

# Динамика модифицированной модели нейронной сети Кропотова-Пахомова

Черных Г. А.<sup>1)</sup>, Письмак Ю. М.<sup>2)</sup>

<sup>1)</sup>Санкт-Петербургский государственный университет, chernykh@gc7511.spb.edu

<sup>2)</sup>Санкт-Петербургский государственный университет, pismaк@jp7821.spb.edu

**Аннотация.** Исследуются динамические режимы в модели реалистической нейронной сети Кропотова-Пахомова, модифицированной с целью получить нетривиальную устойчивую динамику сети в широких областях параметров. Эволюция сети во времени оказывается более устойчивой, если параллельно с применением принципа Богданова-Хэбба<sup>1</sup> в модель ввести механизм, обеспечивающий ослабление взаимодействия между нейронами по мере усиления общей активности сети. Изучены эффекты кластеризации в сети в периодическом режиме. Исследованы наиболее характерные особенности нейросетевых процессов в непериодическом режиме. Обнаружены явления аномальных длиннопериодных резонансов, возникающие в ситуации особой чувствительности нейронов к внешнему стимулу.

## Введение

В области исследования нейронных сетей в настоящее время отчетливо выделяются два основных направления. Доминирующим является прикладное, так как нейросетевые конструкции позволяют решать массу практических задач. Второе направление состоит в детальном моделировании физиологических особенностей работы мозга и нервной системы. Отсюда разделение сетей на искусственные (см. [2–5]) и реалистические (см. [6–8]).

Промежуточную нишу с прикладной стороны занимают исследования динамики искусственных нейронных сетей и моделей сложных природных и социальных процессов (см. [9–12]). Последние часто именуется сложными сетями (complex networks). С противоположной от прикладного направления стороны в нише находятся нейросетевые модели, обладающие той или иной степенью абстрактности (см. [13–15]) по сравнению с реалистическими моделями «низкого уровня». Модифицированную модель Кропотова-Пахомова можно отнести именно к этой группе нейронных сетей. Оригинальная модель была предложена в серии работ [16–18] с целью исследования поведения ансамблей нервных клеток, возбуждаемых внешними

---

<sup>1</sup>Находясь в финальной стадии работы над статьей, авторы обнаружили книгу [1], в которой высказаны идеи, дающие основание называть известный принцип Хэбба принципом Богданова-Хэбба. В приложении 1 приведена соответствующая выдержка из [1].

сигналами. Обладая достаточной простотой, модель отражает важные характеристики динамики реальных нейронов (см., например, [19]), такие как возвращение мембранного потенциала к исходному состоянию после разряда нейрона, наличие порога генерации потенциала действия, кратковременный спад синаптической эффективности в ответ на кратковременную стимуляцию пресинаптического окончания и др.

Согласно [16, 17] сеть может осуществлять пространственную фильтрацию входного сигнала, что, в частности, проявляется в способности нейронов объединяться в группы клеток (кластеры), синергично возбуждаемых или тормозимых постоянным по времени внешним стимулом. Как далее будет показано, свойство нейронной сети распадаться на кластеры является одним из основных факторов, определяющих ее динамику. В периодическом режиме это приводит к формированию жесткой структуры межкластерных связей, а в неперiodическом — оказывает влияние на статистические распределения динамических характеристик сети и устойчивость ее эволюции во времени.

Оригинальная модель при параметрах, соответствующих динамике реальных нейронов, ввиду наличия диссипативных свойств нетривиально эволюционирует только при наличии внешнего воздействия. Некоторые переходные процессы были рассмотрены в [18]. В [20] построено аналитическое решение для симметричной сети, стимулируемой постоянным по времени сигналом. Для получения устойчивой динамики сети (без воздействия на нее извне) в исходную модель были внесены некоторые изменения [21].

Подчеркнем, что мы ставим целью не моделирование какого-либо реального физического объекта, а изучение внутренних динамических свойств самоорганизующихся нейронных сетей, применяемых в качестве инструмента исследования сложных систем и процессов. Модель Кропотова-Пахомова представляется для этого достаточно удачным выбором, поскольку строилась авторами для прикладных исследований, а в модифицированном виде проявляет некоторые важные свойства, общие для данного типа нейронных сетей.

Важность построения и исследования подобных абстрактных моделей сложных систем отмечается в книге [22]:

Благодаря неотделимости сложной системы от внешней среды одной модели самой по себе недостаточно для отражения реальности, даже, если она очень детализирована. Наоборот, модель не может совершенно потерять связь с реальностью, даже находясь на высоком абстрактном уровне. Поэтому требуются модели на различных уровнях обобщения и, несомненно, конструирование разноуровневых моделей является неотъемлемой частью процесса познания сложной сущности природы.<sup>2</sup>

Модифицированная модель, как и ее оригинальный вариант, относится к классу однослойных полносвязных сетей и, будучи заданной системой ко-

---

<sup>2</sup>Переведено с английского авторами статьи.

