

**Обучаемая модель
кратковременной памяти
человека на основе нейронной
сети**

к.т.н., м.н.с. лаб. Физиологии движений ИФ РАН

Ляховецкий В.А.

e-mail: v_la2002@mail.ru

Случайные последовательности

методика

Лист А4 расположен под сенсорным экраном. В шести его клетках, выбранных случайным образом, располагаются цифры, по которым экспериментатор последовательно перемещает правую или левую руку испытуемого. Испытуемый, находящийся с завязанными глазами, должен запомнить и немедленно воспроизвести движения руки.



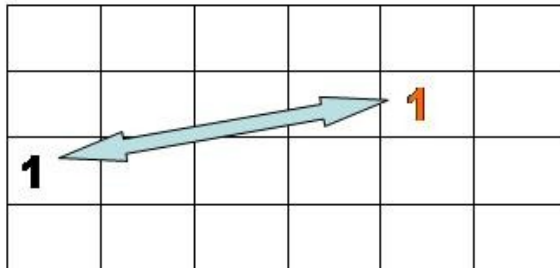
			3		6
4					
				5	
		1		2	

Ляховецкий, Боброва, 2009, 2010

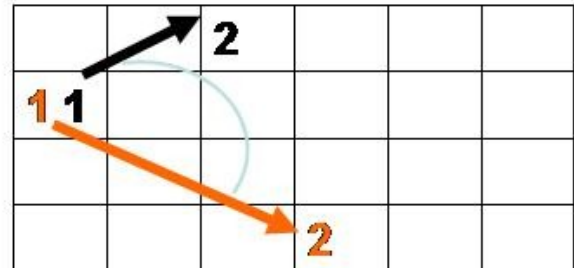
Обработка результатов

Проводилась на основе гипотезы о существовании двух типов кодирования: положений (позиционное) и движений (векторное).

ошибки положения

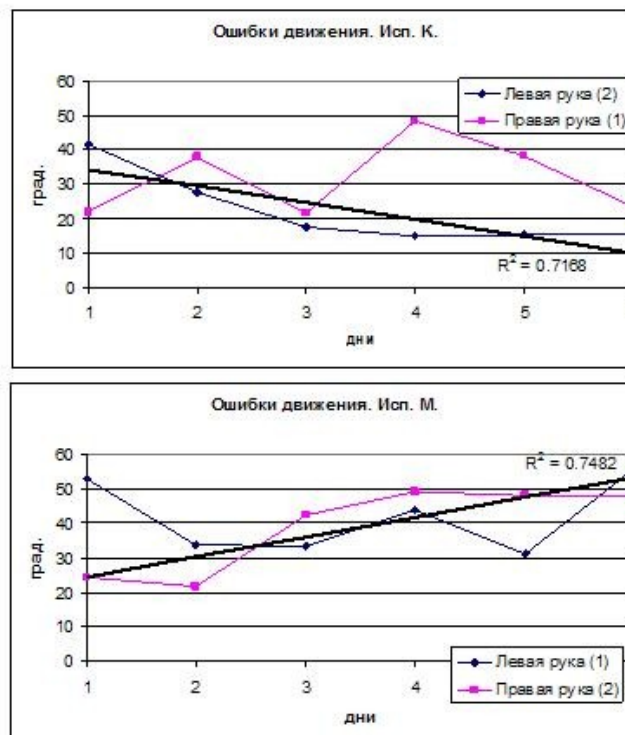


ошибки движения



В одной группе испытуемые начинали работать правой рукой, а продолжали левой, в другой группе – наоборот. Анализировали ошибки испытуемых при кратковременном заучивании, многократном (3 серии) запоминании и воспроизведении в отсутствии зрительной обратной связи одной и той же последовательности движений правой или левой руки, и долговременном обучении, 6 дней по 6 различных серий каждый день для правой или левой руки.

