\underline{a}/T\}^\infty. \quad (6.10)$$

Существует принципиальная разница в значениях  $\{\underline{a}/T\}^+$  и  $\{\underline{a}/T\}^\infty$ . Каждое значение теста из  $\{\underline{a}/T\}^+$  определяется единственным образом в рамках



каждого домена, в то время как значения из  $\{\underline{t}/T\}^\infty$  определяются не единственным образом (рис. 6.11).

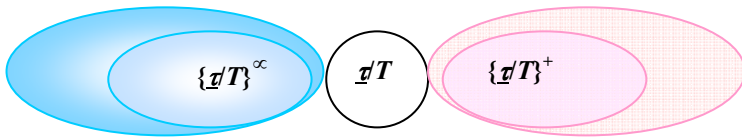


Рис. 6.11 – Распространение семантической волны в рамках орграфа  $G(\tau)$

Приведем пример. Пусть задано  $\underline{t}/T = \text{Возраст}/B4 = [31; 40]$ . На основе описания орграфа  $G(\text{Возраст})$  приведенного в главе 2 получим:

$\{\underline{t}/T\}^+ = \{\text{Возраст}/B4 = [31; 40]; \text{Возраст}/B5 = \text{средних лет}; \text{Возраст}/B6 = \text{трудоспособный}\}$

$\{\underline{t}/T\}^\infty = \{\text{Возраст}/B1 = [31; 40]; \text{Возраст}/B2 = \text{молодой; средних лет}; \text{Возраст}/B3 = \text{молодой; средних лет}\}$ .

Как видим, значения тестов «Возраст/B2» и «Возраст/B3» получаются неоднозначными. На рис. 6.12 показаны две ветки семантической волны (пунктирные стрелки): одна (красная ветка) образована замыканием, другая (синяя ветка) – нисходящей-восходящей детерминацией. Исходное значение  $\underline{t}/T$  теста «Возраст» задано в рамках домена «B4».

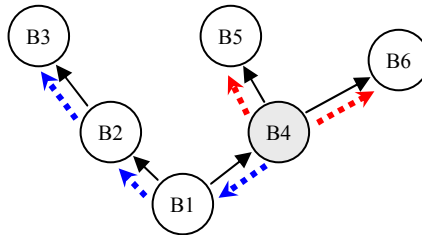


Рис. 6.12 – Семантическая волна на основе орграфа  $G(\text{Возраст})$

Таким образом, можно считать, что любой квант информации  $\underline{t}/T$  порождает соответствующую семантическую волновую функцию в рамках орграфа доменов  $G(\tau)$ . Данная волновая функция вступает во взаимодействие с волновыми функциями других тестов сообщения  $\langle \{\underline{t}/T\} \rangle$ .

Орграф доменов является одним из вариантов орграфа набросков, поэтому все приведенные выше рассуждения относятся и к любому орграфу набросков. Так локализацией образа (модели) можно считать

фиксацию отдельного наброска образа (чистого состояния). В результате делокализации происходит волнообразный переход ко всем связанным наброскам. Поскольку закритические наброски (в частности, финитные наброски) принадлежат области неоднозначной интерпретации (орграфы набросков разных образов пересекаются), то их возбуждение приводит к ассоциативному возбуждению других образов (других орграфов набросков). Таким образом, любой орграф набросков порождает соответствующую волновую функцию на компактном носителе.

Распространение семантической волны для произвольного наброска  $P$  в рамках множества орграфов набросков  $\{Gs(W)\}$  можно схематически изобразить следующим образом:

$$\forall P \in Gs(W), \quad P \rightarrow \{P\}^+, \{P\}^\infty \rightarrow \{Gs\}_W \rightarrow \{Gs\}_{W'}, \quad (6.11)$$

где  $\{Gs\}_W$  – все орграфы набросков образа  $W$ , с которыми совпадают те или иные наброски из  $\{P\}^+$  или  $\{P\}^\infty$ ;  $\{Gs\}_{W'}$  – все орграфы набросков образов  $W'$ , с которыми пересекаются те или иные закритические наброски орграфа набросков образа  $W$ .

Распространение семантической волны показывает, как на основании очень приблизительного наброска (например, любого критического наброска) происходит восстановление целостного образа, т.е. волновой функции образа.

Взаимодействие таких «размазанных» объектов как орграфы набросков сильно усложняет когнитивную картину, порождая так называемые запутанные состояния. Уточним суть этого понятия.

Все объекты квантово-семантической реальности будем именовать QS-объектами. По аналогии с квантовой корреляцией между QS-объектами (например, орграфами доменов или набросков, системопаттернами) могут возникать *квантово-семантические корреляции* (QS-корреляции), а состояния участвующих QS-объектов можно назвать *запутанными* (сцепленными). Применим данные понятия к анализу функционирования системопаттернов.

Как мы уже знаем, любую математическую модель, любую трансформацию реальности и любую когнитивно-поведенческую операцию можно представить с помощью множества системопаттернов

$$k = \{f/\mu: \{a/A\} \rightarrow \{b/B\}\} \cup P_k,$$

где  $f/\mu$  – системопаттерны с известным механизмом реализации  $\mu$ ,  $P_k$  – правила композиции системопаттернов. Поскольку каждый информационный квант  $\underline{a}/T \in \{a/A\}$  порождает волновую функцию вида (6.10), то взаимодействие всех волновых функций для квантов  $\{a/A\}$  системопаттерна  $f/\mu$  порождает волну взаимодействий на разных уровнях

общности. Рассмотрим вопрос о том, что является результатом такого взаимодействия и какую роль играют QS-корреляции? Другими словами, что является результатом действия (квантово-семантического) оператора локализации решения?

Прежде чем искать ответ в рамках квантовой семантики рассмотрим квантовые эффекты от столкновения частиц. Дело в том, что в большинстве реакций при одних и тех же начальных частицах с соответствующими амплитудами вероятности рождаются сразу все типы конечных частиц, которые в принципе возможны. И все они разлетаются в виде квантово-запутанных сферических волн. И пока эти волны не долетели до стенок вакуумной трубы, нельзя сказать, кто именно родился. Лишь взаимодействие с детектором, т.е. сам процесс измерения, вытаскивает какую-то одну возможность из всего бесконечного списка (т.е. конечное состояние локализуется в этом списке). И вот тогда говорят, что в этом конкретном эксперименте родился именно этот набор частиц.

В рамках квантовой семантики в качестве примера рассмотрим вычисление пути при равномерном прямолинейном движении:  $s = v * t$ . Казалось бы, нет ничего проще, подставляем в формулу значения скорости  $v$  и времени  $t$  и получаем путь  $s$ . Однако в рамках рассматриваемой парадигмы взаимодействия волновых функций на базе орграфов доменов  $G(v)$ ,  $G(t)$  и  $G(s)$  автоматы вычислительной среды запускают все сопряженные системопаттерны вида:  $\{f/\mu: v/V, t/T \rightarrow s/S\}$ , т.е. имеет место *квантово-семантический параллелизм*. На рис. 6.13 схематично показано подобное взаимодействие.

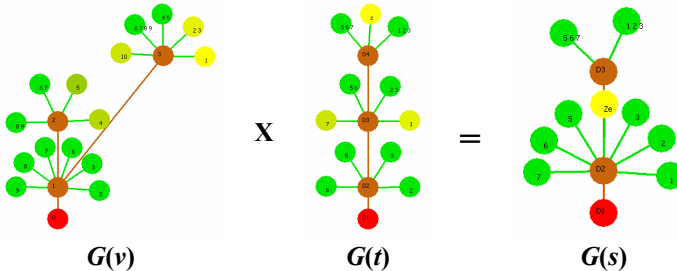


Рис. 6.13 – Пример квантово-семантического системопаттерна

Разные механизмы вычислений  $\{f/\mu\}$  требуют разных ресурсов и, в частности, времени. Если устремить длину иерархии конфигураторов скорости и времени (числоформ) к бесконечности (подобная операция описана в главе 2), то на выходе получим стремящийся к бесконечности ряд ответов, причем ответы поступят скорее всего не одновременно, а в

течении определенного промежутка времени. В результате действия оператора локализации решения в какой-то момент времени произойдет останов. Фрагмент полученного ряда и будет принят в качестве ответа (осуществилось *квантово-семантическое измерение* – процедура определения состояния QS-объекта). Таким образом, до момента квантово-семантического измерения (QS-измерения) результат находится в состоянии суперпозиции, т.е. в любом промежуточном динамическом вероятностном состоянии между всеми значениями  $\{\underline{g}/S\}$ . Собственно QS-измерение и является оператором локализации решения.

Пусть  $\underline{v} = 61$ , а  $\underline{t} = 6$ . Пример возможного ответа – QS-измерения (в терминах конфигурирования пути): <’очень большой’ (ответ получен на 1-ой сек), ‘больше 300’ (на 3 сек), ‘[330, 370]’ (на 8 сек), ‘примерно 360’ (на 12 сек),...>.

Как и в случае со столкновением частиц исходная ситуация (значения  $\underline{v}$  и  $\underline{t}$ ) может быть одна и та же, а ответ – локализованный ряд значений пути – каждый раз может быть разным (зависит от контекста и ресурсов в момент решения). Соответственно, «материализовавшиеся» вычислительные траектории также разные. При этом сорт частиц – домены тестов (уровень общности) – в ответе также заранее нельзя установить. Потенциально возможны любые домены – сорта.

Видимо, можно говорить о наличии *сцепленных состояний* QS-объектов  $G(v)$ ,  $G(t)$  и  $G(s)$ . В этом случае, изменяя состояние любых двух, можно управлять состоянием третьего. Сцепленные состояния орграфов доменов можно моделировать, в частности, с помощью онтологических соглашений, реализующих автоматизмы (см. главу 2).

*Примечание.* На самом деле необходимость в QS-измерении возникает только в моменты коммуникации. Примером такой коммуникации является речь. Иначе говоря, любое взаимодействие является каналом декогеренции, или, что по сути одно и то же, каналом обмена информацией. Именно обмен информацией рассматривается в экзистенциальной интерпретации квантовой механики как причина изменения любых состояний [65]. Другими словами, в этой системе сохраняются альтернативные варианты развития событий, и только для самого наблюдателя реализуется один из них. В остальных интеллектуальных актах необходимости в измерении, по всей видимости, нет. Достаточно допустить наличие сдвигов состояний квантово-семантической неопределённости QS-объектов к какому-либо определённом состоянию.

Рассмотрим одно из возможных квантово-семантических состояний орграфа доменов теста. Пусть для теста «Тест» имеется исходный домен  $D1 = \{\text{Черный } ^1; \text{Белый } ^0\}$ . Предположим, что в процессе созревания орграфа доменов между любыми двумя соседними элементами домена

может возникать новый элемент со «средним качеством». В нашем случае таким элементом будет «Серый» или « $\frac{1}{2}$ ». Новые элементы приводят к появлению новых доменов. Введем в рассмотрение домен  $D2 = \{\text{Черный}^{\wedge 1}; \text{Серый}^{\wedge \frac{1}{2}}; \text{Белый}^{\wedge 0}\}$ . Проблема теперь заключается в том, как организовать обобщающую связь ( $D2 \rightarrow D1$ ). Если следовать формализму главы 2, то возможны два варианта орграфа:

$Tecm^{\wedge T1} \{D1 \{\text{Черный}^{\wedge 1}; \text{Белый}^{\wedge 0}\} D2 \{\text{Черный}^{\wedge 1}; \text{Серый}^{\wedge \frac{1}{2}}; \text{Белый}^{\wedge 0}\}\}; G_1(Tecm) = \{D2 \rightarrow D1\}$ ;

$Tecm^{\wedge T2} \{D1' \{\text{Черный}^{\wedge 1}; \text{Белый}^{\wedge 0}\} D2 \{\text{Черный}^{\wedge 1}; \text{Серый}^{\wedge \frac{1}{2}}; \text{Белый}^{\wedge 0}\}\}; G_2(Tecm) = \{D2 \rightarrow D1'\}$ .

В рамках первого варианта имеем:

$D2.\{\text{Черный}; \text{Серый}\} \rightarrow D1.\{\text{Черный}\}$ ;  $D2.\{\text{Белый}\} \rightarrow D1.\{\text{Белый}\}$ .

В рамках второго варианта имеем:

$D2.\{\text{Белый}; \text{Серый}\} \rightarrow D1'.\{\text{Белый}\}$ ;  $D2.\{\text{Черный}\} \rightarrow D1'.\{\text{Черный}\}$ .

Таким образом, в обоих вариантах орграфа составы доменов одинаковы, отличаются только обобщающие связи. Орграфы  $G_1(Tecm)$  и  $G_2(Tecm)$  будем называть *орграфами-близнецами* (совпадают элементы доменов, но различаются связи). Введем в рассмотрение новый тип орграфа доменов:

$Tecm^{\wedge T3} \{D1'' \{\text{Черный}^{\wedge 1}; \text{Белый}^{\wedge 0}\} D2 \{\text{Черный}^{\wedge 1}; \text{Серый}^{\wedge \frac{1}{2}}; \text{Белый}^{\wedge 0}\}\}; G_3(Tecm) = \{D2 \rightarrow D1''\}$ .

В орграфе  $G_3(Tecm)$  имеет место суперпозиция обобщающих связей для элемента « $D2.$ Серый», что выражается в наличии соответствующей ссылки у каждого элемента домена  $D1''$ . Домен  $D1''$  и в целом орграф  $G_3(Tecm)$  будем относить к *QS-объектам*, находящимся в состоянии суперпозиции. Домен  $D2$  имеет обычный (локализованный) тип. На рис. 6.14 показаны все три орграфа (стрелки обозначают схему обобщения).

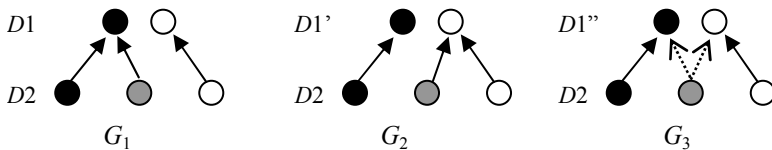


Рис. 6.14 – Характер связей обобщения нового элемента

Акцентируем внимание на том, что один *QS-орграф*  $G_3(Tecm)$  заменяет смесь орграфов  $G_1(Tecm)$  и  $G_2(Tecm)$ . В процессе *QS-декогеренции* (*QS-измерения*) орграф  $G_3(Tecm)$  может перейти в любой из орграфов  $G_1(Tecm)$ ,

$G_2(Tect)$ , причем в разных ситуациях (контекстах) это могут быть разные орграфы.

По аналогии с фотоном света можно ввести характеристику «поляризация» связи элемента домена. Таким образом, у связей существуют поляризационные степени свободы. В случае, когда поляризация связи неопределенна, то есть когда оба варианта связи находятся в состоянии суперпозиции, волновой вектор связи элемента «Серый  $\wedge^{1/2}$ » имеет вид:

$$|\psi\rangle = 1/\sqrt{2}(|0\rangle + |1\rangle), \quad (6.12)$$

где  $|0\rangle$  и  $|1\rangle$  — компоненты, обозначающие два вида связей соответственно (элементы домена  $D1$ ), а  $1/\sqrt{2}$  — нормировочный множитель, обеспечивающий, чтобы суммарная вероятность реализации связи равнялась единице. Волновые вектора крайних элементов имеют вид:

$$|\psi_0\rangle = |0\rangle, \quad |\psi_1\rangle = |1\rangle.$$

Мера запутанности, равная 0, соответствует наличию в системе только сепарабельных состояний с одной компонентой в векторе состояния.

До QS-измерения поляризации связи, состояния поляризации находятся в суперпозиции, т.е. их просто не существует как локальных характеристик связи. В ходе QS-измерения анализатор выделяет из суперпозиции, определяемой выражением (6.12), либо компоненту  $|0\rangle$ , либо компоненту  $|1\rangle$ . Нелокальный объект становится локальным только в момент QS-измерения.

Рассмотрим возможный сценарий деления элементов домена  $D2$ . Пусть при делении пары элементов «Черный» и «Серый» образуется новый элемент «Темно серый  $\wedge^{3/4}$ » и новый домен  $D3$  {Черный  $\wedge^1$ ; Темно серый  $\wedge^{3/4}$ ; Серый  $\wedge^{1/2}$ ; Белый  $\wedge^0$ }, а при делении пары элементов «Серый» и «Белый» образуется новый элемент «Светло серый  $\wedge^{1/4}$ » и новый домен  $D4$  {Черный  $\wedge^1$ ; Серый  $\wedge^{1/2}$ ; Светло серый  $\wedge^{1/4}$ ; Белый  $\wedge^0$ }. На рис. 6.15 показан результирующий QS-орграф  $G(Tect)$ .

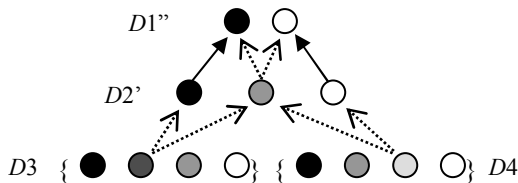


Рис. 6.15 – Результирующий QS-орграф  
Конфигуратор QS-орграфа  $G(Tect)$  имеет следующий вид:

$Tect \wedge T \{$

$D1'' \{ \text{Черный} \wedge 1/2; \text{Белый} \wedge 0/2 \}$

$D2' \# D3 D4 \{ \text{Черный} \wedge 1/4; \text{Серый} \wedge 1/2 \wedge 3/4 \wedge 1/4; \text{Белый} \wedge 0/4 \}$

$D3 \# \{ \text{Черный} \wedge 1; \text{Темно серый} \wedge 3/4; \text{Серый} \wedge 1/2; \text{Белый} \wedge 0 \}$

$D4 \{ \text{Черный} \wedge 1; \text{Серый} \wedge 1/2; \text{Светло серый} \wedge 1/4; \text{Белый} \wedge 0 \}$

$G(Tect) = \{ D3 \rightarrow D2'; D4 \rightarrow D2'; D2' \rightarrow D1'' \}$ .

Легко видеть, что домены  $D1''$  и  $D2'$  относятся к типу  $QS$ -доменов (имеется неоднозначность связей), а домены  $D3$  и  $D4$  имеют обычный тип. При  $QS$ -декогеренции оргграф  $G(Tect)$  переходит в смесь из 8 обычных оргграфов-близнецов ( $8 = 2^3$ ). Знак '#' без алиаса после домена  $D3$  поставлен для того, чтобы отменить вертикальное обобщение по умолчанию от домена  $D4$  (такой знак следует ставить после каждого нового базового домена).

Важной особенностью оргграфа  $G(Tect)$  является наличие двух базовых доменов  $D3$  и  $D4$ . В предыдущих главах мы исключали подобную возможность. Тесты, оргграфы которых содержат несколько базовых доменов, назовем *ультратестами*. Ультратесты не обязательно относятся к типу  $QS$ -объектов. Банк тестов, который содержит ультратесты и/или  $QS$ -оргграфы, назовем *расширенным банком тестов*.

Автоматизмами среды для оргграфа  $G(Tect)$  могут быть построены структурно-завершенные оргграфы  $G^+(Tect)$  и  $G^{++}(Tect)$ , которые также относятся к  $QS$ -объектам.

*Примечание.* Ясно, что при делении элементов домена  $D2$  могли образоваться не два новых домена, а один:  $D5 \{ \text{Черный} \wedge 1; \text{Темно серый} \wedge 3/4; \text{Серый} \wedge 1/2; \text{Светло серый} \wedge 1/4; \text{Белый} \wedge 0 \}$ . Соответствующий оргграф имел бы вид:  $G(Tect) = \{ D5 \rightarrow D2' \rightarrow D1'' \}$ .

Аналогично созреванию оргграфа доменов качественной оппозиции «Черное – белое» происходит созревание тестов других оппозиций, например, «Холодное – горячее», «Истина – ложь» и т.д.

Аналогичная  $QS$ -измерению картина имеет место и с квантовыми битами или кубитами в рамках квантовых вычислений. *Кубит* (Quantum-bit или  $Q$ -бит) квантовый физический объект или ячейка квантовой единицы информации (емкость квантовой ячейки – два бита классической информации), которая может до момента квантового измерения находиться в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между «0» и «1» [33].

*Примечание:* суперпозицию поляризации связей в  $QS$ -оргграфе можно моделировать с помощью кубитов. В этом аспекте применимы методы квантовой информатики и квантовой логики.

По аналогии с кубитом можно рассматривать  $QS$ -элемент домена теста – ячейку квантово-семантической единицы информации, которая до

момента нисходящей детерминации (конкретизации, дезагрегации, дефаззификации), т.е. *QS*-измерения может находиться в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между всеми соответствующими значениями домена-родителя, определяемыми связями обобщения. Каждый элемент не базового домена является, по-сути, острием конуса обобщения, содержащего все элементы всех доменов, которые обобщаются в данный элемент. Весь конус обобщения и представляет собой суперпозицию значений. При реализации нисходящей детерминации (фрагмента *QS*-декогеренции) необходимо предусмотреть случайный выбор и колебательный процесс.

По аналогии с кубитом можно рассматривать *QS*-домен теста  $\tau$ , который до момента *QS*-измерения может находиться в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между всеми значениями домена и/или значениями связей обобщения. На таком же квантово-семантическом уровне находится *QS*-набросок образа (состоит из информационных квантов – элементов и их связей, включая связи обобщения).

Следующий квантово-семантической уровень – *QS*-орграф доменов теста  $\tau$ . до момента *QS*-измерения он находится в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между всеми доменами орграфа  $G(\tau)$  и с учетом суперпозиции поляризации обобщающих связей. На таком же квантово-семантическом уровне находится – *QS*-орграф набросков  $W$  (образа, ситуации): до момента *QS*-измерения он находится в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между всеми набросками орграфа  $G_s(W)$  и с учетом суперпозиции поляризации обобщающих связей.

Следующий квантово-семантической уровень – *QS*-системопаттерн  $f$ : до момента *QS*-измерения он находится в состоянии суперпозиции, т. е. в любом промежуточном динамическом вероятностном состоянии между всеми сопряженными системопаттернами вида  $\{f/\mu: \{a/A\} \rightarrow \{b/B\}\}$ . Разновидностями *QS*-системопаттерна являются *QS*-синдром и *QS*-вероятностная закономерность. Они находятся в состоянии суперпозиции со своими сопряженными предельными синдромами (предельными вероятностными закономерностями). Одной из разновидностей *QS*-системопаттерна является *QS*-радикал (см. главу 7). *QS*-радикал можно интерпретировать как (био)физический агент (в рамках телесной онтологии).

Следующий квантово-семантической уровень – *QS*-синдромные модели знаний и *QS*-вероятностные модели знаний: до момента *QS*-измерения они находятся в состоянии суперпозиции, т. е. в любом промежуточном



динамическом вероятностном состоянии между всеми синдромами (вероятностными закономерностями) модели знаний, а также синдромами сопряженных моделей.

Следующий квантово-семантической уровень – *QS-функциональная система (QS-ФС)* в рамках телесной онтологии. Простейшей ФС является среда радикалов (среда биофизических агентов). Каждая синдромная, вероятностная или образная модель знаний порождает физическую или виртуальную среду радикалов (глава 7). Более того, каждая модель знаний порождает множество дуальных моделей знаний, т.е. множество ФС, решающих одну и ту же задачу. До момента *QS*-измерения *QS-ФС* находится в состоянии суперпозиции, т.е. в любом промежуточном динамическом вероятностном состоянии между всеми ФС, которые могут решить данную задачу (например, разными средами радикалов, решающими одну и ту же задачу).

Наивысшим квантово-семантическим уровнем можно считать – *QS-феноменологическое пространство (QS-ФП)*.

Таким образом, можно говорить о стратификации *QS*-уровней феноменологического пространства включающего, в частности: *QS*-связи; *QS*-элементы доменов тестов; *QS*-домены тестов, *QS*-наброски образов, *QS*-орграфы доменов тестов, *QS*-орграфы набросков, *QS*-системопаттерны, *QS*-модели знаний; *QS*-радикалы, *QS*-ФС и *QS*-ФП.

Важно отметить, что в рамках каждого *QS*-объекта существуют экстремальные структуры, которые являются результатом когнитивной самоорганизации (например, предельные синдромы, экстремальные пограничные слои набросков, критические пути и т.д.). Данные экстремальные структуры определяют преимущественные вероятностные состояния.

У каждого *QS*-объекта, как и у каждой сложной системы, есть локальные (проявленные) свойства и нелокальные (суперпозиция состояний). Та часть *QS*-объекта, которая наиболее сильно взаимодействует с окружением, переходит в проявленное, локальное состояние, а та часть, которая взаимодействует с окружением слабо, остается в нелокальном, запутанном состоянии. Если какие-либо тесты не участвуют в построении моделей знаний, то их орграфы доменов находятся в состоянии суперпозиции или в чистом *QS*-состоянии.

Смешанно-запутанные состояния возникают при взаимодействии объектов феноменологического пространства друг с другом, что приводит к частичной потере когерентности. Эти состояния можно охарактеризовать соотношением классических и *QS*-корреляций, или, иначе говоря, выраженностью и классических, и квантово-семантических свойств.

Приведенная выше схема раскрывает возможный механизм функционирования феноменологического пространства. Более того, можно говорить о построении достаточно полной *синтетической телесно-феноменологической онтологии*.

Таким образом, имеем разные масштабные уровни квантовых и квантово-семантических явлений (разные пространства событий, разные уровни реальности). Видимо, можно считать, что квантовый и квантово-семантический принцип взаимодействия непрерывно транслируется во всех своих звеньях – от молекулярного уровня до когнитивной системы в целом.

Приведенные выше рассуждения показывают принципиальную ограниченность методик выявления знаний у экспертов. В процессе *QS*-измерения (попытки вербализации опыта) вербализуется какой-то (из бесчисленного множества) фрагмент опыта (знаний). В других экспериментах (при других контекстах) реализуются другие «крупинки» опыта.

Таким образом, на каждом квантово-семантическом уровне действуют законы подобные основным законам квантовой физики [65, 103]:

- *суперпозиция* двух или более дискретных чистых квантовых состояний;

- *квантовый параллелизм*, когда квантовая операция (например, логическая) выполняется одновременно над всеми квантовыми состояниями;

- такое явление, как «запутывание» (*квантовая нелокальность*), т. е. упорядоченное взаимодействие на расстоянии удаленных квантовых объектов;

- *квантовое измерение* – процедура определения квантового состояния квантового объекта.

Подведем итог.

Следуя концепции *QS*-декогеренции можно предположить, что в основе феноменологической реальности лежит нелокализованная и «невидимая» квантово-семантическая реальность, которая становится локализованной и «видимой» в ходе происходящего при взаимодействии обмена информацией и сопутствующей этому процессу фиксации (смысловых) состояний. Можно утверждать, что *QS*-декогеренция и есть тот универсальный механизм, который переводит суперпозиционное квантово-семантическое состояние в смешанное, проявленное, наблюдаемое, классическое.

*QS*-запутывание орграфов набросков позволяет лучше понять фундаментальное значение одного из самых замечательных природных когнитивных феноменов – способности выполнять широкий круг физических и психических задач без каких-либо измерений или

вычислений. На эту способность человека обращает внимание и Л. Заде (*теория впечатления* Л.Заде дает способы для получения быстрой и полной информации об окружающем мире без проведения точных измерений).

Взаимопроникающих феноменологических пространств (миров) в любой момент времени счетное множество. Они могут взаимодействовать между собой, так как имеется общее ядро (следствие эволюционного развития и обучения). Расширенный банк тестов  $\{G(\tau)\}$  может служить координатной системой любого феноменологического (ментального) пространства – времени. Каждое пространство имеет свою координатную систему.

Вероятностное представление классического мира возникает в квантовой семантике не потому, что мы чего-то не знаем о системе, а потому, что до измерения у нее нет каких-либо определенных характеристик. В целостной системе продолжает существовать суперпозиция возможных (смысловых) состояний, и только одно из них, для некоторой конкретной подсистемы, реализуется в ходе взаимодействия наблюдателя с реальностью (в ходе QS-измерения).

Таким образом, характеристики объекта реальности «создаются» наблюдателем; вне акта наблюдения состояние любого объекта во многом является неопределенным. Способ наблюдения является фильтром, который извлекает из состояния, существовавшего до измерения, одну из содержащихся в нем возможностей.

Представленное исследование (главы 1 – 6) позволяет подойти к разработке «(Виртуального) самоорганизующегося, квантово-семантического нейрокомпьютера (интеллектуального вычислителя)». Фрагменты концепции такого вычислителя будет рассмотрены в главе 7.

## ГЛАВА 7. МОДЕЛИ ПОЗНАНИЯ И УПРАВЛЕНИЯ. КОГНИТИВНЫЕ СИСТЕМЫ

Синергетическая концепция в теории управления опирается на идею целенаправленной самоорганизации сложных систем путем формирования и преднамеренного ввода в исходные модели системы и поддержания с помощью управляющих воздействий желаемых притягивающих инвариантных многообразий в пространстве состояний системы, на которых естественные свойства объекта наилучшим образом согласуются с требованиями задачи управления [81]. В рамках развиваемого подхода притягивающие многообразия формируются на основе предельных синдромов и предельных вероятностных закономерностей.

Комплексная реализация методов на основе принципа предельных обобщений может привести к изменению феноменологической и алгоритмической базы ситуационных и когнитивных центров, когнитивных ядер госпитальных, телемедицинских, образовательных и иных систем.

### 7.1 Синдромный принцип управления слабоформализованными ситуациями и системами

Синдромный принцип управления вытекает из базовых принципов и опирается на синергетический принцип подчинения, который определяет «подстройку» многочисленных факторов ситуации к параметрам порядка ситуации. Процедура основывается на синдромных и вероятностных моделях знаний. Модели знаний, в свою очередь, формируются на основе базы прецедентов и онтологии предметной области. Каждый синдром может выступать как в роли диагностического критерия, так и в роли цели управления, задавая параметры порядка целевой ситуации. Таким образом, разрешается проблема целеполагания. Реализация синдромного принципа управления может быть выполнена с помощью среды радикалов, т.е. конкурентной среды, что в полной мере отвечает синергетическому подходу (см. раздел 7.2).

Напомним, что для любой синдромной модели знаний  $\{S\}$  можно построить сопряженную предельную модель  $\{S^*\}$ . Можно также найти все предельные синдромы на всех уровнях общности для каждой ситуации  $\alpha \in \Omega$ . Их объединение представляет собой полную предельную синдромную модель знаний в рамках контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ , которую мы обозначили  $\{S^*\}_{Full}$ . Модель  $\{S^*\}_{Full}$  доминирует все другие модели знаний. По аналогии с моделью  $\{S^*\}_{Full}$  строится полная предельная вероятностная модель знаний  $\{R^*\}_{Full}$ .

С точки зрения синергетики накопление знаний – это нелинейный процесс, и система знаний представляет сложную самоорганизующуюся систему, в которой один единственный новый факт может вызвать значительные подвижки в привычной структуре представлений. Таким фактом может быть, например, созревание нового домена в орграфе доменов теста, что может привести к полной перестройке моделей знаний.

Наиболее важной особенностью когнитивной самоорганизации является то, что в процессе упорядочивания (созревания) происходит резкое уменьшение системной информации за счет ее свертывания до предельных структур, моделей, включая синдромные и вероятностные модели знаний, экстремальные пограничные слои набросков, критические орграфы доменов тестов, критические пути в потоках системопаттернов (средах радикалов) и т.д. Свертывание информации без снижения эффективности системы (и даже с ее повышением) является главным эволюционным приобретением когнитивных динамических систем.

