

На правах рукописи

СОЛОВЬЕВА КСЕНИЯ ПАВЛОВНА

**ВЫЧИСЛИТЕЛЬНЫЕ ИССЛЕДОВАНИЯ СИСТЕМНЫХ И ЧАСТНЫХ  
ПРИНЦИПОВ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ В  
РЕКУРРЕНТНЫХ НЕЙРОННЫХ СИСТЕМАХ**

Специальность: 03.01.02 – Биофизика

**АВТОРЕФЕРАТ**  
диссертации на соискание ученой степени  
кандидата биологических наук

Москва – 2019

Работа прошла апробацию в департаменте молекулярной и биологической физики федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (государственный университет)».

Научный руководитель: Виталий Львович Дунин-Барковский, доктор физико-математических наук, профессор, заведующий Отделом нейроинформатики Центра оптико-нейронных технологий НИИСИ РАН, г. Москва

Ведущая организация: Научно-исследовательский технологический Центр нейротехнологий Федерального государственного автономного образовательного учреждения высшего профессионального образования «Южный федеральный университет»

Защита диссертации состоится «8» июля 2019 г. в 11-00 на заседании диссертационного совета ФБМФ 03.01.02.002 при Московском физико-техническом институте по адресу: 141701, Московская обл., г.Долгопрудный, Институтский пер., д.9, МФТИ.

С диссертацией можно ознакомиться в библиотеке и на сайте Московского физико-технического института (государственного университета) <https://mipt.ru/education/post-graduate/soiskateli-biologicheskie-nauki.php>

Работа представлена «26» марта 2019 г. в Аттестационную комиссию федерального государственного автономного образовательного учреждения высшего образования «Московский физико-технический институт (государственный университет)» для рассмотрения советом по защите диссертаций на соискание ученой степени кандидата наук в соответствии с п.3.1 ст. 4 Федерального закона «О науке и государственной научно - технической политике»

## ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

**Актуальность проблемы.** Механизмы работы мозга долгое время считались самой сложной проблемой естествознания. До сих пор даже в специальной литературе встречаются утверждения типа: «Чем больше мы знаем о мозге, тем меньше понимаем, как он работает» [Firesein S. What science wants to know. - Scientific American, 2012, Vol. 306, No. 4, p. 10.]. На фоне такого пессимизма неожиданно важную роль в прогрессе науки о мозге сыграли успехи в построении систем искусственного интеллекта (ИИ).

До 2012 г. прогресс в улучшении характеристик качества функционирования систем ИИ был очень медленным. Резкий скачок в улучшении этих характеристик произошел в связи с переходом от строго логических систем ИИ к системам, основанным на многослойных конструкциях нейроподобных элементов. Это явление часто называют революцией глубокого обучения (DL) [Sejnowski T. The Deep Learning Revolution. MIT Press, 2018].

В результате описанного прогресса технические конструкции за короткое время в 2012-2018 гг. стали проявлять всё более и более сложные интеллектуальные способности. На начальном этапе DL революции прогресс обеспечивался применением метода обратного распространения ошибок. Про этот метод с самого начала было понятно, что он не может быть реализован в биологических нейронных сетях [Liao Q., Leibo J.Z., Poggio T. How Important is Weight Symmetry in Backpropagation? - arXiv:1510.05067]. Другая линия разработки нейронных систем была связана с нейронными моделями конкретных структур мозга: мозжечка, гиппокампа, коры больших полушарий. Эти модели обеспечивают решения сравнительно небольшого спектра задач переработки информации. Настоящая работа посвящена проработке отдельных элементов и подсистем нейронных конструкций с целью получения новых архитектурных элементов нейроморфного интеллекта, основанных на принципах работы мозга.

Таким образом, с одной стороны, что для понимания естественных нейронных систем необходимо открытие каких-то новых принципов. Однако, в ходе работ, в которых делались попытки с одной стороны уйти от механизмов обратного распространения ошибок, а с другой стороны делались попытки реализации более сложных функций на классических нейронных конструкциях, было обнаружено, что объединение предыдущих подходов и использование высокопроизводительных параллельных систем обработки информации представляются необходимым и, вероятно, достаточным для понимания и воспроизведения функций нейронных систем.

За последние три года системы искусственного интеллекта стали применяться практически во всех технологических процессах и используются в устройствах телекоммуникации и индустрии развлечения и продолжают проникать практически во все сферы деятельности человека. Это быстрое изменение техногенной среды в первую очередь обусловлено применением для систем искусственного интеллекта нейроподобных конструкций.

Таким образом, задача выяснения общих принципов функционирования нейронных систем, составляющая ядро настоящей работы, является остро актуальной.

**Цель диссертационной работы:** исследование внутренней динамики и информационных процессов в нейронных системах с рекуррентным взаимодействием на физиологически-адекватных для поставленных задач моделях нейронов и нейронных сетей.

Для достижения поставленной цели в ходе работы решались следующие задачи:

1. Исследование условий возникновения в детальной нейронной модели режима пачечных пейсмекеров за счет взаимодействия возбуждения поверхностной мембраны нейрона с зависящей от эндоплазматического ретикулума динамикой концентрации внутриклеточного кальция и исследование нейронных сетей, обеспечивающих системную генерацию паттернов пачечной ритмики (по типу центрального генератора дыхательной ритмики).
2. Разработка схем создания аттракторных нейронных сетей и анализ их возможной информационной нагрузки.
3. Исследование свойств самоорганизующегося отображения активности сенсоров в состоянии рекуррентной нейронной сети, содержащей непрерывный одномерный («бугорковый») аттрактор.
4. Формулировка общего принципа работы нейронных систем, для понимания причин эффективности информационно-вычислительных конструкций, состоящих из нейроподобных элементов, и демонстрация возможности применения сформулированного принципа для предложения новых эффективных нейроморфных архитектур аналогичных биологическим прототипам.

**Методы исследования.** Во всех разделах работы использовались методы, соответствующие характеру поставленных задач. В главе 2 это – детальные модели типа Ходжкина-Хаксли с учётом динамики проницаемости ионных каналов, в главах 3, 4, 5 – бинарные нейроны типа Мак-Каллока-Питтса). Соответственно, в главе 2 использовался аппарат обыкновенных дифференциальных уравнений, в главе 3 – комбинаторные методы аналитических оценок. Кроме того, для реализации разработанных исследовательских моделей в имитационных экспериментах применялось программирование в среде Matlab.

#### **Научная новизна работы.**

1. Впервые исследована модель функционирования пейсмекерных нейронов и взаимотормозящих нейронных пулов на основе механизма кальций-зависимого высвобождения кальция ( $\text{Ca}^{++}$ - induced  $\text{Ca}^{++}$  release, CICR) из эндоплазматического ретикулума нейронов на уровне динамики проницаемости ионных каналов.

2. Впервые предложена схема создания аттракторных нейронных сетей, содержащая непрерывные одномерные («бугорковые») аттракторы с числом состояний аттрактора, значительно превосходящим число нейронов в сети и произведена вычислительная и аналитическая оценка информационной ёмкости системы.
3. Впервые исследованы свойства самоорганизующегося отображения активности рецепторов в состояния рекуррентной нейронной сети, содержащей непрерывный одномерный («бугорковый») аттрактор.
4. Впервые сформулирован общий принцип работы нейронных систем, объясняющий эффективность информационно-вычислительных конструкций, состоящих из нейроподобных элементов, названный Принципом Павлова (ПП), и на нейронной вычислительной модели продемонстрирована практическая применимость ПП для создания новых типов сетей глубокого обучения.