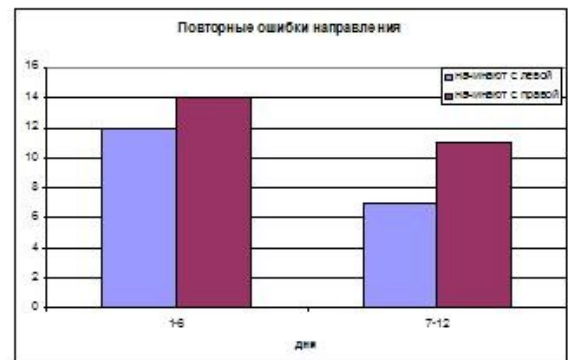
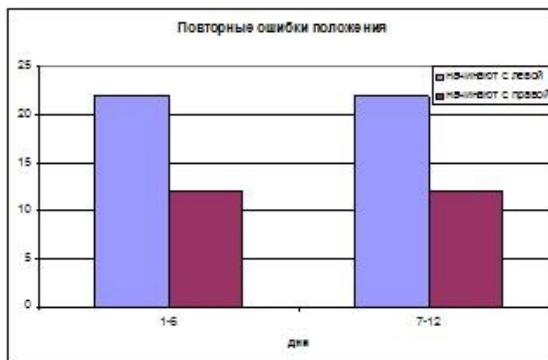
Долговременное обучение. Динамика величины ошибок.



При долговременном обучении сложному моторному навыку ошибки не всегда убывают, тем более линейно.

Скопин и др., 2008

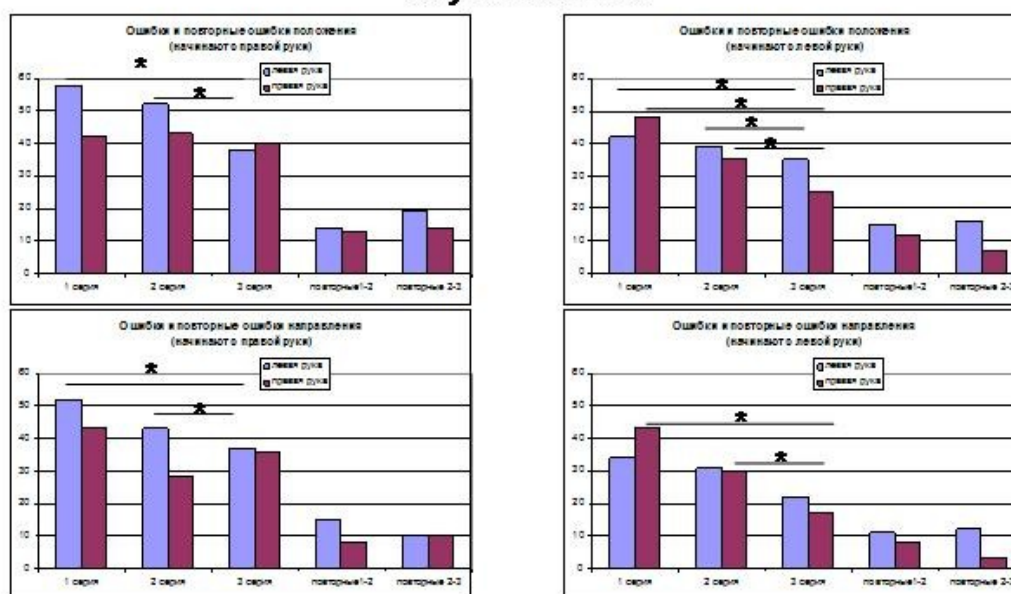
Долговременное обучение. Повторные ошибки.



Для двух последовательных серий движений выделяли одинаковые повторные ошибки по положению (испытуемый дважды неправильно воспроизводил положение руки с расхождением не более, чем на 2 см), и по направлению движения (испытуемый дважды неправильно воспроизводил направление движения руки с расхождением не более, чем на 15°).

В обеих группах при долговременном обучении для всех серий движений число повторных ошибок положения и направления значительно отличалось от случайного.

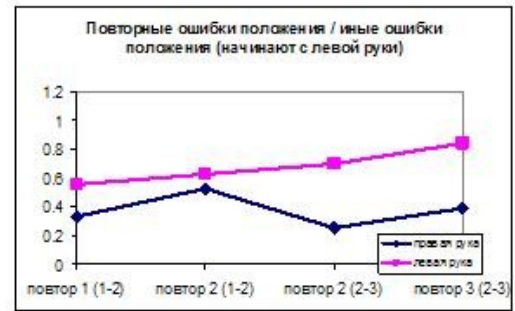
Ошибки положения и движения. Кратковременное заучивание.



При кратковременном заучивании происходит лишь обучение в группе, начинающей с левой руки, и перенос навыка с левой руки на правую. Во всех сериях число повторных ошибок отличается от случайного.

Ляховецкий, Боброва, 2010

Повторные ошибки. Кратковременное заучивание.