Предельные синдромы, претендующие на статус инвариантов, можно рассматривать в качестве параметров порядка развития сложных ситуаций. Действительно, смена одного предельного синдрома другим может привести к качественному изменению поведения системы или ситуации, которое вызвано сменой заключения  $\underline{z}/Z$ . Что касается *непредельных* синдромов, то их смена также может привести к качественному изменению поведения системы. Можно считать, что такие синдромы формируют *предпараметры порядка*. Наличие разных синдромов означает конкуренцию параметров порядка [181].

Факт возможности смены качественного поведения системы (ситуации, процесса) путем целенаправленного изменения текущих синдромов и лежит в основе синдромного принципа управления. Параметры порядка содержатся в некоторых (предельных) синдромах – это значения тестов  $\{\underline{z}/T\}_s$ . В ряде случаев данных для определения синдромов может оказаться недостаточно, тогда управление осуществляется на уровне вероятностных закономерностей (предвестников) из  $\{R\}$ .

Основные задачи синдромного управления:

- накопление опыта решения заданного класса задач в ходе когнитивной самоорганизации (meta-cognition);
- диагностика состояния (развития) ситуаций действительности;
- активное прогнозирование будущего;
- выбор целесообразного управленческого решения – одного из конкурирующих наборов параметров порядка.

Опишем феноменологию и алгоритм реализации синдромного управления.

Эволюцию (динамику) сложной системы или ситуации в рамках контекста  $\langle \Omega(Z), \{G(\tau)\}, \{S\}, \{R\} \rangle$  можно описать как последовательную

смену множеств синдромов и предвестников в рамках некоторых фиксированных моделей знаний  $\{S\}$  и  $\{R\}$ . Другими словами, *состояние системы* (ситуации) в момент времени  $t$  в рамках заданного контекста – это множество синдромов и предвестников в момент времени  $t$ , т.е.  $\{S\}_t$  &  $\{R\}_t$ , причем все синдромы из  $\{S\}_t$  указывают на одно и то же заключение  $z$  (требование непротиворечивости состояния и модели знаний).

*Траектория системы* – это последовательность состояний. В процессе эволюции системы (ситуации) состав синдромов может изменяться без изменения заключения  $z$ , а может скачком изменяться и  $z$ . На максимальном уровне обобщения (стратегическом уровне) траекторию можно описывать последовательностью изменений заключений  $z$ . Таким образом, состояние системы (ситуации, процесса) в фиксированный момент времени  $t$  в рамках контекста  $\langle \Omega(Z), \{G(\tau)\}, \{S\}, \{R\} \rangle$  можно описать орграфом набросков:

$$\{\underline{z}/T\}_t \Rightarrow \{S\}_t \& \{R\}_t \Rightarrow (\underline{z}/Z)_t, \text{ где } \{S\}_t \subset \{S\}, \{R\}_t \subset \{R\}. \quad (7.1)$$

В более общей форме, учитывая информацию предвестников, орграф набросков можно представить так:

$$\{\underline{z}/T\}_t \Rightarrow \{S\}_t \& \{R\}_t \Rightarrow \{J_z \underline{z}/Z\}_t, \text{ где } \{S\}_t \subset \{S\}, \{R\}_t \subset \{R\}, \quad (7.2)$$

где  $J_z$  – оценка возможности того или иного заключения (исхода). Одно заключение должно иметь оценку «необходимость», остальные – «возможность». Наличие нескольких заключений в рамках  $\{J_z \underline{z}/Z\}_t$  свидетельствует о неустойчивости текущего состояния процесса и возможном приближении бифуркации процесса (смены заключений).

На рис. 7.1 показан пример временной динамики изменения синдромов и заключений.

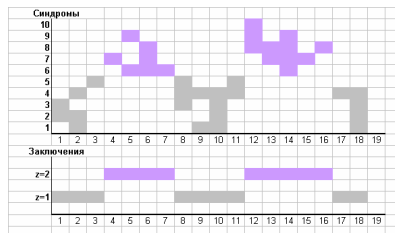


Рис. 7.1 - Пример временной динамики ситуации

Из рисунка ясно, что синдромы  $\{S_1, \dots, S_5\}$  определяют заключение  $z = 1$  (например, «норма»), соответственно, синдромы  $\{S_6, \dots, S_{10}\}$  определяют заключение  $z = 2$  (например, «отклонение»). В моменты времени  $t = 4, 8, 12, 17$  скачкообразно меняется характер процесса. Фрагмент траектории

процесса на синдромном уровне будет иметь следующий вид (индексы за скобками – такты времени):

$$\{S_2, S_3\}_1 \rightarrow \{S_1, S_2, S_4\}_2 \rightarrow \{S_5\}_3 \rightarrow \{S_7\}_4 \rightarrow \{S_6, S_9\}_5 \rightarrow \dots$$

Траектория процесса на уровне заключений имеет такой вид:

$$(z = 1)_{1-3} \rightarrow (z = 2)_{4-7} \rightarrow (z = 1)_{8-11} \rightarrow (z = 2)_{12-16} \rightarrow (z = 1)_{17-18}$$

Изменение состава синдромов и предвестников в состоянии может производиться целенаправленно – назовем это *управлением*, а может и спонтанно, например в результате действия естественных процессов деградации (старения) или воздействия дестабилизирующих внешних факторов.

Пусть  $S = (\{\underline{x}/T\} \rightarrow \{\underline{z}/Z\})$  – произвольный синдром. Ограничим для простоты множество  $Z$  тремя заключениями:  $z_1$  – благоприятное,  $z_2$  – неблагоприятное и  $z_3$  – неопределенное. Заключение  $z_3$  можно отнести к морфологии процесса. Всем артефактам базы прецедентов  $\Omega$  присваивается заключение  $z_3$ . Соответственно модель знаний  $\{S\}$  представима в виде:

$$\{S\} = \{S\}_1 \cup \{S\}_2 \cup \{S\}_3, \quad (7.3)$$

где  $\{S\}_j$  – фрагмент синдромной модели знаний, отвечающий  $z_j$  ( $j=1,2,3$ ). Аналогично, вероятностная модель  $\{R\}$  представима в виде:

$$\{R\} = \{R\}_1 \cup \{R\}_2 \cup \{R\}_3. \quad (7.4)$$

*Примечание.* В ряде случаев можно ограничиться только двумя заключениями:  $z_1$  – благоприятное,  $z_2$  – неблагоприятное. В других случаях список заключений может быть большим, например, это могут быть уровни помощи пациентам  $\{УП_1, \dots, УП_n\}$  [134].

Каждая из компонент моделей знаний  $\{S\}_j$  и  $\{R\}_j$  ( $j=1,2,3$ ), в свою очередь, делится на две непересекающиеся части: директивную зону (верхний индекс  $d$ ) и зону возможности (верхний индекс  $p$ ), а именно:

$$\{S\}_j = \{S\}_j^d \cup \{S\}_j^p; \quad \{R\}_j = \{R\}_j^d \cup \{R\}_j^p \quad (j=1,2,3). \quad (7.5)$$

Синдромам и закономерностям директивной зоны удовлетворяют все прецеденты с соответствующим заключением, чего нельзя сказать о синдромах и закономерностях зоны возможности. Представляется обоснованным стремление когнитивной системы в первую очередь выбирать для управления синдромы и закономерности именно директивных зон  $\{S\}_1^d$  и  $\{R\}_1^d$ , как наиболее узнаваемых в рамках имеющегося опыта. Следует, однако, отметить, что директивные зоны существуют не всегда.

Базовый домен заключений  $Z = \{1, 2, 3\}$  порождает три листка  $Z1 = \{1; -1\} = \{1, 4\}$ ,  $Z2 = \{2; -2\} = \{2, 5\}$ ,  $Z3 = \{3; -3\} = \{3, 6\}$  и, соответственно, три дуальные синдромные модели знаний (см. главу 5). На рис 7.2 показан результирующий оргграф доменов.

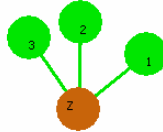


Рис. 7.2 – Оргграф доменов теста «Заключение»

Домены-листья порождают три дополнительных финитных наброска в оргграфе набросков (7.1), а именно:

$$\{\underline{z}/T\}_t \Rightarrow \{S\}_t \& \{R\}_t \Rightarrow \{\underline{z}/Z; \underline{z}/Z1; \underline{z}/Z2; \underline{z}/Z3\}_t \quad (7.6)$$

Множество значений тестов  $\{\underline{z}/T\}$  может включать в себя факты, т.е. такие значения тестов, которые по каким-либо причинам изменить нельзя, например, значение теста «Пол». Основной причиной невозможности изменения значений некоторых тестов является нехватка ресурсов – финансов, времени, мотивации (энергии) и т.д. Если изучается развитие ситуации после чрезвычайного события или катастрофы, например инфаркта миокарда, инсульта или техногенной катастрофы, то характеристики события (катастрофы) являются фактами.

Если существует синдром, который опирается только на факты, то такой синдром назовем *фатальным*. Наличие фатального синдрома в состоянии приводит к невозможности изменения качества процесса, т.е. заключения. Если фатальных синдромов нет, то управление, переводящее в благоприятный режим (нужный режим), в принципе возможно. Таким образом, перед началом управления необходимо убедиться в отсутствии фатальных синдромов неблагоприятного течения процесса. Их отсутствие будем трактовать как *потенциальную достижимость целей управления*, а именно: выход на благоприятный режим функционирования (развития ситуации) и его закрепление.

В ряде случаев фатальные синдромы могут быть преодолены путем сознательного провоцирования кризисов или катастроф. Действительно, чтобы совершить переход от безнадежно старой реальности, превратившейся в хаос, к реальности новой необходимо поломать устой прежнего порядка («Будьте реалистами! Требуйте невозможного!») - лозунг студентов Парижа в 1968 году). В результате может произойти катастрофическое разрешение структурного кризиса, в процессе которого будет преодолен фатальный синдром (желательный сценарий). Однако это



очень болезненный путь. На практике такой путь может означать хирургическое вмешательство, сокращение штата сотрудников, ликвидацию предприятия, военный конфликт, распад государства и т.д.

Фатальные синдромы зачастую определяют развитие ситуации по типу «*фатальной воронки*», когда независимо от решения ситуация движется к катастрофе, последствия которой не всегда предсказуемы. Конкретной причиной деструкции может послужить все что угодно, но каким бы ни был в реальности «спусковой механизм», сам процесс, один раз запущенный, скорее всего, приобретет необратимо «лавинный» характер, развиваясь по известному «принципу домино». «Лавина деструкции» может остановиться, лишь дойдя до самых основ системы (что может означать гибель системы).

Варианты развития ситуации, отвечающие  $z_2$  и  $z_3$ , следует отнести к неблагоприятным в широком понимании этого слова (с разной степенью неблагоприятности).

Суть синдромного управления (СУ), следовательно, состоит:

- во-первых, в изучении или *приобретении опыта* решения подобных задач (формирование синдромных и вероятностных моделей знаний, умений);

- во-вторых (для заданной ситуации), в *выборе стратегии, т.е. целевых синдромов* (и/или вероятностных закономерностей), содержащих параметры порядка, и плана целенаправленного изменения значений тестов, входящих в целевые синдромы (закономерности), но не являющихся фактами;

- в-третьих, в *доказательстве возможности* реализации стратегии путем исключения фатальных синдромов неблагоприятного развития и обоснования наличия необходимых ресурсов для достижения целей;

- в-четвертых, в *контроле и минимизации рисков* возникновения угрожающих состояний (осложнений, катастроф, кризисов, революций) в переходный период (вполне может оказаться так, что без разрушения или коренной трансформации старых структур невозможно выйти на благоприятный режим);

- в-пятых, в *стабилизации* «хороших» синдромов и устранении неблагоприятных предвестников (после того как достигнуто благоприятное развитие ситуации).

Синдромное управление обеспечивает либо стабилизацию состояния (для  $z_1$ ), либо достижение желательных синдромов (вероятностных закономерностей), отвечающих  $z_1$ . Оно существенным образом опирается на синергетический принцип подчинения: достижение параметров порядка  $\{\underline{z}\}_S$  приведет к автоматическому изменению (подстройке) подчиненных параметров  $\{\underline{z}\}'$ . Когнитивная система, опираясь на приобретенный опыт и

сосредотачиваясь на целевых параметрах порядка, минимизирует энергию на выработку и контроль реализации управления.

Вопросы приобретения опыта, т.е. формирования синдромных и вероятностных моделей знаний, были рассмотрены в главах 5 и 6.

Собственно (стратегическое) управление можно представить в виде:

$$U = \cup_S \{ \underline{t}'/T \rightarrow \underline{t}/T \}_S \vee \cup_R \{ \underline{t}''/T \rightarrow \underline{t}/T \}_R, \quad S \in \{S\}_U, \quad R \in \{R\}_U, \quad (7.7)$$

где  $\{S\}_U$  – множество целевых синдромов, а  $\{R\}_U$  – множество целевых вероятностных закономерностей. Управление  $U$  может включать только  $S$  или только  $R$  (данных или ресурсов для контроля достижения  $S$  может быть недостаточно). При очень неблагоприятном текущем развитии ситуации в качестве промежуточных целевых синдромов могут быть выбраны синдромы  $S(\{\underline{t}/T\}, z_3)$ , отвечающие промежуточному (неопределенному) состоянию.

Пусть  $\{\underline{t}/T\}_U$  – совокупное множество различных целевых значений тестов, входящих в  $\{S\}_U$  и  $\{R\}_U$  и не достигнутых к текущему моменту управления. Будем считать, что все тесты, входящие в  $\{\tau\}_U$ , различны. Это необходимое условие согласованности управления (нельзя требовать одновременного достижения разных значений одного и того же теста). Эквивалентное, но не избыточное, как (7.5), представление стратегического управления следующее:

$$U = \{ \underline{t}'/T \rightarrow \underline{t}/T \}_U. \quad (7.8)$$

Рассмотрим некоторые важные характеристики синдромного управления. Пусть для любого  $\tau \in \{\tau\}_U$  время  $t_\tau$  означает время достижения значения  $\underline{t}/T$  (без перерегулирования, т.е. без колебательного процесса), тогда время достижения всех целей управления  $U$  определяется выражением

$$(\underline{t}/A)_U = \max \{ (\underline{t}/A)_\tau \mid \tau \in \{\tau\}_U, A \in G(t) \}. \quad (7.9)$$

Время  $t_U$  может быть параметром управления, например, задающим ограничение:  $t_U < t^*_U$ . В этом случае его можно рассматривать в качестве *ожидаемого времени выхода на благоприятный режим*. Поскольку время  $t_U$  может измеряться месяцами и даже годами, то в качестве этапных ограничений могут выступать времена  $\{t_\tau\}$ . На основе  $\{t_\tau\}$  определяются также времена реализации синдромов  $\{t_S\}$  и времена реализации предвестников  $\{t_R\}$  из  $\{S\}_U$  и  $\{R\}_U$ , а именно:

$$(\underline{t}/A)_S = \max \{ (\underline{t}/A)_\tau \mid \tau \in \{\tau\}_S \cap \{\tau\}_U, A \in G(t) \} \quad (7.10)$$

$$(\underline{t}/A)_R = \max \{ (\underline{t}/A)_\tau \mid \tau \in \{\tau\}_R \cap \{\tau\}_U, A \in G(t) \} \quad (7.11)$$

Важнейшей характеристикой любого управления является энергия, затраченная на реализацию управления (затраты). Пусть для любого  $\tau \in \{\tau\}_U$  энергия  $e_\tau$  означает обобщенную энергию (затраты) достижения значения  $\underline{t}T$  (без перерегулирования), тогда энергия достижения всех целей управления  $U$  является некоторой функцией от затрат  $\{e_\tau\}$ , а именно:

$$(\underline{e}/E)_U = F_U(\{(\underline{e}/E)_\tau \mid \tau \in \{\tau\}_U, E_\tau \in G(e_\tau), E_U \in G(e_U)\}). \quad (7.12)$$

В некоторых простых случаях это может быть аддитивная функция. В более сложных случаях необходимо привлекать конструкции квантовой семантики, в частности квантово-семантическое запутывание орграфов  $\{G(e_\tau)\}$  и  $G(e_U)$ .

Очевидно стремление минимизировать энергию  $(e/E)_\tau$  для любого  $\tau$ . Предположим, что в рамках некоторого домена орграфа  $G(e_\tau)$  затраты описываются тремя значениями {Большие  $\wedge 3$ ; Средние  $\wedge 2$ ; Низкие  $\wedge 1$ }. Для каждого  $\tau \in \{\tau\}_U$  сформируем базу прецедентов  $\Omega(Z_{e,\tau})$ , где  $Z_{e,\tau} = \{1, 2, 3\}$  и контекст  $\langle \Omega(Z_{e,\tau}), \{G(\tau)\}, G(e_\tau) \rangle$ . В рамках указанных контекстов формируются модели знаний  $\{S\}_{e,\tau}$  и  $\{R\}_{e,\tau}$ . Остается выбрать синдромы и/или вероятностные закономерности, которые отвечают низким затратам. Для новых синдромов цикл оптимизации по затратам повторяется (до тех пор, пока удастся сформировать базы прецедентов). В итоге получается самоподобная иерархическая процедура оптимизации затрат, напоминающая по принципам построения (физический) фрактал. После ряда реализаций энергетически оптимальные синдромные схемы разных уровней управления сохраняются в памяти, формируя «критический путь» в виде автоматизма вычислительной среды. Любой критический путь саморазвертывается в процессе выполнения, привлекая все более тонкие механизмы регулирования. За верхним уровнем управления остается только функция запуска автоматизма типа «критический путь».

На основе  $\{e_\tau\}$  определяются затраты на реализацию синдромов  $\{e_S\}$  и затраты на реализации предвестников  $\{e_R\}$  из  $\{S\}_U$  и  $\{R\}_U$ , а именно:

$$(\underline{e}/E)_S = F_S(\{(\underline{e}/E)_\tau \mid \tau \in \{\tau\}_S \cap \{\tau\}_U, E_\tau \in G(e_\tau), E_S \in G(e_S)\}). \quad (7.13)$$

$$(\underline{e}/E)_R = F_R(\{(\underline{e}/E)_\tau \mid \tau \in \{\tau\}_R \cap \{\tau\}_U, E_\tau \in G(e_\tau), E_R \in G(e_R)\}). \quad (7.14)$$

Приведенные оценки следует рассматривать и как глобальные (результатирующие), и как локальные в рамках любого кванта времени.

Выражения (7.9) – (7.14) позволяют рассмотреть возможный сценарий отбора синдромов  $\{S\}_U$  и предвестников  $\{R\}_U$ .

На множествах синдромов  $\{S\}_1$  и предвестников  $\{R\}_1$  сформируем оценки минимальных затрат и времен достижения каждого синдрома и предвестника:  $\{(\underline{e}/E)_S\}$ ,  $\{(\underline{e}/E)_R\}$ ,  $\{(\underline{t}/A)_S\}$ ,  $\{(\underline{t}/A)_R\}$ . Подобные оценки

могут быть частью опыта, т.е. неотъемлемыми характеристиками синдромов и предвестников (наравне с весами). Каждая оценка представляет собой орграф доменов, один из вариантов которого в формате квантовой семантики (QS-формате) представлен ниже:

*Оценка*  $(e_s, t_s, \dots)$  {  
 2 {Приемлемая  $\wedge^1$  2; Неприемлемая  $\wedge^2$  3}  
 1 {Большая  $\wedge^3$ ; Средняя  $\wedge^2$ ; Низкая  $\wedge^1$ }}

Ясно, что созревание такого орграфа идет от оппозиции {Приемлемая; Неприемлемая}, поэтому связь « $\wedge^2$ » у элементов домена «2» является мерцающей, многозначной (см. главу 6). В ходе QS-декогеренции многозначность связи может быть устранена (порождаются два орграфа-близнеца). В плане анализа эволюции интеллекта важно отметить, что все орграфы могут содержать только «мыслительный код» (без термов) и даже только направления информационных потоков, а именно:

*Оценка*  $(e_s, t_s, \dots)$  {2 { $\wedge^1$  2;  $\wedge^2$  3} 1 { $\wedge^3$ ;  $\wedge^2$ ;  $\wedge^1$ }}

Такое представление соответствует каноническому орграфу и просто реализуется с помощью нейронных ансамблей.

Исходя из оценок, формируются множества недоминируемых (по затратам и времени) альтернатив  $\{S\}_1^*$  и  $\{R\}_1^*$ . И, наконец, с привлечением весов (например, директивных зон моделей знаний  $\{S\}_1^d$  и  $\{R\}_1^d$ ) и эмоциональных оценок формируются множества  $\{S\}_U$  и  $\{R\}_U$ . Схематично процесс формирования стратегии можно отобразить так

$$\{S\}_1, \{R\}_1 \rightarrow \{S\}_1^*, \{R\}_1^* \rightarrow \{S\}_U, \{R\}_U \quad (7.15)$$

Какие домены тестов «Оценка» (энергии, времени) используются в каждой конкретной ситуации принятия решения зависит от контекста ситуации.

При реализации стратегии в рамках синергетической парадигмы контроль многочисленных процессов в переходный период (от неблагоприятной фазы к благоприятной) не является обязательным. Однако в ряде ПрО такой контроль желателен и даже необходим. Пример – медицина, где необходимо предупреждать осложнения.

Для контроля рисков осложнений, кризисов, катастроф в переходный период введем тест «Риск» с помощью орграфа (в QS-формате)

*Риск* {2 {Приемлемый  $\wedge^1$  2; Неприемлемый  $\wedge^2$  3} 1 {Большой  $\wedge^3$ ; Средний  $\wedge^2$ ; Низкий  $\wedge^1$ }}

Мы намеренно стремились сохранить схожесть с орграфом теста «Оценка», чтобы показать возможность однотипных решений (тиражирование

решений легко объяснить с точки зрения эволюции). Более детализированный оргграф доменов теста «Риск» был приведен в главе 4.

Введем тест «Осложнение» (угрожающие состояния, кризисы, катастрофы) с помощью оргграфа доменов

*Осложнение*  $\{2 \{Тяжелое О. \wedge 1 \dots; Среднее О. \wedge k \dots; Легкое О. \wedge n \dots\} 1 \{O_1 \wedge 1; O_2 \wedge 2; O_3 \wedge 3; \dots\}\}$

Очевидно, в каждой ПрО будет свой список осложнений (угрожающих состояний, кризисов), которые желательно или необходимо контролировать в переходный период. Для каждого осложнения создадим базу прецедентов  $\Omega(Z_o, O_j)$ , где  $Z_o = \{1 - \text{не произошло}; 2 - \text{произошло}\}$ ;  $O_j - j$ -е осложнение (кризис). В рамках каждого контекста  $\langle \Omega(Z_o, O_j), \{G(\tau)\} \rangle$  формируются модели знаний  $\{S\}_{o_j}$  и  $\{R\}_{o_j}$ . Поскольку

$$\{S\}_{o_j} = \{S\}_{1,o_j} \cup \{S\}_{2,o_j}; \quad \{R\}_{o_j} = \{R\}_{1,o_j} \cup \{R\}_{2,o_j}, \quad (7.16)$$

то в переходный период необходимо избежать реализации синдромов  $\{S\}_{2,o_j}$  и предвестников  $\{R\}_{2,o_j}$ , которые выступают в качестве этапных ограничений. Ясно, что в первую очередь следует избегать синдромов и закономерностей директивных зон  $\{S\}_{1,o_j}^d$  и  $\{R\}_{1,o_j}^d$ . Оргграф теста «Осложнение» позволяет ситуативно изменять уровень общности контролируемых угроз.

Пример динамики ситуации. Если по горизонтальной оси отложить время, а по вертикальной – шкалу «благоприятная – неблагоприятная», то динамика ситуации может быть отображена так, как показано на рис. 7.3.

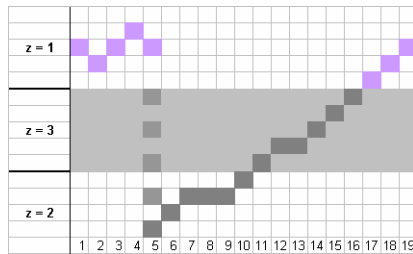


Рис. 7.3 - Пример временной динамики с ударным воздействием

Из рисунка видно, что в момент времени  $t = 5$  произошло скачкообразное ухудшение ситуации (катастрофа, например, сосудистая – инсульт или инфаркт, травма). С помощью синдромного управления удалось постепенно перевести ситуацию в зону неопределенности, а в момент  $t = 17$

ситуация снова вышла на благоприятный режим. Текущая задача – закрепить благоприятный режим (например, укрепить здоровье).

Стратегия стабилизации и закрепления благоприятной эволюции (повышение устойчивости) заключается, в частности, в стремлении увеличить число благоприятных синдромов. Общее число таких синдромов дает синдромная модель знаний в сочетании с методом выделения групп совместимых синдромов (синдром может принадлежать нескольким группам). Подобное увеличение можно интерпретировать как повышение гарантии благоприятной эволюции. Число синдромов можно интерпретировать как *индекс устойчивости*. Чем он меньше, тем ситуация (процесс) ближе к точке бифуркации. На рис. 7.4 для рассмотренного выше примера (рис. 7.1) показано число синдромов в каждый такт времени.

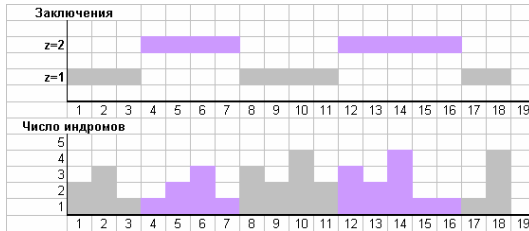


Рис. 7.4 - Пример временной динамики

Из рисунка видно, что в моменты переключения фаз процесса число синдромов минимально, индекс устойчивости равен 1 (почти всегда). Это может свидетельствовать о нахождении ситуации (системы) в зоне бифуркации, т.е. сильной неустойчивости (случай безударной динамики). На рис. 7.3 приведен пример динамики с первоначальным ударным воздействием (резким скачком), а затем с постепенной эволюцией.

Если увеличить число благоприятных синдромов не удастся, то необходимо стремиться максимально замедлить процесс деградации, т.е. постепенного исчезновения благоприятных синдромов. Неуправляемое изменение состава синдромов происходит в результате естественной эволюции саморазвивающейся сложной системы (ситуации, процесса):  $\{\underline{z}/T \rightarrow \underline{z}''/T\}$ . Стабилизация имеющихся синдромов заключается в стабилизации значения каждого теста (не являющегося фактом) из множества  $\{\underline{z}/T\}_S$  для  $S \in \{S\}_\alpha$ . «Слабыми звеньями» или *критическими звеньями* (элементами) являются те значения тестов  $\underline{z}/T$ , стабилизация которых наиболее проблематична (с точки зрения имеющихся ресурсов и мотивации), следовательно, от их стабилизации и будет зависеть *устойчивость* благоприятного развития процесса (ситуации). Например,

одним из тестов «хорошего» синдрома у пациента может быть следующий « $\underline{z}/T$  = Семейное положение? состоит в браке», следовательно, от продолжительности сохранения брака будет зависеть, в частности, судьба данного синдрома, и если брак «висит на волоске», то это критическое звено в стабилизации благоприятного течения процесса.

Перманентно должны выполняться *отдельные программы* по стабилизации критических звеньев «хороших» синдромов, что обеспечит, по крайней мере, неухудшение устойчивости благоприятного течения процесса (например, владение иностранным языком требует постоянного совершенствования; для поддержания здоровья нужно заниматься спортом и т.д.). На это могут потребоваться ресурсы не меньше, а может и больше, чем на достижение новых синдромов. Каждая программа стабилизации значения какого-либо теста также формируется на основе синдромного принципа управления: строятся модели знаний, выбираются и реализуются целевые синдромы и/или вероятностные закономерности (и т.д. по индукции). В качестве множества заключений  $Z$  может использоваться оппозиция: {1 - Показатель стабильный; 2 - Показатель ухудшается}.

Важным аспектом стабилизации является устранение или недопущение предвестников наступления неблагоприятной фазы развития ситуации.

Приведем для примера один из вариантов формализации *индексов устойчивости* (ИУ) благоприятной динамики ситуации. Пусть  $\sigma_\tau$  – ИУ значения теста  $\underline{z}/T$ ;  $\sigma_S$  – ИУ синдрома  $S$ ;  $\sigma_\alpha$  – ИУ ситуации  $\alpha$ . Положим

$$\sigma_\tau \in (0, 1],$$

$$\sigma_S = \sum_\tau (\sigma_\tau | \tau \in \{\tau\}_S) / |\{\tau\}_S|,$$

$$\sigma_\alpha = \sum_S (\sigma_S | S \in \{S\}_\alpha) / (|\{S\}_\alpha| + |\{R\}_{2,\alpha}|),$$

где  $\{S\}_\alpha$  – текущее множество «хороших» синдромов для ситуации  $\alpha$ ;  $\{R\}_{2,\alpha}$  – множество текущих предвестников неблагоприятного исхода. Все  $\sigma \in (0, 1]$ . Максимальная устойчивость значения теста, синдрома, ситуации имеет место тогда, когда  $\sigma = 1$ . Конфигуратор теста «Индекс устойчивости» может быть следующий (в QS-формате):

*Индекс устойчивости*  $\wedge \sigma \{$

3 {Приемлемый  $\wedge^3$  2; Неприемлемый  $\wedge^1$  2}

2 {Низкий  $\wedge^1$  (0; p1]; Средний  $\wedge^2$  (p1; p2]; Большой  $\wedge^3$  (p2; 1]}

1 {(0, 1]}}

На основе орграфа  $G(\sigma) = \{1 \rightarrow 2 \rightarrow 3\}$  может быть построен структурно-завершенный орграф  $G^+(\sigma)$ .

Прогнозируемые значения индекса устойчивости тех или иных синдромов (вероятностных закономерностей) могут служить основанием для их выбора или, наоборот, отказа от их реализации.

Если текущее состояние процесса неблагоприятное ( $z_2$  или  $z_3$ ), то необходимо выбрать в качестве цели один из достижимых благоприятных синдромов (если таковые имеются) и реализовать управление  $U$ , т.е. обеспечить целевое изменение значений тестов. Разные достижимые синдромы дают разные варианты (альтернативы) управления. Возможный сценарий выбора того или иного варианта рассмотрен выше. Следует отметить, что достижимость синдрома может существенно зависеть от мотивации того, кто будет осуществлять непосредственное управление, поэтому потенциальная достижимость еще не означает реальной достижимости. Это необходимо учитывать. Следует выбирать такие целевые синдромы, достижимость которых меньше всего зависит от субъективной мотивации.

Покажем, что СПУ позволяет реализовать иерархический самоподобный процесс управления.

Пусть  $S(\{\underline{z}T\}, z_1)$  – целевой синдром. Любое целевое значение теста  $\underline{z}T \in \{\underline{z}T\}_S$  определяет оппозицию заключений  $Z_\tau = \{z_{\underline{z}}; z_{-\underline{z}}\}$  и новый контекст  $\langle \Omega(Z_\tau), \{G(\tau)\} \rangle$ . В рамках данного контекста определяется модель знаний  $\{S\}_\tau$ . Соответственно встает задача выбора целевого синдрома 2-го уровня иерархии для достижения цели  $z_{\underline{z}}$ . Общее количество задач 2-го уровня, связанных с синдромом  $S(\{\underline{z}T\}, z_1)$ , не превышает  $|\{\underline{z}T\}_S|$  (некоторые значения тестов могут быть уже достигнуты). Каждый целевой синдром 2-го уровня аналогичным образом порождает определенное число задач 3-го уровня и т.д. Взаимоувязка и последовательность решения всех задач определяется имеющимися ресурсами. По аналогии с синдромами строится иерархия управлений на основе вероятностных закономерностей. В итоге получили иерархический самоподобный процесс управления.