**Научно-практическая значимость.** Полученные данные вносят существенный вклад в понимание механизмов работы мозга, в том числе и за счёт расширения арсенала нейронных моделей, которые могут быть применены при интерпретации экспериментальных данных. На основе разработанных моделей предложены новые варианты искусственных нейронных сетей глубокого обучения. Результаты диссертационного исследования использованы при выполнении НИР и НИОКР по темам:

1. Грант РФФИ 13-07-01004 «Построение и анализ нейронных сетей с нульмерными, одномерными и многомерными аттракторами», руководитель - д.ф.-м.н., проф. Дунин-Барковский В.Л. 2012-2014 гг. (на базе НИИСИ РАН);
2. Грант РФФИ 14-19-01698 "Создание функциональных прототипов электронных синапсов и построение на их основе модели нейроморфной вычислительной системы", руководитель - к.т.н. Зенкевич А.В. 2014 г. (на базе МФТИ(ГУ));
3. Грант РФФИ 14-18-03773 "Одаренность: когнитивные процессы и модели социо-культурной реализации", руководитель – член-корр. РАН Ушаков Д.В. (на базе ИП РАН);
4. Грант РФФИ 16-07-01059 «Вычислительное исследование системных и частных принципов информационной архитектуры нейронных систем», руководитель - д.ф.-м.н., проф. Дунин-Барковский В.Л. 2015-2017 гг. (на базе НИИСИ РАН);
5. Грант фонда ФГБУ «Фонд содействия развитию малых форм предприятий в научно-технической сфере» по теме «Разработка прототипа программно-аппаратного модуля для обучения вхождению в эффективные психофизиологические состояния и контролю над их сохранением на основе нейрообратной связи.» Заявка № НТИ-26474, Проект № 26070, грантополучатель ООО "Образовательные технологии";
6. Грант фонда ФГБУ «Фонд содействия развитию малых форм предприятий в научно-технической сфере» по теме «Разработка

прототипа аппаратно-программного модуля нейроинтерфейсного взаимодействия с виртуальной реальностью». Заявка № ЦМИТ-Н-16807, Проект 22854, Грантополучатель ООО «Системы виртуальной реальности»

**Положения диссертационного исследования, выносимые на защиту.**

1. Разработана детальная модель функционирования пейсмекерных нейронов и взаимотормозящих нейронных пулов на основе кальций-зависимого высвобождения кальция из эндоплазматического ретикулума нейронов на уровне динамики проницаемости ионных каналов.
2. Предложена схема создания аттракторных нейронных сетей, содержащая непрерывные одномерные («бугорковые») аттракторы с числом состояний аттрактора, значительно превосходящей число нейронов в сети. На основе вычислительных экспериментов и аналитической оценки установлено, что число состояний непрерывного («бугоркового») аттрактора в рекуррентных нейронных системах растёт как  $\sim N^2$ . Численная оценка константы пропорциональности даёт величины порядка  $1000N$  для  $N=10^6$ .
3. Предложена схема самоорганизующегося отображения активности рецепторов в состояния рекуррентной нейронной сети, содержащей непрерывный одномерный («бугорковый») аттрактор. Данное отображение обладает свойствами топологической связности, масштабируемости и помехоустойчивости.
4. Сформулирован общий принцип работы нейронных систем, объясняющий эффективность информационно-вычислительных конструкций, состоящих из нейроподобных элементов и на основе этого принципа предложен новый вариант искусственной нейронной сети глубокого обучения.

**Апробация работы.** Результаты работы представлялись на Всероссийских и международных научных конференциях и семинарах, в том числе: научном семинаре Barrow Neurological Institute, (Phoenix, USA, 2012); международном конгрессе "Global Future 2045", (Линкольн-центр в Нью-Йорке, США, 2013); московской осенней психологической школе студентов, аспирантов и молодых учёных, (Москва, 2013); научном семинаре Института нейроинформатики Университета и Политехнического института Цюриха INI-402-0826-00L, (Цюрих, декабрь 2014); 16-ой, 18-ой и 20-й Международных конференциях «Нейроинформатика» (г. Москва, Россия в 2014, 2016 и в 2018 гг.); 3-ей и 4-ой Всероссийских с международным участием конференциях "Гиппокамп и память: норма и патология", (Пущино, 2015 и 2018 гг.); международной конференции FIERCES on BICA. First International Early Research Career Enhancement School on Biologically Inspired Cognitive Architectures, (Москва, 2016); международной конференции BICA - 2017, (Москва, гостиница «Балчуг», 2017); научном семинаре НИИ молекулярной электроники <http://www.niime.ru/press-center/news/440-nauchnyy-seminar-neyromorfnye-sistemy-i-ikh-realizatsiya-proshel-v-niime/>,

НИИМЭ, (Москва, 24 января 2018); международной конференции Computational Intelligence in Bioinformatics and Computational Biology 2018 (Saint Louis, Missouri, USA, 31 мая – 2 июня 2018); международной конференции 26th Nonlinear Dynamics of Electronic Systems Conference (Acireale, Italy, 11- 13 июня 2018); международной конференции Volga Neuroscience 2018 (Нижний Новгород – Самара – Нижний Новгород, 22 – 27 июля 2018).

**Публикации.** По теме диссертации опубликовано 15 печатных работ, среди них: 5 в журналах, рекомендованных ВАК, 4 в полнотекстовых сборниках научных трудов Всероссийских и международных научно-технических конференций, 4 в тезисах международных научных конференций, 4 в международных журналах, индексируемых в базах данных Scopus и WoS.

**Структура диссертации.** Работа состоит из введения, пяти глав, заключения, списка сокращений и списка литературы. Общее количество страниц: 152. Работа содержит 66 иллюстраций, 3 таблицы и 2 приложения; список литературы включает 146 наименований.

## СОДЕРЖАНИЕ РАБОТЫ

**Во введении** обоснована актуальность темы диссертационной работы, определена цель и сформулированы задачи исследования. Раскрыта новизна полученных результатов, их научная и практическая ценность. Представлена краткая аннотация всех глав диссертации.

**Первая глава** посвящена обзору литературы по фундаментальным принципам информационных процессов работы мозга и искусственных нейронных систем в контексте реверберационных нейронных процессов и рекуррентных нейронных систем. Рекуррентность трактуется в широком смысле, как способность использовать выход системы или отдельных ее частей в продолжающемся функционировании той же системы.

**Во второй главе** проведено исследование функциональных возможностей взаимодействия возбуждения поверхностной мембраны нейрона с зависящей от эндоплазматического ретикула динамикой внутриклеточного кальция; приведено математическое описание единичных моделей нейрона и синапса и модели нейронной сети дыхательных нейронов.

В первом подразделе рассмотрена детальная (на уровне нейронных каналов) точечная модель единичного пейсмейкерного дыхательного нейрона (модель генерации пачечной ритмики), основанная на взаимодействии каналов внешней мембраны нейронов и динамики внутриклеточной концентрации кальция, связанной с эндоплазматическим ретикулом. Помимо основной клеточной мембраны внутри клетки имеются мембраны, ограничивающие внутреннее пространство эндоплазматического ретикула (ЭР). В модели считаем, что мембрана ЭР также описывается единым мембранным потенциалом (не совпадающим с мембранным потенциалом

основной клеточной мембраны). Изменения проводимостей описываются уравнениями типа Ходжкина - Хаксли.