нечноразностных уравнений, развивается в дискретном времени. Методология работы с модифицированной сетью состояла в следующем. В начальный момент времени бралась сеть с нулевыми связями и тривиальными значениями динамических переменных, затем в течение определенного времени она подвергалась воздействию случайного сигнала, после чего исследовалась динамика сети без внешнего стимулирования. Характер первичного сигнала, как и воздействие на сеть в процессе установившейся динамики, могут повлиять на динамику сети только в областях фазовых переходов и так называемых *смешанных* фаз (см. раздел 4), где внешние импульсы способны перевести сеть из одного динамического режима в другой. Вдали от областей высокой чувствительности установившаяся динамика сети при отсутствии стимулирования не зависит от типа внешнего воздействия, которое оказывалось на нее ранее. Именно такие динамические режимы подробно рассмотрены в настоящей работе. Кроме того, в исследованных режимах количество нейронов также не играло существенной роли, а именно, выбиралось такое их количество, при котором дальнейшее увеличение числа элементов сети уже не влияло на характер динамики.

Для устойчивого эволюционирования модифицированной нейронной сети в определенном динамическом режиме в ней должен сохраняться соответствующий этому режиму уровень беспорядка, проявляющийся в рассогласованном поведении нейронов и различных численных значениях межнейронных связей. Максимальная упорядоченность сети происходит в случае полной синхронизации нейронов, что приводит к так называемому *обнулению*, в результате которого динамика сети становится тривиальной: все нейроны неактивны, а значения межнейронных связей стремятся к нулю. Максимальный уровень беспорядка достигается во время воздействия на нейронную сеть случайными импульсами. После снятия сигнала, а в некоторых случаях еще до момента его отключения, в сети начинаются переходные процессы, связанные с ее упорядочиванием, в результате которых нейронная сеть либо попадает в один из устойчивых динамических режимов, либо обнуляется.

В периодическом режиме нейронная сеть обладает структурой связей, которая формируется в результате разбиения сети на группы синфазно-колеблющихся нейронов. Зависимости всех динамических переменных модели от времени в периодическом режиме, а также тип соответствующей этому режиму структуры связей, могут быть аналитически найдены по временным зависимостям активностей нейронов<sup>3</sup>. Частоты колебаний разных нейронов сети в периодическом режиме могут отличаться друг от друга, но количество этих частот ограничено и в проведенных экспериментах не превышало двух.

При непериодическом динамическом режиме нейроны сети могут переключаться с одной частоты колебаний на другую. Количество допустимых частот также ограничено. При этом существуют доминирующие частоты, на которых нейроны находятся наибольшее время. Промежутки времени,

<sup>3</sup> Активность нейрона является одной из его динамических переменных

в течение которых нейрон колеблется с доминирующей частотой и не переходит на другие частоты, подчиняются кусочно-степенному распределению и по величине могут на два-три порядка превосходить период колебаний. В результате при непериодическом режиме в нейронной сети могут появляться короткоживущие структуры, подобные тем, которые образуются в периодических режимах.

Необходимо отметить, что вышеупомянутые колебания нейронов являются результатом модификации модели и не имеют аналога в оригинальном варианте. Именно благодаря наличию этих высокочастотных колебаний, выполняющих роль своеобразной подпитки диссипативной сети, последняя может устойчиво эволюционировать независимо от внешнего воздействия. Однако интерес представляют не сами колебания, а их модуляции. Если исключить из рассмотрения высокочастотную составляющую, то каждому периодическому режиму будет соответствовать свое стационарное состояние с определенной структурой связей.

Интересно, что при определенных условиях в некоторых областях высокой чувствительности сети к внешним воздействиям в ней возникают *длиннопериодные колебания*, период которых может в тысячи раз превосходить период высокочастотных колебаний нейронов. Обнаружены как собственные, так и вынужденные длиннопериодные колебания. В обоих случаях поведение большинства нейронов не изменяется, и они образуют своеобразный фон, на котором оставшаяся малая часть нейронов эволюционирует по сложному закону. Длиннопериодные колебания представляют собой модулированные по фазе или частоте высокочастотные колебания. Вынужденные колебания появляются при воздействии на один или несколько нейронов сети периодическими импульсами. Замечательно, что период вынуждающих импульсов должен быть по порядку величины равен периоду высокочастотных колебаний. Таким образом, мы имеем явление, в котором период вынужденных колебаний на несколько порядков превосходит периоды вынуждающих импульсов и собственных колебаний.

Модель Кропотова-Пахомова в модифицированном варианте представляет собой весьма обширный объект исследования. В настоящей работе отражен лишь некоторый характерный срез ее свойств, позволяющий получить представление о наиболее важных динамических особенностях сети.

## 1 Формулировка модели

**Оригинальная модель.** Модель Кропотова-Пахомова представляет собой группу из  $n$  взаимодействующих нейронов. Сеть эволюционирует в дискретном времени  $k$ , а изменение ее динамических переменных при переходе от одного момента времени к последующему задается набором рекуррентных соотношений

$$P_i(k+1) = (1 - \alpha)P_i(k) + \sum_{j=1}^n W_{ij}(k)N_j(k) - \beta N_i(k) + S_i(k), \quad (1)$$

$$W_{ij}(k) = (x_i^1(k) + x_i^2(k)) W_{ij}^0, \quad (2)$$

$$x_i^1(k+1) = (1 - A_1)x_i^1(k) + B_1 N_i(k) + C_1, \quad (3)$$

$$x_i^2(k+1) = (1 - A_2)x_i^2(k) - B_2 N_i(k) + C_2, \quad (4)$$

$$N_i(k) = \theta(P_i(k) - h_i), \quad (5)$$

$$\theta(x) = 0 \text{ при } x \leq 0, \quad \theta(x) = 1 \text{ при } x > 0. \quad (6)$$