Поскольку в качестве цели управления может выбираться не один синдром или вероятностная закономерность, а несколько совместимых синдромов и закономерностей, то в таком случае количество задач управления 2-го уровня не превышает значение (верхняя оценка):

$$M_2 = \Pi_S |\{\tau\}_S| + \Pi_R |\{\tau\}_R|, \quad S \in \{S\}_U, \quad R \in \{R\}_U. \quad (7.17)$$

Точное количество задач 2-го уровня составляет  $|\{\tau\}_U|$ .

Вопросы поиска совместимых синдромов в процессе формирования управления достаточно подробно рассмотрены в главе 5.

Ниже приведен результирующий алгоритм синдромного управления (полный вариант).



*Алгоритм 7.1 – Синдромное управление*

*Дано:*  $Z$  – множество заключений, в котором  $z_1$  – выделенное (благоприятное) состояние или исход; ситуация действительности  $\alpha$  (известно только, что ее динамика может оцениваться шкалой  $Z$ ).

*Найти:* 1) заключение  $z_\alpha$ ; 2) управление  $U$ , которое либо переводит ситуацию  $\alpha$  в благоприятный режим (при  $z_\alpha > 1$ ), либо стабилизирует благоприятный режим (при  $z_\alpha = 1$ ).

*Выполнить:* 1) реализовать управление  $U$ ; 2) включить в опыт (сохранить) найденные эффективные когнитивно-поведенческие структуры.

*Тело алгоритма:*

*Этап I. Приобретение знаний и умений (опыта – meta-cognition).*

*Шаг 1.* Создание или модификация Банка тестов  $\{G(\tau)\}$ .

*Шаг 2.* Создание, модификация Базы прецедентов  $\Omega = \{\alpha\{\underline{z}/T\}, \underline{z}/Z\}$ .

*Шаг 3.* Построение предельной синдромной модели знаний  $\{S^*\}_{Full}$  и предельной вероятностной модели знаний  $\{R^*\}_{Full}$  в рамках контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ .

*Шаг 4.* Создание физических (телесных) или виртуальных иерархических функциональных систем и автоматизмов среды (включая критические пути), обеспечивающих функционирование моделей знаний.

*Этап II. Применение опыта для оценки ситуации и выработки управления.*

*Шаг 5.* Оценка состояния ситуации  $\alpha\{\underline{z}/T\}, \underline{z}/Z$  – определение синдромов  $\{S\}_\alpha$  и заключения  $z_\alpha$ .

*Шаг 6.* Если  $z_\alpha > 1$ , то: 1) исключение фатальных синдромов неблагоприятного развития; 2) определение части или всего множества достижимых (по ресурсам) синдромов и предвестников, обеспечивающих выход на благоприятное развитие ситуации. В случае неуспеха решения любой из задач *Выход* с заключением «Активный мониторинг».

*Шаг 7.* Если  $z_\alpha > 1$ , то: 1) формирование множеств запрещенных в переходный период синдромов  $\{S\}_{2,of}$  и предвестников  $\{R\}_{2,of}$  осложнений, кризисов; 2) формирование управления  $U$  на множестве достижимых синдромов и предвестников с учетом ограничений на качество переходного периода и сохраненных в опыте «критических путей».

*Шаг 8.* Если  $z_\alpha = 1$ , то: 1) определение части или всего множества достижимых (по ресурсам) синдромов и вероятностных закономерностей, обеспечивающих улучшение или стабилизацию благоприятного развития ситуации; 2) определение предвестников развития неблагоприятной фазы (исхода); 3) формирование управления  $U$  на множестве достижимых

синдромов и вероятностных закономерностей с учетом нейтрализации (или замедления) появления предвестников неблагоприятной фазы, а также сохраненных в опыте «критических путей». Если  $U$  не сформировано, то *Выход* с заключением «Мониторинг».

*Шаг 9.* 1) реализация управления  $U$ ; 2) если  $U$  приемлемо по ресурсам и имеет ярко выраженную положительную эмоциональную оценку, то сохранение  $U$  в качестве «критического пути».

*Шаг 10.* В регламентированные моменты времени возврат к *Шагу 5*. Если на *Этапе II* происходила корректировка или контекста или моделей знаний или автоматизмов среды (критических путей), то возврат к *Шагу 1*.

Окончание алгоритма (*Выход*).

Сокращенный вариант алгоритма синдромного управления подразумевает, что *Этап I* (meta-cognition) выполнен заранее и модели знаний известны, следовательно, он включает только *Этап II* (cognition).

Приведенный алгоритм СУ не учитывает возможность фальсификации синдромной модели знаний, т.е. глубокое переосмысление опыта. Этот вопрос подробно рассматривался в главе 5.

В следующем разделе будет рассмотрена схема реализации алгоритма синдромного управления с использованием иерархии функциональных систем (сред радикалов или агентных систем), включая физическую (телесную) реализацию критических путей.

В завершение раздела приведем названия некоторых ПрО, где с успехом может применяться синдромный принцип управления: «Здоровье», «Преступность», «Наркомания», «Бизнес», «Урожайность», «Качество образования», «Инвестиции», «Выборы», «Экология», «Инвалидность», «Безработица», «Технологии», «Рейтинг (вузов)», «Политика», «Когнитивный (административный, бизнес) центр», «Центр когнитивных образовательных программ и технологий» и т.д.

## **7.2 Среды радикалов. Синтетическая телесно-феноменологическая онтология**

Для придания активности и «телесности» моделям знаний синдромные и вероятностные модели предлагается преобразовывать в среду радикалов [147]. Понятие «радикал» было введено для символического моделирования проблемных областей сложных систем [189]. Под *радикалом* понимается любая функциональная система, имеющая два доступных извне состояния: *активное* и *пассивное*. Активный радикал функционирует согласно своему предназначению, а пассивный радикал нет. Он как бы выключен.

Множество радикалов со связями между собой является *средой радикалов* (СР). Вопросами активирования среды радикалов занимаются системы, которые называются *активаторы*. Система всех активных радикалов среды радикалов в данный момент времени образует *системоквант*, который определяет квант поведения среды в этот момент. Интеллектуальная система должна определять такие системокванты поведения сложной системы, которые обеспечивают функциональную устойчивость системы на протяжении всего ее жизненного цикла [190].

*Примечание.* Развиваемая далее конкретизация понятий «Радикал» и «Среда радикалов» является уникальной и не коррелирует ни с одной известной моделью.

Операция ультраоснащения для каждого синдрома и вероятностной закономерности создает радикал в виде нейроморфного ансамбля: радикал-синдром (S-радикал) и радикал-закономерность (R-радикал) [147]. Задачей радикала (биофизического агента) в активном состоянии является установление наличия или отсутствия соответствующего синдрома (предвестника) у анализируемой ситуации.

Таким образом, каждая модель знаний порождает среду радикалов. В определенном смысле разные среды являются конкурирующими. Активность среды радикалов в процессе решения целевой задачи для нового прецедента зависит от выбранного сценария. Сценарии различаются, в частности, порядком применения некоторых критериев оптимальности, а также допустимой мощностью системоквантов (допустимая мощность – внешний по отношению к СР управляющий параметр). Сохраненная в памяти активность СР является разновидностью «критического пути».

Под *функциональной системой* понимается комплекс избирательно вовлеченных компонентов, взаимоотношения между которыми принимают характер синергизма, обеспечивающего получение системой фокусированного полезного результата [11, 25]. Базовый синергетический принцип формирования любой функциональной системы – принцип взаимодействия ее компонентов на получение общего результата. Предполагается, что каждый элемент системы не только привносит свой индивидуальный микровклад в дело достижения приспособительного эффекта, но и параллельно с этим облегчает реализацию аналогичных привнесений другими членами данной интеграции. Среда радикалов позволяет моделировать простейшую форму процесса взаимодействия компонентов в ходе достижения ими общего результата (получение заключения).

Таким образом, с концептуальной точки зрения можно считать, что радикал является *телом синдрома* (закономерности), а среда радикалов

является *телом* (когнитивно-поведенческой) функциональной системы, решающей Z-задачу. Каждая Z-задача формирует собственную функциональную систему, а поскольку задачи (управления, принятия решений) образуют иерархию, то и функциональные системы также образуют иерархию.

Отметим некоторые важные свойства, которыми должны обладать радикалы и среды радикалов в рамках развиваемой парадигмы.

Во-первых, радикалы могут взаимодействовать между собой, влияя на процессы активации/деактивации в среде радикалов. В результате самоорганизации под влиянием принципа предельных обобщений в СР могут образовываться критические структуры, в частности критические пути – последовательности системоквантов. Один физический нейрон может входить в разные радикалы.

Во-вторых, радикалы могут взаимодействовать между собой как точно, так и распределенно с помощью полей и/или сред. Действительно, в современном естествознании феномен жизни объясняется через когерентное состояние вещества тканей живых существ, которые образуют активную среду, формирующую самоподдерживающийся электромагнитный каркас организма. Естественно предположить, что данный каркас играет важную роль и в деятельности когнитивных функциональных систем (сред радикалов), обеспечивая их делокализацию. В результате границы таких систем неопределенно расширяются по сравнению с физическим телом, что позволяет, в частности, объяснить многие Psi-феномены (телепатии, суггестии и т.д.) [127].

Постулируем, что нейроны обладают «фоновой» разрядной активностью. Феномен поведенческой специализации при этом возникает не как процесс генерации молчащими нейронами избирательных фазных активаций, приуроченных к определенному этапу системопаттерна, а в итоге фрагментации в такие периоды постоянно присутствующего «фонового потока» биоэлектрических разрядов (как переход от пачечно-групповой формы импульсации к регулярной и наоборот) [25]. Это важно, прежде всего, для физического моделирования.

В-третьих, в результате QS-рекогеренции среду радикалов можно представить как *суперпозицию радикалов* (неразделяемый смысл). Точно так же иерархию функциональных систем можно представить в виде суперпозиции функциональных систем (сред радикалов, сред биофизических агентов). QS-рекогеренция объясняет практическую невозможность вербализации точной временной последовательности системоквантов, т.е. детального алгоритма (логики) выработки решения или управления. В большинстве случаев можно только с определенной вероятностью установить, что обработала та или иная Z-задача (да и то не

всегда, учитывая дуальные модели знаний). В главе 6 были введены понятия QS-радикал, QS-ФС.

Среду радикалов или функциональную систему будем реализовывать с помощью нейроморфной сети (НМС). Синдромный подход обеспечивает логическую прозрачность построения подобной нейроморфной сети, а реализация предельной модели знаний обеспечивает предельные когнитивные характеристики сети.

На рис. 7.5 представлена общая архитектура нейронного ансамбля, реализующего радикал-синдром  $S_{ij}$  и функцию распределенного информационного обмена [147].

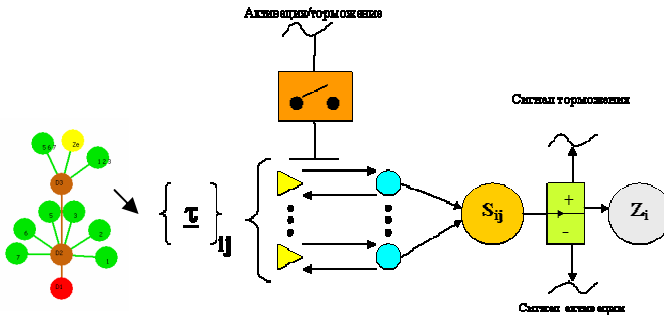








Рис.7.5 – Модель нейронного ансамбля, реализующего радикал-синдром

Опишем назначение и механизм работы каждого нейрона:

	<p>- <i>Активирующий/деактивирующий нейрон (АД-нейрон)</i>. Сигнал управления на этот нейрон может подаваться напрямую. Кроме того, нейрон осуществляет непрерывный мониторинг волновых процессов, проходящих в среде, на предмет наличия специфичной для этого нейрона частоты <math>\lambda</math> в спектре сигнала. Отдельный сигнал несет признак 1/0 - активации/деактивации. АД-нейрон или запускает, или прекращает работу всего <i>нейронного <math>\lambda</math>-ансамбля</i> (частота <math>\lambda</math> является уникальным кодом ансамбля). Подобный командный нейрон имеет и среда радикалов. Запускающий нейрон модулирует активность всех нейронов радикала, что следует считать необходимым и достаточным условием их идентификации в качестве элементов определенной функциональной системы.</p>
	<p>- <i>Нейроны проверки наличия значений тестов <math>\{\tau/T\}_{ij}</math> (<math>\tau</math>-Нейроны)</i>. Все они активируются АД-<math>\lambda</math>-нейроном. Каждый такой нейрон генерирует непрерывный сигнал «1», если он получил на входе</p>

	требуемое значение $\underline{z}/T$ (домены принадлежат орграфу доменов $G(\tau)$ ). Значение теста $\tau$ -нейрон ищет на множестве результатов, предоставляемых поисковой нейронной сетью (кратко – поисковым нейроном). Если требуемого значения теста нет, то $\tau$ -нейрон генерирует сигнал «0».
	- <i>Поисковые нейроны</i> (П-нейроны). Каждый поисковый нейрон (поисковая сеть) связан с соответствующим $\tau$ -нейроном и находится с ним в непрерывном взаимодействии. П-нейроны активизируются/деактивизируются АД- $\lambda$ -нейроном одновременно с $\tau$ -нейронами. После активизации каждый П-нейрон выполняет собственную программу поиска значений соответствующего теста $\tau$ . В простейшем случае, выводится запрос ( $\tau?$ ) на общую панель управления нейроморфной сети, после чего П-нейрон переходит в режим ожидания. После того как получено какое-то значение $\underline{z}$ происходит его замыкание ( $\underline{z}$ ) <sup>+</sup> в рамках нейронной сети орграфа доменов $G(\tau)$ и отправляется сигнал завершения работы $\tau$ -нейрону.
	- <i>Нейрон, реализующий логику синдрома</i> (S-нейрон). Если все поступившие сигналы от $\tau$ -нейронов «1», то генерируется «1», в противном случае – «0».
	- <i>Нейрон завершения работы <math>\lambda</math>-ансамбля</i> (З-нейрон). Если сигнал от S-нейрона «0», то в зависимости от общего характера управления работой НМС передается или адресный сигнал активации другого радикала или генерируется распределенный сигнал активации со спектром $[\lambda_1, \lambda_2]$ (спектр зависит от установок). Если сигнал от S-нейрона «1», то передается широкополосный сигнал торможения (деактивации) на все АД-нейроны радикалов и активизируется Z-нейрон.
	- <i>Нейрон фиксации результата</i> (Z-нейрон). Передает в центр управления НМС код заключения $\underline{z}/Z$ , соответствующий синдрому $S_{ij}$ .

По аналогии с НМС радикала – синдрома строится НМС радикала-закономерности (предвестника).

Как видим, в рамках радикала каждый нейрон обладает ярко выраженным свойством *системоспецифичности*. Другими словами, нейроны работают не как детекторы элементарных сенсорных или моторных признаков, а отражают в своей активности процесс формирования полноценных функциональных систем «организменного» уровня. Следует отметить также высокую *пластичность* среды радикалов в

целом (избыточность путей достижения цели). Пластичность в полной мере проявляет себя в условиях наличия препятствий достижения цели.

ФС «Модель знаний – Среда радикалов» в самом прямом смысле слова выступает как система функций (процессов), а не как набор морфологически определенных, взаимодействующих друг с другом элементов. На рис. 7.6 представлены примеры простых синдромных моделей знаний, формируемых методом предельных обобщений.

id_str	t1	t2	t3	ds	weight
1	1			1	5
2	2			1	4
3	3		6	1	1
4	4	7		2	3
5	5		7	2	2
6		6		2	1
7			1	2	4
8	6	2	3	3	2

id_str	t1	t2	t3	ds	weight
1	1			1	5
2	2			1	4
3	3			1	1
4	4			2	3
5	5			2	2
6		1		2	1
7		2		2	4
8		3		3	2
9		4		3	5
10			1	3	1
11			2	3	4
12			3	3	2

а)

б)

Рис.7.6 – Примеры синдромных моделей знаний

Синдромы содержат значения тестов t1, t2, t3, заключение ds и вес – weight. Вес синдрома определяется, например, количеством прецедентов, которые он охватывает. Для каждого синдрома автоматически создается радикал -нейронный  $\lambda$ -ансамбль (рис. 7.5). Итоговая нейроморфная сеть функциональной системы (среды радикалов) состоит из командного АД-ФС-нейрона и нейронных  $\lambda$ -ансамблей всех синдромов модели знаний, включая, возможно, связи между  $\lambda$ -ансамблями (радикалами).

На рис. 7.7 показаны примеры сред радикалов, соответствующие моделям знаний рис. 7.6. В центре каждого рисунка находится командный АД-ФС-нейрон, который активирует/деактивирует саму ФС (среду радикалов). Подобный командный нейрон есть у каждой ФС.

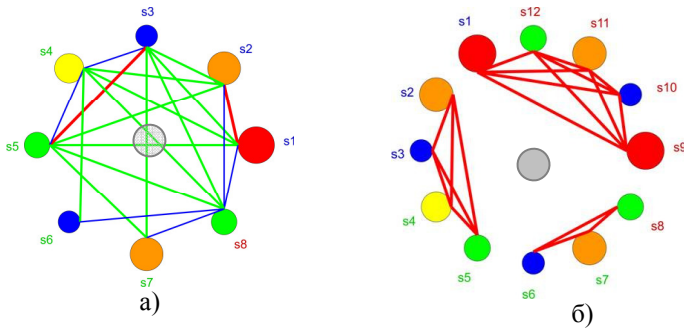


Рис. 7.7 – Примеры сред радикалов

Каждый нейронный  $\lambda$ -ансамбль (радикал) представлен на рис. 7.7 закрашенным кругом. Размер и цвет круга соответствуют его весу (красный – максимальный вес). Цвет номера синдрома соответствует коду заключения. Связь (проводимость) между радикалами отражает степень пересечения множеств тестов, фигурирующих в каждом синдроме (цвет связи соответствует уровню проводимости). Максимальная проводимость равна «1» (изображена красным цветом). Чем выше проводимость, тем выше приоритет активации синдрома. Связи между синдромами могут быть симметричны и асимметричны. Пример асимметричной связи (модель знаний 7.6.а):  $\text{Вес}(s3 \rightarrow s1) = 1$  (тесты  $s3$  полностью покрывают тесты  $s1$ ), а  $\text{Вес}(s1 \rightarrow s3)=0,5$  (тесты  $s1$  лишь на половину покрывают тесты  $s3$ ). Проводимость задается, например, таким конфигуратором (в QS-формате):

*Проводимость* {  
 3 { Достаточная ^3 2; Низкая ^0 1 2}  
 2 { Отсутствует ^0 [0; 0]; Низкая ^1 (0; p1]; Средняя ^2 (p1; p2]; Высокая ^3 (p2; 1]}  
 1 {[0, 1]} }

На рис. 7.8 показан пример учебной программы, автоматически формирующей НМС модели знаний с асимметричными связями. Программа позволяет моделировать все функции НМС [144].

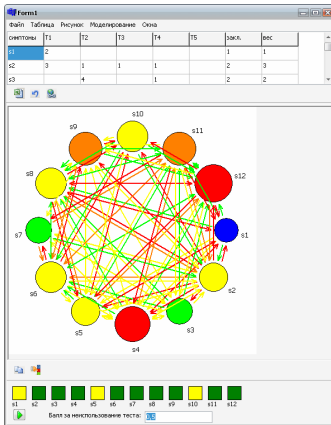


Рис. 7.8 – НМС модели знаний с асимметричными связями

Внизу диалогового окна показан текущий системоквант – линейка радикалов, в которой цветом обозначено состояние каждого радикала, и результат его работы в текущий момент времени (белый – еще не работал; зеленый – запрашивает данные; салатный – нуждается в данных, но их пока нет; желтый – отработал, но нет заключения; красный – заключение получено).

На рис. 7.9 показано точечное взаимодействие двух радикалов (ненулевая проводимость), а на рис. 7.10 показано полностью распределенное взаимодействие радикалов в среде радикалов. Характер работы НМС существенно зависит от общей энергии, выделяемой центром



управления НМС (АД-ФС-нейроном): при максимальной энергии одновременно запускаются все радикалы (рис. 7.10); при минимальной – в каждый момент времени функционирует только один радикал; возможны промежуточные варианты.

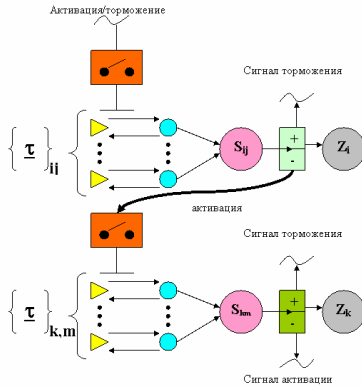


Рис. 7.9 – Точечное взаимодействие двух радикалов

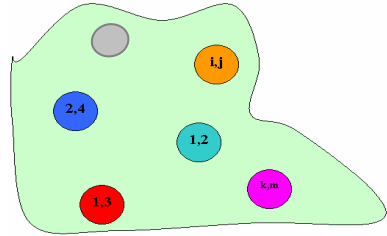


Рис. 7.10 – Распределенное взаимодействие в среде радикалов

Запуск НМС любого уровня осуществляется как точно, так и с помощью внешнего волнового поля, содержащего те или иные  $\lambda$  частоты (в пределе все  $\lambda$  частоты). Связи (проводимости) между радикалами могут определяться не только контекстом  $\langle \Omega(Z), \{G(\tau)\} \rangle$  (т.е. априорными весами синдромов и предвестников), но и устанавливаться в процессе обучения. В общем случае, вычислительная траектория как последовательность системоквантов (одновременно работающих радикалов) определяется ситуативно.

Между радикалами одной ФС есть одновременно и кооперация, и конкуренция за ресурсы (энергию). Другими словами, есть все условия для возникновения самоорганизации вычислительных путей (последовательностей системоквантов) в рамках каждой ФС. Самоорганизация приводит к формированию «критических путей». Подчеркнем, что данные процессы не зависят от самой модели знаний и определяются исключительно «физикой» системы, ее телесностью. Ясно, что одна и та же Z-задача может решаться разными ФС, разными последовательностями системоквантов с разной энергией. Эволюционно выигрывают те системы, которые каждую Z-задачу решают с меньшей (психической) энергией.

В разделе 5.5 был рассмотрен пример «Оценка течения заболевания (после инфаркта миокарда)». Для данного примера была построена минимальная синдромная модель знаний, содержащая 39 синдромов (полная модель содержит более 4000 синдромов). На рис. 7.11 приведено окно программы, с помощью которой создается и эксплуатируется среда радикалов, соответствующая данной модели знаний [144, 153].

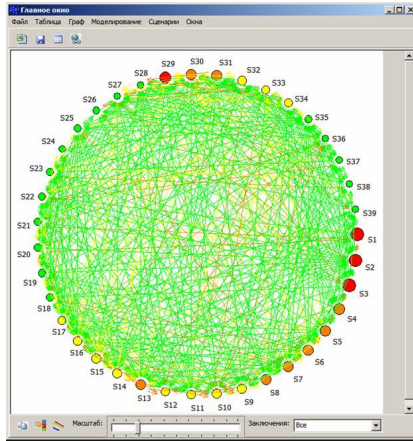


Рис. 7.11 – Среда радикалов для оценки течения заболевания

На рис. 7.12 показаны связи и характеристика отдельного радикала-синдрома.

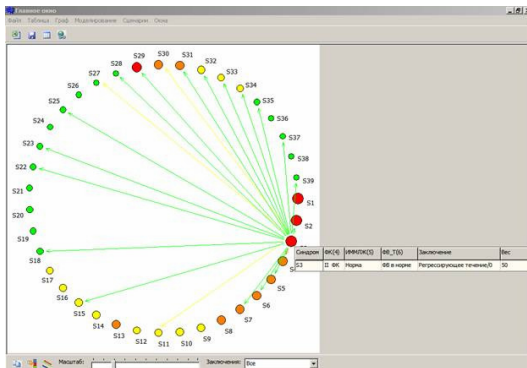


Рис. 7.12 – Связи и характеристика отдельного радикала

П. Баком была высказана гипотеза о критической природе функционирования мозга [193]. Можно предположить, что самоорганизация мозга в критическое состояние основывается на усилении связей с соседями тех нейронов (радикалов, ФС), которые часто активируются, с одной стороны, и «выводе» в рефлекторные дуги тех групп нейронов (радикалов, ФС), которые активируются слишком часто, с другой. Эта пара процессов на разных уровнях общности, наряду с самоорганизацией предельных моделей знаний, вполне укладывается в общую схему устройства критически самоорганизующихся систем.

В общем случае у любой ФС есть свой запускающий нейрон. В результате получаем ансамбль ФС с похожим принципом функционирования. Каждая ФС в ансамбле решает свою целевую  $Z$ -задачу. В итоге получаем иерархию ансамблей НМС с все возрастающим уровнем общности задач. Пример типичной иерархии функциональных систем показан на рис. 7.13.

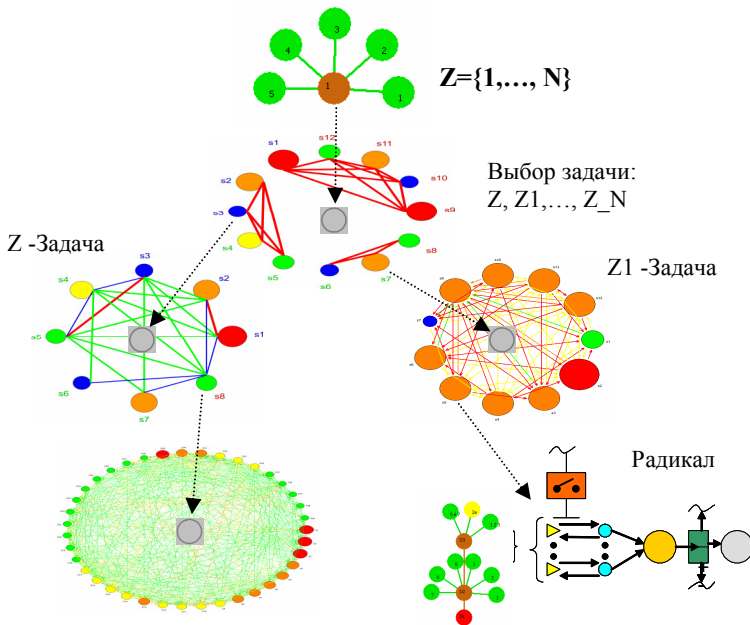


Рис. 7.13 – Иерархия функциональных систем (сред радикалов, агентов)

Первоначально решается вопрос о том, какую из задач  $\{Z, Z_1, \dots, Z_N\}$  решать, так как любая  $Z$ -задача порождает  $N$  дуальных задач (см. главу 5).

После выбора задачи запускается ФС, решающая данную задачу. Процесс детализации управления (принятия решения) самоподобно повторяется до тех пор, пока всеми запускаемыми ансамблями не станут радикалы (или биофизические агенты).

Представленная модель может объяснить механизм возникновения иерархии нейронных ансамблей, предназначенных для решения широкого круга задач. По-видимому, эволюция всех природных (когнитивных) систем идет в направлении увеличения иерархичности, позволяющей при переходе с уровня на уровень свертывать информацию [29].

Таким образом, искомая среда радикалов (функциональная система) строится автоматически на основе заданного множества ситуаций действительности (прецедентов) с известными заключениями (исходами), а также Банка элементарных тестов.

Среда радикалов может использоваться в двух режимах: 1) режим диагностики состояния ситуации, процесса; 2) режим управления.

В режиме диагностики каждый активный радикал осуществляет поиск значений необходимых тестов. Вначале значения тестов ищутся в рабочей зоне (общая витрина данных), если там нет значения какого-либо теста, то выводится (посылается по известным каналам связи) соответствующий запрос эксперту. Если радикал отработал, но заключение не найдено, то одновременно активизируются все радикалы, проводимость которых превышает некоторый порог. В общем случае имеются разные сценарии развития активности в среде радикалов. Сценарии различаются последовательностью применения некоторых критериев. На рис. 7.14 приведен пример диалогового окна для формирования разных сценариев.

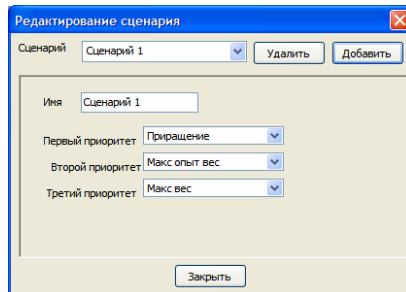


Рис. 7.14 – Пример формирования сценария активности СР

Использовались, в частности, следующие критерии: 1) минимум ввода новых тестов от пользователя (Приращение); 2) максимальный опытный вес синдрома (формируется в ходе обучения СР); 3) максимальный

априорный вес (Макс вес); 4) минимальная стоимость исследований; 5) минимальный дискомфорт (травматичность, болевые ощущения) для клиента; 6) минимальное время проведения исследований; 7) случайный выбор (по умолчанию). Каждый критерий последовательно уменьшает многозначность в выборе активности радикалов в рамках следующего системокванта. Некоторые радикалы могут быть исключены полностью. Среда радикалов может обучаться, первоначально активизируя наиболее результативные в прошлом радикалы (зависит от сценария активности).

Одновременно работающие радикалы образуют текущий системоквант. Один из примеров отображения системокванта – линейка радикалов на рис. 7.8. Поскольку процессы управления могут одновременно выполняться для разных систем (например, пациентов), то среды радикалов должны функционировать в многозадачном режиме. Каждая задача характеризуется своим текущим системоквантом. Пользователь (специалист) должен иметь возможность в диалоговом режиме просматривать и корректировать ход решения каждой задачи. На рис 7.15 показан пример создания новой задачи («Оценка и прогноз течения заболевания»), а на рис. 7.16 показан текущий системоквант для одной из задач (расшифровка цвета радикалов приведена при описании рис. 7.8).

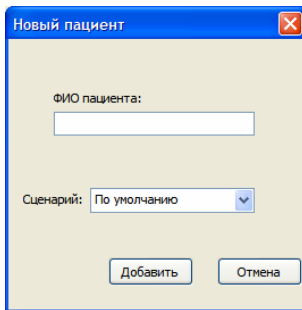


Рис. 7.15 – Создание задачи

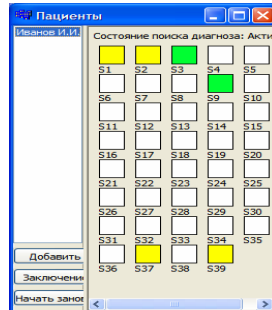


Рис. 7.16 – Системоквант задачи

В режиме управления задаются целевые синдромы и/или вероятностные закономерности, которые запускают *радикалы-цели*. Каждый радикал-цель активен до тех пор, пока активен соответствующий целевой синдром (закономерность). Радикалы-цели можно рассматривать как *центры активации «потребностей»* системы до момента удовлетворения потребностей. Рассогласование текущих значений тестов и целевых значений, которые активируются радикалами-целями, создает *отрицательные «эмоции»*. Последние являются движущей силой преобразований (целенаправленного поведения). Все остальные радикалы

работают в режиме диагностики состояния. Радикалы-цели создают, по сути, информационно-энергетические импульсы (глава 3).