Эти уравнения имеют специфический вид для разных каналов. Для нашей задачи использованы модели медленного натриевого (Na persistent, NaP), быстрого натриевого (Na), кальциевого (P-канала) канала, калиевого задержанного выпрямляющего тока, проводимого К-каналом (K), калиевого А-канала, калиевого канала, управляемого внутриклеточной концентрацией ионов кальция  $Ca^{2+}$  – SK-канал. Для него используем уравнения:

$$g_{SK} = g_{0SK} (m_{SK})^2; \quad dm_{SK} / dt = (m_{\infty SK} - m_{SK}) / \tau_{SK}; \quad m_{\infty SK} = [Ca^{2+}]^2 / ([Ca^{2+}]^2 + (SKhalf)^2)$$

(SKhalf – значение концентрации ионов кальция  $[Ca^{2+}]$ , при котором величина  $m_{\infty SK}$  равна  $1/2$ ).

Для внутренней мембраны использованы модели хлорного и специфического кальциевого канала:

$$g_{CaR} = g_{0CaR} / \left( 1 + \exp \left( - \left( [Ca^{2+}] - Ca_0 \right) / DCa \right) \right),$$

здесь  $g_{0CaR}$  – максимальная проводимость,  $Ca_0$  – пороговое значение концентрации ионов кальция,  $DCa$  – ширина переходной зоны (по концентрации). Проводимость хлорного канала задаётся уравнениями:

$$g_{CIR} = g_{0CIR} m_{CIR}; \quad dm_{CIR} / dt = (m_{\infty CIR} - m_{CIR}) / \tau_{CIR};$$

$$m_{\infty CIR} = 1 / (1 + \exp((V_R - V_{CIR}) / \Delta_{CIR})).$$

Изменение концентрации ионов кальция  $Ca^{2+}$  описывается феноменологическим уравнением:

$$d[Ca^{2+}] / dt = k_1 (I_{Ca} + I_{CaR}) - k_2 [Ca^{2+}].$$

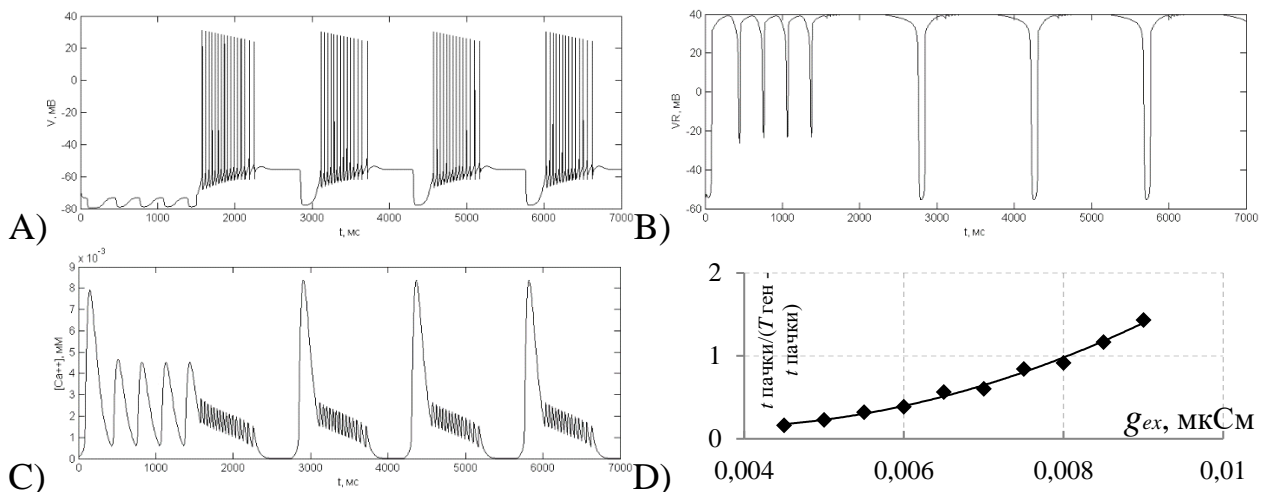


Рисунок 1. А: Зависимость мембранного потенциала нейрона от времени. В: Зависимость внутриклеточной концентрации ионов Ca от времени. С: Зависимость потенциала внутренней мембраны (мембраны ЭР) от времени. В момент времени  $t=1.5$  с включается ток стимуляции. D: зависимость отношения длительности пачки к остальной части периода генерации пачек от  $g_{ex}$ .

Исследована зависимость характера картины электрической активности нейрона от величины постоянного возбуждающего сигнала (рисунок 1), а также от уровня величины шума возбуждающего сигнала и показано, что характер картины активности нейрона определяется средним уровнем постоянной компоненты возбуждения и максимальным уровнем значения возбуждения.



Обнаружен и исследован диапазон величины внешнего входного сигнала, в котором нейрон является пачечным пейсмейкером. Исследована также зависимость характера картины электрической активности нейрона от параметров каналов мембраны ЭР.

Во втором подразделе второй главы исследуется нейросетевая модель дыхательной ритмики, построенная на основе моделей нейронов, изученных в первом подразделе. Главными элементами дыхательного центра являются нейроны, разделённые на два множества (пула) – нейроны вдоха и выдоха. Нейроны каждого множества генерируют импульсы, дающие вклады в соответствующие сигналы. Сигнал каждого типа (вдоха и выдоха) имеет вид почти периодических пачек импульсов. Число импульсов в пачке (или средняя частота импульсов в пачке) определяет интенсивность сигнала, а частота пачек определяет частоту сигнала, между нейронами вдоха и выдоха установлены тормозящие связи. Нейроны одного пула не связаны между собой. На каждый нейрон поступает возбуждающий сигнал от внешнего источника. Тормозящее влияние нейронов разных пулов друг на друга задаётся формулой

$\mathbf{g}_{in} = k_{in} \mathbf{Inh} \mathbf{f}_{in}$ , где  $\mathbf{g}_{in}$ ,  $\mathbf{f}_{in}$  – вектор-столбцы, составленные из величин  $g_{in}$ ,  $f_{in}$  для всех нейронов,  $\mathbf{Inh}$  – матрица тормозящих связей,  $k_{in}$  – нормировочная постоянная, задающая среднюю силу связей. Ненулевые элементы матрицы связей – положительные случайные величины со средним, равным 1. Для задания тормозящей функции  $\mathbf{f}_{in}$  для каждого нейрона используются уравнения  $df_{ac}/dt = \alpha_0 \theta(V - V_0) (1 - f_{ac}) - \alpha_1 \theta(V_0 - V) f_{ac}$ ,  $df_{in}/dt = \alpha_1 \alpha_2 f_{ac} - \alpha_2 f_{in}$ , где  $f_{ac}$  – функция активации нейрона, заметно отличная от нуля только во время и некоторое время после генерации импульса, константы  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$  определяют обратные характерные времена изменения функций активации и торможения,  $V_0$  – напряжение, достижение которого означает генерацию импульса.

В этой части подробно рассматриваются ритмические режимы работы всей нейронной сети и анализируется поведение отдельных нейронов во время работы сети в разных режимах.