Здесь индекс  $i = 1, \dots, n$  нумерует нейроны сети;  $P_i(k)$ ,  $N_i(k)$  и  $h_i$  — потенциалы, активности и пороги активизации нейронов;  $S_i(k)$  — внешние стимулы;  $W_{ij}(k)$  и  $W_{ij}^0$  — матрица связей и матрица постоянных коэффициентов;  $x_i^1(k)$  и  $x_i^2(k)$  — так называемые активаторы и депрессанты, сумма которых определяет эффективность синаптических связей нейронов;  $\alpha$ ,  $\beta$ ,  $A_{1(2)}$ ,  $B_{1(2)}$ ,  $C_{1(2)}$  — параметры модели, причем  $\alpha$  и  $A_{1(2)}$  берутся из интервала  $[0, 1]$ , поскольку от них зависят степени диссипации потенциалов, депрессантов и активаторов. Нейрон с номером  $i$  считается активным, если  $N_i(k) = 1$ , и неактивным при  $N_i(k) = 0$ .

В момент времени  $k = 0$  задаются начальные значения переменных модели, матрица  $W_{ij}^0$  и пороги  $h_i$ . Далее сеть развивается в зависимости от выбранных параметров, внешних сигналов  $S_i(k)$  и начальных условий.

В своей оригинальной формулировке модель не имеет устойчивых непериодических динамических режимов при отсутствии внешнего воздействия и параметрах, соответствующих поведению реальных нейронов. После прекращения стимулирования сети потенциалы  $P_i(k)$  затухают, и через некоторое время все нейроны становятся неактивными, за исключением отдельных вырожденных случаев выбора параметров, когда, напротив, вся сеть постоянно активна, либо разбита на группы поочередно активизирующихся нейронов.

Диссипативные процессы в сети необходимо компенсировать. При этом нужен и второй механизм, который бы удерживал сеть от стремления к полностью активному состоянию.

**Принцип Богданова-Хэбба.** «Подпитка», компенсирующая диссипацию, должна регулироваться внутренней динамикой сети. Одним из естественных путей является введение в модель правила обучения Богданова-Хэбба [23], согласно которому увеличиваются связи одновременно активных нейронов

$$V_{ij}(k+1) = V_{ij}(k) + \nu N_i(k) N_j(k), \quad (7)$$

где  $V_{ij}(k)$  — матрица связей,  $\nu$  — числовой коэффициент.

Будем использовать не классическое выражение (7), а видоизмененный принцип Богданова-Хэбба [21]

$$V_{ij}(k+1) = V_{ij}(k) + \nu N_i(k) N_j(k-1). \quad (8)$$

За счет введенной временной задержки такое правило обучения позволяет получить несимметричную матрицу связей и делает возможным запись и

воспроизведение сетью последовательностей образов. При произвольных задержках будем иметь

$$V_{ij}(k+1) = V_{ij}(k) + \nu \sum_{\{m\}} N_i(k) N_j(k-m), \quad (9)$$

где  $\{m\}$  — множество задержек.

Сделаем матрицу  $W_{ij}^0$  зависящей от времени, введем диссипацию матричных элементов и применим принцип (9). Тогда выражение (2) переписывается в виде

$$W_{ij}(k) = (x_i^1(k) + x_i^2(k)) W_{ij}^0(k) \quad (10)$$

$$W_{ij}^0(k+1) = (1 - \mu) W_{ij}^0(k) + \nu \sum_{\{m\}} N_i(k) N_j(k-m). \quad (11)$$

Значение параметра диссипации  $\mu$  также должно быть в интервале  $[0, 1]$ .

**Охлаждение.** Из выражения (1) следует, что при росте количества элементов в сети и примерном сохранении процентного соотношения активных и неактивных нейронов потенциалы последних будут получать большие приращения. Это обстоятельство в совокупности с использованием принципа Богданова-Хэбба, увеличивающего значения связей, может привести к потере устойчивости динамики.

Решить проблему сильной чувствительности динамики сети к количеству нейронов и увеличить устойчивость можно, поделив второе слагаемое в (1) на величину

$$\sum_{j=1}^n N_j(k) + 1. \quad (12)$$

Эта процедура условно названа *охлаждением* нейронной сети.

**Модифицированная модель.** Используя исходные определения (1)–(6) и изменения оригинальной модели, предложенные в двух предыдущих пунктах, запишем окончательные соотношения, определяющие модифицированную модель Кропотова-Пахомова

$$P_i(k+1) = (1 - \alpha) P_i(k) + \frac{\sum_{j=1}^n W_{ij}(k) N_j(k)}{\sum_{j=1}^n N_j(k) + 1} - \beta N_i(k) + S_i(k), \quad (13)$$

$$W_{ij}(k) = (x_i^1(k) + x_i^2(k)) W_{ij}^0(k), \quad (14)$$

$$W_{ij}^0(k+1) = (1 - \mu) W_{ij}^0(k) + \nu \sum_{\{m\}} N_i(k) N_j(k-m), \quad (15)$$

$$x_i^1(k+1) = (1 - A_1) x_i^1(k) + B_1 N_i(k) + C_1, \quad (16)$$

$$x_i^2(k+1) = (1 - A_2)x_i^2(k) - B_2N_i(k) + C_2, \quad (17)$$

$$N_i(k) = \theta(P_i(k) - h_i). \quad (18)$$

Все дальнейшее изложение будет относиться только к модифицированной модели.