Каждая НМС, реализующая синдром или закономерность, ведет себя, по сути, как интеллектуальный агент (биофизический агент). Особенно важна поисковая функция. Она может быть сколь угодно сложной. В целом НМС активно пытается найти нужную информацию, чтобы установить наличие синдрома (закономерности) и, соответственно, получить заключение. Отметим, что данная задача может решаться сколь угодно долго (до тех пор, пока не поступит сигнал «отбоя» или не будет найдено решение). На рис. 7.17 показана витрина данных в ИС «Оценка и прогноз течения заболевания». Пять верхних строк (выделены желтым цветом) – текущие запрашиваемые радикалами значения тестов. Нижние две строки (выделены серым цветом) – «зависшие» тесты, т.е. тесты, на которые пользователь (агент) до настоящего времени не дал ответа. Эти «зависшие» тесты определяют «зависшие» радикалы, которые активированы, но не могут завершить работу без нужных данных.

Тест	Значение	Требуют
T9(ФВ_5)	23 сыктене	S39
T12(Пок)		S37
T14(Тяж./напр тр)		S37,S38
T15(ГБ)		S38
T20(Работа/const)		S39
T4(ФК)	III ФК	S1
T5(ФМЛДК)	Ув. ФМЛДК	
T6(ФВ_Т)	ФВ сыктене	
T1(Гр. Инв)/const)		S1,S2
T2(МА)		S1,S2,S38

Выводить запрос автоматически

Рис. 7.17 – Пример витрины данных для CP

Наличие распределенного обмена управляющими воздействиями как между радикалами, так и между ФС обуславливает возможность управления «извне». Другими словами, сигналы управления любого уровня иерархии могут исходить от внешних источников, например других когнитивных систем (агентов). Точно так же с помощью поля возможна передача во внешнюю среду информации о работающей Z-задаче («передача мыслей»). Следовательно, предлагаемые модели позволяют объяснить ряд Psi-эффектов, например, таких как телепатия (передача мыслей на расстояние) или суггестия (внушение) [127]. Единственное требование:  $\lambda$  частоты соответствующих функциональных систем разных

агентов (когнитивных систем) должны быть синхронизированы, например в процессе взаимообучения. Часть  $\lambda$ -частот ФС, по-видимому, обусловлена эволюцией и поэтому является общей для данного вида когнитивных систем.

Теперь, после того как рассмотрена физическая реализация моделей знаний, можно сказать, что настоящее исследование содержит, в частности, описание *синтетической телесно-феноменологической онтологии*. Банк тестов и модели знаний отвечают за феноменологию, а иерархия функциональных систем (сред радикалов, агентов) отвечает за «телесность». Оставшаяся часть главы дополняет эту онтологию.

Синтетическая онтология является фундаментом синдромного принципа управления. Ее развитие является приоритетной метакогнитивной технологией.

На рис. 7.18 показан завершающий этап помощи врачу в диагностике и планировании реабилитации в рамках задачи «Оценка и прогноз течения заболевания» [159, 144].

Пациент: Петров П.П.

Заключение: Прогрессирующее течение(Неблагоприятное)

Синдромы которые дали заключение:

Синдромы	ИФ/ИПК	ФВ_С	Работа/const
S39	III ФК	ФВ снижена	Не работают

Исключенные синдромы:

Исключить:	Синдром	Гр. Инв/const	Инд. КДО	ФК	ИФ/ИПК	ФВ_Т	Вид тр	Инд ЛТ	ФВ_С	Q, не Q	ЖЗ	Лок	Образов/const	Тактич/тр	ПБ
<input type="checkbox"/>	S1	не инв	Нет	Норма											
<input type="checkbox"/>	S2	не инв	Нет		II ФК										
<input type="checkbox"/>	S3				II ФК	Норма			ФВ в норме						
<input type="checkbox"/>	S4				Норма				ФВ в норме	Фин./Унств.					
<input type="checkbox"/>	S5				II ФК					Норма				ФВ в норме	
<input type="checkbox"/>	S6				II ФК					Норма				ФВ в норме	Q
<input type="checkbox"/>	S7				II ФК					Норма				Q	Нет

Изменения в тестах принадлежат к Благоприятным исходам:  
ФВ\_С="ФВ снижена" -> "ФВ в норме"

Для достижения заключения: Регрессирующее течение  
Рекомендованы целиевые Синдром (S2): \*\*\*\*\*  
Убедиться в наличии или возможности коррекции след. значений тестов:  
Тест Гр. Инв/const="не инв"  
Тест МАС="Нет"  
Тест ФК="III ФК" -> "II ФК"

Закрыть

Рис. 7.18 – Пример реализации синдромного управления

В процессе диалога специалист (врач) может исключить недостижимые с его точки зрения синдромы. Партнерская система предлагает те целевые синдромы, которые менее затратны по ресурсам (оценки затрат берутся из базы знаний или выясняются экспертным путем).

Если наделить радикал способностью реализовывать произвольный системопаттерн  $f/\mu$ , то среда радикалов превратится в среду (когнитивных,

биофизических, интеллектуальных) агентов.

Синтетическая онтология с синдромным управлением описывают, по сути, возможный принцип работы (виртуального) *синергетического, квантово-семантического нейрокompьютера*.

В заключение данного раздела проведем некоторые аналогии развиваемого подхода с нейрофизиологическими представлениями [11, 25, 177]. Для большего сходства среду нейронов будем называть «мозгом».

Нейроны работают не как детекторы элементарных сенсорных или моторных признаков, а отражают в своей активности процесс формирования полноценных функциональных систем «организменного» уровня. Другими словами, нейроны обладают ярко выраженным свойством *системоспецифичности*. Общий принцип формирования каждой интеграции (на уровне системопаттерна или на уровне ФС) определяется как принцип избирательного взаимодействия всех ее составляющих и коллективной саморегуляции на получение «полезного для системы приспособительного результата», например решения определенной Z-задачи. Нейроны способны радикальным образом менять тип своей функциональной специализации в зависимости от Z-задачи.

Нейроны, системопаттерны и функциональные системы в целом обладают «фоновой» разрядной активностью. Феномен поведенческой специализации при этом возникает не как процесс генерации молчащими клетками (паттернами, ФС) избирательных фазных активаций, приуроченных к определенному этапу системопаттерна (потока паттернов), а в итоге фрагментации в такие периоды постоянно присутствующего «фоновых потока» энергетических разрядов (как переход от пачечно-групповой формы импульсации к регулярной и наоборот). «Фоновая» импульсная активность нейронов, паттернов и ФС является инструментом оперантного научения [25]. Вероятность стабильного воспроизведения соответствующей формы импульсации нейрона, системопаттерна, ФС в рамках разных интеграций существенно различна.

Активность элементов среды всех уровней дискретна, что позволяет рассматривать системоквант активности среды (активность в интервал времени). Любая активность среды на системном уровне представима в виде потока системопаттернов или на более высоком уровне обобщения – потока системоквантов. Базовый феномен системопаттерна и ФС – феномен QS-запутывания – проявляется в форме возникновения у большого числа дистантно расположенных нейронов характерных фазных активаций. Лежащие в разных структурах «мозга» нервные клетки, обладая выраженной индивидуальностью и «метаболической» изолированностью друг от друга, в определенные периоды времени начинают работать как единое целое.



Радикал является простейшим системным нейронным образованием (интеграцией), проявляющим целенаправленное поведение. Среда радикалов реализует *голографический принцип*: каждый элемент среды отражает в своей активности состояние ее конечного результата.

Среда радикалов (системопаттернов) демонстрирует огромное разнообразие активностей при решении одной и той же задачи (высокую пластичность). В результате степень сходства составов интеграций (избирательно активирующихся нейронных ансамблей) последовательно развертывающихся стереотипных когнитивно-поведенческих актов (системопаттернов) не равна единице. Это совпадает с принципом вариативности функциональных систем, который является одним из основных положений теории П.К. Анохина [11].

В силу огромной избыточности имеет место высокая стабильность ФС в ситуации свободного выбора, в сочетании со способностью к быстрой реорганизации в условиях любого воздействия, затрудняющего достижение результата.

Учитывая наличие в среде радикалов директивной зоны и зоны возможностей, можно говорить о принадлежности нейрона к стабильной или нестабильной составляющей функциональной системы. Кроме того, существуют критические пути (эффективные последовательности системоквантов), что также накладывает отпечаток на характер активности нейронов. По мере формирования критических путей в рамках ФС отмечается постепенное снижение степени выраженности поведенческой специализации нейронов, не входящих в критический путь или радикалы директивной зоны. Таким образом, наряду с относительно стабильным «ядром», в состав любой ФС входит и значительная часть нейронов, образующих своего рода лабильную периферию.

Связки «синдром – радикал», «закономерность – радикал», «системопаттерн – радикал», «модель знаний – среда радикалов» характеризуют когнитивно-поведенческий уровень. Процесс эстафетного распространения в среде нейронов волны возбуждения от афферентного ее звена в сторону исполнительных (моторных) структур является лишь одним из способов организации нейронных систем когнитивно-поведенческого уровня («сверху вниз»). Существует и прямо противоположный способ организации нейронных систем («снизу вверх») как результат самоорганизации и QS-запутывания поведенческих элементов (эффект памяти), например первичных системопаттернов (так называемое «резонансное возбуждение», возникающее в виде когерентного согласованного поведения многокомпонентных систем, с последующим положительным подкреплением). Сформировавшиеся навыки (автоматизмы поведенческого уровня) приводят к появлению

системопаттернов уже когнитивно-поведенческого уровня. В этом месте происходит взаимодействие *трех парадигм: кибернетической, синергетической и квантовой (квантово-семантической)*.

Имеются определенные аналогии с так называемой пейсмекерной теорией организации мозговой деятельности [176]. В основе ее лежит гипотеза о существовании в центральных структурах головного мозга человека и животных особых, так называемых «мотивационных» центров, клетки которых способны при возникновении внутренних потребностей организма оказывать генерализованные возбуждающие влияния на нейроны других областей ЦНС. Действительно, как радикалы, так и ФС в целом запускаются системоспецифичными нейронами (АД-нейронами), которые, вероятно, можно отнести к типу «пейсмекеров». Учитывая иерархию ФС, можно сделать вывод о том, что структуры «мозга» выстраиваются по определенному градиенту возбудимости.

На когнитивном уровне возможно многократное «проигрывание» какой-либо целенаправленной деятельности, не доводя дело до развертывания исполнительных механизмов, до попытки получения результата. Данная возможность будет проиллюстрирована ниже при рассмотрении Многоцелевого банка знаний.

### 7.3 Задачи познания ПрО. Поиск закономерностей

Задача познания предметной области сводится к задаче усиления логической эмпирической теории за счет обнаружения аксиом на множестве эмпирических данных. В практическом плане задача состоит в автоматическом выделении семантических образов произвольных причинно-связанных процессов из слабоструктурированных информационных потоков, отображающих реальные физические и когнитивные процессы.

Причинно-следственные связи, в частности, позволяют:

- определять условия и время попадания траекторий процессов в поглощающее или инвариантное множество;
- строить информационные множества;
- предсказывать катастрофы и бифуркации процессов (например, инфаркты, инсульты и т.д.);
- выявлять параметры порядка при анализе путей наиболее безопасного выхода из критического состояния после катастрофы;
- проводить анализ асимптотического поведения системы;
- определять дрейф структурных параметров системы.

Определим базовые структуры закономерностей (по аналогии с [12]).

Утверждение «в ситуации действительности  $\alpha$  вслед за множеством

событий  $\{c\}$  на временном интервале  $\delta/\Lambda$  наблюдается событие  $e$ » запишем следующим образом:  $J_{+1} \text{Exp}(\alpha, \{c\}, \delta/\Lambda, e)$ .

Утверждение «в ситуации действительности  $\alpha$  вслед за множеством событий  $\{c\}$  на временном интервале  $\delta/\Lambda$  не наблюдается событие  $e$ » запишем следующим образом:  $J_{-1} \text{Exp}(\alpha, \{c\}, \delta/\Lambda, e)$ .

Введем  $J$ -атомарные формулы с предикатным символом РС (возможная причина):

$J_{+1} \text{PC}(\{c\}, \delta/\Lambda, \{\underline{d}/T\}, e)$  будем интерпретировать как « $\{c\}$  является возможной причиной наступления  $e$  на интервале  $\delta/\Lambda$  при условии выполнения  $\{\underline{d}/T\}$ » (+Правило для РС).

$J_{\vee} \text{PC}(\{c\}, \delta/\Lambda, \{\underline{d}/T\}, e)$  будем интерпретировать как «Хотя бы одно из событий  $\{c\}$  является возможной причиной наступления  $e$  на интервале  $\delta/\Lambda$  при условии выполнения  $\{\underline{d}/T\}$ » ( $\vee$ Правило для РС).

$J_L \text{PC}(L(\{c\}), e)$  будем интерпретировать как «Факт  $L(\{c\})$  является возможной причиной наступления  $e$ » ( $L$ Правило для РС).

$J_{-1} \text{PC}(\{c\}, \delta/\Lambda, \{\underline{d}/T\}, e)$  будем интерпретировать как « $\{c\}$  является возможной причиной ненаступления  $e$  на интервале  $\delta/\Lambda$  при условии выполнения  $\{\underline{d}/T\}$ » (-Правило для РС).

$J_0 \text{PC}(\{c\}, \delta/\Lambda, \{\underline{d}/T\}, e)$  будем интерпретировать как «сведения о причинно-следственных отношениях между  $\{c\}$  и  $e$  на интервале  $\delta/\Lambda$  при условии выполнения  $\{\underline{d}/T\}$  противоречивы» (0Правило для РС).

$J_{?} \text{PC}(\{c\}, \delta/\Lambda, \{\underline{d}/T\}, e)$  будем интерпретировать как «сведений о причинно-следственных отношениях между  $\{c\}$  и  $e$  на интервале  $\delta/\Lambda$  при условии выполнения  $\{\underline{d}/T\}$  недостаточно» (?Правило для РС).

Условия  $\{\underline{d}/T\}$  действуют на множестве ситуаций действительности  $\Omega$ . Условия  $\{\underline{d}/T\}$  могут быть фильтром как ситуаций  $\alpha$ , так и событий  $\{c\}$ . Условия  $\{\underline{d}/T\}$  могут отсутствовать (в таком случае получим безусловные правила).

Пример фильтра только для ситуаций действительности (из медицинской практики):  $\{\underline{d}/T\} = \{\text{Пол?Ж; Ds? Бронхит; Возраст/В2? 40-55}\}$ . Пример фильтра только для событий:  $\{\underline{d}/T\} = \{\text{Тип тестов в событиях } \{c\}? \text{ Биохимия крови}\}$ .

Тот факт, что ситуация действительности  $\alpha$  удовлетворяет условиям  $\{\underline{d}/T\}$ , будем записывать с помощью нотации:  $\alpha \nabla \{\underline{d}/T\}$ . Аналогично будем записывать сведения о том, что события  $\{c\}$  удовлетворяют условиям  $\{\underline{d}/T\}$ :  $\{c\} \nabla \{\underline{d}/T\}$

Дадим неформальное описание «+Правил» для РС.

Пусть

- существует хотя бы одна ситуация действительности, удовлетворяющая  $\{\underline{d}/T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими

$\{\underline{d}T\}$ , на интервале  $\delta/A$  наблюдается событие  $e$ ;  
 - не существует ни одной ситуации действительности, удовлетворяющей  $\{\underline{d}T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{d}T\}$ , на интервале  $\delta/A$  не наблюдается событие  $e$ .

Тогда события  $\{c\}$  являются возможной причиной наступления события  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{d}T\}$ .

Формальная запись «+Правила» будет иметь следующий вид:

$$\exists \alpha \forall \{\underline{d}T\} (J_{+1} \text{Exp} (\alpha, \{c\} \nabla \{\underline{d}T\}, \delta/A, e)) \& \neg \exists \beta \nabla \{\underline{d}T\} (J_{-1} \text{Exp} (\beta, \{c\} \nabla \{\underline{d}T\}, \delta/A, e)) \rightarrow J_{+1} \text{PC} (\{c\}, \delta/A, \{\underline{d}T\}, e).$$

Дадим неформальное описание «vПравила» для РС.

Пусть

- существует хотя бы одна ситуация действительности, удовлетворяющая  $\{\underline{d}T\}$ , где вслед за каким либо событием из  $\{c\}$ , удовлетворяющим  $\{\underline{d}T\}$ , на интервале  $\delta/A$  наблюдается событие  $e$ ;  
 - не существует ни одной ситуации действительности, удовлетворяющей  $\{\underline{d}T\}$ , где вслед за каким либо событием из  $\{c\}$ , удовлетворяющим  $\{\underline{d}T\}$ , на интервале  $\delta/A$  не наблюдается событие  $e$ .

Тогда множество событий  $\{c\}$  содержит возможные причины наступления события  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{d}T\}$ .

Формальная запись «vПравила» будет иметь следующий вид:

$$\exists \alpha \nabla \{\underline{d}T\} \exists c' \in \{c\} \nabla \{\underline{d}T\} (J_{+1} \text{Exp} (\alpha, c', \delta/A, e)) \& \neg \exists \beta \nabla \{\underline{d}T\} \forall c' \in \{c\} \nabla \{\underline{d}T\} (J_{-1} \text{Exp} (\beta, c', \delta/A, e)) \rightarrow J_v \text{PC} (\{c\}, \delta/A, \{\underline{d}T\}, e).$$

Неформальное описание «-Правила» для РС.

Пусть

- существует хотя бы одна ситуация действительности, удовлетворяющая  $\{\underline{d}T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{d}T\}$ , на интервале  $\delta/A$  не наблюдается событие  $e$ ;  
 - не существует ни одной ситуации действительности, удовлетворяющей  $\{\underline{d}T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{d}T\}$ , на интервале  $\delta/A$  наблюдается событие  $e$ .

Тогда события  $\{c\}$  являются возможной причиной *ненаступления* события  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{d}T\}$ .

Формальная запись «-Правила» будет иметь следующий вид:

$$\exists \alpha \nabla \{\underline{d}T\} (J_{-1} \text{Exp} (\alpha, \{c\} \nabla \{\underline{d}T\}, \delta/A, e)) \& \neg \exists \beta \nabla \{\underline{d}T\} (J_{+1} \text{Exp} (\beta, \{c\} \nabla \{\underline{d}T\}, \delta/A, e)) \rightarrow J_{-1} \text{PC} (\{c\}, \delta/A, \{\underline{d}T\}, e).$$

Неформальное описание «0Правила» для РС.

Пусть

- существует хотя бы одна ситуация действительности, удовлетворяющая  $\{\underline{t}/T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{t}/T\}$ , на интервале  $\delta/A$  не наблюдается событие  $e$ ;

- существует хотя бы одна ситуация действительности, удовлетворяющая  $\{\underline{t}/T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{t}/T\}$ , на интервале  $\delta/A$  наблюдается событие  $e$ .

Тогда сведения о причинно-следственных отношениях между  $\{c\}$  и  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{t}/T\}$  противоречивы.

Формальная запись «0Правила» будет иметь следующий вид:

$$\exists \alpha \nabla \{\underline{t}/T\} (J_{-1} \text{Exp} (\alpha, \{c\} \nabla \{\underline{t}/T\}, \delta/A, e)) \& \exists \beta \nabla \{\underline{t}/T\} (J_{+1} \text{Exp} (\beta, \{c\} \nabla \{\underline{t}/T\}, \delta/A, e)) \rightarrow J_0 \text{PC} (\{c\}, \delta/A, \{\underline{t}/T\}, e).$$

Рассмотрим теперь неформальное описание «?Правила» для РС.

Пусть

- не существует ни одной ситуации действительности, удовлетворяющей  $\{\underline{t}/T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{t}/T\}$ , на интервале  $\delta/A$  наблюдается событие  $e$ ;

- не существует ни одной ситуации действительности, удовлетворяющей  $\{\underline{t}/T\}$ , где вслед за событиями  $\{c\}$ , удовлетворяющими  $\{\underline{t}/T\}$ , на интервале  $\delta/A$  не наблюдается событие  $e$ .

Тогда сведений о причинно-следственных отношениях между  $\{c\}$  и  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{t}/T\}$  недостаточно.

Формальная запись «?Правила» будет иметь следующий вид:

$$\neg \exists \alpha \nabla \{\underline{t}/T\} (J_{-1} \text{Exp} (\alpha, \{c\} \nabla \{\underline{t}/T\}, \delta/A, e)) \& \neg \exists \beta \nabla \{\underline{t}/T\} (J_{+1} \text{Exp} (\beta, \{c\} \nabla \{\underline{t}/T\}, \delta/A, e)) \rightarrow J_? \text{PC} (\{c\}, \delta/A, \{\underline{t}/T\}, e).$$

Все операции необходимо рассматривать с учетом модели описания действительности.

Рассмотрим пример с безусловными правилами. Пусть множество прецедентов  $\Omega$  задано таблицей 7.1.

Таблица 7.1 - Множество прецедентов  $\Omega = \{\alpha_1, \alpha_2\}$

№ $\alpha$	Код $e$	Событие $e$
1	$c_1$	<АДс/D1? 310, $\underline{t}/A1$ ? 12.45>
1	$c_2$	<ОНМК/{Имеется; Отс}? Имеется, $\underline{t}/A1$ ? 15.10>
2	$c_3$	<АДс/D1? 270, $\underline{t}/A1$ ? 08.20>
2	$c_4$	<Гол.боль/{Имеется; Отс}? Имеется, $\underline{t}/A1$ ? 08.20>
2	$c_5$	<ОНМК/{Имеется; Отс}? Имеется, $\underline{t}/A1$ ? 18.00>

#АДс – АД систолическое; ОНМК – острое нарушение мозгового кровообращения.

Введем для АДс домен  $D3 = \{\text{Предельно низкое; Низкое; Норма; Высокое; Предельно высокое}\}$  (Домен  $D2$  содержит интервальные элементы). Сформируем оргграф доменов:  $D1 \rightarrow D2 \rightarrow D3$ . С помощью встроенных функций вычислительной среды все события, в том числе события  $c_1$  и  $c_3$ , автоматически преобразуются в события более высокого уровня общности. В частности, будут сгенерированы события:

$$c_1(D3, A1) = \langle \text{АДс/}D3? \text{ Пред.высокое, } \underline{t}/A1? 12.45 \rangle;$$

$$c_3(D3, A1) = \langle \text{АДс/}D3? \text{ Пред.высокое, } \underline{t}/A1? 08.20 \rangle.$$

Также автоматизмами вычислительной среды из любого события  $\langle \underline{t}/T, \underline{t}/A \rangle$  формируется событие-факт «Имело место  $\underline{t}/T$ ». Получим, в частности, такие события-факты:

$$c_1(D3, A1) \rightarrow ef_1(D3) = \langle \text{Имело место АДс/}D3? \text{ Пред.высокое} \rangle;$$

$$c_3(D3, A1) \rightarrow ef_3(D3) = \langle \text{Имело место АДс/}D3? \text{ Пред.высокое} \rangle;$$

$$c_2 \rightarrow ef_2 = \langle \text{Имело место ОНМК} \rangle; c_5 \rightarrow ef_5 = \langle \text{Имело место ОНМК} \rangle.$$

Очевидно, события-факты  $ef_1(D3)$  и  $ef_3(D3)$ , а также  $ef_2$  и  $ef_5$  эквивалентны между собой. Введем следующие обозначения:

$$ef_1(D3) = ef_3(D3) = e_{AD}; \quad ef_2 = ef_5 = e_{ОНМК}.$$

Если задать  $\delta_1 = 4$  часа, тогда справедливы утверждения:

$$J_{+1} \text{Exp}(1, c_1, \delta_1, c_2); J_{-1} \text{Exp}(2, c_3, \delta_1, c_5); \quad J_{+1} \text{Exp}(1, e_{AD}, \delta_1, e_{ОНМК});$$

$$J_{-1} \text{Exp}(2, e_{AD}, \delta_1, e_{ОНМК})$$

Согласно 0Правилу будем иметь:  $J_0PC(e_{AD}, \delta_1, e_{ОНМК})$ , т.е. данные противоречивы.

Зададим  $\delta_2 = 12$  часов, тогда справедливы утверждения:

$$J_{+1} \text{Exp}(1, e_{AD}, \delta_2, e_{ОНМК}) \text{ и } J_{+1} \text{Exp}(2, e_{AD}, \delta_2, e_{ОНМК}).$$

Согласно +Правилу будем иметь:  $J_{+1} PC(e_{AD}, \delta_2, e_{ОНМК})$ , т.е. событие-факт  $e_{AD}$  является возможной причиной события-факта  $e_{ОНМК}$  на интервале  $\delta = 12$  часов (в рамках заданного множества прецедентов  $\Omega = \{\alpha_1, \alpha_2\}$ ).

Аналогично может быть сделан вывод, что событие-факт  $e_{ГолБоль}$  является возможной причиной события-факта  $e_{ОНМК}$  на интервале  $\delta = 12$  часов, т.е.  $J_{+1} PC(e_{ГолБоль}, \delta_2, e_{ОНМК})$ .

Заметим, что события-факты

$$c_1(D1, A1) \rightarrow ef_1(D1) = \langle \text{Имело место АДс/}D1? 310 \rangle;$$

$$c_3(D1, A1) \rightarrow ef_3(D1) = \langle \text{Имело место АДс/}D1? 270 \rangle,$$

которые непосредственно вытекают из эмпирических данных, формально не являются эквивалентными и только переход на более высокий уровень общности позволяет сделать определенные выводы о причинной связи.

Поскольку связка событий  $\{e_{АД}, e_{ГолБольш}\}$  также является возможной причиной события-факта  $e_{ОНМК}$  на интервале  $\delta = 12$  часов, т.е.  $J_{+1}PC(\{e_{АД}, e_{ГолБольш}\}, \delta_2, e_{ОНМК})$ , то целесообразно рассмотреть задачу нахождения максимального множества  $\{c\}$ , которое является возможной причиной события  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{t}/T\}$ . Для такого предельного множества введем обозначение  $\{c\}_{\delta e, \{\underline{t}/T\}}$ . Формально данная задача запишется следующим образом:

$$\forall \delta/A, \{\underline{t}/T\}, e \quad \{c\}_{\delta e, \{\underline{t}/T\}} = \max_{\Omega} \{\{c\} \mid J_{+1}PC(\{c\}, \delta/A, \{\underline{t}/T\}, e)\}. \quad (7.18)$$

Необходимо отметить, что состав предельного множества формируется с учетом всех допустимых обобщений тестов или, другими словами, наличия автоматизмов вычислительной среды. Возможен следующий результат оптимизации:  $\{c\}_{\delta e, \{\underline{t}/T\}} = \emptyset$ . Для рассматриваемого примера

$$\{c\}_{\delta=12, ОНМК} = \{e_{АД}, e_{ГолБольш}\}.$$

Другая экстремальная задача состоит в нахождении минимального периода  $\delta/A$ , на котором множество событий  $\{c\}$  является возможной причиной  $e$  при условии выполнения  $\{\underline{t}/T\}$ . Формально данная задача запишется следующим образом:

$$\forall \{\underline{t}/T\}, \{c\} \nabla \{\underline{t}/T\}, e \quad \delta/A_{\{c\}, e, \{\underline{t}/T\}} = \min_{\Omega} \{\delta/A \mid J_{+1}PC(\{c\}, \delta/A, \{\underline{t}/T\}, e)\} \quad (7.19)$$

Для рассматриваемого примера получим:

$$\delta/\{c\} = \{e_{АД}, e_{ГолБольш}\}, ОНМК) = 9.40.$$

Рассмотрим задачу нахождения неизбыточного множества  $\{c\}$ , содержащего все возможные причины события  $e$  на интервале  $\delta/A$  при условии выполнения  $\{\underline{t}/T\}$ . Для такого предельного множества введем обозначение  $\{c\}_{\delta e, \{\underline{t}/T\}}^*$ . Формально данная задача запишется следующим образом:

$$\forall \delta/A, \{\underline{t}/T\}, e \quad \{c\}_{\delta e, \{\underline{t}/T\}}^* = \min_{\Omega} \{\{c\} \mid J_{\vee}PC(\{c\}, \delta/A, \{\underline{t}/T\}, e)\}. \quad (7.20)$$

Информационное множество, введенное в главе 3, позволяет выявить вероятностные закономерности, а на их основе аномальные события, например нарушения динамики ряда.

#### 7.4 Когнитивные и метакогнитивные информационные системы

Под *когнитивными информационными системами* (КИС) будем понимать приложения, удовлетворяющие трем условиям:

- ядром приложения является Многоцелевой банк знаний (МБкЗ);
- в рамках приложения функционируют иерархические среды

радикалов, колонии виртуальных роботов и/или мультиагентные системы, а точнее – консилумы *когнитивных агентов*;

- реализуются ключевые когнитивные и метакогнитивные феномены.

КИС ориентированы на помощь человеку в постановке задач, на решение плохо формализованных профессиональных и творческих задач, на выявление и эффективное использование своего когнитивного потенциала. Одна из основных задач КИС – это обеспечение (повышение) *функциональной устойчивости* профессиональной деятельности специалиста. КИС относятся к классу истинно партнерских систем.

Ясно, что состав обязательных к реализации когнитивных и метакогнитивных феноменов будет со временем расширяться, но такие два феномена как «созревание» и «дозревание» решений и инструментов принятия решений являются основополагающими. Важным феноменом является «интуиция». В целом приоритетными являются метакогнитивные технологии, так как они направлены на совершенствование механизмов принятия решений и управления.

Колонии виртуальных роботов могут имитировать функционирование иммунной системы, пчелиный рой или муравьиную колонию (например, в задачах оптимизации или обеспечения безопасности). Среды радикалов, колонии роботов, агентные системы характеризуют разные уровни когнитивной организации (распределенный интеллект): агенты управляют средами радикалов и колониями роботов, роботы могут управлять средами радикалов (функциональными системами). Примем, что агенты могут испытывать «эмоции», которые зависят от степени удовлетворения их потребностей (примитивными эмоциями могут обладать радикалы, среды радикалов, роботы и колонии роботов). Следовательно, КИС может обладать довольно развитой «эмоциональной сферой».

Основой КИС является синтетическая телесно-феноменологическая онтология, а партнерские функции КИС реализуются, в частности, с использованием синдромного принципа управления. Именно поэтому вопросы разработки КИС входят в проблематику принципа предельных обобщений и отражают один из важнейших прикладных аспектов данного принципа.

Сферы применения КИС:

- критические технологии (критические информационные системы и критические системы управления), включая технологии «экстренных вычислений» (Urgent Computing) [30] (медицина, космос, транспорт, энергетика, ЧС, военные технологии, системы безопасности и т.д.);

- когнитивные технологии (развивающего) обучения и когнитивные тренажеры;

- приложения Business Intelligence;



- автономные адаптивные (мобильные) системы [61], роботы; сенсорные поля;
  - интеллектуальный GRID [114]; ситуационные и когнитивные центры [100]; виртуальные организации и кафедры; виртуальные миры [17];
  - социальные сети и экспертные сообщества [164];
  - интеллектуальные персональные системы (секретарь, консультант, «учитель») [5];
  - системы искусственного, распределенного или гибридного интеллекта.
- Пример специализированной медицинской КИС «Оценка и прогноз течения заболевания» был рассмотрен выше.

*Глобальная КИС* зародилась и развивается на основе Интернета (в нынешней стадии – проект «Всемирный мозг» [165]). Одна из главных целей глобальной КИС – максимальное удовлетворение информационных потребностей пользователей.

К недостаткам КИС можно отнести большую сложность разработки и сопровождения. Однако выраженный положительный эффект в большинстве случаев оправдывает затраты.

Ниже рассматриваются различные аспекты построения и функционирования КИС, которые связаны с ППО.

#### *7.4.1 Особенности поддержки принятия решений в рамках КИС*

Среди всех текущих моделей знаний  $Z$ -задачи может не быть корректной, т.е. верной для всех ситуаций действительности  $\text{PrO}$ , а не только для ситуаций из  $\Omega$ . Например, всегда, когда модель знаний содержит непрерывный базовый домен какого-либо теста, она не является корректной (это наиболее частая причина).