В частности, рассмотрено два режима работы сети: на основе работы пейсмейкерных нейронов с короткими и длинными пачками импульсов (когда длительность пачки меньше и больше длительности паузы между пачками).

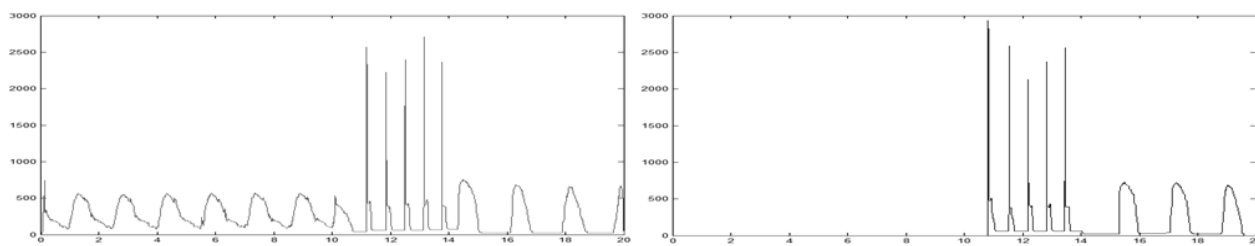


Рисунок 2. Частоты первого (слева) и второго (справа) нейронных пулов при импульсном воздействии дантроленом с 10-ой по 14-ю секунды.

Наконец, в главе рассмотрено модельное воздействие дантроленом (блокатора кальциевого канала эндоплазматической мембраны) в постоянном

и импульсном режиме. В результате, оказывается, что в некоторых случаях, импульсное воздействие дантроленом может перевести систему дыхательных нейронов, которая не генерирует дыхательную активность в режим, когда в системе генерируется нормальная дыхательная динамика (рисунок 2).

**Третья глава** посвящена исследованию одного из вероятных механизмов представления непрерывных величин в нервной системе – непрерывным («бугорковым») аттракторам. Рассматриваются вопросы формирования связей нейронных сетей, содержащих непрерывные аттракторы, и их информационной емкости. Предложена схема создания «предформированных» аттракторных сетей с длиной непрерывного аттрактора, в несколько раз превышающего число нейронов в сети, обладающая физиологическим правдоподобием.

Модельная сеть представляет собой  $N$  связанных пороговых бинарных нейронов Мак-Каллока-Питтса. Иначе говоря, состояние сети характеризуется  $N$ -мерным вектором  $\mathbf{X}$  со значениями координат «0» или «1».

Связи нейронов друг с другом определяются симметричной матрицей  $\mathbf{T}$  ( $N \times N$ ), значение которой задаётся при помощи метода «молекулярных меток» (см. далее). В сети задана асинхронная случайная циклическая динамика. В последовательные такты времени случайно выбирается порядок обхода всех  $N$  нейронов и производится  $N$  обновлений состояний нейронов по формуле:

$$x_i(t+1) = \begin{cases} 1, & \text{при } (\mathbf{T}_{i,j}, \mathbf{X}(t)) \geq \theta \\ 0, & \text{при } (\mathbf{T}_{i,j}, \mathbf{X}(t)) < \theta \end{cases}, \text{ для } i = \Xi(t, 1), \dots, \Xi(t, N), \text{ где } \theta - \text{порог, } \Xi(t, j) -$$

случайная (независимая в разные моменты  $t$ ) перестановка индексов  $\{1, \dots, N\}$ . «Молекулярные метки» мы считаем элементами упорядоченного кольцевого множества, из  $M$  элементов, то есть, 1-ый и  $M$ -ый элемент считаются соседними. На множестве введена метрика. Если разница номеров двух меток меньше или равна  $d$ , то расстояние между метками равняется нулю, в противном случае - одному. Мы будем использовать  $M = kN$ , где  $k$  – натуральное число, а  $N$  – число нейронов в сети.

Молекулярные метки случайно распределяются между нейронами, так, что каждый нейрон получает по  $k$  молекулярных меток с попарной разницей номеров больше некоторого числа  $D$  (рисунок 3).

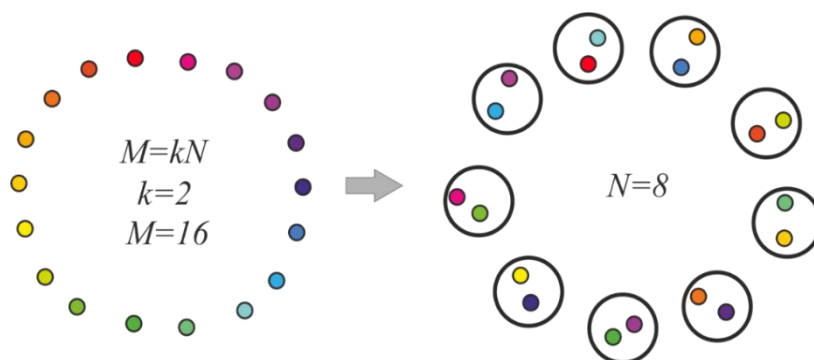


Рисунок 3. Схема распределения  $M$  «молекулярных меток» (слева) между  $N$  нейронами (справа). На приведённом примере  $k = 2$ , близкими цветами обозначены метки с

близкими номерами.

Между нейронами, у которых есть метки с нулевым расстоянием (с разницей номеров меньше  $d$ ) образуются положительные связи, равные единице, в противном случае отрицательные, равные  $-\sigma$ . В такой нейронной сети имеется множество устойчивых состояний активности. Для изучения структуры этого множества используется вычислительный эксперимент по «динамическому методу» - когда к существующему закону динамики добавляется еще один закон динамики, благодаря которому сеть не может долго задерживаться в любом из своих устойчивых состояний, а вынуждена переходить в соседнее устойчивое состояние (механизм утомления). В нашем случае, в состоянии возбуждения нейрона его порог возрастает на фиксированную величину  $\Delta\theta$ , и все время убывает экспоненциально до начального значения порога  $\theta_0$  с постоянной времени  $\tau$ :

$\theta(t) = (\theta + \Delta\theta \cdot \partial(t-t_i) - \theta_0) e^{-\frac{t-t_i}{\tau}} + \theta_0$ , где  $t_i$  - время  $i$ -го момента перехода нейрона в возбуждённое состояние, а  $\partial(t-t_i) = 1$ , при  $t = t_i$ ; 0, при  $t \neq t_i$ .

Для визуализации состояний нейронной сети во времени используется абсолютная развертка и относительная развертка (L-plot). И в том и в другом случае по оси абсцисс откладывается время, а цветом пикселя расстояние по Хеммингу между состояниями. В абсолютном представлении по оси ординат указывается номер состояния, с которым происходит сравнение. А в относительном - в качестве состояний, с которыми происходит сравнение активности в момент времени  $t$ , мы берем состояния сети в моменты, отстоящие от данного на время от  $-\tau$  до  $\tau$ .

При помощи абсолютной визуализации мы определяем максимальное значение  $k$ , при котором в сети еще поддерживается непрерывное перемещение активности по всему аттракторному множеству, а при помощи абсолютного — мы определяем точное значение  $k$  когда это возможно (рисунок 4).