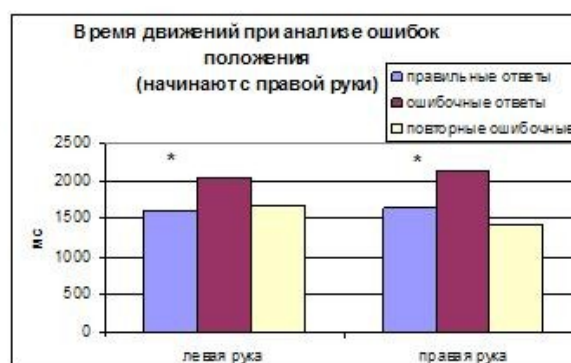
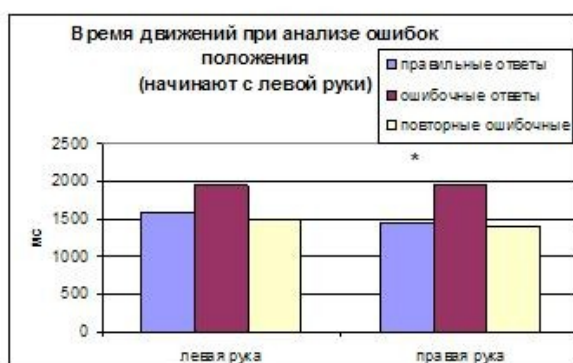


Каждая повторная ошибка характеризуется парой времен воспроизведения, каждую повторную ошибку можно сравнивать с иными ошибками в предыдущей и последующей серии движений.

Время правильных и ошибочных движений

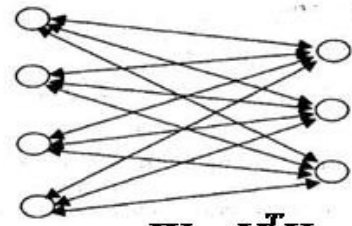
Среднее время движений, связанных с повторными ошибками, значительно меньше времени движений, связанных с иными ошибками, и не отличается от времени правильно запомненных движений.

По-видимому, при кратковременном обучении испытуемые уверены в правильности таких ошибочных движений и не пытаются корректировать их при повторном заучивании последовательности движений руки, то есть не могут отличить такие неверные ответы от правильных.



Гетероассоциативная сеть

Сохраняет пары объектов. Может восстановить объект, когда ассоциированный с ним объект предлагается ей в качестве подсказки. В сети два слоя элементов — по одному для каждого из объектов пары. Оба слоя соединяются двунаправленными связями, активность может передаваться по связям в обоих направлениях.



$$W = X^T Y$$

(для одной пары объектов)

$$W = \sum_i X_i^T Y_i$$

(для i пар объектов)

$$W = \sum_i X_i^T X_{i+1}$$

(для последовательности)

Kosko, 1988

Гетероассоциативная сеть

- 1) На первый слой подается бинаризованное представление объекта X .
 - 2) Распространяется активность на второй слой. Вход элемента второго слоя равен
 - А. $net_j = \sum_i x_i w_{ij}$
 - Б. $net_j = \sum_i x_i w_{ij} - \varphi_j$
 - 3) Вычисляется новое состояние для каждого элемента второго слоя. $y_j = f(net_j)$
 - 4) Распространяется активность на первый слой. Вход элемента первого слоя равен
 - А. $net_i = \sum_j y_j w_{ji}$
 - Б. $net_i = \sum_j y_j w_{ji} - \theta_i$
 - 5) Вычисляется новое состояние для каждого элемента первого слоя. $x_i = f(net_i)$
- Двустороннее распространение сигналов активности повторяется до тех пор, пока не будет достигнуто устойчивое состояние.
- Максимальная емкость сети:
- А. $0.5 \cdot n \cdot \log_2 n$
 - Б. $1 \div 2^n$

Для автоассоциативной сети число итераций зависит от того, насколько близок стимул к прототипу (Andersen et al., 1977).

Обучающее правило сети

Вычисление начальной матрицы весов

- Хеббовское обучение (a discrete approximation of Hebbian learning, in particular a type of Grossberg outstar avalanche, Kosko, 1988);
- Разбиение на подсети (Constantini et al., 2003);
- АВМ - асимметричная матрица весов (Xu, Leung, He, 1994);
- Exponential BAM (Jeng et al., 1990);

Изменение внутреннего представления информации

- Ортогонализация запоминаемых векторов (Wang et al., 1990);
- Использование помехоустойчивого кода;
- Переописание запоминаемых векторов (например, разбиением на чанки).