Если корректная модель знаний не построена, то это означает, что качественное решение целевой  $Z$ -задачи в текущий момент времени не найдено. Активный поиск решения откладывается до «лучших времен», но сама задача не пропадает, она переходит как бы в фоновый режим, постоянно отслеживая изменения онтологии  $\{G(\tau)\}$ .

По прошествии некоторого времени и решения новых целевых задач может измениться онтология, т.е. фрагменты Банка тестов  $\{G(\tau)\}$ , в частности, могут добавиться новые вершины в оргграфы доменов тестов. При любом изменении  $\{G(\tau)\}$  автоматизмами КИС (в фоновом режиме) предпринимается попытка заново решить все нерешенные ранее задачи, т.е. для каждой из них строятся новые (предельные) модели знаний. Для какой-то из  $Z$ -задач очередная попытка может оказаться успешной, т.е. будет найдено качественное решение (предположительно – корректная модель знаний). Это можно интерпретировать как «озарение». Фоновый процесс

поиска качественного решения можно интерпретировать как один из вариантов «созревания» решения. Поскольку решение означает обрушение старых моделей знаний и возникновение новых, то «озарение» можно интерпретировать как когнитивную катастрофу. Появление лишь одной новой вершины в орграфе доменов какого-либо теста (одного из сотен или тысяч тестов) способно привести к «озарению».

В прикладном плане разработана архитектура и реализован прототип Многоцелевого банка знаний (на примере клинической медицины), который поддерживает встроенную процедуру постоянного пересмотра всех решаемых ранее задач, приводя к периодическому обновлению синдромных и вероятностных моделей знаний (задачи диагностики, прогнозирования, оптимизации лечения) [203]. Факт целесообразности обновления той или иной модели знаний сообщается куратору задачи, а изменения вступают в силу только после его разрешения. Показано, что в процессе эксплуатации системы частота обновлений моделей знаний («озарений») со временем немонотонно, но уменьшается. Функции базы прецедентов выполняет госпитальная информационная система [4]. Схожая технология реализована и в рамках специализированной КИС «Оценка и прогноз течения заболевания» (рассмотрена выше).

Полномасштабная реализация процессов созревания когнитивных структур в активном (инициируется агентами) и фоновом режимах будет означать, что КИС обладает собственным целенаправленным поведением. Главной целью КИС является нахождение качественных решений, в частности моделей знаний (возможно, в диалоге с пользователем).

Выделим ключевые аспекты и феноменологические операторы получения качественных решений (управлений) в рамках КИС, а также наметим пути их реализации.

Любое решение или закон управления имеет несколько уровней общности, а именно:

1. Осознание или осмысление системной парадигмы управления или исследования.
2. Построение системного решения на основе принятой парадигмы.
3. Развертывание системного решения в частные прикладные решения (образы, портреты) такие как: статистический портрет, информационный образ, граф причинно-следственных связей, модели подпроцессов и подсистем и т.д.

Синдромный принцип управления является примером парадигмы управления, на базе которой формируется системное решение – целевые синдромы и предвестники. Далее системное решение самоподобным образом разворачивается в частные решения.

С позиции наблюдателя любой реальный процесс выработки решения

протекает, как правило, путем чередования активной и пассивной фаз и редко ограничивается одной фазой. В активной фазе процесс является циклическим с контролем времени выполнения этапов. В пассивной фазе процесс протекает «подсознательно» с непредсказуемым временем окончания. Таким образом, в самом общем виде один макроцикл выработки решения можно феноменологически отобразить комбинацией четырех операторов - оператора *расширения*  $\lambda$ , оператора *локализации решения*  $\vartheta$ , оператора *дозревания решения и инструментов*  $\chi$  и оператора *оценки истинности информации*  $\gamma$ , а именно:

$$X^* = (\vartheta(\lambda(X)) \vee \chi(\lambda(X))) \equiv (\vartheta \vee \chi)(\lambda(X)), \quad (7.21)$$

при условии  $E(\lambda) > E_I^*$ ,  $\chi(\lambda(X)) = \text{истина}$ ,

где  $X$  – первичная (априорная) информация;  $E$  – энергия - движущая сила когнитивного процесса;  $I$  – цель;  $X^*$  – результирующее решение. В качестве решения может выступать теория, образ (например, художественный), системопаттерн (например, математическая модель), функциональная система (например, для решения Z-задачи) и т.д.

С помощью композиции операторов  $\lambda$  и  $\vartheta$  реализуется активная фаза, а с помощью оператора дозревания  $\chi$  реализуется пассивная фаза (хотя процесс и в этом случае запускается активным способом – с помощью  $\lambda$ ). Связка “ $\vee$ ” – это «неисключающее ИЛИ», что подразумевает возможность параллельного протекания активных и пассивных процессов.

Задачей оператора расширения  $\lambda$  является получение максимума достоверной информации (с точки зрения цели  $I$  и с учетом оператора  $\gamma$ ) о прошлом, настоящем и будущем поведении системы, используя доступные ресурсы и априорную информацию  $X$ . В результате действия оператора расширения происходит резкое возрастание информативности всей системы, зачастую сравнимое с «информационным взрывом». Локализация  $\vartheta$  реализует феномен семантического синтеза (частично феномен локализации был рассмотрен в разделе «Квантовая семантика», глава 6).

Для успешной реализации  $\lambda$  необходимо создать *критическое усилие*. Данное требование выражается условием:  $E(\lambda) > E_I^*$ . Если данное условие не выполняется, то качественное решение (или просто решение) вряд ли может быть получено. Решение задачи прекращается, как только исчезает движущая сила - энергия. Параметр  $E_I^*$  существенно зависит от физики системы, осуществляющей выработку решения, в частности от наличия у системы «врожденных» и/или приобретенных автоматизмов (синергий), связанных с получением и преобразованием информации. Чем больше автоматизмов задействовано в процессе выработки решения, тем меньше требуется энергии (меньше значение параметра  $E_I^*$ ), следовательно, тем

больше шансов получить качественное решение. В первом приближении справедлива следующая оценка:  $|\lambda(X)| \sim E(\lambda)$ .

Главной особенностью оператора  $\chi$  является созревание не только решения, но также инструментов (meta-cognition), т.е. всех операторов  $\lambda, \vartheta, \chi, \gamma$  и т.д. Если в процессе перманентного созревания возник новый инструмент или решение, то это означает бифуркацию процесса – относительно плавную или скачкообразную смену технологии выработки решения и пересмотр (переосмысление) ранее принятых решений. Оператор  $\chi$  объединяет все когда-либо решавшиеся задачи, и «озарения» могут происходить для «давно забытых» задач.

В терминах модели ситуации действительности макроцикл (7.21) может быть представлен следующим образом:

$$\{\underline{u}/U\}^* = (\vartheta \vee \chi) (\lambda(\{\underline{u}/T\})), \quad (7.22)$$

при условии:  $\{\underline{e}/E\}_\lambda > \{\underline{e}/E\}_T^*$ ;  $\gamma(\lambda(\{\underline{u}/T\})) = \text{истина}$ ,

где  $\{\underline{e}/E\}_\lambda$  – параметры, описывающие критическое усилие.

Рассмотрим основные этапы, способы формирования и возбуждения внутренних сил взаимодействия, которые могли бы породить в семиотическом пространстве КИС устойчивые диссипативные структуры, приводящие к решению. Предварительно сделаем три предположения относительно видов, источников и способов переноса энергии:

- энергия бывает *целевой* и *фоновой*; целевую энергию генерирует агент-диспетчер задач КИС (когнитивный агент), однако большая часть процессов подпитывается фоновой энергией и не контролируется диспетчером; фоновая энергия формируется, во многом, из остатков целевой активности (энергии) в прошлом;

- фоновые информационные процессы протекают всегда, даже при отсутствии текущей целевой энергии; в результате протекания фоновых и целевых процессов непрерывно активируются/деактивируются или формируются новые динамические паттерны (системопаттерны  $f/\mu$ ); большая часть новых паттернов относится к разряду «короткоживущих» (неустойчивых); «долгоживущие» паттерны имеют шансы стать автоматизмами; паттерны передают энергию друг другу, формируя целенаправленные потоки паттернов;

- паттерны формируют орграфы набросков всех типов (реализуют механизм *созревания* орграфов), обеспечивают циркуляцию информации в рамках орграфов набросков (восходящую и нисходящую), а также квантово-семантическое запутывание разных орграфов; на макроуровне паттерны обеспечивают *созревание* и функционирование, в частности, иерархии сред радикалов/агентов (умений КИС).

Основные этапы выработки решения следующие [154]:

(А) *Фиксация цели* – Z-задачи и критериев достижения цели  $\{f/\mu\}_0$ . Генерация агентом-диспетчером целевой энергии, создающей облако притяжения (активности) в окрестности Z-задачи (это могут быть радикалы-цели). Факт наличия целевой энергии означает, что данная Z-задача находится в *фокусе внимания* диспетчера КИС (активная фаза).

(В) *Загрузка (дозагрузка)* информации извне, повышение концентрации потоков информации в окрестности решаемой задачи. Требуется обеспечить «критическую массу» информации в локальной информационной окрестности задачи и «критическую энергию загрузки».

(С) *Интенсивное перемешивание* информации (организация максимального взаимодействия динамических паттернов между собой, QS-запутывание орграфов набросков) до момента возникновения самоорганизации (когерентности). Требуется обеспечить «критическую энергию перемешивания».

(D) *Рефлексия* на предмет проверки условия когерентности (возникающее свойство самодостаточности). Если когерентности нет, то возврат к этапам (В) и (С), т.е. дозагрузка информации и увеличение энергии перемешивания, если это возможно. Если дополнительная целевая энергия отсутствует или уменьшается, то Выход: решение в текущий момент не найдено. Потеря фокуса внимания агента-диспетчера. Если целевая энергия превышала некоторый порог, то следы задачи остаются в фоновом семиотическом пространстве КИС.

(Е) *Созревание* – самоорганизация («воронка - фильтр»): модификация или формирование новых орграфов набросков, структур и системопаттернов на основе имеющейся информации и имеющейся целевой энергии (или даже значительно меньшей целевой энергии, так как часть процессов может питаться фоновой энергией). Решениями являются некоторые орграфы набросков, категориальные структуры и/или системопаттерны вида  $\{f/\mu\}$ . Созревание – самый продолжительный этап и может длиться часы, дни, месяцы и даже годы.

(F) *Рефлексия* на предмет удовлетворения решения критериям  $\{f/\mu\}_0$ . Если получено частичное решение, то возврат к пункту (В).

(G) *Локализация* окончательного решения: выбор подмножества решений максимального уровня общности.

(H) *Рефлексия* на предмет пересмотра всего множества целей диспетчера в связи с принятым решением (пересматривается выделение целевых энергий на актуальные цели).

(I) *Дозревание* всех решаемых когда-либо Z-задач и инструментов в режиме фоновой активности (неконтролируемой диспетчером). Факт появления удовлетворительного (нового) решения какой-либо «старой» Z-задачи автоматически привлекает внимание диспетчера, т.е. внезапно

возникает фокус внимания (происходит «озарение»). На этом этапе происходит также модернизация всех инструментов принятия решения.

Этапы (А) – (Е) можно отнести к фазе расширения, а этапы (Е) – (G) – к фазе локализации. Этап (Е) – переходный этап. Вообще говоря, процесс решения задачи может быть бесконечным, если достаточно энергии. Потеря фокуса внимания происходит только в двух случаях: либо найдено удовлетворительное решение, либо иссякла целевая энергия в текущий момент (на этапе (D)). Суммарный приток энергии к задаче должен быть достаточным не только для погашения роста энтропии в фазе расширения, но и для ее уменьшения, что усиливает порядок в системе. Возникающее чувство самодостаточности на этапе D может означать появление «ниточки» системного решения (концептуального решения) и главное – появление уверенности в том, что решение существует. Таким образом, главной задачей при синтезе решения в активной фазе является обеспечение когерентности на заключительном этапе фазы расширения. Решение может быть мультистабильным (мерцающим), если имеется несколько зон когнитивной стабильности изучаемой ситуации.

Важнейшим этапом является пассивная фаза (I) – фаза *дозревания* всех решаемых когда-либо Z-задач, а также инструментов принятия решений (meta-cognition). Если удовлетворительное решение не было найдено в активной фазе за контролируемое время, то оно может «дозреть» в пассивной фазе, но время появления решения заранее предсказать нельзя.

Образной иллюстрацией процесса выработки решения может быть «подъем в гору и спуск с горы». Решение находится по ту стороны горы, внизу. Подъем на гору соответствует фазе расширения, вершина – моменту возникновения когерентности (доказательству существования решения и появлению системного решения), спуск – фазе локализации решения. Основная энергия тратится при подъеме в гору.

Как уже отмечалось ранее, при перемешивании согласно методу предельных обобщений и модели многоуровневых набросков происходит генерирование колоссального числа разноуровневых описаний, набросков (заполняется семиотическое пространство, пространство смыслов) и, соответственно, переходных слоев. В переходных слоях происходит перестройка картины мира с одного масштаба на другой. Здесь встречаются разные законы, действуют смешанные языки, рождаются новые смыслы. Отсюда актуальность методов связывания, сращивания, соединения набросков образов и решений в переходных слоях. Наброски решений, полученные на одних уровнях общности, передаются на нижележащие уровни для детализации, конкретизации. Взаимодействия реализуются через динамические паттерны, которые пронизывают все семиотическое пространство КИС. И чем больше энергия, которая

перераспределяется в результате образования новых паттернов, тем быстрее идут процессы энергоинформационного обмена между слоями, описаниями, задачами и, следовательно, тем сильнее когерентность в соответствующей системе. Во многом *перемешивание* отвечает концепции динамической теории информации [187] и области Джокера [99].

Приведем простой пример «созревания» и «дозревания» решений в КИС. Предположим, для пациента Петрова сегодня решалась  $Z$ -задача «Прогноз течения заболевания». В соответствии с полученным прогнозом врач назначил ему лечение. Пусть по прошествии какого-то периода времени (например, месяца) существенно изменился контекст  $\langle \Omega(Z), \{G(\tau)\}, \{S\}, \{R\} \rangle$  и КИС заново перерешала в фоновом режиме все решенные ранее  $Z$ -задачи, включая задачу Петрова. Вполне может оказаться так, что прогноз Петрову в рамках новых моделей знаний ошибочен. КИС обязана проинформировать врача об этом факте, чтобы тот скорректировал лечение (если это еще возможно). В рамках традиционных приложений такой сценарий реализовать нельзя.

Обозначим через  $\Psi$  *оператор трансформации* действительности, или *оператор перехода*. Оператор перехода (движения) зависит от решения  $X^*$  и действующих возмущений  $\nu$ . Он отвечает на вопрос: «Как изменится действительность в результате реализации выбранной стратегии  $X^*$  и действия возмущений  $\nu$ ?».

Обозначим через  $\pi$  *оператор рефлексии*. Рефлексия является обобщением обратной связи. Она приводит к целостному представлению, знанию о целях, содержании, формах, способах и средствах своей деятельности [117]. Оператор рефлексии в ряде случаев обеспечивает возможность полного или частичного «отката» назад после применения оператора  $\Psi$ -перехода. В сочетании с оператором перехода он позволяет реализовать поступательно-возвратное развитие ситуации (исследования, теории).

Оператор рефлексии предполагает знание *оператора оценки свойств информации*  $\sigma^t$  некоторым множеством агентов  $\{Ag\}$ . Оператор  $\sigma^t$  формирует множество значений  $\{\underline{p}/P\}_t$  – количественных и качественных параметров всесторонней оценки некоторой информацией в момент времени  $t$  (например, принятого в прошлом решения). Зависимость от времени определяется тем, что многие характеристики, например, такие как ценность информации (конъюнктурная ценность, прогностическая ценность, асимптотическая ценность), эффективность и обоснованность решения, изменяются со временем. Применительно к решению (7.21) можно записать:

$$\{\underline{p}/P\}_t = \sigma^t(X^*, \alpha, \{Ag\}) = \sigma(X^*, t, \alpha, \{Ag\}), \quad (7.23)$$

где  $\alpha$  – ситуация действительности (в динамике). В развернутом виде, с учетом (7.22), можно записать:

$$\{p/P\}_{\underline{t}/A} = \sigma(\{\underline{u}/U\}^*, \underline{t}/A, \{\langle J_\tau \underline{t}/T, J_t \underline{t}'/A \rangle\}_\omega, \{Ag\}) \quad (7.24)$$

Каждый тест в (7.24), включая время, задается своим конфигуратором. Рефлексивный прогноз параметров  $\{p/P\}_{\underline{t}/A}$  может оказывать существенное влияние на оператор локализации решения  $\mathcal{A}$ .

Общую феноменологию когнитивного и метакогнитивного управления можно описать следующей гиперциклической схемой (рис. 7.19):

$$X(t+1) = \lambda(\chi(\pi(\Psi(\mathcal{A}(X(t)), v))))), \quad \chi(X) = \text{истина}, \quad (7.25)$$

где  $t$  – такты времени ( $\underline{t}/A$ ). Подчеркнем, что все операторы в (7.25) зависят от цели управления  $I$  (принятия решений) и имеющихся ресурсов. Ресурсы могут изменяться скачкообразно.

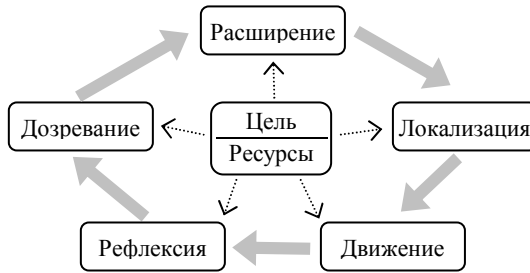


Рис. 7.19 – Гиперциклическая схема когнитивного и метакогнитивного управления

Приведенная схема отражает чередование активной и пассивной фаз управления. Она содержит в себе элементы кибернетики (активная фаза), синергетики (самоорганизация в процессе созревания и дозревания) и квантово-семантической теории (QS-запутывание структур в процессе созревания и дозревания). Метакогнитивное управление направлено на совершенствование механизмов управления, так как если регуляторные механизмы недостаточно сформированы, то увеличение ресурсов не сможет повлиять на эффективность решения задач. Как правило, когнитивное и метакогнитивное управление реализуются параллельно.

Типичные цели управления  $I$  на языке нелинейной динамики и синдромного принципа управления могут быть сформулированы следующим образом:

- Завести траекторию системы в заданное поглощающее множество



$$W = \{S\}_W \& \{R\}_W;$$

- Обеспечить асимптотическое движение системы к притягивающему или инвариантному множеству  $W = \{S\}_W \& \{R\}_W$ .

#### 7.4.2 Инструменты поддержки решений в рамках КИС

Многие важные инструменты были рассмотрены в главе 3. Например, основным инструментом реализации оператора расширения  $\lambda$  можно считать построение информационного множества. Ситуации  $\beta$  из информационного множества должны охватывать прошлое, настоящее и будущее развитие ситуации  $\alpha$ . Сходство или различие двух ситуаций  $\alpha$  и  $\beta$  более точно можно определить с помощью понятия «орбиты ситуации действительности».

Для прогнозирования развития любой системы или ситуации необходимо знать точно или приблизительно оператор эволюции [34, 35]. Обозначим через  $O(\{\underline{c}/C\})$  окрестность множества  $\{\underline{c}/C\}$ . Тот факт, что некоторая ситуация  $\alpha$  удовлетворяет условиям  $O(\{\underline{c}/C\})$ , будем записывать следующим образом:  $\alpha \nabla O(\{\underline{c}/C\})$ .

*Эмпирический оператор эволюции* (трансформации) произвольной ситуации действительности  $\phi^{\underline{t}/A}()$  определяет значения заданных тестов  $\{a/A\}$  в момент времени  $\underline{t}/A$ , используя для этого базу прецедентов  $\Omega$ :

$$\phi^{\underline{t}/A}(\{a/A\}/O(\{\underline{c}/C\})) = \cup_{\alpha \in \Omega} \{f/\mu: \underline{t}/A, \{<J_r \underline{t}/T, J_t \underline{t}/A>\}_\alpha \rightarrow \{J_a \underline{a}/A\}_\alpha | \alpha \nabla O(\{\underline{c}/C\})\}, \quad (7.26)$$

где  $f/\mu \in k_C$  (банк вычислительных моделей);  $\underline{t}'/A = \underline{t}/A$  или  $\underline{t}'/A \in [0, \underline{t}/A]$  или  $\underline{t}'/A \leq \underline{t}/A$  (выбор варианта зависит от  $\{a/A\}$ ). Результат представляет собой мультимножество. Кроме результата, оператор эволюции позволяет дать развернутое пояснение каждому результату, что порой не менее важно, чем сам результат. Отметим также, что множество прецедентов  $\Omega$  может содержать как реальные, так и модельные ситуации. Начальные значения и параметры подобия разных ситуаций действительности содержатся во множестве  $O(\{\underline{c}/C\})$ .

Свертка частных результатов позволяет во многих случаях повысить устойчивость общего результата. Если правую часть оператора эволюции обозначить  $\cup_{\alpha \in \Omega} \{a/A\}_\alpha$ , то *оператор эволюции со сверткой значений* запишется следующим образом:

$$\phi^{\underline{t}/A}(\{a/A\}/O(\{\underline{c}/C\})) \equiv g/\mu_g: \underline{t}/A, \cup_{\alpha \in \Omega} \{J_a \underline{a}/A\}_\alpha \rightarrow \{J_a \underline{a}/A\}_{\underline{t}/A}. \quad (7.27)$$

где  $g/\mu_g \in k_C$ . Эмпирический оператор эволюции позволяет реализовать принцип опережающего отражения действительности.

Модель вычислительных знаний  $k_C$  представима в виде [154]:

$$k_C = \{f/\mu: k^1 \rightarrow k^2 \mid \mu \in \{\mu\}_f\} \cup P_k, \quad (7.28)$$

где  $f/\mu$  – системопаттерны, реализующие те или иные математические модели (задачи/цели);  $\{\mu\}_f$  – разные механизмы реализации системопаттернов (со своей энергетикой и ресурсами);  $k^1$  – входные данные задачи (описание информационной среды и задание);  $k^2$  – выходные данные задачи (содержит либо искомые параметры, либо критерии достижения цели);  $P_k$  – правила композиции схем задач, т.е. правила, описывающие способы объединения системопаттернов (в результате образуются потоки, вихри, циклы).

Обозначим элементарные тесты через  $a, b, c, \{a\}$  и т.д. Пусть  $W(\{c/C\})$  – некоторое многообразие на множестве результатов тестов  $\{c/C\}$ . Модель знаний  $k_C$  в развернутом виде представима следующим образом:

$$k_C = \{f/\mu: \{J_b b/B\} \rightarrow \{J_a a/A\} \mid \{c/C\} \in W(\{c/C\}), \mu \in \{\mu\}_f\}, \quad (7.29)$$

где  $f/\mu$  – системопаттерны, реализующие те или иные математические модели. Системопаттерны могут возникать (созревать) в результате QS-запутывания орграфов набросков разных типов. Примером системопаттерна является эмпирический оператор эволюции.

Модель (7.29) содержит как эмпирические, так и фундаментальные закономерности, например, по химии, физике, биологии и т.д. [13, 76, 78, 79, 141]. В общем случае, модель  $k_C$  имеет иерархическую структуру, детализация которой будет приведена в следующем разделе.

Для обозначения функциональных систем можно использовать следующие нотации:

$$g/\mu_g: \{\} \rightarrow \{J_z z/Z\}, \mu_g \in \{\mu_g\} = \{S\} \vee \{R\} \text{ – среда радикалов; } \quad (7.30)$$

$$g/\mu_g: \{\} \rightarrow \{J_z z/Z\}, \mu_g \in \{\mu_g\} = \{f/\mu_f\} \text{ – множество паттернов. } \quad (7.31)$$

Особенностью нотаций является отсутствие спецификаций входных данных. Это связано с тем, что разные механизмы  $\mu_g$  могут использовать различные входные данные.

Оператор перехода  $\Psi$  может быть смоделирован на основе одноименного фрагмента вычислительных знаний:

$$\Psi_k = \{f/\mu: \{<J_\tau \underline{t}/T, J_i \underline{t}/A>\} \rightarrow \{<J_\tau \underline{t}/T, J_i \underline{t}/A>\}' \mid \mu \in \{\mu\}_f\} \cup P_k. \quad (7.32)$$

События в модели (7.32) учитывают как управляющие воздействия, так и возмущения вида  $<J_v \underline{v}/V, J_i \underline{t}/A>$ , которые имели место в период  $[t/A, (t/A)']$ .

Базы прецедентов  $\{\Omega\}$ , банк тестов  $\{G(\tau)\}$  и модель вычислительных знаний  $k_C$  образуют простейшую модель предметной области:  $A = \langle \{\Omega\},$

$\{G(\tau)\}$ ,  $k_c$ , на основе которой может функционировать КИС. С использованием модели ПрО вычисляются, в частности, синдромные и вероятностные модели знаний, а значит, реализуется синдромный принцип управления.

Рассмотрим одну из возможных схем работы оператора расширения  $\lambda$ , которая базируется на модели ПрО  $\langle \{\Omega\}, \{G(\tau)\}, k \rangle$  [157].

Пусть  $\Sigma$  – рабочая область, в которой хранятся все известные значения тестов. Для помещения результатов тестов в рабочую область будем использовать нотацию  $\{\underline{b}/B\} \mapsto \Sigma$  (при выполнении данной операции применяется механизм разрешения противоречий). Через  $\Sigma_f$  обозначим вспомогательную область, в которую будем помещать пары  $\langle f/\mu, \{\underline{b}/B\} \rangle$ , если к значениям тестов  $\{\underline{b}/B\}$  применялось отображение  $f/\mu$ . Для помещения данных во вспомогательную область также будем использовать нотацию  $\langle f/\mu, \{\underline{b}/B\} \rangle \mapsto \Sigma_f$ . Новые ситуации могут поступать в  $\{\Omega\}$  непрерывно. Самоорганизация знаний на основе метода предельных обобщений приводит к непрерывной модификации модели ПрО.

Одна из схем работы оператора расширения  $\lambda$  может быть следующей:

Вход:  $\{\underline{d}/D\}$ ;

Модель ПрО:  $\langle \{\Omega\}, \{G(\tau)\}, k \rangle$ ;

Инициализация:  $\Sigma_f = \emptyset$ ;  $\{\underline{d}/D\} \mapsto \Sigma$ ;

В цикле, пока не наступит стабилизация  $\Sigma$ :

1. Если  $\{\underline{c}/C\} \in \Sigma$  и  $\{\underline{c}/C\} \in W_f(\{c/C\})$  и  $\{\underline{b}/B\} \in \Sigma$  и  $\langle f/\mu, \{\underline{b}/B\} \rangle \notin \Sigma_f$ , то выполнить:  $f/\mu: \{\underline{b}/B\} \rightarrow \{\underline{a}/A\}$ ,  $\mu \in \{\mu\}_f$ ,  $f \in k$ ;

2.  $\{\underline{a}/A\} \mapsto \Sigma$ ;  $\langle f/\mu, \{\underline{b}/B\} \rangle \mapsto \Sigma_f$ ;

3.  $\forall \underline{a}/A \in \Sigma$ ,  $\underline{a}/A \rightarrow \{\underline{a}/A\}^+$ ;  $\{\underline{a}/A\}^+ \mapsto \Sigma$ ;

4. Если  $\{\underline{b}\}/B \in \Sigma$  и  $\langle g, \{\underline{b}\}/B \rangle \notin \Sigma_f$ , то выполнить:

$g: \{\underline{b}\}/B \rightarrow \neg(B \setminus \{\underline{b}\}/B)$ ;  $\neg(B \setminus \{\underline{b}\}/B) \mapsto \Sigma$ ;  $\langle g, \{\underline{b}\}/B \rangle \mapsto \Sigma_f$ ;

5. Если  $\exists f/\mu: \{\underline{b}/B\} \rightarrow \underline{a}/A$  на  $W_f(\{c/C\})$ ,  $\underline{a}/A \in OЗ_f$  и  $\neg \underline{a}/A \in \Sigma$  и  $\{\underline{c}/C\} \in \Sigma$  и  $\langle h_f, \neg \underline{a}/A \rangle \notin \Sigma_f$ , то выполнить:

$h_f: \neg \underline{a}/A \rightarrow \neg_f \underline{a}/A = \neg \{\{\underline{b}/B\} \mid f/\mu: \{\underline{b}/B\} \rightarrow \underline{a}/A, \mu \in \{\mu\}_f,$

$\{\underline{c}/C\} \in W_f(\{c/C\}), f \in k\}$ ;  $\neg_f \underline{a}/A \mapsto \Sigma$ ;  $\langle h_f, \neg \underline{a}/A \rangle \mapsto \Sigma_f$ ;

Окончание цикла.

Выход:  $\Sigma$ .

Приведенная выше схема базируется, по сути, на двух основных законах логики: modus ponens и modus tollens (рассуждение от противного). В следующем разделе будет рассмотрен пример процедурной реализации модели ПрО и механизмов вывода. Эффективным инструментом поддержки решения является консилиум когнитивных агентов [134, 135]. Разработано несколько прототипов программ.

### 7.4.3 Пример реализации Многоцелевого банка знаний

Приведем пример процедурной реализации модели вычислительных знаний (7.29) в области ПрО «Клиническая медицина». Учитывая, что профессиональная лексика и Банк тестов  $\{G(\tau)\}$  описываются с помощью синтаксиса лексических деревьев [4, 134], решено системопаттерны (отображения)  $f/\mu$  также описывать с помощью синтаксиса лексических деревьев. Реализовано несколько вариантов такого представления. В рамках первого варианта описание произвольного функционального отображения (ФО) состоит из четырех блоков: блок условий применимости ФО –  $\{\underline{c}/C\}$ ; блок входных параметров (тестов) –  $\{\underline{b}/B\}$ ; блок выходных параметров (тестов) –  $\{\underline{a}/A\}$ ; блок описания методов обработки данных (механизмов  $\{\mu\}$ ). Аббревиатура ‘ФО’ задает тип лексического дерева (зарезервированное сочетание символов). Ниже приведена общая схема записи ФО:

```
<Полное имя ФО> {ФО_<Сокр_имя_ФО>
  [Условие {cond_
    <тесты условия - {c/C}>}]
  Вход {inp_
    <входные тесты - {b/B}>}
  Выход {out_
    <выходные тесты - {a/A}>}
  [Методы {met_
    <методы обработки>}]}
```

Предполагается, что в качестве тестов выступают тесты из Банка тестов  $\{G(\tau)\}$ , хотя это и не является обязательным требованием. Методов обработки данных может быть несколько. В число методов может входить создание специального (когнитивного) агента или виртуального робота, который будет решать все вопросы, связанные с реализацией ФО (добычей, обработкой и представлением данных). В ряде случаев исходные данные могут пересылаться удаленным агентам (консультантам, экспертам) для принятия решения. Пример записи ФО (СИ – сердечный индекс):

```
Тип гемодинамики {ФО_ТипГемо
  Вход {inp_
    Пол/ПОЛ?
    Возраст/N?
    СИ? }
  Выход {out_
```

```

    Тип_Гемодинам }
    Методы {met_
        fo_TG()}
} ТипГемо,

```

где fo\_TG() – скрипт. Знак вопроса в конце теста означает, что каким-либо образом должно быть введено значение данного теста (с консоли или найдено в БД). Скрипты могут быть описаны в том же файле, что и множество отображений.