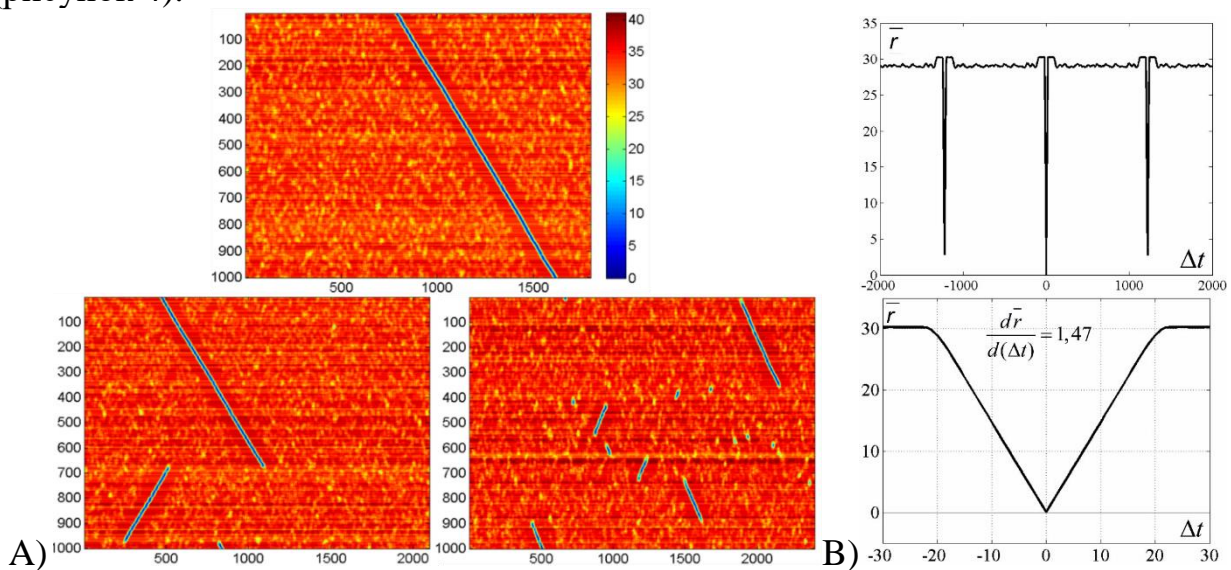


Рисунок 4. А: Абсолютная развёртка активности нейронной сети по состояниям непрерывного аттрактора в зависимости от времени. По горизонтальной оси отмечен номер точки аттрактора, по вертикальной шаг по времени, цветом обозначено расстояние по Хеммингу от текущего состояния

нейронной сети до всех обозначенных точек аттрактора.  $N=300$ ,  $D=80$ ,  $d=12$ ,  $\sigma = 3$ . (верхний)  $k=6$ ,  $M=1800$ , (нижний левый)  $k=7$ ,  $M=2100$ , (нижний правый)  $k=8$ ,  $M=2400$ . В: Иллюстрация для метода определения длины непрерывного аттрактора при помощи графика  $L$ -plot. (вверху) График усреднённого по времени сечения графика  $L$ -plot (внизу) Верхний график в диапазоне  $\Delta t$  от -30 до 30.  $N=300$ ,  $M=900$ ,  $T=1225$ , где  $T$  - период движения активности по непрерывному аттрактору.

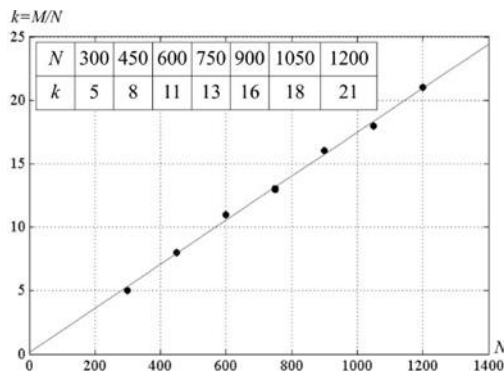


Рисунок 5. Зависимость максимальной длины непрерывного аттрактора от числа нейронов в сети.

В ходе вычислительных экспериментов была исследована зависимость относительной максимальной длины непрерывного аттрактора  $k=M/N$  от числа нейронов в сети  $N$  (рисунок 5). Если экстраполировать полученную прямую до характерного числа нейронов в кортикальной колонке  $N \approx 10000 \div 50000$ ,

получим  $k \approx 150 \div 850$ .

Также была произведена аналитическая оценка максимальной длины непрерывного аттрактора. В результате была получена формула для критического значения  $k$ :

$$k_c < \frac{N}{2\delta \sqrt[15]{5N}}, \text{ где коэффициент } \sqrt[15]{5N} \text{ порядка 1 и меняется очень медленно}$$

с  $N$ . Таким образом, уравнение дает практически линейную зависимость  $k_c$  от  $N$ . В частности, мы имеем:

при  $L=15$ ,  $N=300$ ,  $\sqrt[15]{5 \cdot 300} \approx 1.62$ ,  $k=7.7$  (в эксперименте,  $k=5$ );

при  $L=15$ ,  $N=1200$ ,  $\sqrt[15]{5 \cdot 1200} \approx 1.79$ ,  $k=28$  (в эксперименте,  $k=21$ ).

**В четвертой главе** исследуется простая модельная нейронная система, в которой осуществляется самоорганизующееся отображение многомерных рецепторных сигналов  $R$ -мерного пространства (где  $R$  – количество рецепторов) на состояния сети из  $N$  взаимодействующих нейронов. Рекуррентная сеть из  $N$  взаимодействующих нейронов обладает одномерным непрерывным аттрактором («бугорковый аттрактор»). Модельная система состоит из двух связанных между собой частей, условно называемых далее рецепторы и нейроны рекуррентной сети (или просто «нейроны») (рисунок 6). Состояние среды характеризуется  $R$ -мерным вектором  $\mathbf{V}$ , с вещественными координатами в диапазоне от нуля до

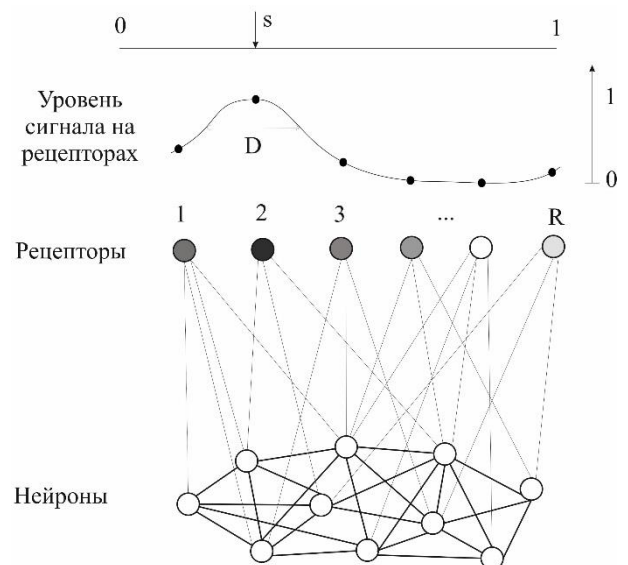


Рисунок 6. Схема выбора сигнала на рецепторах и соединения нейронов рекуррентной сети.

единицы. Рецепторы между собой никак не связаны. Каждый рецептор связан с каждым нейроном рекуррентной сети. Матрица связей рецепторов с нейронами  $\mathbf{W} (N \times R)$ .