Обучение сети

- Методы линейного программирования (Wang et al., 1990);
- Генетические алгоритмы (Shen, Cruz, 2005);
- Pseudo-Relaxation learning algorithm for BAM (PRLAB) (Oh, Kothari, 1994);
- Quick learning algorithm (QLBAM) (Hattori et al., 1993);

QLBAM

1. Матрица весов W вычисляется по формуле, предложенной Коско.

2. Для первого слоя

$$\Delta W_{nm} = -\lambda(S_{xn}^{(k)} - \xi x_n^{(k)})y_m^{(k)}/(M+1), \text{ если } S_{xn}^{(k)}x_n^{(k)} \leq 0$$

$$\Delta \theta_{xn} = \lambda(S_{xn}^{(k)} - \xi x_n^{(k)})/(M+1), \text{ если } S_{xn}^{(k)}x_n^{(k)} \leq 0$$

$$W_{nm} = W_{nm} + \Delta W_{nm}$$

$$\theta_{xn} = \theta_{xn} + \Delta \theta_{xn}$$

$S_{xn}^{(k)} = \text{SUM}_m (W_{nm} y_m^{(k)} - \theta_{xn})$, k – число векторов, ξ – положительная нормирующая константа, λ – фактор релаксации, M – число нейронов второго слоя.

3. Для второго слоя матрица весов корректируется аналогичным образом.

Модель

На основе гетероассоциативной нейронной сети Коско.

•**Модель 1. Векторное кодирование.**

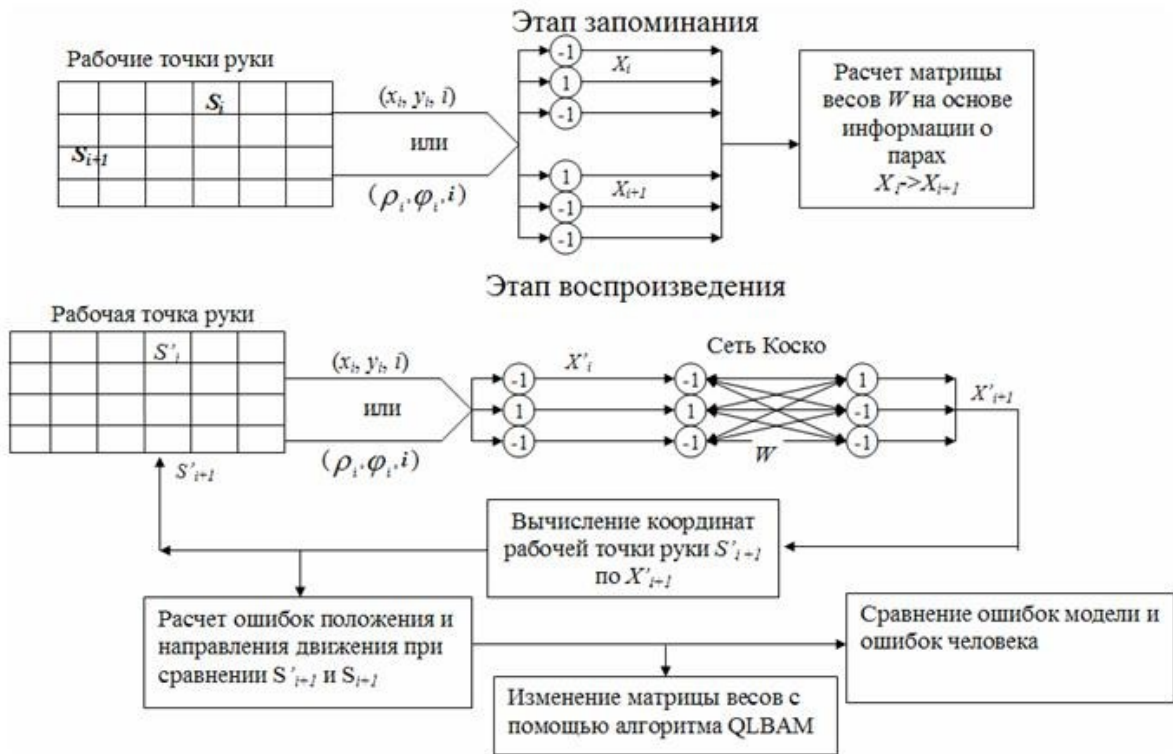
Элементами векторов являлись α и ρ . Для хранения в бинаризованной форме требуется 3 и 5 элементов соответственно (+3 элемента на номер движения).