Унификация синтаксиса всех объектов модели Про позволяет формировать значения входных тестов с помощью произвольных лексических деревьев. Ниже представлено отображение «Диагностика состояния НЕЗДОРОВЬЯ», которое использует ЛД ‘Жалобы’ из файла Жалобы.tris; ЛД ‘Анамнез\_кард’ из файла Анамнез.tris; ЛД ‘Об\_но\_кард’ из файла ОБ\_но.tris (файлы \*.tris содержат описание профессиональной лексики).

*Диагностика состояния НЕЗДОРОВЬЯ* {ФО\_ДСН

```

Входные данные {inr_
    Осмотр врача {
        Жалобы          {Жалобы: Жалобы}
        Анамнез         {Анамнез: Анамнез_кард}
        Объективно     {ОБ_но: Об_но_кард}}
    Пол/ПОЛ?}
Выходные данные {out_
    Заключение/ЗАКЛ}
Методы {met_
    Эксперт А {}
    Агент Ag {}
    Вычислительный алгоритм {}}
}
{ПОЛ
    |М; Ж|}}
{ЗАКЛ
    Здоров
    Болен {Ds}}

```

Приведенный фрагмент содержит описание двух доменов в виде лексических деревьев: ПОЛ и ЗАКЛ. Если заключение формирует Эксперт, то он обязан его представить в формате ЗАКЛ. Экранная форма для просмотра и редактирования данного системопаттерна (отображения) представлена на рис. 7.20.

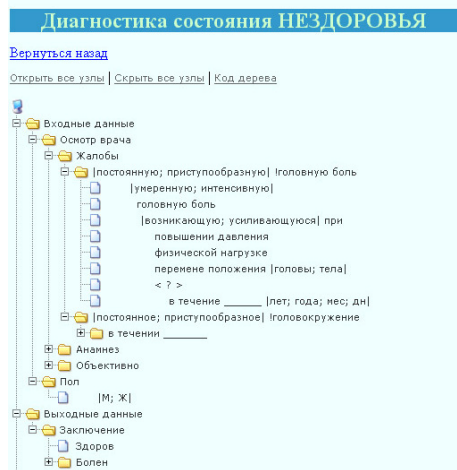


Рис. 7.20 – Пример системопаттерна с лексикой

Второй вариант описания ФО более экономный (в квадратных скобках необязательные параметры):

[Имя\_ФО:] {Тест}<sub>1</sub> → {Тест}<sub>2</sub> [(Правило\_1 [, ..., Правило\_K])]

Пример:

Пол/ПОЛ, Возраст/N, СИ -> Тип\_Гемодинам (fo\_TG()).

Для типов ЛД: ТЕСТ, ФО и ПРО сочетание символов ‘->’ является функциональным, т.е. зарезервированным. Правило (скрипт) может быть опущено в том случае, если отображение является продукцией вида «Если  $A=a$ , то  $B=b$ ». В последнем случае ФО имеет вид:  $A?a \rightarrow B?b$ .

Первый вариант представления ФО является более универсальным, чем второй. Он позволяет сколь угодно подробно описывать все компоненты, включая домены, способы отображения и хранения выходной информации. Достоинством второго варианта является лаконичность.

Некоторые ФО могут объединяться в семантические группы, соответствующие той или иной предметной области. Каждая такая группа имеет название, которое формируется следующим образом: Про\_<СеманТя>. Ниже приводится общий синтаксис семантической группы отображений (паттернов):

<Имя предметной области> {ПРО\_<СеманТя>

[Определения {def\_  
 <Сокращенное имя теста> - <Полное имя теста> [{Ссылка}]  
 ....}]  
 [Банк тестов: *Файл1* [, *Файл2*,...]]  
 <ФО\_1>; ...; <ФО\_N>} [ПрО]

Аббревиатура 'ПРО' задает тип лексического дерева и является зарезервированной; {Ссылка} позволяет сослаться на ЛД, описывающее тест в Банке тестов. Объединение ФО в семантические группы позволяет процессору реализовать дополнительные аналитические и вычислительные функции, в частности, строить логические схемы ПрО, осуществлять прямой и обратный вывод и т.д.

Предметные области могут образовывать сложные иерархии. Пример иерархии ПрО:

ПрО {{...{}}...{...{...{}}...{}}

Каждая пара фигурных скобок означает новое ПрО. Любая вложенная ПрО наследует определения, соглашения, тесты и системопаттерны всех родительских ПрО. Видимость любого ФО ограничена ближайшими окаймляющими фигурными скобками (наименьшей ПрО).

В качестве примера семантической группы отображений приведем фрагмент ПрО «Эхокардиография» [134]:

*ПрО "Эхокардиография"* {ПРО\_ЭхоКГ

Определения {def\_

КДР - Конечный диастолический размер

КСР - Конечный систолический размер

КДО - Конечный диастолический объем

КСО - Конечный систолический объем

УО - Ударный объем {ТЕСТЫ: УдарОбъем}

ФВ - Фракция выброса {ТЕСТЫ: ФракВыброса}

СИ - Сердечный индекс {ТЕСТЫ: СердИндекс}

СокрСпособМиок - Сократительная способность миокарда}

#ТЕСТЫ – файл, содержащий Банк тестов.

Рост, Вес -> Площ\_Тела (fo\_S\_Tela())

RR -> ЧСС (fo\_RR())

КСР -> КСО (fo\_KSO())

КДР -> КДО (fo\_KDO())

КДО, КСО -> УО, ФВ (fo\_UO\_FV ())

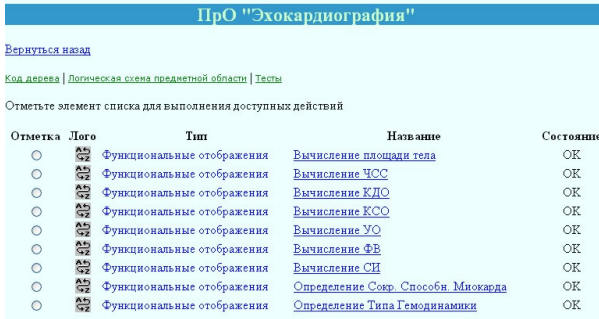
Площ\_Тела, ЧСС, УО -> СИ (fo\_SI ())

ФВ -> СокрСпособМиок (fo\_CCM())

Пол/ПОЛ, Возраст/N, СИ -> Тип\_Гемодинам (fo\_TG()).

} ПРО\_ЭхоКГ

Допускается использование разных вариантов описания системопаттернов. На рисунке 7.21 приведено окно просмотра содержимого объекта «Предметная область» для случая, когда каждое ФО представлено в соответствии с первым вариантом формата (каждое ФО имеет имя).



The screenshot shows a web interface for 'PrO "Эхокардиография"'. It includes navigation links like 'Вернуться назад', 'Код дерева', 'Логическая схема предметной области', and 'Тесты'. Below these is a table with columns: 'Отметка', 'Лого', 'Тип', 'Название', and 'Состояние'. The table lists ten items, all of type 'Функциональные отображения', with various medical test names and a status of 'OK'.

Отметка	Лого	Тип	Название	Состояние
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление площади тела</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление ЧСС</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление КДО</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление КСО</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление УО</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление ФЕ</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Вычисление СИ</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Определение Сокр. Способи Миокарда</a>	OK
<input type="radio"/>		Функциональные отображения	<a href="#">Определение Типа Гемодинамики</a>	OK

Рис. 7.21 – Окно просмотра содержимого ПрО

С ЛД типа 'ПрО' связаны дополнительные функции, в частности:

- «Код дерева»;
- «Логическая схема предметной области»;
- «Тесты».

При выборе пункта «Код дерева» открывается дополнительное окно, в котором показывается исходное лексическое дерево выбранного ФО или всей ПрО. При выборе пункта «Логическая схема предметной области» открывается дополнительное окно, в котором показывается или логическая схема группы ФО, или причинно-следственный граф. При выборе пункта меню «Тесты» открывается окно, в котором перечислены все тесты данного ПрО (рис. 7.22). Краткие названия тестов выделены жирным шрифтом. Если в разделе «Определения» ПрО дается их расшифровка, то она приводится в скобках. Вычисляемые тесты выделены подчеркиванием.

Пользователь имеет возможность отметить интересующие его тесты и выполнить прямой или обратный логический вывод. При прямом логическом выводе будет показан список тестов, значения которых можно получить, зная значения отмеченных тестов (с учетом преобразования доменов в рамках банка тестов). Пример приведен на рисунке 7.23.а. При обратном логическом выводе будет получен список первичных тестов, значения которых необходимы для вычисления значений отмеченных тестов (рис. 7.23.б). При желании пользователь может произвести необходимые вычисления.



Важно отметить, что прямой и обратный логический вывод осуществляются без использования каких-либо методов (механизмов), указанных в отображениях. Для согласования доменов тестов в разных ФО ключевую роль играет Банк тестов  $\{G(\tau)\}$ . Если домен у какого-либо теста не указан, то по умолчанию используется базовый домен.

**Про "Эхокардиография"**

[Снять метки](#)  
[Логическая схема](#)

**Входные данные:**

- Рост
- Вес
- RR
- КДР (Конечный диастолический размер)
- КСР (Конечный систолический размер)
- КДО (Конечный диастолический объем)
- КСО (Конечный систолический объем)
- Площ Тела
- ЧСС
- УО (Ударный объем)
- ФВ (Фракция выброса)
- Пол
- Возраст
- СИ (Сердечный индекс)
- Сокр.Способ.Миок (Сократительная способность миокарда)
- Тип Гемодинам

---

**Выберите тип вывода:**

- Прямой вывод
- Обратный вывод

Рис. 7.22 – Все тесты Про

**Про "Эхокардиография"**

[Снять метки](#)  
[Логическая схема](#)

**На основе параметров :**

- Рост
- Вес

**можно получить:**

- Площ Тела

**а) Прямой вывод**

**Про "Эхокардиография"**

[Снять метки](#)  
[Логическая схема](#)

**Для получения:**

- Тип Гемодинам

**необходимо:**

- Рост
- Вес
- RR
- КДР (Конечный диастолический размер)
- КСР (Конечный систолический размер)
- Пол
- Возраст

б) Обратный вывод  
Рис. 7.23 – Вывод

Фрагмент общего списка объектов МБкЗ приведен на рис. 7.24.

Многоцелевой банк знаний				
Отметьте элемент списка для выполнения доступных действий				
Отметка	Лого	Тип	Название	Состояние
<input type="radio"/>		Папка с файлами	<a href="#">АЛГОРИТМЫ ДИАГНОСТИКИ (СИНДРОМЫ)</a>	OK
<input type="radio"/>		<a href="#">Алгоритм диф. диагностики</a>	<a href="#">АЛД "Боль в груди"</a>	OK
<input type="radio"/>		<a href="#">Алгоритм диф. диагностики</a>	<a href="#">АЛД "Острая боль в эпигастриальной области"</a>	OK
<input type="radio"/>		<a href="#">Алгоритм диф. диагностики</a>	<a href="#">АЛД «Острый живот»</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">АНАМНЕЗ</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">ЖАЛОБЫ НА</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">МОЛОЧНЫЕ ЖЕЛЕЗЫ</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">Орган</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">ЛАБОРАТОРНЫЕ ИССЛЕДОВАНИЯ</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">НАЗНАЧЕНИЯ</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">ОБЪЕКТИВНЫЙ СТАТУС</a>	OK
<input type="radio"/>		Лексическое дерево	<a href="#">РЕКОМЕНДАЦИИ</a>	OK

Рис. 7.24 – Фрагмент общего списка объектов МБкЗ

Банк тестов  $\{G(\tau)\}$  является частью МБкЗ. Для просмотра и редактирования разных его фрагментов могут создаваться разные приложения. Так, для создания и редактирования лексических деревьев применяется лексический процессор [4]. Для моделирования и редактирования орграфов доменов тестов применяется специальное приложение [144]. Важной компонентой МБкЗ являются различные базы прецедентов  $\Omega$ , включая базы медицинских карт, базы изображений и сигналов. Важнейшими функциями баз прецедентов являются: построение орграфов набросков (глава 6); построение (предельных) синдромных и вероятностных моделей знаний (глава 5); классификация нового прецедента. К числу важных структур медицинских МБкЗ относятся «Модели пациента» и «Модели заболеваний» [134].

Многоцелевой банк знаний служит ядром госпитальных и телемедицинских систем третьего поколения (когнитивных информационных систем). Учитывая сложность разработки МБкЗ, целесообразно создавать общее ядро для многих госпитальных/телемедицинских систем, используя, например, технологию «облаков». С ядром взаимодействуют когнитивные агенты (создаются для каждого пациента). На основе МБкЗ производятся расчеты вторичных параметров (тестов) в рамках различных документов (записей) медицинской карты. С использованием объектов МБкЗ (орграфов доменов тестов, профессиональной лексики, алгоритмов диагностики и прогнозирования, системопаттернов) строится интеллектуальный интерфейс медицинских приложений [134]. Важнейшую роль МБкЗ и когнитивные агенты играют в автоматизированном контроле качества медицинского обслуживания, в частности, мониторинге выполнения стандартов обслуживания [134]. С помощью распределенного МБкЗ, опирающегося на синтетическую телесно-феноменологическую онтологию и синдромное управление, наиболее просто и естественно осуществляется передача профессиональных знаний и умений.

## **7.5 Концепция «Виртуальной сплошной среды» и Субъективная реальность**

Рассмотрим конструктивный феноменологический набросок концепции *Виртуальной сплошной средой* (ВСС), которая в полной мере опирается на парадигму предельных обобщений. Данная концепция (модель) может служить отправной точкой для проведения углубленных формализованных исследований когнитивных архитектур и процессов, в том числе в рамках КИС.

Постулируем ключевые аспекты функционирования ВСС:

1. ВСС неразрывно связана с физическим носителем, образуя единую Систему (организм). Система реализует целенаправленное поведение (когнитивно-поведенческий уровень).

2. ВСС обладает врожденными и приобретенными *автоматизмами*, действующими постоянно и обеспечивающими создание, развитие и разрушение когнитивных структур ВСС. Автоматизмы являются разновидностью системопаттернов.

С ВСС связано неоднородное энергетическое поле, которое питает все процессы. Энергия в любой локальной области ВСС складывается из фоновой и целевой энергий. Имеет место конкуренция за энергию.

3. Первоосновой ВСС является *Хаос*. Хаос поддерживается непрерывным и всеобъемлющим действием автоматизмов, а также открытостью среды. Хаос ВСС полностью неустраним.

В первом приближении (при QS-декогеренции) Хаос – очень большое число полностью или частично изолированных друг от друга структур, находящихся в различных динамических состояниях (протоструктуры, структуры, субструктуры, обломки структур и т.д.). Спонтанное взаимодействие (запутывание) разных структур может приводить к зарождению и созреванию новых структур.

Состав структур и автоматизмов ВСС описывается синтетической телесно-феноменологической онтологией. Первичными структурами являются базы прецедентов, орграфы набросков, включая орграфы доменов тестов, а также элементарные системопаттерны (врожденные и приобретенные).

С биофизической точки зрения нейроны среды обладают «фоновой» хаотической разрядной активностью (имеет место квантово-волновой дуализм). Феномен поведенческой специализации при этом возникает в результате фрагментации постоянно присутствующего «фонового потока» энергетических разрядов.

4. В рамках ВСС могут образовываться устойчивые системные образования – *ментальные многообразия*. Ментальные многообразия формируются, в частности, любыми контекстами  $\langle \Omega(Z), \{G(\tau)\}, \{S\}, \{R\} \rangle$  или Z-задачами, а также моделями ПрО  $\langle \{\Omega\}, \{G(\tau)\}, k \rangle$ . Активизация фрагмента многообразия может приводить к активизации всего многообразия (имеет место конкурентное перетягивание энергии, приводящее к торможению других структур). Ментальные многообразия имеют самоподобную структуру и могут притягиваться, делиться, объединяться.

*Размерность* ментального Z-многообразия (когнитивного ресурса) определяется числом тестов (координат), с помощью которых описывается  $\Omega(Z)$ . Совокупность всех тестов определяет *размерность когнитивного*

*пространства* или ВСС. *Сложность* ментального  $Z$ -многообразия определяется произведением числа доменов всех тестов  $\{G(\tau)\}_Z$ , а также мощностью *ресурсов активности* ПрО, т.е.  $k$ , включая модели знаний. Другими словами, сложность ментального  $Z$ -многообразия (когнитивного ресурса) определяется как высокой степенью дифференцированности, так и высокой степенью иерархической интегрированности (синтетической онтологией).

Высокая степень сложности  $Z$ -многообразия проявляется в способности субъекта создавать многомерные модели реальности, выделяя в ней множество взаимосвязанных сторон, включая конкурирующие наборы параметров порядка. Напротив, низкая степень сложности  $Z$ -многообразия свидетельствует о понимании и интерпретации субъектом происходящего (в рамках  $Z$ -задачи) на основе весьма упрощенных моделей и фиксации одних и тех же сторон действительности из-за использования ограниченного набора измерений (тестов) и ограниченного ресурса активности. Мощности когнитивного ресурса может быть недостаточно для реконструкции адекватной модели задачи, что ведет к ошибкам.

За счет аккумуляции и удержания значительной энергии в ментальных многообразиях наиболее интенсивно идут процессы самоорганизации и QS-запутывания, которые приводят к возникновению критичных образований. Примеры: предельные синдромные и вероятностные модели знаний, экстремальные пограничные слои набросков, критические пути и т.д.

5. В областях ВСС с повышенной энергией (например, ментальных многообразиях) созревают *субструктуры высоких порядков*, в частности иерархии функциональных систем в виде сред радикалов, колоний роботов, многоагентных систем. Разновидностью субструктур являются иерархии системопаттернов, иерархии синдромных и вероятностных моделей знаний.

Порядок субструктур определяется уровнем иерархии. Субструктуры относятся к наиболее энергоемким сущностям ВСС (фоновой энергии для их зарождения и возбуждения, как правило, недостаточно). Субструктуры отвечают за реализацию системных свойств ВСС, обеспечивая жизнедеятельность и безопасность Системы.

6. Функционирование ВСС удовлетворяет *Принципу Баланса Энтропий* (неравновесное состояние). Баланс энтропий обеспечивает *фоновую энергию*, питающую автоматизмы и системопаттерны ВСС. С любым фрагментом ВСС связано *энтропийное многообразие*.

*Энтропия* – мера несовершенства структур и процессов ВСС в конкретных ситуациях действительности (например, несовершенство текущих моделей знаний, неудовлетворенные потребности радикалов,

роботов, агентов и т.д.). Критериев, аспектов (мер) несовершенства бесконечно много, так же как бесконечно много разных ситуаций.

*Рефлексия* – это актуализация некоторого подмножества множества энтропий, т.е. текущая оценка степени несовершенства тех или иных качеств ВСС. С помощью рефлексии осуществляется корректировка Баланса Энтропий.

7. В энергетическом поле ВСС может возникать *резонатор* – энергетический сгусток (энергетический резонанс), являющийся источником *целевой энергии*. В области действия резонатора формируются или возбуждаются субструктуры высоких порядков, а также «высвечиваются» энергетическим лучом группы QS-запутанных набросков образов. Субструктуры находятся в возбужденном состоянии до тех пор, пока подпитываются целевой энергией. В каждый момент времени активизируемые резонатором субструктуры и наброски образов не покрывают всей ВСС. Резонаторы имеют разную мощность.

Одним из решающих факторов самоорганизации системы – перехода ее элементов из разобщенного состояния в режим кооперативного взаимодействия – является наличие особого, общего по отношению ко всем ним энергетического источника. Его роль определяется как функция оказания на систему неспецифической (неорганизационной) формы влияния. Структуры перестают быть независимыми, между ними возникают корреляция и соответственно когерентные, согласованные действия или так называемое «резонансное возбуждение» (одно из центральных понятий синергетики). Собственно термин «Резонатор» и олицетворяет такой энергетический источник.

Потоки целевой энергии во многом формируются синдромным принципом управления, который задает иерархию целей и, соответственно, импульсы: «текущее состояние → целевое состояние». Импульсы определяют направление и мощность потока целевой энергии (движение резонатора). Величина энергии во многом предопределяет достижимость целей системопаттерна.

8. Мощный резонатор может сформировать *характеристическое направление* – область ВСС с направленным движением повышенной фоновой энергии (след резонатора). В рамках характеристических направлений возникают *хаотические ансамбли* протоструктур и структур ВСС. Внутри хаотических ансамблей происходит интенсивное перемешивание и QS-запутывание структур, что приводит к возникновению и созреванию новых структур, субструктур и системопаттернов (следствие локальной самоорганизации). «Долгоживущие» системопаттерны могут стать автоматизмами ВСС, «долгоживущие» субструктуры могут стать функциональными системами

ВСС.

Хаотические ансамбли могут разрушиться при снижении уровня энергии, но при достаточной энергии и зрелости формирующихся структур они могут превратиться в ментальные многообразия.

Процессы самоорганизации и QS-запутывания в ментальных многообразиях идут более интенсивно, чем в хаотических ансамблях за счет аккумуляции большей энергии. Следовательно, средняя энтропия ментального многообразия ниже средней энтропии хаотического ансамбля. Энтропия последнего ниже энтропии окружающего хаоса. Критические структуры образуются только в ментальных многообразиях.

Первичный хаос, хаотические ансамбли, ментальные многообразия и область резонатора задают динамическую, самоподобную (фрактальную) структурированность и топологию ВСС. Последовательные стадии созревания структур ВСС определяют *системогенез* структур и функциональных систем когнитивно-поведенческого уровня.

9. С ВСС связан *дуализм* «локальное – нелокальное»: одновременно с наличием полностью или частично локализованных сущностей ВСС проявляет свойства *активной среды*, в которой происходят автоволновые процессы и процессы QS-запутывания. Параметры активности среды могут сохраняться в виде «информационных кодов». Благодаря кодам и QS-запутыванию активность среды может воспроизводиться. «Информационный код» характерен для начальной стадии формирования системопаттерна и функциональной системы. QS-запутывание является наивысшей стадией созревания системопаттерна и функциональной системы, которая переводит их в разряд автоматизмов.

QS-рекогеренция позволяет считать, что резонатор порождает неразделяемую смысловую суперпозицию возникающих субструктур высоких порядков и «высвечиваемых» набросков образов.

10. ВСС является открытой средой. Процесс восприятия и *понимания* (осмысления) *нового явления, образа* включает в себя три основных этапа: 1) построение орграфов набросков образа, явления; 2) построение синдромно-вероятностных моделей эмпирических знаний в рамках положительного контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ , где  $Z = \{\text{образ}^+; \text{образ}^-\}$ ; 3) QS-запутывание орграфов нового образа с орграфами других образов и явлений (в результате такого запутывания возникают системопаттерны).

До момента осмысления поток информации от сенсоров Системы представляет собой не более чем фотографию действительности.

Повторяющаяся синхронность или последовательность активизации набросков разных образов при достаточном уровне энергии может приводить к QS-запутыванию соответствующих орграфов набросков, а также к появлению новых устойчивых или неустойчивых образов.

11. Множество разных ВСС при определенных условиях взаимодействия может породить *когнитивный эфир* – специфическое информационное (хаотическое) пространство, в которое погружены все ВСС из множества и посредством которого осуществляется прямое (парное) и распределенное взаимодействие между ВСС.

Рассмотрим некоторые следствия из базовых положений (постулатов).

ВСС является *предельным переходом* (QS-рекогеренцией) от множества структур и процессов к континууму структур и процессов. Подобный предельный переход в методологическом плане позволяет «оторваться» от физической основы: будь то нейронная сеть, социальная среда (мультиагентная среда) или другая физическая среда.

ВСС нельзя «выключить» (пока существует Система) и, скорее всего, нельзя «перенести» на другой физический носитель. ВСС обладает диссипацией, инерционностью и параметризацией (внутренним) временем. ВСС постоянно воссоздает свое разнообразие, питаемое незамкнутостью Системы. Для когерентного поведения ВСС должна быть сильно неравновесной. Синергетические эффекты могут проявляться только в локальных областях ВСС (например, в рамках хаотических ансамблей или ментальных многообразий). QS-эффекты могут охватывать всю ВСС, а возможно и когнитивный эфир.

Свойства ВСС проявляются в конкретной ситуации действительности и существенно зависят от текущего Баланса энтропий, а также направления движения и мощности резонатора (множества импульсов). Формальное описание импульсов приведено в главе 3.

Субструктуры высоких порядков и «подсвечиваемые» наброски образов в зоне действия резонатора ментально воспринимаются как «сознание» (осознаваемое движение мысли). Область ВСС вне действия резонатора можно трактовать как «подсознание» (четкой границы между ними нет). Результат «сознание» получает целиком, без последовательного "проигрывания" мыслей, что затрудняет или делает невозможной вербализацию многих актов мышления. Возникновение сложных целостных свойств в субструктурах достигается посредством самоорганизации и QS-запутывания в критическое состояние. Важно отметить, что «сознание» не может контролировать всю ВСС.

Примеры субструктур наивысших порядков: предельные синдромные и вероятностные модели знаний; экстремальные пограничные слои набросков в орграфах набросков; экстремальные потоки системопаттернов (синергии), в частности, «критические пути» в средах радикалов и т.д.

Пример структур возрастающих порядков: орграф доменов теста → структурно-завершенный орграф доменов → предельный структурно-завершенный орграф доменов → минимальный предельный орграф

доменов (содержит только те домены, которые используются в предельных моделях знаний). Банк минимальных предельных орграфов доменов тестов целесообразно использовать при передаче знаний (в процессе обучения).

Созревание протоструктур и структур ВСС в основном осуществляется в фоновом режиме (аналогия – на уровне «подсознания»), т.е. вне области действия резонатора. Важнейшую роль в созревании играют характеристические направления, в рамках которых возникают *хаотические ансамбли*. Последние обеспечивают зарождение и развитие протоструктур и структур, в частности системопаттернов.

Вследствие неоднородности энергетического поля в ВСС могут возникать «энергетические ловушки» – замкнутые ментальные многообразия («шары»), между структурами которых и остальной частью ВСС нет QS-запутывания, а значит, нет ассоциативных переходов между орграфами набросков образов. Подобные «шары» наиболее часто возникают и «раздуваются» в процессе обучения, если оно оторвано от жизненных реалий. Со временем одни «шары» могут «схлопнуться» (информация забывается), другие «лопаются», сливаясь в единое смысловое пространство ВСС (предпочтительный сценарий). Наличие больших «шаров» в ВСС может приводить к ощущению «раздвоения».

Воображаемые образы, «сновидения» и «фантазии» являются следствием или дефокусировки энергии резонатора (вторичных волн), или неконтролируемых локальных всплесков фоновой энергии, в том числе при отсутствии целевой энергии. Всплески фоновой энергии (следов от многих резонаторов) могут достигать силы резонатора, что приводит к неожиданному появлению фокуса внимания к определенным (новым) образам или субструктурам.

Наличие «когнитивного эфира» позволяет объяснить многие Psi-феномены, в частности телепатию (передачу мыслей на расстояние) и суггестию (внушение).

Энтропия несовершенства моделей знаний складывается из суммы ошибок применения моделей (фальсификаций моделей) и суммы отказов применения моделей, например в случае, если тип входных данных не поддерживается моделью. Из-за отказов применения вычислительные мономодели (дифференциальные, интервальные, нечеткие, лингвистические и т.д.) обладают значительной энтропией (несовершенством). QS-запутывание орграфов набросков и орграфов доменов устраняет этот недостаток.

В определенной степени ВСС можно рассматривать в качестве конструктивной модели *субъективной реальности*, на тему которой ведется немало дискуссий [17, 19, 21, 29, 32, 38, 39, 53]. Приведем в этой связи высказывание философа Дубровского Д. И. [57]: «...Важно



подчеркнуть, что проводимые в этом плане исследования имеют стратегическую цель: выяснение существенных и необходимых свойств того типа самоорганизации, который создает представленность для системы информации в форме субъективной реальности и способность оперировать ею (в «чистом» виде), что имеет первостепенное значение для создания новых направлений разработки ИИ». С ним солидарен физиолог Иваницкий А. М.: «Главная загадка природы: как на основе работы мозга возникают субъективные переживания» [66, 67].

Концепция ВСС позволяет на модельном уровне рассмотреть прямую и обратную задачи, а именно: показать, как на основе «физики» возникают субъективные переживания (прямая задача), и, наоборот, как процессы мышления интерпретируются на «физическом» уровне. Ответы дают синтетическая телесно-феноменологическая онтология, синдромный принцип управления и постулаты ВСС.

Зарождающаяся глобальная КИС на основе Интернета, по-видимому, в недалекой перспективе сможет удовлетворить всем постулатам ВСС. Одни из главных постулатов – первичный Хаос, невозможность «выключения», невозможность полного контроля над процессами, функционирование мультиагентных сред – уже реализованы. Технологии Semantic Web и Интеллектуальный Грид [114] приблизят реализацию других постулатов.

Предложенный подход отрицает взгляд на интеллект как на «интеллектуальный» решатель задач. ВСС – это прежде всего «вечный двигатель», реактор, в котором протекают бесчисленные процессы трансформации и созревания определенных структур, рождаются и «бушуют» вихри паттернов. Концепция ВСС дает необходимую системную интерпретацию многих феноменов мышления.

## **7.6 Образовательные технологии на основе принципа предельных обобщений. Когнитивные тренажеры**

Стремительное развитие компьютерных и телекоммуникационных технологий привело человечество к порогу когнитивной эры и «экономики знаний». Ни одна отрасль в истории человечества не знала до сих пор таких темпов развития. Однако на фоне ускоряющегося технологического прогресса все более очевидной становится растущая неадекватность системы образования. Когнитивная экономика, экономика знаний предполагают глубокую модернизацию (смену) парадигмы обучения.

Одним из краеугольных положений новой парадигмы должно стать формирование у субъекта *способности к разноуровневым обобщениям*. Сформировав выраженные способности к обобщению, можно ожидать (согласно принципу подчинения), что и другие профессиональные качества

получат адекватное развитие. Другими словами, «умение обобщать» является одним из параметров порядка современного образования.

Приведем цитату из работы ведущих специалистов Института прикладной математики им. М.В. Келдыша РАН [100]: «...диагностика состояния больного, судя по медицинским руководствам, требует определения от 400 до 1000 параметров. При этом разные области медицины «говорят на разных языках», вкладывая в одни и те же термины разный смысл. Однако врач в состоянии оперировать в пространстве характеристик и признаков, размерности, не превышающей 5 – 7. Какие же это признаки? Очевидно, опытный диагност, в отличие от начинающего, среди всего пространства параметров выделяет «главные», «нужные», «подходящие». Собственно, умение выделять подобные «параметры порядка» для разных заболеваний и состояний организма и является результатом профессиональной деятельности. В ходе работы происходит самоорганизация в информационном пространстве врача, позволяющая отделять главное от второстепенного. *Динамика этого процесса плохо понята и изучена*, поэтому и не удается учить врачей быстро и хорошо» (курсивом выделено мною).

Авторы процитированной работы делают заключение: «По-видимому, ключевая способность человека, позволившая опередить остальные виды – удивительная способность быстро выявлять параметры порядка в разных ситуациях (естественно, отбрасывая лишнее), следить за ними, а также быстро менять поведенческие стратегии в зависимости от них».

Приведенные цитаты как нельзя лучше иллюстрируют предлагаемую новую парадигму образования: *научить осмысленно (целенаправленно) выделять параметры порядка в самых разных ситуациях действительности и использовать найденные параметры для управления.*

Выделение параметров порядка в реальной жизни – творческий процесс, требующий высокой квалификации и профессионального опыта. Как правило, такой процесс протекает на подсознательном уровне, что затрудняет или делает невозможной его вербализацию. Обучить способности выделять параметры порядка значит научить быстро оценивать ситуацию и принимать качественные решения в своей профессиональной области, а чем выше чувство компетентности, тем больше у обучаемого появляется ощущение, что он выбрал «свое дело», следовательно, усиливается мотивация к обучению. Недаром Отто фон Бисмарк еще в XIX веке говорил, что «войны выигрывают не генералы, а школьные учителя».