Рекуррентная сеть состоит из  $N$  связанных симметричными связями бинарных нейронов. Нейронам присвоены номера от 1 до  $N$ . Состояние нейронной сети характеризуется  $N$ -мерным вектором  $\mathbf{X}$  со значениями координат «0» или «1». Матрица связей нейронов между собой  $\mathbf{T} (N \times N)$ . Нейроны расположены по кольцу так, что первый и  $N$ -й нейроны являются соседними. Каждый нейрон связан с  $L$  соседями «слева» и  $L$  соседями «справа» положительными связями, с весом равным единице, и отрицательными связями, с весом  $-\sigma$ , со всеми остальными нейронами. В сети задана асинхронная случайная циклическая динамика. В последовательные такты времени случайно выбирается порядок обхода всех  $N$  нейронов и производится  $N$  обновлений состояний нейронов по формуле:

$$x_i(t+1) = \begin{cases} 1, & \text{при } (\mathbf{W}_{i\cdot}, \mathbf{V}) + (\mathbf{T}_{i\cdot}, \mathbf{X}(t)) \geq \theta \\ 0, & \text{при } (\mathbf{W}_{i\cdot}, \mathbf{V}) + (\mathbf{T}_{i\cdot}, \mathbf{X}(t)) < \theta \end{cases}, \quad \text{для } i = \Xi(t, 1), \dots, \Xi(t, N), \text{ где } \theta -$$

порог,  $\mathbf{W}_{ig}$  – строка матрицы связей нейронов с рецепторами, соответствующая  $i$ -му нейрону, а  $\Xi(t, j)$  – случайная (независимая в разные моменты  $t$ ) перестановка индексов  $\{1, \dots, N\}$ .

Конструкция работает в режимах обучения и тестирования. Обучение системы состоит в многократном последовательном повторении следующей процедуры:

Подается вход на рецепторы  $\mathbf{V}$ . В каждый момент выбирается случайно и равновероятно действительное число  $s$  в полуинтервале  $[0, 1)$ . Возбуждения рецепторов при этом определяются значениями функции  $G(D, s, v)$ , где  $G$  вычисляется по формуле:

$$G = \max_p \left[ \exp \left( - \left( \frac{v - (s + p)R}{D} \right)^2 \right) \right],$$

в точках  $v = \{1, 2, K, R\}$  (рис. 1),  $p$  принимает значения  $-1$ ,  $0$  или  $1$ .

Изначально активность всех нейронов рекуррентной сети равна нулю. Под воздействием сигнала на рецепторах, в рекуррентной нейронной сети запускается асинхронная динамика, приводящая сеть в одно из ее устойчивых состояний.

Для всех нейронов, активных в данной точке аттрактора, связи со всеми рецепторами изменяются по правилу Хебба:  $\mathbf{W}_{i\cdot} = \mathbf{W}_{i\cdot} + \eta((\mathbf{XV}^T)_{i\cdot} - \mathbf{W}_{i\cdot})$ , где  $i$  – номер нейрона,  $\mathbf{W}_{ig}$  – строка матрицы связей нейронов с рецепторами, соответствующая  $i$ -му нейрону,  $\mathbf{V}$  – сигнал на рецепторах,  $\mathbf{X}$  – вектор установившегося состояния в сети нейронов и  $\eta$  – коэффициент обучения.

Тестирование системы производится следующим образом: 1. Последовательно, с фиксированным шагом (0.01–0.001) задаются значения параметра сигнала  $s$  от 0 до 1.2. То же, что и в процедуре обучения.



В результате имитационного моделирования было выявлено, что при значении параметра  $\mu = L/N = D/R$ ,  $0.1 \leq \mu < 0.3$  отображение сходится и данная конструкция обладает следующими свойствами. 1. Свойство масштабируемости представления отдельных участков отображаемого сигнала. Были проведены модельные эксперименты, в которых сигнал первой половины входного многообразия ( $s < 1/2$ ) подавался в 1, 2, 4 и 10 раз чаще, чем второй ( $s > 1/2$ ) (рисунок 7). При этом область, на которую отображается первая половина рецептивного поля, увеличивается согласно увеличению частоты ее использования. 2. Свойство устойчивости к шуму отображения. При одном и том же протоколе обучения, отображение сигнала с пятикратной амплитудой белого шума (при  $A=5$ ) принимало правильный вид (рисунок 8А), тогда как отображение сигнала с десятикратной амплитудой белого шума (при  $A=10$ ) уже не принимало правильного вида (рисунок 8В). Если уровень шума меняется вдоль входной кривой ( $A = A(s)$ ), то благодаря быстрой сходимости отображения в участках, соответствующих низкому уровню шума, образуется “центр кристаллизации” и все отображение сходится быстрее (рисунок 8). 3. Свойство связности отображения в сетях, содержащих непрерывный аттрактор из нескольких компонент связности. Здесь рассматривается отображение пространства рецепторов, содержащее одно однопараметрически заданное кольцевое множество, на нейронную сеть, содержащую два кольцевых непрерывных аттрактора на одних и тех же нейронах сети. В результате было получено, что отображение происходит на одну из двух компонент связности (рисунок 9).

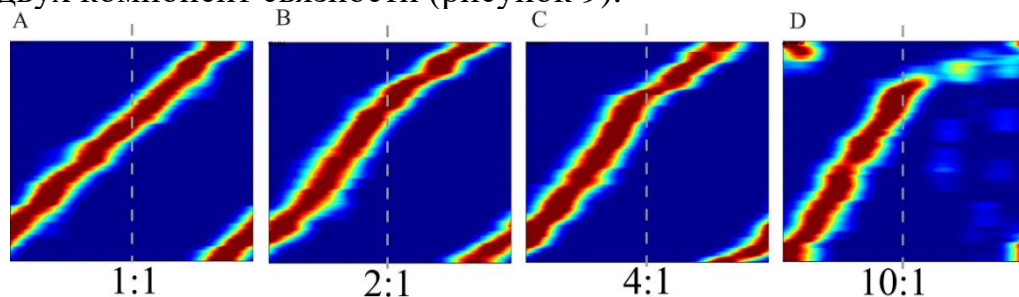


Рисунок 7. Матрица связей нейронов с рецепторами  $W$ , для не равновероятного использования первой и второй половины значения входного сигнала. Соотношения вероятностей сигнала в первой половине к сигналу во второй половине: (А) 1:1; (В) 2:1; (С) 4:1; (D) 10:1. Число нейронов  $N=300$ , число рецепторов  $R=300$ .