## 2 Выбор параметров

Характер динамики нейронной сети существенно определяется ее многочисленными параметрами. Мы ограничим область исследования теми значениями параметров, при которых модифицированная модель сохраняет некоторые важные свойства оригинальной модели.

1. Пороги активации нейронов (18) должны удовлетворять условию  $h_i \geq 0$ . Это исключает возможность активизации каждого отдельно взятого нейрона сети в отсутствие внешнего стимула и сигналов от других нейронов. Положим

$$h_i = 0, \quad i = 1, \dots, n.$$

2. Параметры  $A_{1(2)}, B_{1(2)}, C_{1(2)}$  в (16), (17) подбираются таким образом, чтобы эффективности синаптических связей *активных* нейронов уменьшались, а после перехода последних в неактивные состояния восстанавливались (см. рис. 1). Во всех экспериментах над моделью полагаем

$$A_1 = 0.4, \quad A_2 = B_1 = C_1 = 0.2, \quad B_2 = 0.5, \quad C_2 = 0.1. \quad (19)$$

Легко показать, что при таких значениях параметров и условии  $N_i(k) = 0, \forall k > C, C \in \mathbb{N}$

$$\lim_{k \rightarrow \infty} (x_i^1(k) + x_i^2(k)) = 1. \quad (20)$$

3. Постоянство матрицы  $W_{ij}^0(k)$  исходной модели, где изменения связей зависят только от их эффективностей  $x_i^1(k) + x_i^2(k)$ , в модифицированной версии отражено условием

$$\frac{\langle \Delta W_{ij}^0(k) \rangle}{\max W_{ij}^0(k) - \min W_{ij}^0(k)} \ll \frac{\langle \Delta(x_i^1(k) + x_i^2(k)) \rangle}{\max(x_i^1(k) + x_i^2(k)) - \min(x_i^1(k) + x_i^2(k))}, \quad (21)$$

где  $\Delta f(k) = f(k+1) - f(k)$ , а средние  $\langle \rangle$  и экстремумы вычисляются по одному и тому же достаточно широкому интервалу  $[k_1, k_2]$ . Выполнение соотношения (21) обеспечивается надлежащим выбором значений параметров  $A_{1(2)}, B_{1(2)}, C_{1(2)}$  (эти параметры уже определены в (19)),  $\mu$  и  $\nu$ . При исследовании динамических режимов модели полагаем

$$\mu = 0.001, \quad \nu = 0.1. \quad (22)$$

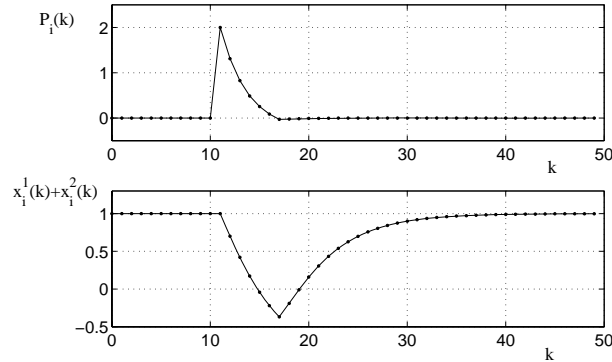


Рис. 1. Потенциал и эффективность синаптических связей нейрона до и после воздействия на него внешним сигналом  $S_i(10) = 2$  ( $S_i(k) = 0$  при  $k \neq 10$ ) при параметрах  $A_{1(2)}$ ,  $B_{1(2)}$ ,  $C_{1(2)}$ , заданных (19).

На рис. 2<sup>4</sup> показано, что влияние эволюции матричного элемента  $W_{ij}^0(k)$  на соответствующую эффективную связь  $W_{ij}(k)$  при условии (21) можно сравнить с медленной амплитудной модуляцией высокочастотных осцилляций, вызываемых быстрым изменением эффективности связей  $x_i^1(k) + x_i^2(k)$ .

Начальные условия выбраны тривиальными

$$P_i(0) = 0, \quad W_{ij}(0) = W_{ij}^0(0) = 0, \quad x_i^1(0) = x_i^2(0) = 0, \\ i, j = 1, \dots, n.$$

Для ввода в динамические режимы нейронная сеть подвергалась внешнему воздействию. В течение первых 2000 шагов в каждый момент времени на один случайно выбранный нейрон подавался сигнал величины 0.5.

Характеристики устойчивых динамических режимов практически не зависят от начальных условий, силы импульсов и количества воздействуемых за один шаг нейронов. Характер первичного стимулирования сети может повлиять только на сам факт появления устойчивого режима. К примеру, величины и частоты следования импульсов  $S_i(k)$  должны быть достаточными для активизации нейронов.