Параметры α и ρ имеют различный физический смысл, хранятся в рамках двух подсетей, между которыми отсутствуют связи (Constantini et al., 2003). Для расчета симметричной матрицы весов $W\rho$, используемой при запоминании последовательности длин перемещений, применялся алгоритм расчета, предложенный в (Kosko, 1988). Для расчета асимметричной матрицы весов $W\alpha$, используемой при запоминании последовательности направлений перемещений, применялся алгоритм расчета, предложенный в (Xu, Leung, He, 1994).

•**Модель 2. Позиционное кодирование.**

Элементами векторов являлись x и y . Поскольку $x \in [1,6]$ и $y \in [1,4]$, для хранения их в бинаризованной форме требуется 3 и 2 элемента соответственно (+3 элемента на номер движения) (Kosko, 1988).

Схема модели

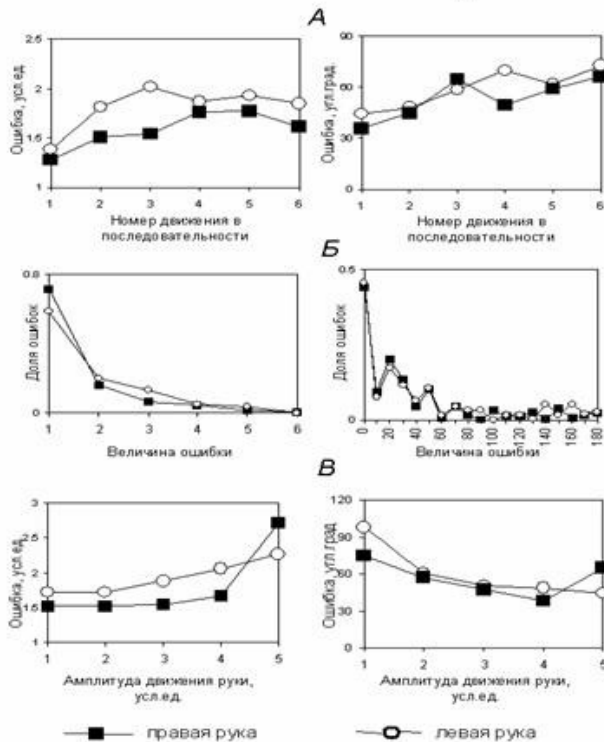


Ошибки

человека

Ошибки положения

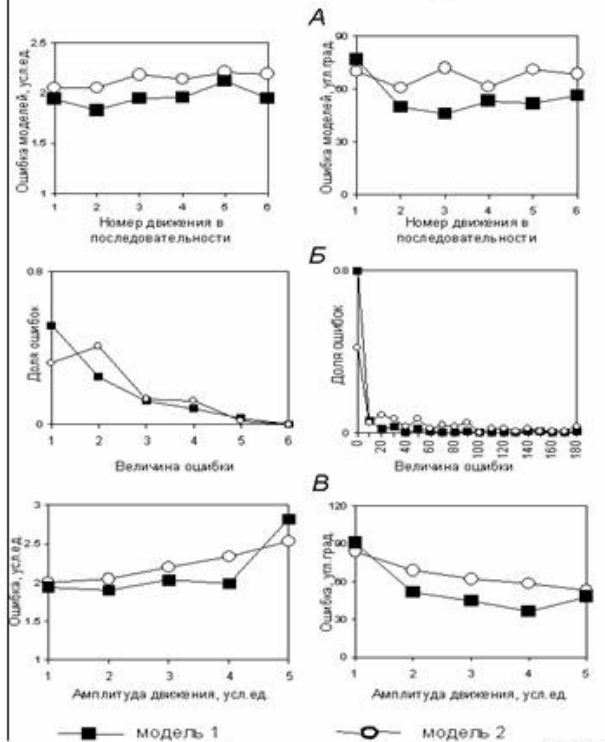
Ошибки движения



модели

Ошибки положения

Ошибки движения



Ляховецкий, Боброва, 2009

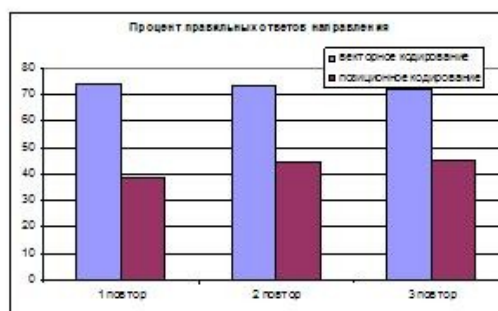
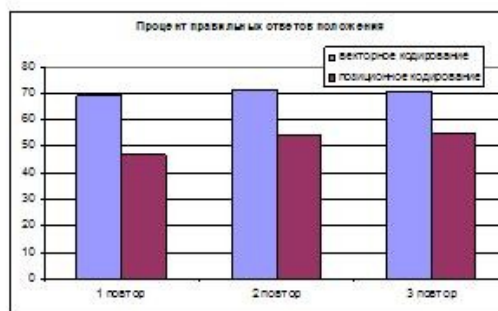