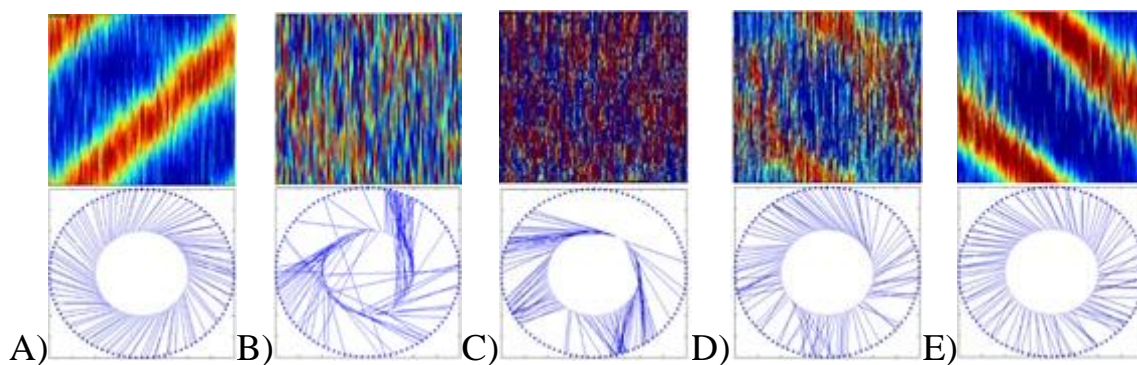


Рисунок 8. На верхних рисунках показана матрица  $W$ , на нижних – кольцевой график соответствий параметра  $s$  входного множества и точки непрерывного аттрактора рекуррентной сети. (A) Отображение сигнала с амплитудой белого шума, наложенного на сигнал, в 5 раз превышающей амплитуду сигнала. 1000 итераций обучения с коэффициентом обучения  $\eta = 0.1$ , затем 1000 итераций с  $\eta = 0.01$  и затем ещё 20000 итераций с  $\eta = 0.001$ ; (B) Амплитуда шума, в 10 раз превышающая амплитуду входного сигнала. 1000 итераций обучения с коэффициентом обучения  $\eta = 0.1$ , затем 1000 итераций с  $\eta = 0.01$  и затем ещё 20000 итераций с  $\eta = 0.001$ . (C), (D), (E) Три стадии обучения системы, отображающей сигнал с уровнем шума, меняющимся синусоидально от нулевого уровня (в центре рецепторного сигнала) до уровня амплитуды шума, в десять раз превышающей уровень сигнала (по краям рецепторного сигнала). На верхних рисунках изображена матрица  $W$ , на нижних – кольцевой график соответствий параметра  $s$  входного множества и точки непрерывного аттрактора рекуррентной сети. (C) Первая стадия. 1000 итераций с коэффициентом обучения  $\eta = 0.1$ , начиная со случайной матрицы  $W$ ; (D) Вторая стадия. 1000 итераций с коэффициентом обучения  $\eta = 0.01$  после окончания первой стадии; (E) третья стадия. 20000 итераций с коэффициентом обучения  $\eta = 0.001$  после окончания второй стадии, изображенной на рисунке 7, D. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

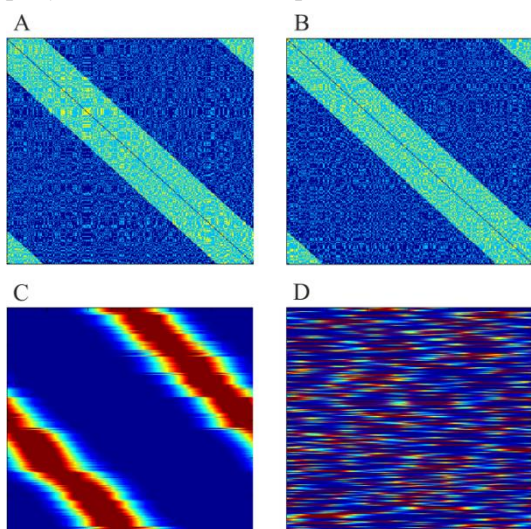


Рисунок 9. (A) Матрица связей нейронов  $T$  в первой нумерации; (B) Матрица связей нейронов  $T$  во второй нумерации. На (A) и (B) цветом отмечены силы связей: тёмно-синее  $-10$ , голубое  $1$ , жёлтое  $2$ ; (C) Матрица  $W$  после обучения в первой нумерации; (D) Вид той же матрицы  $W$  после обучения во второй нумерации. Число нейронов  $N = 300$ , число рецепторов  $R = 300$ .

В пятой главе формулируется общий принцип работы нейронных систем, который делает понятным эффективность информационно-вычислительных конструкций, состоящих из нейроподобных элементов, на основе обобщения всей совокупности экспериментальных данных об условных и инструментальных рефlekсах и данных вычислительных экспериментов с адаптивными нейроподобными системами.

### ПРИНЦИП ПАВЛОВА (ПП)

*Сеть нейронов, где каждая связь постепенно изменяется как функция локально доступных компонент сигналов ошибки и активности, соединяемых связью нейронов, приходит в процессе функционирования сети к безошибочной работе.*

Под нейроном в ПП фигурирует функциональный элемент, осуществляющий операцию:  $Y(t+1) = S(\sum_{i=1}^N \sigma_i(t) \cdot X_i(t) - \theta)$ , здесь  $Y(t+1)$  – значение выходной переменной нейрона в момент  $(t+1)$ ,  $S(\cdot)$  – вход-выходная функция нейрона,  $\sigma_i(t)$  и  $X_i(t)$  – соответственно, значения веса входной переменной с индексом  $i$  ( $i = 1, \dots, N$ ) и самой входной переменной в момент времени  $t$ ;  $\theta$  – порог нейрона.

При этом связи между нейронами медленно меняются под влиянием вектора ошибки:  $\sigma_i(t+1) = \sigma_i(t) + \varepsilon \cdot F(\sum_{k=1}^K E_k(t) \cdot \gamma_{ki} \cdot g(Y(t), X_i(t)))$ , здесь  $K$  – компонентный вектор ошибки,  $E$  представляет разность между тем выходным вектором, который желательно иметь для данного входа на нейронную сеть и тем значением выходного вектора, которое реально наблюдается. Существенно, что множители  $\gamma_{ki}$  – постоянны. Они выбираются для сети с помощью фиксированного (случайного) однократного выбора. Следует отметить, что ПП справедлив для большого числа типов нейронов (простые, сложные, импульсные, непрерывные, и т.д.).

Далее в пятой главе приводится пример конструкции, использующей Принцип Павлова. Исследуется решение классической задачи – обучение нейронной сети тому, является ли сигнал на входе нейронной сети симметричным относительно центральной линии. В качестве сигнала берется строка из 100 нулей и единиц. Сигнал считается симметричным, если первые пятьдесят символов совпадают со вторыми пятьюдесятью символами. Такая задача представляется интересной, поскольку предоставляет возможность работать с практически бесконечным числом объектов обучения. В данном случае, их не меньше, чем  $2^{50} \approx 10^{15}$ . В нашем случае, и обучающие и «экзаменационные» примеры представляли собой неповторяющиеся последовательности строчек из 0 и 1.

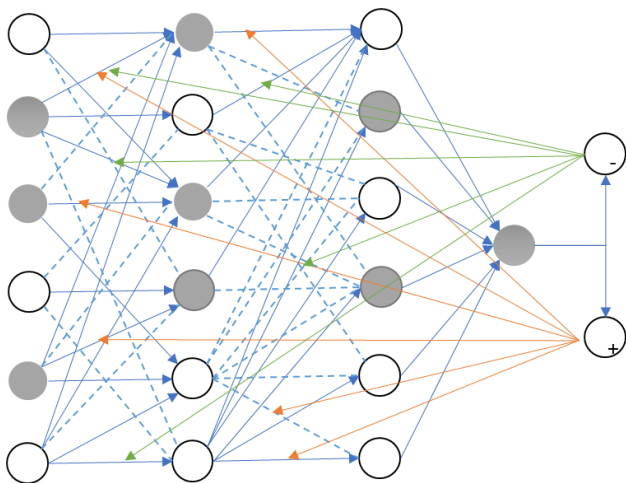


Рисунок 10. Архитектура сети, использующей Принцип Павлова. Оранжевым и зеленым обозначены обратные связи от разных нейронов ошибки. Белым обозначены нейроны, не активные в данном цикле активности.