Количество нейронов во всех экспериментах над моделью, результаты которых приведены в настоящей работе, равно 64.

<sup>4</sup>На графиках рис. 2 точками отмечены только значения функций, и кривые лишь кажутся сплошными, ввиду того, что рассматривается большой интервал времени. Эту особенность нужно учитывать и в дальнейшем при рассмотрении аналогичных графиков однозначных функций.



### Динамика нейронной сети Кропотова-Пахомова

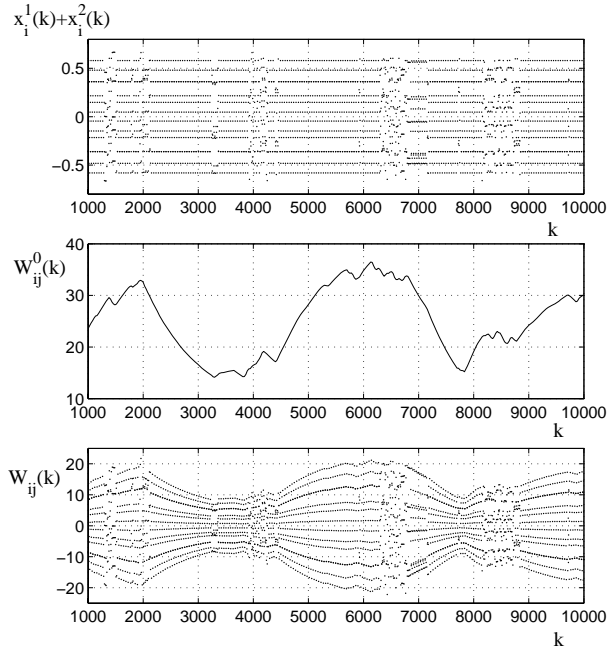


Рис. 2. Зависимости от времени эффективности связей  $x_i^1(k) + x_i^2(k)$   $i$ -го нейрона сети, связи  $W_{ij}^0(k)$ , идущей от  $j$ -го нейрона к  $i$ -му, и соответствующей эффективной связи  $W_{ij}(k)$  в неперiodическом динамическом режиме при значениях параметров, заданных (19) и (22).

### 3 Типы динамических режимов

Перед рассмотрением типов динамических режимов модифицированной модели нейронной сети Кропотова-Пахомова введем несколько определений.

Функция  $f(k)$ , заданная на множестве целых чисел, называется *периодической на интервале*  $[k_1, k_2]$ , если

$$f(k + mT) = f(k); \quad \forall m \in \mathbb{Z} : k + mT \in [k_1, k_2], \quad (23)$$

где  $T$  — период функции  $f(k)$ .

*Средняя активность* нейронной сети — это функция дискретного времени  $k$ , заданная выражением

$$\langle N(k) \rangle \equiv \sum_{i=1}^n N_i(k) / n. \quad (24)$$

Аналогично вводятся средние величины и для других динамических переменных модели (13)–(18).

Состояние нейронной сети называется *нулевым*, если все ее нейроны неактивны

$$N_i(k) = 0, \quad i = 1, 2, \dots, n.$$

*Обнуление* — переход нейронной сети в нулевое состояние из состояния с хотя бы одним активным нейроном.

Отметим, что после обнуления при отсутствии внешнего воздействия на сеть потенциалы и связи нейронов стремятся к нулевым значениям, и динамика нейронной сети становится *тривиальной*.

Отправным пунктом для определения типов динамических режимов нейронной сети являются зависимости активностей ее нейронов от времени.

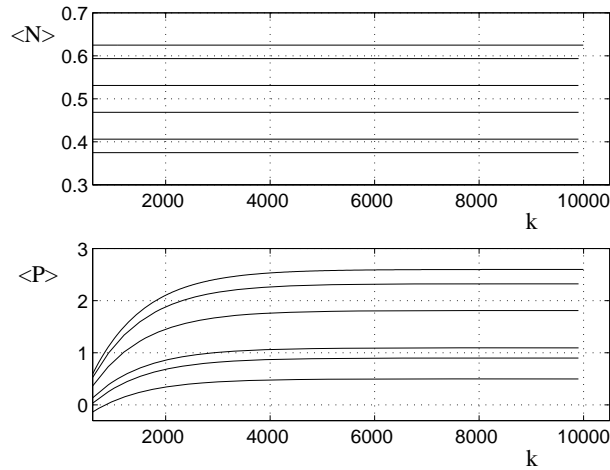


Рис. 3. Зависимости от времени средних активности и потенциала в процессе перехода нейронной сети в установившийся периодический динамический режим. Период колебаний  $T = 6$ .

Будем считать *периодическим* динамический режим, в котором активности всех нейронов сети являются периодическими функциями времени согласно определению (23), иначе динамический режим считаем *непериодическим*. Подчеркнем, что введенное определение периодических режимов допускает непериодичность остальных динамических переменных модели. С момента наступления периодического режима начинается переходный процесс к состоянию, где все динамические переменные модели периодичны (см. рис. 3).

Назовем *установившимся* периодический динамический режим, в котором *все* динамические переменные модели суть периодические функции времени.