Собственно весь материал книги посвящен когнитивным технологиям выделения параметров порядка и управления на их основе. Важной особенностью данных технологий с образовательной точки зрения является

их концептуальная (методологическая) ясность и относительная простота. Благодаря этим качествам предлагаемые технологии легко освоят как студенты технических, так и гуманитарных специальностей (и даже школьники). Процедурная реализация построения предельных моделей знаний, которые содержат конкурентные наборы параметров порядка, также не вызывает особых затруднений, однако в этой части гуманитариям лучше использовать стандартные инструменты. Студентам ИТ-направлений полезно самим попрактиковаться в написании соответствующих программ. Средства реализации могут быть самими разными – от электронных таблиц до «облачных» (Web) технологий.

Приведем набросок специализированной когнитивной программы (этапов) повышения профессиональных качеств специалистов нового типа:

- формулирование  $Z$ -задач, например, «Выявить когнитивные и психологические факторы, отличающие специалистов от новичков в профессиях с высоким уровнем неопределенности» или «Разработка и анализ системы автоматического распознавания пола людей по изображению лица»;

- формирование Банка тестов  $\{G(\tau)\}$  для описания прецедентов из  $\Omega(Z)$  (главным критерием является развитость орграфов доменов);

- формирование структурно-завершенного Банка тестов  $\{G^+(\tau)\}$  на базе  $\{G(\tau)\}$ ; построение канонических орграфов доменов тестов, расчет энтропии орграфов доменов;

- моделирование созревания орграфов доменов тестов и нелокальных квантово-семантических эффектов, в частности QS-запутывания орграфов; построение расширенного Банка тестов;

- для заданного множества значений тестов  $\{\underline{g}/T\}$  построение замыкания  $\{\underline{g}/T\}^+$ , верхнего  $\{\underline{g}/T\}^{\wedge}$  и нижнего  $\{\underline{g}/T\}^{-}$  пределов;

- для заданного значения теста  $\underline{g}/T$  построение волновой функции на базе орграфа доменов  $G(\tau)$  (см. квантовую семантику); построение конуса обобщения;

- формирование базы прецедентов  $\Omega(Z)$ ; определение общего числа описаний (набросков в орграфе набросков) и числа финитных описаний (набросков);

- построение информационных диаграмм тестов (доменов тестов) на  $\Omega(Z)$ ; вычисление информативности значений тестов с использованием орграфа доменов информативности;

- формирование обобщенных запросов к базе прецедентов  $\Omega$  на основе банка тестов; формирование ответов разных уровней общности;

- реализация операторов обобщенного формального анализа понятий в рамках контекстов  $\langle \Omega(Z), \{G(\tau)\} \rangle$  и  $\langle \Omega(Z), \{S\} \rangle$ ; формирование моделей понятий (концептов) на основе положительного контекста;

- построение полных (предельных) синдромных и вероятностных моделей эмпирических знаний в рамках контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ ;
- построение минимальных предельных синдромных и вероятностных моделей эмпирических знаний в рамках контекста  $\langle \Omega(Z), \{G(\tau)\} \rangle$ ;
- определение критических описаний базы прецедентов;
- формирование дуальных моделей знаний в рамках контекста  $\langle \Omega(Z_j), \{G(\tau)\} \rangle$ , где  $j=1, \dots, N$ ;
- построение моделей исключительного сходства прецедентов на основе предельных синдромов;
- построение моделей знаний на основе временных рядов; прогнозирование многомерных временных рядов на основе моделей знаний;
- применение метода «Штрихкода» для квантования непрерывных базовых доменов;
- оценка параметрической устойчивости синдромных и вероятностных моделей знаний;
- выделение директивной зоны (обязательной для всех прецедентов) и зоны возможностей в моделях знаний;
- построение орграфов набросков разного типа для визуальных образов и сигналов –  $G_s(W)$ ;
- построение базы визуальных образов  $\Omega(\{W\})$  или базы сигналов;
- построение экстремального пограничного слоя набросков для заданного образа из базы прецедентов;
- построение предельных синдромных перцептивных моделей знаний на основе орграфов набросков образов и базы прецедентов  $\Omega(\{W\})$ ;
- решение задачи классификации для новой ситуации или нового образа на основе базы прецедентов  $\Omega$ ; вычисление мер сходства;
- построение иерархии или орграфа набросков математических моделей;
- построение орбит динамических систем на разных уровнях общности (в рамках разных фазовых пространств, формируемых Банком тестов);
- построение множества системопаттернов для заданной формулы;
- построение модели ПрО на основе системопаттернов и Банка тестов (например, в формате лексических деревьев или в XML формате);
- реализация прямого и обратного вывода на базе модели ПрО (Многоцелевого банка знаний);
- построение и анализ работы сред радикалов (на основе моделей знаний) при разных критериях оптимальности и разных схемах активизации среды (распределенная, локализованная, смешанная);
- построение и анализ работы агентной системы на основе Многоцелевого банка знаний ПрО;

- реализация интеллектуального интерфейса на основе Многоцелевого банка знаний, включая диалог на ограниченном естественном или профессиональном языке;
- оценка энергетических и временных затрат на реализацию синдромов и вероятностных закономерностей;
- реализация синдромного принципа управления текущей ситуацией (диагностика состояния процесса, выработка и обоснование управления);
- формирование критических путей решения Z-задачи (критической последовательности системоквантов радикалов или системопаттернов);
- определение индекса устойчивости текущей фазы процесса;
- реализация функций автоматического контроля качества работы специалиста (оператора, врача) с использованием когнитивных агентов;
- реализация консилиумов когнитивных агентов для решения Z-задачи, включая сопряжение банков тестов, моделей знаний и т.д.;
- создание когнитивных тренажеров для заданного класса Z-задач (задачи диагностики, прогнозирования, управления);
- проектирование и разработка интеллектуальных персональных систем по направлениям (обучение, медицина, секретари, собеседники и т.д.);
- построение когнитивной информационной системы, обеспечивающей поддержку решения заданного класса Z-задач;
- реализация методов психосемантики на основе парадигмы предельных обобщений (например, метода семантического дифференциала).

Эффективным средством интенсификации обучения являются когнитивные тренажеры. Они полностью адекватны мышлению специалиста, поэтому у студентов не возникает психологического дискомфорта. Функциональные задачи когнитивных тренажеров в целом совпадают с представленной выше программой. Исходные данные для многих учебных приложений брались из библиотеки [201].

Предложенная концепция модернизации обучения является обобщением многолетнего опыта преподавания на кафедре медицинской кибернетики и вычислительной техники Днепропетровской государственной медицинской академии, а также на кафедре информационных технологий и кибернетики (позже – специализированных компьютерных систем) Украинского государственного химико-технологического университета.

Парадигма предельных обобщений задает общую направленность пути улучшения образования. В этом плане книгу можно рассматривать как проект системы повышения эффективности практического интеллекта.

## ЛИТЕРАТУРА

1. Авдеева З. К. Когнитивное моделирование для решения задач управления слабоструктурированными системами (ситуациями) / З. К. Авдеева, С. А. Коврига, Д. И. Макаренко // UBS. – 2007. – №16. – С. 26 – 39.
2. Автоматическое порождение гипотез в интеллектуальных системах / Под общ. ред. В. К. Финна. – М. : Книжный дом «Либроком», 2009.
3. Александрович Ю. С. Оценочные и прогностические шкалы в медицине критических состояний / Ю. С. Александрович, В. И. Гордеев. – СПб : Изд-во «Сотис», 2007. – 140 с.
4. Алпатов А. П. Госпитальные информационные системы: архитектура, модели, решения / А. П. Алпатов, Ю. А. Прокопчук, В. В. Костра. – Дн-ск : УГХТУ, 2005. – 257 с.
5. Алпатов А. П. Информационные технологии в образовании и здравоохранении / А. П. Алпатов, Ю. А. Прокопчук, О. В. Юденко, С. В. Хорошилов. – Дн-ск : ИТМ НАНУ, 2008. – 287 с.
6. Алпатов А. Опыт эксплуатации информационной системы участковой поликлиники / А. Алпатов, С. Горбаненко, Е. Каминская, Ю. Прокопчук, В. Харченко, Я. Шкрюм // Укр. журн. мед. техніки і технології. –1999. – № 1-2. – С. 63 – 74.
7. Алпатов А. П. Статистическое моделирование зависимостей энергетических возможностей и стоимости носителей от их технических характеристик / А. П. Алпатов, В. И. Кузнецов, А. П. Сарычев // Межведомственный сборник «Системные технологии», 2007. – С. 20 – 33.
8. Алюшин А. Л. Темпомиры: Скорость восприятия и шкалы времени / А. Л. Алюшин, Е. Н. Князева. – М. : Издательство ЛКИ, 2008. – 240 с.
9. Амосов Н. Н. Алгоритмы разума / Н. Н. Амосов. – К. : Наукова думка, 1979. – 221 с.
10. Андон П. И. Логические модели интеллектуальных информационных систем / П. И. Андон, А. Е. Яшунин, В. А. Резниченко. – Киев : Наукова думка, 1999. – 396 с.
11. Анохин П. К. Принципиальные вопросы общей теории функциональных систем / П. К. Анохин. – М. : Наука. 1971.
12. Аншаков О. М. ДСМ-метод и модификационные вычисления / О. М. Аншаков // Искусственный интеллект и принятие решений. – 2008. – №1. – С. 55 – 79.
13. Апонин Ю. М. Иерархия моделей математической биологии и численно-аналитические методы их исследования / Ю. М. Апонин, Е. А. Апонина // Математическая биология и биоинформатика. – 2007. – Т. 2, №2. – С. 347 – 360.
14. Аракелов К. С. Моделирование динамики запутанных квантовых состояний / К. С. Аракелов, Ю. И. Ожигов // Вестник Московского университета. Серия 15 : Вычислительная математика и кибернетика. – 2008. – № 4. – С. 34 –39.
15. Артемьева И. Л. Иерархическая модель онтологии физической химии. Часть I. Модель метаонтологии “Сущности”/ И. Л. Артемьева, В. А. Цветников, В. А. Реутов. – Владивосток : ИАПУ ДВО РАН, 2001. – 24 с. – (Препринт / ИАПУ ДВО РАН ; № 11).
16. Артемьева И. Л. Спецификации задач в терминах онтологии химии и методы их решения / И. Л. Артемьева, Н. В. Рештаненко. – Владивосток : ИАПУ ДВО РАН, 2006. – 36 с.

17. Архитектура виртуальных миров : коллективная монография. 2-ое издание / под редакцией М. Б. Игнатьева, А. В. Никитина, А. Е. Войскунского. – Санкт-Петербург, 2009.
18. Ателье ER, Калининград [Электронный ресурс] – Режим доступа [www.metaphor.ru](http://www.metaphor.ru).
19. Баксанский О. Е. Когнитивный образ мира / О. Е. Баксанский, Е. Н. Кучер. – М. : РООИ «Реабилитация», 2010. – 224 с.
20. Белинский А. В. Квантовая нелокальность и отсутствие априорных значений измеряемых величин в экспериментах с фотонами / А. В. Белинский // УФН. – Т. 173. № 8. – 2003. – С. 905.
21. Белов Ю. А. Память человека и мышление – образное и символическое. Концептуальное модельное представление / Ю. А. Белов, З. Л. Рабинович // Доклады Академии Наук (РФ) : Секция «Информатика». – 2009. – Т. 427, № 6. – С. 761 – 764.
22. Бериша А. М. Методы обнаружения знаний в «зашумленных» базах данных / А. М. Бериша, В. Н. Вагин, А. В. Куликов, М. В. Фомина // Известия РАН. Теория и системы управления. – 2005. – №6. – С. 143 – 158.
23. Бескова И. А. Феномен сознания / И. А. Бескова, И. А. Герасимова, И. П. Меркулов. – Изд-во : Прогресс-Традиция, 2010. – 367 с.
24. Блинова В. Г. Языки представления химических структур в интеллектуальных системах для конструирования лекарств / В. Г. Блинова, Д. А. Добрынин // НТИ, сер.2. – 2000. – №6.
25. Бобровников Л. В. Основы квантовой синергетики функциональных систем. Ч. I / Л. В. Бобровников. – М. : Российская академия естественных наук, 2008. – 187 с.
26. Бондаренко М. Ф. Теория интеллекта. Учебник для вузов / М. Ф. Бондаренко, Ю. П. Шабанов-Кушнаренко. – Харьков : СМИТ, 2007. – 576 с.
27. Борисов В. В. Нечеткие модели и сети / В. В. Борисов, В. В. Круглов, В. В. Федулов. – М. : Горячая линия – Телеком, 2007. – 284 с.
28. Боумейстер Д. Физика квантовой информатики / Д. Боумейстер, А. Экерт, А. Цайлингер. – М. : "Постмаркет", 2002. – 376 с.
29. Буданов В. Г. Квантовая механика и проблема сознания: перспектива междисциплинарного сотрудничества / В. Г. Буданов, И. А. Герасимова // Эпистемология и философия науки. – 2005. – №4. – С. 56 – 63.
30. Бухановский А. В. Интеллектуальные высокопроизводительные программные комплексы моделирования сложных систем: концепция, архитектура и примеры реализации / А. В. Бухановский, С. В. Ковальчук, С. В. Марьин // Изв. Вузов. Приборостроение. – 2009. – Т. 52, № 10. – С. 5 – 24.
31. Вагин В. Н. Дедукция и обобщение в системах принятия решений / Вагин В. Н. – М. : Наука, 1988. – 383 с.
32. Вайнцвайг М. Н. Архитектура мыслящей системы и нейронные сети / М. Н. Вайнцвайг, М. П. Полякова // Сб. Интеллектуальные процессы и их моделирование. Информационные сети. – М. : ИППИ РАН, 1994. – С. 132 – 151.
33. Валиев К. А. Квантовые компьютеры и квантовые вычисления / К. А. Валиев // УФН. – 2005. – Т.175, №3.
34. Валькман Ю. Р. Модельно-параметрическое пространство – средство представления знаний исследователей сложных систем / Ю. Р. Валькман,

- А. Ю. Рыхальский // УСиМ. – 2009. – №1. – С. 20 – 30.
35. Валькман Ю. Р. О моделировании образного мышления в компьютерных технологиях: общие закономерности мышления / Ю. Р. Валькман, В. С. Быков // Знания–Диалог–Решение : международная конференция KDS-2005, 2005, Варна, Болгария : сб. трудов. – София : ИТА, 2005. – Т.2. – С. 37 – 45.
  36. Васильев В. И. Интеллектуальные системы управления. Теория и практика / В. И. Васильев, Б. Г. Ильясов. – М : URSS, 2009. – 392 с.
  37. Васильев В. И. Принцип редукции в задачах обнаружения закономерностей / В. И. Васильев, А. И. Шевченко, С. Н. Эш. – Донецк : ИПШ «Наука і освіта», 2009. – 340 с.
  38. Величковский Б. М. Когнитивная наука: Основы психологии познания : в 2 т. Т. 1 / Б. М. Величковский. – М. : Смысл : Издательский центр «Академия», 2006. – 448 с.
  39. Витяев Е. Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов : моногр. / Е. Е. Витяев. – Новосибирск : НГУ, 2006. – 293 с.
  40. Воронков Г. С. Механизмы решения задач в элементарном сенсориуме: нейронные механизмы опознания и сенсорного обучения / Г. С. Воронков // Нейрокомпьютеры: разработка и применение. – 2004. – № 2-3. – С. 92 – 100.
  41. Гаврилова Т. А. Базы знаний интеллектуальных систем / Т. А. Гаврилова, В. Ф. Хорошевский. – СПб : Питер, 2000. – 384 с.
  42. Гарднер Г. Структура разума. Теория множественного интеллекта / Г. Гарднер. – М., 2007.
  43. Генкин А. А. Новая информационная технология анализа медицинских данных (программный комплекс ОМИС) / А. А. Генкин. – СПб. : Политехника, 1999. – 191 с.
  44. Гладун В. П. Процессы формирования новых знаний / В. П. Гладун. – София : СД "Педагогъ", 1994. – 192 с.
  45. Гладун В. П. Партнерство с компьютером. Человеко-машинные целеустремленные системы / В. П. Гладун. – Киев : Port-Royal, 2000. – 128 с.
  46. Гостев В. И. Нечеткие регуляторы в системах автоматического управления / В. И. Гостев. – К : «Радиоаматор», 2008. – 972 с.
  47. Гольдштейн С. Л. Разрешение проблемных ситуаций при поддержке систем, основанных на знаниях / С. Л. Гольдштейн, А. Г. Кудрявцев. – Екатеринбург : ИД «Пироговъ», 2006. – 218 с.
  48. Горбань П. А. Нейросетевой анализ структуры индивидуального пространства смыслов / П. А. Горбань // «Нейрокомпьютеры»: разработка, применение. – 2002. – № 4. – С. 14 – 19.
  49. Горбань А. Н. Нейроинформатика / А. Н. Горбань, В. Л. Дунин-Барковский, А. Н. Кирдин и др. – Новосибирск : Наука. Сибирское предприятие РАН, 1998. – 296 с.
  50. Горский Ю. М. Гомеостатика: гармония в игре противоречий / Ю. М. Горский, А. М. Степанов, А. Г. Теслинов. – Иркутск : Издательство «Репроцентр А1», 2008. – 634 с.
  51. Горский Д. П. Обобщение и познание / Д. П. Горский. – М. : Мысль, 1985. – 208 с.
  52. Давиденко А. М. Новые методы изучения действующих производств и их возможности / А. М. Давиденко, М. Д. Кац // Восточно-Европейский журнал передовых технологий. – 2004. – №6 (12). – С. 189 – 193.
  53. Деннетт Д. С. Виды психики: на пути к пониманию сознания / Д. С. Деннетт ;



- пер. с англ. А. Веретенникова. – М. : Идея-Пресс, 2004. – 184 с.
54. Дернер Д. Логика неудачи: Стратегическое мышление в сложных ситуациях : пер. с нем. / Д. Дернер. – М. : Смысл, 1997. – 237 с.
  55. Дрюк М. А. Химическая теория резонанса и проблемы формирования философской рациональности нового типа / М. А. Дрюк. – М. : Изд. РУДН, 2003. – 181 с.
  56. Дубровин В. И. Интеллектуальные средства диагностики и прогнозирования надежности авиадвигателей / В. И. Дубровин, С. А. Субботин, А. В. Богуслаев, В. К. Яценко. – Запорожье : ОАО "Мотор Сич", 2003. – 279 с.
  57. Дубровский Д. И. Сознание, мозг, искусственный интеллект / Д. И. Дубровский. – М. : Стратегия-Центр, 2007. – 272 с.
  58. Досин Д. Г. Интеллектуальні системи, базовані на онтологіях / Д. Г. Досин, В. В. Литвин, Ю. В. Нікольський, В. В. Пасічник. – Львів : «Цивілізація», 2009. – 414 с.
  59. Дюк В. А. Информационные технологии в медико-биологических исследованиях / В. А. Дюк, В. Л. Эмануэль. – СПб : Питер, 2003. – 528 с.
  60. Евин И. А. Синергетика мозга и синергетика искусства / И. А. Евин. – М. : РХД, 2005.
  61. Жданов А. А. Автономный искусственный интеллект / А. А. Жданов. – М. : БИНОМ. Лаборатория знаний, 2008. – 359 с.
  62. Загоруйко Н. Г. Прикладные методы анализа данных и знаний / Н. Г. Загоруйко. – Новосибирск : ИМ СО РАН, 1999. – 270 с.
  63. Заде Л. Понятие лингвистической переменной и его применение к принятию приближенных решений / Л. Заде. – М. : Мир, 1976. – 165 с.
  64. Заполовский Н. И. Разработка и исследование системы управления электроприводом переменного тока с использованием методов нечеткой логики / Н. И. Заполовский, В. И. Носков, Н. В. Мезенцев, Н. В. Горбач // Вестник НТУ "ХПИ". Тематический выпуск : Информатики и моделирование. – Харьков : НТУ "ХПИ". – 2006. – № 23. – С. 53 – 60.
  65. Зурек В. Декогеренция и переход от квантового мира к классическому / В. Зурек // Los Alamos Science. – 2002. – №27.
  66. Иваницкий А. М. Главная загадка природы: как на основе работы мозга возникают субъективные переживания / А. М. Иваницкий // Психологический журнал. – 1999. – Т. 20, № 3. – С. 93 – 104.
  67. Иваницкий А. М. Мозговая основа субъективных переживаний: гипотеза информационного синтеза / А. М. Иваницкий // Журнал высшей нервной деятельности. – 1996. – Т. 46, № 2.
  68. Ивашев С. П. Системное квантование мыслительной деятельности человека / С. П. Ивашев. – Волгоград : Волг. ГМУ, 2005. – 229 с.
  69. Игнатов Д. И. Бикластеризация объектно-признаковых данных на основе решеток замкнутых множеств / Д. И. Игнатов, С. О. Кузнецов // 12-ая национальная конференция по искусственному интеллекту с международным участием : сб. трудов. – М. : Физматлит, 2008. – Т.1. – С. 175 – 182.
  70. Илюхина В. А. Иерархический принцип организации и феномен интеграции разных по скоростям мозговых процессов в механизмах познавательной деятельности человека / В. А. Илюхина // IV-й конференция по когнитивным наукам, июнь, 2010 г., г. Томск : тезисы докладов. – Томск : Томский государственный университет, 2010. – Т. 2. – С. 300.

71. Интеллектуальные системы автоматического управления / под ред. И. М. Макарова, В. М. Лохина. – М. : ФИЗМАТЛИТ, 2001. – 576 с.
72. Итеративное агрегирование и его применение в планировании / Под ред. Л. М. Дудкина. – М. : Экономика, 1979. – 328 с.
73. Каляев И. А. Модели и алгоритмы коллективного управления в группах роботов / И. А. Каляев, А. Р. Гайдук, С. Г. Капустян. – М. : Физматлит, 2009.
74. Карпов А. В. Психология метакогнитивных процессов личности / А. В. Карпов, И. М. Скитяева. – М. : Изд-во «Институт психологии РАН», 2005. – 352 с.
75. Кац М. Д. Разработка методов идентификации и субоптимизации для управления технологическими процессами малотоннажной химии : автореферат диссертации на соискание учёной степени доктора технических наук / М. Д. Кац. – Харьков, 1992. – 40 с.
76. Качанова Т. Л. Метатехнология системных реконструкций / Т. Л. Качанова, Б. Ф. Фомин. – СПб : Изд-во СПбГЭТУ «ЛЭТИ», 2001. – 336 с.
77. Киселева Н. Н. Компьютерное конструирование неорганических соединений. Использование баз данных и методов искусственного интеллекта / Н. Н. Киселева. – М. : Наука, 2005. – 288 с.
78. Клещев А. С. Задачи индуктивного формирования знаний в терминах непримитивных онтологий предметных областей / А. С. Клещев. – Владивосток : ИАПУ ДВО РАН, 2003. – 35 с.– (Препринт/ ИАПУ ДВО РАН ; № 6).
79. Клещев А. С. Математические модели онтологий предметных областей. Ч. 1. Существующие подходы к определению понятия «онтология» / А. С. Клещев, И. Л. Артемьева // Научно-техническая информация : Серия 2. Информационные процессы и системы. – 2001. – № 2. – С. 20 – 27.
80. Клінічні рекомендації для лікарів з питань організації та надання медичної допомоги хворим з гостримі хірургічними захворюваннями органів живота : відомча інструкція. – К. : Дніпро-VAL, 2004. – 353 с.
81. Колесников А. А. Синергетические методы управления сложными системами: Теория системного синтеза / А. А. Колесников. – М. : КомКнига, 2006. – 240 с.
82. Колешко В. М. Когнитивная многомерность – основа синергизма интеллектуальной технологии поиска знаний / В. М. Колешко, А. В. Гулай, В. А. Гулай // Высшэйшая школа. – 2010. – N 2. – С. 20 – 25.
83. Кроновер Р. М. Фракталы и хаос в динамических системах. Основы теории / Р. М. Кроновер. – М. : Постмаркет, 2000. – 353 с.
84. Кузнецов С. О. Методы теории решеток и анализа формальных понятий в машинном обучении / С. О. Кузнецов // Новости искусственного интеллекта. – 2004. – №3. – С. 19 – 31.
85. Лакофф Джордж Женщины, огонь и опасные вещи: Что категории языка говорят нам о мышлении / Джордж Лакофф ; пер. с англ. И. Б. Шатуновского. – М. : Языки славянской культуры, 2004. – 792 с.
86. Ларичев О. И. Теория и методы принятия решений, а также Хроника событий в Волшебных странах : учебник / О. И. Ларичев. – М. : Университетская книга, Логос, 2006. – 392 с.
87. Лачинов В. М. Информодинамика или путь к Миру Открытых систем / В. М. Лачинов, А. О. Поляков. – СПб. : Издательство СПбГТУ, 1999.
88. Лбов Г. С. Устойчивость решающих функций в задачах распознавания образов

- и анализа разнотипной информации / Г. С. Лбов, В. Б. Бериков. – Новосибирск : Изд-во Ин-та математики, 2005. – 218 с.
89. Леонтьев Д. А. Психология смысла: природа, строение и динамика / Д. А. Леонтьев. – М. : Смысл, 2007. – 512 с.
  90. Липков А. И. Проблема художественного воздействия: принцип аттракциона / А. И. Липков. – М. : Наука, 1990.
  91. Литвин В. В. Бази знань інтелектуальних систем підтримки прийняття рішень / В. В. Литвин. – Львів : Вид. Львівської політехніки, 2011. – 240 с.
  92. Литвиненко А. Е. Метод направленного перебора в системах управления и диагностирования / А. Е. Литвиненко. – К. : НБУ им. Вернадского, 2007. – 328 с.
  93. Луценко Е. В. Семантические информационные модели управления агропромышленным комплексом : монография / Е. В. Луценко, В. И. Лойко. – Краснодар : КубГАУ, 2005. – 480 с.
  94. Лукашевич И. П. Роль структурной организации знаний при анализе биохимических данных / И. П. Лукашевич, М. И. Савина // Искусственный интеллект и принятие решений. – 2010. – № 1. – С. 33 – 40.
  95. Люгер Джорж. Ф. Искусственный интеллект. Стратегии и методы решения сложных проблем : 4-е издание : пер. с англ. / Люгер Джорж. Ф. – М. : Издательский дом "Вильямс", 2005. – 864 с.
  96. Ма Ш. Современная теория критических явлений / Ш. Ма. – М. : Мир, 1980. – 298 с.
  97. Макаренко Д. И. Концептуальное моделирование военной безопасности государства : монография / Д. И. Макаренко, Е. Ю. Хрусталева. – М. : Наука, 2008. – 303 с.
  98. Максимов В. И. Когнитивные технологии – от незнания к пониманию / В. И. Максимов // Когнитивный анализ и управление развитием ситуаций (CASC'2001) : 1–я международная конференция : сб. статей. – М. : ИПУ РАН, 2001. – С. 4 – 41.
  99. Малинецкий Г. Г. Математические основы синергетики: Хаос, структуры, вычислительный эксперимент : изд. 5-е / Г. Г. Малинецкий. – М. : Издательство ЛКИ, 2007. – 312 с.
  100. Малинецкий Г. Г. Когнитивные центры как информационные системы для стратегического прогнозирования / Г. Г. Малинецкий, С. К. Маненков, Н. А. Митин. – М. : Препринт ИПМ. – 2010. – № 46.
  101. Махортов С. Д. Математические основы искусственного интеллекта: теория LP-структур для построения и исследования моделей знаний продукционного типа / С. Д. Махортов ; под ред. В. А. Васенина. – М. : Издательство МЦНМО, 2009. – 300 с.
  102. Мельник О. Г. Інноваційні системи економічної діагностики підприємств на засадах індикаторів. Теоретико-методологічні та методичні засади / О. Г. Мельник, І. Б. Олексів, Н. Ю. Подольчак, Р. В. Шуляр. – Львів : Видавництво Львівської політехніки, 2009. – 212 с.
  103. Менский М. Б. Квантовые измерения и декогеренция / М. Б. Менский. – М. : ФИЗМАТЛИТ, 2001.
  104. Менский М. Б. Человек и квантовый мир. Странности квантового мира и тайна сознания / М. Б. Менский. – Фрязино : ВЕК2, 2005. – 320 с.
  105. Моисеев В. Н. Логика Открытого Синтеза : в 2 томах / В. Н. Моисеев. – СПб. :

Издательский дом «Мирь», 2010.