Архитектура сети представляла собой (1) входной слой из 100 рецепторов, на котором появляются сигналы; (2) скрытый слой из 200 или 400 нейронов, выходной нейрон и два нейрона ошибки: «+» – нейрон и «-» – нейрон. Нейрон «+» возбуждается, когда сигнал выходного нейрона 1, а должен был быть 0, а нейрон «-» возбуждается, когда сигнал выходного нейрона 0, а должен был быть 1 (схема сети изображена на рисунке 10). Все нейроны в сети – нейроны Мак-Каллока-Питтса (возбужден/не возбужден). От нейронов ошибки к



каждому синапсу на нейронах скрытого слоя идут связи со случайными (не меняющимися во времени) весами. После серии  $K$  предъявлений входных сигналов меняются веса синапсов на нейронах скрытого слоя и на выходном нейроне. Величина изменения складывается из потенциальных изменений при каждом предъявлении входного сигнала. А это изменение, в свою очередь, определяется тем, какие сигналы были в соответствующем цикле у пре- и пост-синаптического нейронов и того, какой сигнал пришёл от нейронов ошибки. При отсутствии ошибки связи не меняются. При обучении равновероятно выбираются симметричные и несимметричные сигналы.

Данная конструкция обучилась распознавать некоторые симметричные сигналы. После 500000 обучающих входных сигналов сеть получала такую настройку, что при подаче на неё 100000 случайных сигналов, половина из которых была симметричными, сеть «сообщила» о симметричном входе (в среднем, из 10 попыток) 122.7 раза, из них 4.6 срабатываний были ложными. Таким образом, система гарантировано обучается узнавать симметрию 100-битных паттернов, хотя и далеко не всегда. Вычислительный эксперимент также показал, что сеть, у которой связи на нейронах скрытого слоя не обучаются, узнавать симметричные сигналы не обучается вообще. Следует также отметить, что способность к обучению у данной сети возникает только при одном из 50 выборов случайных связей от нейронов ошибки к изменяемым синапсам.

В **заключении** обобщены основные результаты диссертационного исследования.

В **приложениях 1 и 2** представлены основные функции и процедуры кода к главе 2, использованные при вычислительном моделировании.

## **ВЫВОДЫ**

1. Использование детальных моделей нейронов, учитывающих динамику проводимости ионных каналов, позволяет представить вычислительную модель включения процессов CICR в работу нейронной сети в форме, удобной для последующей экспериментальной проверки.
2. Анализ моделей представления данных о внешнем мире состояниями нейронной сети с учетом сведений и возможностей формирования нейронных архитектур позволяет существенно ослабить ограничения на кажущуюся необходимой информационную ёмкость нейронной системы.
3. Применение ансамблевой архитектуры нейронной сети к хорошо зарекомендовавшей себя в практических приложениях конструкции Кохонена позволяет оценить возможности реализации подобной архитектуры в живом мозге.
4. Обобщение принципов получения интеллектуальных компетенций в искусственных нейронных сетях и сведений о пластических перестройках в нервных системах позволяет получить компактное описание механизмов функциональной адаптации нейронных систем.

## СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ ПО ТЕМЕ ДИССЕРТАЦИИ

1. **Solovyeva K.P.** Bump-Attractor Based Self-Organized Maps. – Dynamical Neuroscience XX. Collective Cognition: The Neurophysiology of Social Neuroscience, A Satellite Symposium Immediately Preceding the 42nd Annual Meeting of the Society for Neuroscience. Bethesda, USA, National Institute of Health, 2012, p. 37.
2. **Соловьева К.П.** Формирование самоорганизующихся отображений сенсорных сигналов на непрерывные нейросетевые аттракторы. - Математическая биология и биоинформатика. 2013. Т.8. №1. С. 234-247. URL: [http://www.matbio.org/2013/Solovyeva\\_8\\_234.pdf](http://www.matbio.org/2013/Solovyeva_8_234.pdf).
3. Котов В.Б., **Соловьева К.П.** Моделирование нейронов генерирующих периодические залпы импульсов на основе внутриклеточной динамики концентрации кальция. - Труды НИИСИ РАН, 2013 г., т.3 №2, с. 66-77.
4. **Соловьева К.П.**, Котов В.Б. Нейросетевая модель дыхательной ритмики с учётом процессов, связанных с эндоплазматическим ретикуломом. - Труды НИИСИ РАН, 2013 г., т.3 №2, с. 78-87.
5. Дунин-Барковский В.Л., Карандашев Я.М., **Соловьева К.П.** Об архитектуре вычислительных процессов в мозге. - Труды НИИСИ РАН, 2013 г., т.3 №2, с. 43-49.
6. **Соловьева К.П.**, Гиперкольцевые нейронные аттракторы. - Сборник научных трудов XVI Всероссийской научно-технической конференции с международным участием "Нейроинформатика-2014", часть 1, с. 155-164.
7. **Соловьева К.П.**, Карандашев Я.М. Полиаттракторная топологически-упорядоченная сеть импульсных нейронов. - Сборник научных трудов XVI Всероссийской научно-технической конференции с международным участием "Нейроинформатика-2014", часть 1, с. 165-174.
8. **Solovyeva KP**, Karandashev IM, Zhavoronkov A and Dunin-Barkowski WL (2016) Models of Innate Neural Attractors and Their Applications for Neural Information Processing. Front. Syst. Neurosci. 9:178. doi: 10.3389/fnsys.2015.00178.
9. **Соловьева К.П.**, Щукин Т.Н., Иващенко А.А., Дунин-Барковский В.Л. Базовые принципы нейронной организации. - III Всероссийская конференция с международным участием конференция «Гиппокамп и память: норма и патология», Пущино, 2015, сс. 31-32.
10. **Solovyeva K.P.**, Karandashev Y.M., Dunin-Barkowski W.L. Analytical estimates of efficiency of attractor neural network with inborn connections. – ITM Web of Conferences, 2016, Vol. 6, No. 02009, DOI: 10.1051/itmconf/20160602009, 5 p.
11. Дунин-Барковский В.Л., **Соловьева К.П.** Принцип Павлова в проблеме обратного конструирования мозга. - Нейроинформатика-2016. XVIII Международная научно-техническая конференция. Сборник научных трудов. Часть 1, МИФИ, 2016, сс. 11-23.

12. Dunin-Barkowski W., **Solovyeva K.** Data Formats inside the Brain. – FIERCES ON BICA 2016: First International Early Research Career Enhancement School on Biologically Inspired Cognitive Architectures. MEPHI, 2016, p. 48.
13. Shakirov V.V., **Solovyeva K.P.**, Dunin-Barkowski W. L. Review of state-of-the-art in DL artificial intelligence. - Optical Memory & Neural Networks (Information Optics), 2018, том 27, номер 2.
14. Dunin-Barkowski W.L., **Solovyeva K.P.**, Pavlov Principle and Brain Reverse Engineering.- 2018 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), (St. Louis, MO, 2018). pp. 1-5. doi: 10.1109/CIBCB.2018.8404975. URL: <https://ieeexplore.ieee.org/document/8404975>
15. **Solovyeva K.P.**, Kiselev I.I., Dunin-Barkowski W.L. Stochastic resonance in desynchronization of homogeneous neuronal pool. - Non-Linear Dynamics of Electronic Systems, Acireale, Catania, Italy, June, 10-13, 2018.