Периодические режимы делятся также на *простые* и *сложные* в зависимости от наличия в сети нейронов, активности которых имеют разные периоды колебаний. Если периоды активностей *всех* нейронов сети одинаковы, то периодический режим называется *простым*. Если же в нейронной сети существует хотя бы два нейрона с различными периодами колебаний<sup>5</sup>, то такой периодический режим называется *сложным*.

Зависимости на рис. 3 представляют собой пример простого периодического режима, в котором период колебаний всех нейронов сети равен 6. Причем, имеются нейроны со всевозможными сдвигами фаз, поэтому количество различных значений, принимаемых средней активностью  $\langle N \rangle$  также равно 6. В случае сложного периодического режима ввиду наличия в сети нейронов в разные периоды колебаний количество различных относительных сдвигов фаз сильно возрастает, вследствие чего повышается число значений средней активности.

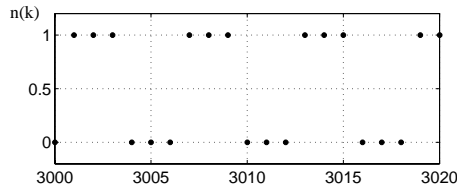


Рис. 4. Зависимость активности одного из нейронов сети от времени в периодическом режиме. Период колебаний  $T = 6$ .

В простейшем случае один период колебаний активности нейрона образован двумя блоками нулей и единиц одинаковой длины (см. рис. 4)

$$\underbrace{11 \dots 1}_{T/2} \underbrace{00 \dots 0}_{T/2}. \quad (25)$$

В нейронной сети реализуются и более сложные формы колебаний, при которых в двоичных последовательностях  $n_i(k)$  чередуются блоки нулей и единиц различных длин.

При рассмотрении динамики нейронной сети целесообразно ввести понятие *устойчивости* динамического режима, под которой будем понимать свойство сети сохранять текущий тип режима неизменным. В качестве меры устойчивости можно взять интервал времени, в течение которого не происходит изменения типа динамического режима при отсутствии действия на сеть внешнего сигнала.

Факт неустойчивости динамического режима легко определяется, если на интервале наблюдения за сетью происходит смена типа динамического поведения сети. Однако, при отсутствии подобных событий однозначный

<sup>5</sup>Говоря далее о колебаниях нейронов, будем понимать соответствующие периодические зависимости активностей нейронов от времени.

ответ можно дать лишь в случае установившегося периодического режима, который является устойчивым.

Понятие устойчивости будет использоваться только для сравнительного анализа различных неперiodических динамических режимов. Строгое же определение устойчивости и ее меры выходит за рамки настоящей работы. Достаточно сказать, что кроме меры устойчивости динамического режима во времени необходимо вводить и меру устойчивости относительно малых изменений параметров модели (см. следующий раздел).

#### 4 Фазы динамических режимов и фазовые переходы

Рассмотрим один из примеров фазовой диаграммы (рис. 5), где представлены все экспериментально наблюдаемые динамические режимы нейронной сети. Пределы изменения параметров

$$0 \leq \beta \leq 3, \quad 0 \leq \alpha \leq 1.$$

В сине-зеленых тонах отмечены периодические режимы, в желто-красных — неперiodические.

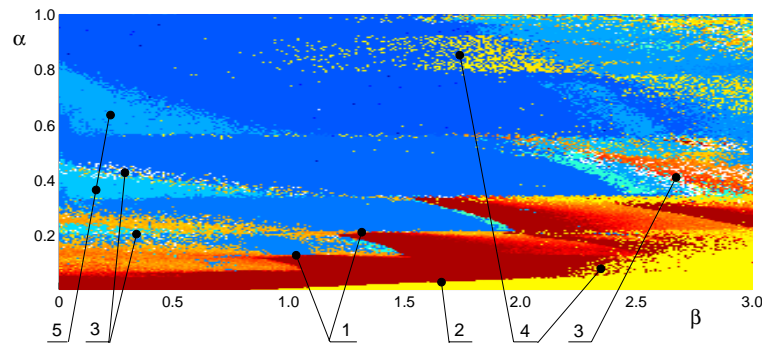


Рис. 5. Фазы динамических режимов. Для периодических режимов (синий-голубой-зеленый) отложены логарифмы периодов колебаний сети (переход от синего ( $T < 10$ ) к зеленому ( $T \sim 300$ ) соответствует увеличению периода колебаний), для неперiodических (желтый-оранжевый-красный) — логарифмы времен жизни сети (в желтых областях сеть обнуляется практически сразу после прекращения накачки, в красных областях время жизни сети превышает  $10^6$  шагов). Белым точкам соответствуют параметры, при которых наблюдаются длиннопериодические колебания с  $T > 500$  (см. сноска 3). Сносками 1 и 2 обозначены фазовые переходы: 1 — периодический  $\leftrightarrow$  устойчивый неперiodический, 2 — устойчивый неперiodический  $\leftrightarrow$  неустойчивый неперiodический. Сносками 4 отмечены примеры смешанных фаз, а сносками 5 — сложных периодических режимов.

Особенностями данной фазовой диаграммы являются наличие *сплошных*

и *смешанных* (сноски 4) фаз. В областях смешанных фаз тип динамического режима существенно зависит от малых изменений параметров и характера сигнала накачки. Важно отметить, что аналогичная ситуация происходит и на *четких* фазовых переходах (сноски 1 и 2) между сплошными фазами. В определенном смысле некоторые области смешанных фаз можно считать «сильно размытыми» фазовыми переходами.