Число итераций сети для различного числа ошибочно запомненных битов

Для обеих схем кодирования сеть достоверно быстрее приходит в устойчивое состояние при правильных ответах (0 ошибочных бит). То есть, хотя информация о числе ошибочных битов не доступна сети, для оценки правильности ответа модель может использовать количество итераций, необходимых для перехода сети в устойчивое состояние.

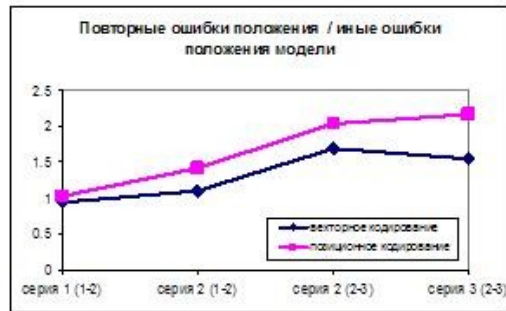
Ош. бит.	0	1	2	3	4	5	6	7	8
Схема 1	1.35	1.6	1.62	1.69	1.75	1.74	1.69	1.72	1.8
Схема 2	1.4	1.6	1.7	1.8	1.8	1.7	-	-	-

Будем корректировать по QLВAM матрицу весов только с помощью тех векторов, которые медленно воспроизводились сетью.

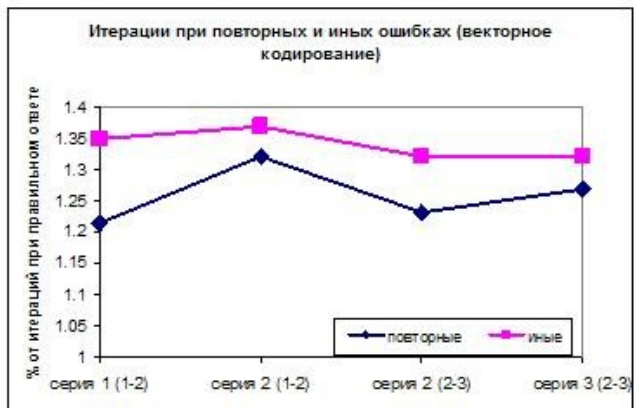
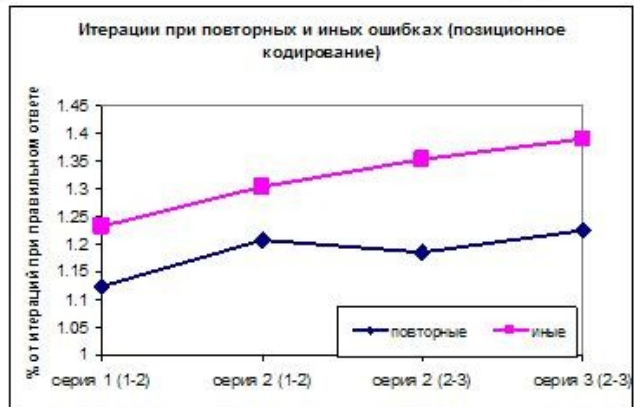
Процент правильных ответов модели



Число повторных ошибок модели



«Время» ошибок модели



Выводы

- Алгоритм QLВAM удобно использовать при моделировании процессов, протекающих в кратковременной памяти при заучивании;
- Феномен повторных ошибок является эмерджентным свойством обучаемой модели кратковременной памяти на основе нейросети.