106. Мюллер Х. Составление ментальных карт. Метод генерации и структурирования идей / Х. Мюллер. – Издательство : ОМЕГА-Л, 2007. – 128 с.
107. Назаренко Г. И. Медицинские информационные системы и искусственный интеллект : науч. пособие / Г. И. Назаренко, Г. С. Осипов. – М. : Медицина XXI, 2003. – 240 с.
108. Найсер У. Познание и реальность. Смысл и принципы когнитивной психологии / У. Найсер. – М. : "Прогресс", 1981. – 230 с.
109. Настенко Е. А. Гемодинамические детерминанты неинвазивной оценки адекватности коронарного кровоснабжения у кардиохирургических больных / Е. А. Настенко // Доповіді Академії Наук України. – 1999. – № 3. – С. 174 – 180.
110. Настенко Е. А. Методы разделения смесей многомерных функциональных характеристик при системном анализе сложных биомедицинских процессов / Е. А. Настенко // VI наук. конф. Асоц. серцево-судинних хірургів України : зб. наук. праць. – К., 1998. – С. 163 – 165.
111. Наумов Л. Б. Болезни органов живота. Алгоритмы дифференциальной диагностики, лечения, врачебно-трудовой экспертизы / Л. Б. Наумов, Р. С. Мамиконян, В. Л. Наумов. – Ереван : «Айастан», 1990. – 456 с.
112. Неделько В. Н. Прогнозирование информационных потребностей операторов особо сложных систем управления / В. Н. Неделько // Искусственный интеллект. – 2002. – №3. – С. 420 – 427.
113. Нечаев Ю. И. Концептуальные основы создания бортовых интеллектуальных систем / Ю. И. Нечаев // Информационно-измерительные и управляющие системы. Ч. 2. Корабельные системы. – 2006. – № 9. – С. 39 – 49.
114. Нечаев Ю. И. Концепция и методологические основы создания интеллектуального базиса ГРИД-систем / Ю. И. Нечаев, А. В. Бухановский, В. Н. Васильев // Научно-технический вестник СПбГУ ИТМО «Технологии высокопроизводительных вычислений и компьютерного моделирования». – 2008. – Вып. 54. – С. 13 – 28.
115. Никаноров С. П. Концептуализация предметных областей / С. П. Никаноров. – М. : Концепт, 2009. – 268 с.
116. Николис Г. Познание сложного / Николис Г., Пригожин И. – М : Мир, 1990.
117. Новиков Д. А. Теория управления организационными системами / Д. А. Новиков. – М. : МПСИ, 2005. – 584 с.
118. Овезгельдыев А. О. Синтез и идентификация моделей многофакторного оценивания и оптимизации / А. О. Овезгельдыев, Э. Г. Петров, К. Э. Петров. – К. : «Наукова думка», 2002. – 163 с.
119. Онищенко М. С. Концепт «свободная страна»: лингвокогнитивный анализ / М. С. Онищенко. – Ульяновск : УлГТУ, 2009. – 201 с.
120. Орлов В. А. Комплексная программа оценки физического состояния и функциональных возможностей организма человека / В. А. Орлов. – М. : Из-во АриНА, 1996. – 75 с.
121. Орлов В. А. Многоцелевой банк знаний. Ч. 1. Концепция и политика / В. А. Орлов, А. С. Клещёв. – Владивосток : ИАПУ ДВО РАН, 2003. – 40 с.
122. Осипенко Г. С. Введение в символический анализ динамических систем / Г. С. Осипенко, Н. Б. Ампилова. – СПб : "Издательство Санкт-Петербургского

- университета", 2005. – 237 с.
123. Осипов Г. С. Лекции по искусственному интеллекту / Г. С. Осипов. – М. : URSS, 2009. – 272 с.
  124. Панченков А. Н. Энтропия-2: Хаотическая механика / А. Н. Панченков. – Нижний Новгород : Издательство общества «Интелсервис», 2002. – 713 с.
  125. Пенроуз Р. Новый ум короля. О компьютерах, мышлении и законах физики / Р. Пенроуз. – М. : Едиториал УРСС, 2003.
  126. Петито Ж. Новые основания качественной физики / Ж. Петито, Б. Смит // Развитие знаний естественных наук и искусственного интеллекта. – Лондон : Питман, 1990. – С. 231 – 249.
  127. Петренко В. Ф. Многомерное сознание: Психосемантическая парадигма / В. Ф. Петренко. – М. : Новый хронограф, 2010. – 440 с.
  128. Переверзев-Орлов В. С. , Трунов В. Г. Синдромный анализ: новые вызовы / В. С. Переверзев-Орлов, В. Г. Трунов // Информационные Процессы (Information Processes), Электронный научный журнал. – 2008. – Т. 8 . – С. 235 – 239.
  129. Пилипенко О. В. Эффективность научно-технических проектов и программ / О. В. Пилипенко, Е. С. Переверзев, А. П. Алпатов, В. Т. Марченко, О. Н. Печенева, П. П. Хорольский. – Днепропетровск : Пороги, 2008. – 509 с.
  130. Попова З. Д. Семантико-когнитивный анализ языка / З. Д. Попова, И. А. Стернин. – Воронеж : «Истоки», 2007. – 252 с.
  131. Потапов А. С. Распознавание образов и машинное восприятие. Общий подход на основе принципа минимальной длины описания / А. С. Потапов. – СПб : Политехника, 2007.
  132. Прангишвили И. В. Энтропийные и другие системные закономерности: Вопросы управления сложными системами / И. В. Прангишвили ; Ин-т проблем управления им. В.А. Трапезникова. – М. : Наука, 2003. – 428 с.
  133. Проблема сознания в философии и науке / под ред. Д. И. Дубровского. – Москва : Канон +, 2009. – 472 с.
  134. Прокопчук Ю. А. Интеллектуальные медицинские системы: формально-логический уровень / Ю. А. Прокопчук. – Днепропетровск : ИТМ НАНУ и НК АУ, 2007. – 259 с.
  135. Прокопчук Ю. А. Когнитивное ядро госпитальных и телемедицинских систем: концепция, модели, приложения / Ю. А. Прокопчук // Укр. ж-л телемедицини та медичної телематики. – 2012. – Т.10, №1. – С. 40 – 46.
  136. Прокопчук Ю. А. Технологический аудит и субоптимизация на основе принципа предельных обобщений / Ю. А. Прокопчук, А. С. Белецкий // Системні технології. Регіональний міжвузівський збірник наукових праць. – Дніпропетровськ, 2012. – Випуск 2 (79). – С. 82 – 87.
  137. Прокопчук Ю. А. Алгоритм синдромного принципа управления / Ю. А. Прокопчук // Системный анализ. Информатика. Управление : Всеукраинская научно-практическая конференция, март, 2012, Запорожье : материалы конференции. – Запорожье : КПУ, 2012. – С. 239 – 240.
  138. Прокопчук Ю. А. Предельные синдромные и вероятностные модели знаний / Ю. А. Прокопчук // Научный вестник Херсонской государственной морской академии. – 2011. – №2(5). – С. 322 – 333.
  139. Прокопчук Ю. А. Квантовая семантика: пути реализации / Ю. А. Прокопчук //

- Математическое и программное обеспечение интеллектуальных систем : IX-й Международная научно-практическая конференция, ноябрь, 2011 г., Днепропетровск : сборник докладов. – Дн-ск : Из-во ДНУ, 2011. – С. 228 – 229.
140. Прокопчук Ю. А. Когнитивное моделирование на основе принципа предельных обобщений: методология, задачи, приложения / Ю. А. Прокопчук // Искусственный интеллект. – 2011. – №3. – С. 82 – 93.
  141. Прокопчук Ю. А. Развитие агентных технологий биомедицины как основы активизации знаний / Ю. А. Прокопчук // Философские проблемы биологии и медицины. Вып. 5. Нормативное и дескриптивное. – М.: ИФ РАН - Принтберри, 2011. – С. 231 – 233.
  142. Прокопчук Ю. А. Самоорганизация и критичность в когнитивных динамических системах / Ю. А. Прокопчук // Advanced Computer Systems and Networks : Design and Application : 5-th International Conference ACSN-2011, September 29 – October 01, 2011, Lviv, Ukraine : Proceedings of the conference. – Львов : Изд. НУ «Львівська політехніка», 2011. – С. 247 – 248.
  143. Прокопчук Ю. А. Синдромный принцип управления слабоформализованными системами и процессами / Ю. А. Прокопчук // Збірник наукових праць «Індуктивне моделювання складних систем». – Київ : МННЦ ІТС НАНУ та МОНУ, 2011. – Вып 3. – С. 152 – 164.
  144. Прокопчук Ю. А. Реализация агентных технологий и среды радикалов в развивающихся информационных системах / Ю. А. Прокопчук, В. Л. Мозолев, С. В. Чубаров, С. В. Шкваря // Информационные технологии в управлении сложными системами : научная конференция, июнь, 2011 г., Днепропетровск : сб. докладов. – Дн-ск : ИТМ – «Свидлер», 2011. – С. 150 – 152.
  145. Прокопчук Ю. А. Модель «созревания» решения и «озарения» / Ю. А. Прокопчук // Нелинейная динамика в когнитивных исследованиях : конференция, май, 2011, Нижний Новгород : труды конференции. – Нижний Новгород : ИПФ РАН, 2011. – С. 155 – 157.
  146. Прокопчук Ю. А. Методология разработки интеллектуальных приложений на основе принципа предельных обобщений / Ю. А. Прокопчук // Вестник Херсонского НТУ. – 2011. – №2(41). – С. 32 – 43.
  147. Прокопчук Ю. А. Построение логически прозрачной нейроморфной сети на основе метода предельных обобщений / Ю. А. Прокопчук, Т. П. Яровая // Інтелектуальні системи прийняття рішень і проблеми обчислювального інтелекту : міжнародна конференція, травень, 2011 р., Євпаторія : збірник наукових праць. – Херсон : ХНТУ, 2011. – Том 1. – С. 296 – 300.
  148. Прокопчук Ю. А. Модели структур виртуальной сплошной среды когнитивных динамических систем / Ю. А. Прокопчук // Нейроинформатика -2011 : XIII Всероссийская научно-техническая конференция, январь, 2011 г., Москва : сб. научных трудов. В 3-х частях. – М. : НИЯУ МИФИ, 2011. – Ч.І. – С. 254 – 263.
  149. Прокопчук Ю. А. Метод штрихкода квантования значений / Ю. А. Прокопчук // Системный анализ. Информатика. Управление : Всеукраинская научно-практическая конференция, март, 2011 г., Запорожье : материалы конференции. – Запорожье : КПУ, 2011. – С. 164 – 165.
  150. Прокопчук Ю. А. Индуктивная модель описания ситуаций действительности / Ю. А. Прокопчук // Збірник наукових праць «Індуктивне моделювання складних

- систем». – Київ : МННЦ ІТС НАНУ та МОНУ, 2010. – Вып 2. – С. 161 – 173.
151. Прокопчук Ю. А. Метод предельных обобщений – новый метод анализа в биологии, лабораторной и клинической медицине / Ю. А. Прокопчук // Постгеномные методы анализа в биологии, лабораторной и клинической медицине : I-ой Международная научно-практическая конференция, ноябрь, 2010 г., Москва: сб. материалов. – М. : Изд-во «Парк-медиа», 2010. – С. 148.
  152. Прокопчук Ю. А. Модель динамической интеллектуальной СППР / Ю. А. Прокопчук // Збірник наукових праць «Питання прикладної математики і математичного моделювання». – Дн-ск : Вид-во ДНУ, 2010. – С. 239 – 251.
  153. Прокопчук Ю. А. Онтологическое и когнитивное моделирование в рамках метода предельных обобщений / Ю. А. Прокопчук, В. Л. Мозолев, В. Л. Шараева // Математическое и программное обеспечение интеллектуальных систем : VII-я Международная научно-практическая конференция, ноябрь, 2010 г., Днепропетровск : Сб. докладов. – Дн-ск : Из-во ДНУ, 2010. – С. 184 – 185.
  154. Прокопчук Ю. А. Автоматизмы когнитивных динамических систем / Ю. А. Прокопчук // Искусственный интеллект. – 2010. – №4. – С. 34 – 43.
  155. Прокопчук Ю. А. Информационная структура теории естественной предметной области / Ю. А. Прокопчук // Вестник Херсонского НТУ. – 2010. – №2(38). – С. 11 – 19.
  156. Прокопчук Ю. А. Архитектура многоцелевых банков знаний в области клинической медицины / Ю. А. Прокопчук // Знання – Онтології – Теорії – 2009 : Всероссийская конференция, октябрь, 2009 г., Новосибирск : Сб. трудов в 2-х томах. – Новосибирск : ЗАО "РИЦ Прайс-Курьер", 2009. – Т.1. – С. 173 – 177.
  157. Прокопчук Ю. А. Интеллектуальное синергетическое управление динамическими системами / Ю. А. Прокопчук // Искусственный интеллект. – 2009. – №4. – С. 12 – 21.
  158. Прокопчук Ю. А. Метод предельных обобщений для решения слабо формализованных задач / Ю. А. Прокопчук // Управляющие системы и машины. – 2009. – №1. – С. 31 – 39.
  159. Прокопчук Ю. А. Оценка течения заболевания на основе метода предельных обобщений / Ю. А. Прокопчук, Е. В. Сергиени, О. А. Харченко, А. В. Татьянаенко, В. А. Челашеев // Математическое и программное обеспечение интеллектуальных систем : VII-й Международная конференция , ноябрь, 2009 г., Днепропетровск : сб. докладов – Дн-ск : Из-во ДНУ, 2009. – С. 235 – 236.
  160. Прокопчук Ю. А. Консилиум интеллектуальных систем / Ю. А. Прокопчук // Искусственный интеллект. – 2007. – №4. – С. 491 – 496.
  161. Прокопчук Ю.А. Компьютерный анализ логических ошибок врача / Ю. А. Прокопчук // Украинский журнал телемедицины и медицинской телематики. – 2007. – Т.5, №1. – С. 85 – 88.
  162. Рабинович З. Л. О естественных механизмах мышления и интеллектуальных ЭВМ / З. Л. Рабинович // Кибернетика и системный анализ. – 2003. – № 5. – С. 82 – 88.
  163. Рабинович З. Л. Концептуальное представление об опознании образов и решении проблем в памяти человека и возможностях его использования в искусственном интеллекте / З. Л. Рабинович// Knowledge – Dialogue – Solution : XI-th International Conference FOI-Commerce, 2005, Sofia : Proceedings. – Sofia,

2005. – Vol. 1. – P. 1 – 8.

164. Райков А. Н. Сетевые ситуационные центры – новая волна / А. Н. Райков // НТИ. Сер. 2. – 2009. – № 11. – С. 10 – 17.
165. Редько В. Г. Эволюция, нейронные сети, интеллект. Модели и концепции эволюционной кибернетики / В. Г. Редько. – М.: КомКнига, 2005.
166. Романовский Ю. М. Математическое моделирование в биофизике / Ю. М. Романовский, Н. В. Степанова, Д. С. Чернавский. – Москва-Ижевск : Институт компьютерных исследований, 2003. – 402 с.
167. Саати Т. Анализ иерархических процессов / Т. Саати – М.: Радио и связь, 1993. – 315 с.
168. Самойлов В. Д. Модельное конструирование компьютерных приложений / В. Д. Самойлов. – К.: Наукова думка, 2007. – 198 с.
169. Сарычев А. П. Идентификация состояний структурно-неопределенных систем / А. П. Сарычев. – Днепропетровск : Ин-т технической механики НАН Украины и НКА Украины, 2008. – 268 с.
170. Сироджа И. Б. Квантовые модели и методы искусственного интеллекта для принятия решений и управления / И. Б. Сироджа. – К.: Наукова думка, 2002. – 490 с.
171. Славин М. Б. Методы системного анализа в медицинских исследованиях / М. Б. Славин. – М.: Медицина, 1989. – 304 с.
172. Смирнов В. А. Логика и клиническая диагностика. Теоретические основы / В. А. Смирнов, А. М. Анисов, Г. П. Арутюнов и др. – М.: Наука, 1994. – 297 с.
173. Соколов А. Ю. Алгебраическое моделирование лингвистических динамических систем / А. Ю. Соколов // Проблемы управления и информатики. – 2000. – №2. – С. 141 – 148.
174. Соколов А. Ю. Проектирование систем управления на ЭВМ (MATLAB/Simulink/Control System) / А. Ю. Соколов, Ю. Н. Соколов, В. М. Ильшюк, М. М. Митрахович, Д. Н. Гайсенюк. – Х.: «ХАИ», 2005. – 582 с.
175. Стефанюк В. Л. Сотрудничающий компьютер: проблемы, теории, приложения / В. Л. Стефанюк, А. В. Жожикашвили. – М.: Наука, 2007. – 274 с.
176. Судаков К. В. Системное квантование процессов жизнедеятельности / К. В. Судаков // Системокванты физиологических процессов / Под ред. К. В. Судакова. – М.: Междунар. гуманитарн. фонд арменоведения, 1997. – С. 9 – 52.
177. Судаков К. В. Системное построение динамических стереотипов головного мозга / К. В. Судаков // Успехи современной биологии. 2008. –Т.128, № 3. – С. 227 – 244
178. Теслинов А. Г. Концептуальное проектирование сложных решений / А. Г. Теслинов. – СПб: "Питер", 2009. – 288 с.
179. Финн В. К. Интеллектуальные системы и общество / В. К. Финн. – Изд. 3-е. – М.: КомКнига, 2007.
180. Хакен Г. Информация и самоорганизация. Макроскопический подход к сложным системам : пер. с англ / Хакен Г. // Синергетика: от прошлого к будущему. – Изд.2. – М.: КомКнига, 2005. – 248 с.
181. Хакен Г. Принципы работы головного мозга: Синергетический подход к деятельности мозга, поведению и пониманию / Хакен Г. – М.: "Пер Сэ", 2001. – 352 с.
182. Хамерофф С. События сознания как оркестрованный выбор пространства / С. Хамерофф, Р. Пенроуз // Парапсихология и психофизика. – 1998. – №2 (26). – С. 85 – 88.



183. Химико-технологические системы. Синтез, оптимизация и управление / Под ред. И. П. Мухлѐнова. – Л. : Химия, 1986. – 416 с.
184. Холодная М. А. Психология интеллекта: парадоксы исследования / М. А. Холодная. – СПб : Питер 2002. – 127 с.
185. Хорошевский В.Ф. OntosMiner: семейство систем извлечения информации из мультязычных коллекций документов / В.Ф. Хорошевский // Труды конференции КИИ-2004, 2004, Тверь, Россия. – Тверь : 2004.
186. Цыганков В. Д. Нейрокомпьютер и мозг : учебное пособие / В. Д. Цыганков. – М. : СИНТЕГ, 2001. – 248 с.
187. Чернавский Д. С. Синергетика и информация: Динамическая теория информации / Чернавский Д. С. – М. : Книжный дом «ЛИБРОКОМ», 2009. – 304 с.
188. Черняховская М. Ю. База знаний для системы интеллектуальной поддержки обследования больных для врача терапевта / М. Ю. Черняховская, А. С. Клещев, Ю. В. Кулаков, С. Б. Оникиенко. – Владивосток : ИАПУ ДВО РАН, 1998. – 46 с. (Препринт / ИАПУ ДВО РАН ; № 16).
189. Чечкин А. В. Радикалы и системокванты интеллектуальных систем / А. В. Чечкин // Моделирование функциональных систем / Под ред. К. В. Судакова, В. А. Викторова. – М.: ЗАО “РИТ-Экспресс”, 2000. – С. 73 – 94.
190. Чечкин А. В. Слабо формальные системы / А. В. Чечкин // Интеллектуальные системы. – 2007. – Т.11, Вып. 1- 4. – С. 137 – 158.
191. Эйген М. Гиперцикл. Принципы самоорганизации макромолекул / М. Эйген, П. Шустер ; Пер. с англ. В. М. Андреева. – М. : Мир, 1982.
192. Яхно В. Г. Базовые модели для описания процессов распознавания и осознания сенсорных сигналов / В. Г. Яхно // Нелинейная динамика в когнитивных исследованиях : Всероссийская конференция, май, 2009 г., Нижний Новгород : Тезисы докладов. – Нижний Новгород : ИПФ РАН, 2009. – С. 196 – 198.
193. Bak P. How nature works: The science of self-organized criticality / Bak P. – Springer-Verlag, New York, Inc., 1996.
194. Biologically Inspired Cognitive Architectures : Papers from the AAAI Fall Symposium, AAAI Technical Report FS-08-04 / ed. A. V. Samsonovich. – AAAI Press, Menlo Park, CA, 2008.
195. Cognitive Modeling and Agent-Based Social Simulation : Papers from the AAAI Workshop, AAAI Technical Report WS-06-02 / eds. M. A. Upal and R. Sun. – AAAI Press, Menlo Park, CA, 2006.
196. Cox Michael T. Metareasoning: Thinking about Thinking / Cox Michael T. and Raja, Anita (Eds.). Cambridge, MA : The MIT Press., 2011.
197. Ganter B. Formal Concept Analysis: Mathematical Foundations / B. Ganter, R. Wille. – Springer, 1999.
198. Ganter B. Pattern Structures and Their Projections / B. Ganter, S.O. Kuznetsov // Proc. 9th Int. Conf. on Conceptual Structures, ICCS'01, G. Stumme and H. Delugach, Eds., Lecture Notes in Artificial Intelligence, 2120, 2001.
199. Hawkins J. On intelligence / J. Hawkins, S. Blakeslee. – Publisher: Times Books, 2004. – 272 p.
200. Kauffman, Stuart. At Home in the Universe / Kauffman, Stuart. – New York: Oxford University Press, 1995.
201. Merz C.J. UCI Repository of Machine Learning Datasets // C.J. Merz,

- P. M. Murphy. – Information and Computer Science University of California, Irvine, CA, 1998. Режим доступа <http://www.ics.uci.edu/~mlearn/MLRepository.html>
202. Nguyen H. S. Discretization Methods in Data Mining / H. S. Nguyen, S. H. Nguyen // *Rough Sets in Knowledge Discovery 1 : Methodology and Application* / Ed. by L. Polkowski, A. Skowron. – Heidelberg : Physica-Verlag, 1998. – 451 – 482.
  203. Prokopchuk Iu. Multi-purpose knowledge bank // *Hospital Information Technology Europe*. – 2010. – Vol 3. – № 3. – P. 44 – 47.
  204. Prokopchuk Iu. Model of Multilevel Pattern Sketches / Iu. Prokopchuk // *3-rd International Conference on Inductive Modelling*, May, 2010, Yevpatoria, Ukraine : Proceeding. – Kyiv : IRTC ITS, 2010. – P. 51 – 54.
  205. Prokopchuk Iu. Method of Limiting Generalizations for Solving Logical and Computing Tasks / Iu. Prokopchuk // *2-nd International Conference on Inductive Modelling*, September, 2008, Kyiv, Ukraine : Proceedings – Kyiv : IRTC ITS, 2008. – P. 62 – 65.
  206. Prokopchuk Yu. Representation of Algorithmic Knowledge in Medical Information Systems / Yu. Prokopchuk, V. Kostra // *Knowledge-Based Intelligent Information and Engineering Systems (KES 2003) : 7-th International Conference*, September, 2003, Oxford, UK : Proceedings, Part II, Series: Lecture Notes in Artificial Intelligence, Vol. 2773. – Berlin : Springer-Verlag Berlin Heidelberg, 2003. – P. 1069 –1074.
  207. Prokopchuk Yu. Technology of construction of open local expert systems attached to his medical documents / Yu. Prokopchuk, V. Kostra // *2-nd European Medical and Biological Engineering Conference (EMBEC'02)*, 2002, Vienna, Austria : Proceedings. – Vienna, 2002. –Vol. 1. – P. 766 – 767.
  208. Prokopchuk Yu. The Architecture of Hospital Information Systems / Yu. Prokopchuk, V. Kostra // *International Conference on Mathematics and Engineering Techniques in Medicine and Biological Sciences (METMBS'2001)*, June, 2001, Las Vegas, Nevada, USA : Proceedings / Ed. Valafar. – Las Vegas : CSREA Press, 2001. – P. 197 – 200.
  209. Samsonovich A. V. The mental state formalism of GMU-BICA / A. V. Samsonovich, K. A. De Jong, A. Kitsantas // *International Journal of Machine Consciousness*. – 2009. – Vol. 1, No. 1. – P. 111 – 130.
  210. Samsonovich A. V. Fundamental principles and mechanisms of the conscious self / A. V. Samsonovich, L. Nadel. – *Cortex* 41, 2005. – P. 669 – 689.
  211. Sterratt D. Principles of computational modelling in neuroscience / D. Sterratt, B. Graham, A. Gillies, D. Willshaw. – Cambridge U. Pr., 2011. – 406 p.
  212. *Handbook of Granular Computing* / Ed. by W. Pedrycz, A. Skowron and V. Kreinovich. – New York: Wiley InterScience, 2008. – 1116 p.

## СОДЕРЖАНИЕ

<b>Предисловие</b>	3
<b>Список сокращений</b>	6
<b>Глава 1. Физическое и феноменологическое описание действительности</b>	7
1.1 Микроуровень и макроуровень действительности	7
1.2 Когнитивная стабильность и нестабильность	10
1.3 Когнитивная самоорганизующаяся критичность	14
1.4 Смысловая реальность. Предельные смыслы	19
1.5 Тесты, системопаттерны, системокванты	22
1.6 Принципы когнитивной парадигмы	24
<b>Глава 2. Элементарные тесты. Орграфы доменов</b>	28
2.1 Базовые определения	28
2.2 Каноническое представление орграфа доменов	47
2.3 Замыкание, верхний и нижний пределы множества данных	51
2.4 Нисходящая детерминация. Субоптимизация	59
2.5 Энтропийные характеристики орграфов доменов	71
2.6 Созревание протоструктур тестов	82
<b>Глава 3. Модель описания ситуаций действительности</b>	86
3.1 Свойства тестов. Лексика. Механизмы реализации тестов	86
3.2 Тест «Время». События	94
3.3 Структуры и их свойства	99
3.4 наброски ситуаций, образов	105
3.5 Модель развития ситуации. Импульсы. Описание ситуации	112
3.6 Информационные множества. Множества обобщения	118
3.7 Сходство ситуаций действительности	124
<b>Глава 4. Конфигураторы тестов</b>	132
4.1 Определение конфигулятора. Примеры	132
4.2 Способы формирования доменов и конфигураторов	134
4.3 Метод «Штрихкода» квантования значений	172
4.4 Реализация базы прецедентов и конфигураторов с помощью электронной таблицы	184
<b>Глава 5. Метод предельных обобщений</b>	187
5.1 Вероятностные закономерности, формальные синдромы, синдромные и вероятностные модели знаний	187
5.2 Применение синдромных и вероятностных моделей знаний	218
5.3 Критические описания базы прецедентов. Синдромные модели знаний для критических описаний	224
5.4 Дуальные синдромные модели знаний	231
5.5 Анализ формальных понятий, бикластеризация и когнитивный	235

подход	
5.6 Пример «Оценка и прогноз течения заболевания»	248
<b>Глава 6. Многоуровневые наброски образов, ситуаций, моделей</b>	251
6.1 Краткий обзор методических подходов к построению набросков	251
6.2 Многоуровневая морфология феномена действительности. Качественная прерывность	257
6.3 Орграфы набросков на основе грануляции и сжатия	259
6.4 Орграф набросков формируемый в процессе диагностического познания (на примере медицины)	270
6.5 Орграфы набросков на основе индикаторов	276
6.6 Орграфы набросков на основе оценочных и прогностических шкал	278
6.7 Квантовая семантика: пути реализации	282
<b>Глава 7. Модели познания и управления. Когнитивные системы</b>	300
7.1 Синдромный принцип управления слабоформализованными ситуациями и системами	300
7.2 Среды радикалов. Синтетическая телесно-феноменологическая онтология	314
7.3 Задачи познания предметной области. Поиск закономерностей	330
7.4 Когнитивные и метакогнитивные информационные системы	335
7.5 Концепция «Виртуальной сплошной среды» и Субъективная реальность	354
7.6 Образовательные технологии на основе принципа предельных обобщений. Когнитивные тренажеры	361
<b>Литература</b>	366
<b>Содержание</b>	379
<b>Contents</b>	381
<b>Abstracts</b>	383

## CONTENTS

<b>Preface</b>	3
<b>List of Abbreviations</b>	6
<b>Chapter 1. Physical and phenomenological description of reality</b>	7
1.1 Micro- and macrolevel of reality	7
1.2 Cognitive stability and instability	10
1.3 Cognitive Self-Organized Criticality	14
1.4 Semantic reality. Limiting meanings	19
1.5 Tests, systemopatterns, and systemoquanta	22
1.6 Principles of cognitive paradigm	24
<b>Chapter 2. Elementary tests. Domain digraphs</b>	28
2.1 Basic definitions	28
2.2 Canonical presentation of a domain digraph	47
2.3 Closure, upper and lower bounds of a data set	51
2.4 Descending determination. Suboptimization	59
2.5 Entropy characteristics of domain digraphs	71
2.6 Test protostructure maturation	82
<b>Chapter 3. Real-world situation description model</b>	86
3.1 Test properties. Lexis. Test implementation mechanisms	86
3.2 Test “Time”. Events	94
3.3 Structures and their properties	99
3.4 Situation sketches	105
3.5 Situation development model. Impulses. Situation description	112
3.6 Information sets. Generalization sets	118
3.7 Similarity of real-world situations	124
<b>Chapter 4. Test configurators</b>	132
4.1 Definition of a configurator. Examples	132
4.2 Domain and configurator formation methods	134
4.3 “Bar-code” quantization method	172
4.4 Spreadsheet implementation of a precedent and configurator base	184
<b>Chapter 5. Method of limiting generalizations</b>	187
5.1 Probabilistic regularities, formal syndromes, syndromic and probabilistic knowledge models	187
5.2 Application of syndromic and probabilistic knowledge models	218
5.3 Critical descriptions of a precedent base. Syndromic knowledge models for critical descriptions	224
5.4 Dual syndromic knowledge models	231
5.5 Analysis of formal notions, biclusterization, and cognitive approach	235
5.6 Example: assessment and prediction of a clinical course	248

<b>Chapter 6. Multilayer sketches of images, situations, and models</b>	251
6.1 Brief analytical overview of systematic approaches to sketching	251
6.2 Multilayer morphology of a real-world phenomenon. Qualitative discontinuity	257
6.3 Sketch digraphs based on granulation and contraction	259
6.4 Sketch digraph formed in the course of diagnostic cognition (by the example of medicine)	270
6.5 Sketch digraphs based on indicators	276
6.6 Sketch digraphs based on evaluation and prognostic scales	278
6.7 Quantum semantics: ways of implementation	282
<b>Chapter 7. Cognition and management models. Cognitive systems</b>	300
7.1 Syndromic principle of weak-formalized situation and system management	300
7.2 Radical media. Synthetic corporeal-phenomenological ontology	314
7.3 Subject domain cognition problems. Search for regularities	330
7.4 Cognitive and metacognitive information systems	335
7.5 Concept of a virtual continuum and subjective reality	354
7.6 Educational technologies based on the principle of limiting generalizations. Cognitive training simulators	361
<b>References</b>	366
<b>Contents</b>	379
<b>Abstracts</b>	383

**Прокопчук Ю. О. Принцип граничних узагальнень: методологія, задачі, прикладення.** Монографія. - Дніпропетровськ: ІТМ НАНУ і НКАУ, 2012.- 384 с.

У монографії описані підходи до формалізації системних уявлень про когнітивні та метакогнітивні процеси на рівні математичних моделей, комп'ютерних архітектур і програм. В основу запропонованих моделей покладено Принцип граничних узагальнень. За рахунок граничної узагальненості моделі виявляються придатними для опису безлічі інших ситуацій, образів, а також трансляції рішень в інші предметні області. Найчастіше в якості узагальнених моделей виступають параметри порядку розвитку складних ситуацій. Приводяться алгоритми побудови таких моделей.

У монографії також знайшли відображення теоретичні і практичні результати багаторічної роботи автора в області створення спеціалізованих інформаційних систем, орієнтованих на роботу з інформацією та знаннями в різних предметних областях: медицині, освіті, техніці, наукових дослідженнях.

Для фахівців та науковців в області когнітивних технологій, інтелектуальних систем, а також студентів комп'ютерних спеціальностей.

**Prokopchuk Iu. A. Principle of Limiting Generalizations: Methodology, Problems, and Applications.** Monograph. – Dnepropetrovsk: Institute of Technical Mechanics of the National Academy of Sciences of Ukraine and the State Space Agency of Ukraine, 2012. - 384 p.

This book formalizes the systems ideas of cognitive and meta-cognitive processes on the level of mathematical models, computer architectures, and computer programs. The models are based of the Principle of Limiting Generalizations. Due to its maximum degree of generality, the model proves to be suitable for the description of a variety of other similar situations and images and for the translation of solutions to other subject domains. In many cases, order parameters of the development of complex situations serve as generalized models. Algorithms for the construction of such models are given.

This book also presents the theoretical and practical results of the many years' work of the author in the development of dedicated information systems for knowledge information processing in various subject domains.

Readership: experts and researchers in the area of cognitive technologies and intelligent systems and computer science students.

Наукове видання

НАЦІОНАЛЬНА АКАДЕМІЯ НАУК УКРАЇНИ  
НАЦІОНАЛЬНЕ КОСМІЧНЕ АГЕНТСТВО УКРАЇНИ  
Інститут технічної механіки

**ПРОКОПЧУК ЮРІЙ ОЛЕКСАНДРОВИЧ**

**Принцип граничних узагальнень: методологія, задачі, прикладення**

Дніпропетровськ, Інститут технічної механіки НАН України і НКА  
України

(російською мовою)

Відповідальний за випуск А. П. Алпатов

Оригінал-макет підготовлено в Інституті технічної механіки НАНУ і НКАУ  
Свідоцтво про внесення суб'єкта видавничої справи до державного реєстру  
видавців, виготовників і розповсюджувачів видавничої продукції Серія ДК  
№634 от 17.10.2001 р. м. Дніпропетровськ, вул. Лешко-Попеля, 15.

Підписано до друку 19.04.12. Формат 70x100/16. Гарн. Times New Roman.  
Обл.-вид. арк. 19.8. Тираж 300 примірн. Замовлення № 531.

Отпечатано в типографії ООО «Барвікс».

Свідоцтво про внесення суб'єкта видавничої справи до державного реєстру  
видавців, виготовників і розповсюджувачів видавничої продукції Серія ДК  
№24 от 25.07.2000 р., 49005, м. Дніпропетровськ, вул. Сімферопольська, 17