Интересный пример представляет собой граница, отмеченная сносками 2 и 4 между устойчивым<sup>6</sup> (темно-красная область) и неустойчивым (желтая область) непериодическими режимами. Граница обладает протяженным интервалом четкого фазового перехода, а также «размытым» участком, представляющим собой область смешанных фаз.

Сноски 4 указывают на примеры особых областей смешанных фаз. Находясь в этих областях, нейронная сеть может эволюционировать весьма длительное время, но малая вариация параметров или незначительный внешний сигнал легко ее обнуляют. То есть устойчивый в смысле большого времени жизни сети динамический режим оказывается неустойчив по отношению к вариациям параметров и внешнему воздействию.

В окрестностях фазовых переходов или в областях смешанных фаз, где нейронная сеть чувствительна к внешнему воздействию, могут появляться различного рода эффекты, одним из которых является возникновение собственных и вынужденных длиннопериодических колебаний. В разделе 6 отмечается, что одним из условий появления вынужденных длиннопериодических колебаний должен быть выбор параметров на границе фазовых переходов. Собственные длиннопериодические колебания нейронной сети, как это видно на диаграмме рис. 5 (сноски 3), также появляются в областях сильной чувствительности сети<sup>7</sup>. Экспериментально показано, что собственные длиннопериодические колебания могут возникнуть и в результате локального по времени воздействия на нейронную сеть, находившуюся до этого в периодическом режиме при параметрах, соответствующих, например, фазовым переходам, отмеченным сносками 1 на рис. 5.

## 5 Энтропия рассинхронизации и энтропия связей

Нейронная сеть, находясь в периодическом динамическом режиме, разбита на группы синфазно колеблющихся нейронов — *кластеры*. Процесс синхронизации сети начинается или уже во время воздействия на нее сигналом накачки, или непосредственно после его отключения. Прекращение эволюции нейронной сети в непериодическом динамическом режиме наступает в результате обнуления, то есть, когда все нейроны одновременно становятся неактивны. Таким образом, в обоих рассмотренных случаях сеть развивается в сторону большей взаимной синхронизации нейронов, а время

<sup>6</sup>Понятия устойчивого динамического режима здесь относительное. В устойчивом режиме сеть эволюционирует гарантированно более  $10^6$  шагов, тогда как в неустойчивом обнуляется после прекращения воздействия на сеть сигналом накачки.

<sup>7</sup>Множество пар параметров (белые точки на рис. 5), соответствующих собственным длиннопериодическим колебаниям составляет менее 2% площади фазовой диаграммы.

жизни сети в непериодическом режиме будет зависеть от степени рассинхронизации и средней скорости ее уменьшения.

Удобной характеристикой является *энтропия рассинхронизации* нейронной сети, которую мы введем следующим образом. Рассмотрим зависимости активностей нейронов от времени  $N_i(k)$  на промежутке  $[k - \Delta k, k]$ , где  $k$  и  $\Delta k$  — некоторые момент и интервал дискретного времени. Эти зависимости представляют собой последовательности нулей и единиц одинаковой длины, то есть двоичные числа. Для удобства будем их записывать в десятичной системе. Таким образом, каждому моменту времени  $k$  и промежутку  $\Delta k$  сопоставлено множество целых десятичных чисел, несущих полную информацию о поведении активностей нейронов сети за последние  $\Delta k + 1$  шагов. На полученном множестве определим энтропию, предварительно выбрав такое покрытие рассматриваемого множества непересекающимися единичными отрезками, при котором количество этих отрезков минимально,

$$S_1(k, \Delta k) = - \sum_{i=1}^n p_i \ln p_i, \quad (26)$$

где  $n$  — количество нейронов в сети,  $p_i$  — вероятность попадания  $i$ -го элемента множества в интервал единичной длины<sup>8</sup>. Величину  $S_1(k, \Delta k)$  назовем энтропией рассинхронизации нейронной сети.

Определим энтропию связей  $W_{ij}^0(k)$  нейронной сети, как

$$S_r(k) = - \sum_{i=1}^{n^2} p_i \ln p_i.$$

Для вычисления вероятностей  $p_i$  необходимо найти минимальное покрытие непересекающимися интервалами длины  $r$  множества, состоящего из  $n^2$  вещественных чисел — матричных элементов матрицы  $W_{ij}^0(k)$ .

Величины  $S_1(k, \Delta k)$  и  $S_r(k)$  дополняют друг друга. С одной стороны, энтропия связей  $S_r(k)$  является функцией только состояния в момент  $k$ , тогда как энтропия рассинхронизации определяется по некоторому интервалу  $\Delta k$ . С другой стороны, как будет показано в разделе 6, кластеризация связей есть следствие кластеризации нейронов, поэтому величина  $S_1(k, \Delta k)$  «быстрее реагирует» на изменения упорядоченности нейронной сети.

Обе энтропии зависят от параметра — длины интервалов покрытия, но нас будет интересовать изменение энтропий со временем, а не их абсолютное значение.

Рассмотрим несколько примеров. На рис. 6 представлены переходные процессы, где нейронная сеть эволюционирует по направлению к устойчивым периодическим режимам, что сопровождается уменьшением обеих

<sup>8</sup>Определение (26) предполагает зависимость энтропии от длины интервалов покрытия, но в данном случае естественно покрывать множество интервалами единичной длины, поскольку покрываемое множество состоит из целых чисел.

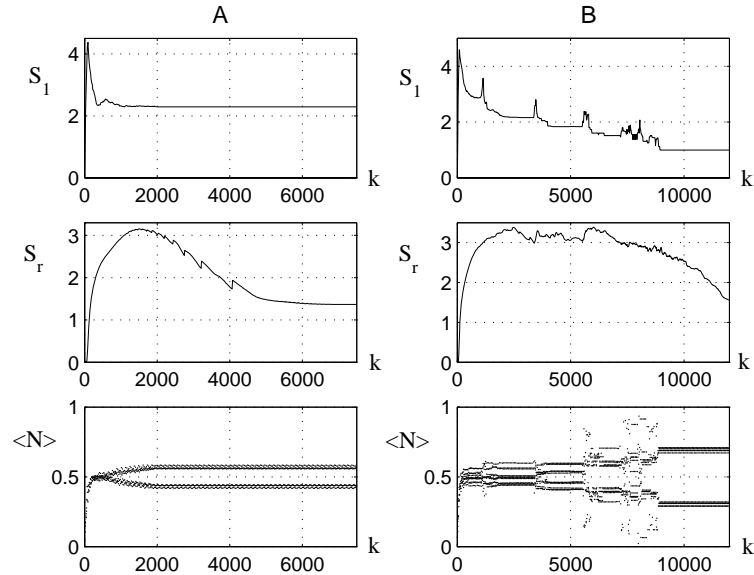


Рис. 6. Энтропия рассинхронизации  $S_1$ , энтропия связей  $S_r$  ( $r = 1.2$ ) и средняя активность  $\langle N \rangle$  в переходных процессах к устойчивым периодическим режимам. Сигнал накачки действует на интервале  $[0, 2000]$ . (А) Почти полная синхронизация нейронов происходит до момента отключения сигнала накачки. (В) Переходный процесс происходит в несколько стадий, формируемых длительными участками периодического режима и короткими — неперидического.

энтропий. Во втором случае переходный процесс содержит участки, на которых сеть находится в периодическом режиме, сменяемых кратковременными неперидическими. На рис. 7 показаны примеры неустойчивого и устойчивого<sup>9</sup> неперидических режимов. По поведению средней активности на рис. 7(А) ясно, что обнуление сети в неперидическом режиме происходит вследствие полной синхронизации всех нейронов или, говоря другими словами, возникновения только одного нейронного кластера, тогда как в периодических режимах таких кластеров несколько, что исключает обнуление. Из приведенных примеров следует, что синхронизация колебаний нейронов предшествует упорядочиванию межнейронных связей.

## 6 Периодический режим

Как только в некоторый момент времени  $k$  нейронная сеть попадает в периодический режим, то есть, когда зависимости активностей всех нейронов

<sup>9</sup>Устойчивость здесь понимается в том смысле, что за время наблюдения не произошло обнуления сети и отсутствуют явные тренды в приведенных зависимостях.

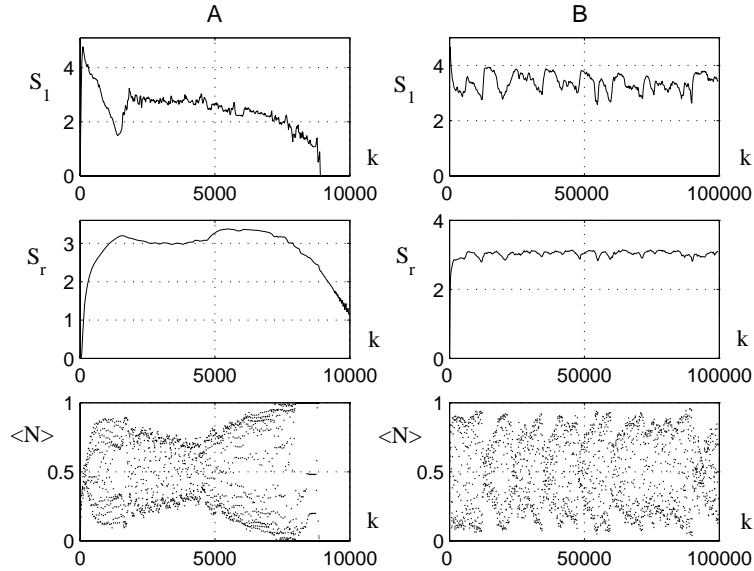


Рис. 7. То же, что и на рис. 6, но для неустойчивого (А) (обнуление при  $k = 8875$ ) и устойчивого (В) непериодических режимов.

от времени становятся периодическими функциями, наступает переходный процесс, асимптотическим состоянием которого является установившийся периодический режим (см. рис. 3), где уже все динамические переменные модели — периодические функции времени. Во время переходного процесса в сети формируется некоторая структура связей. Структура связей и характер колебаний нейронов взаимозависимы и поэтому произвольны. В нейронной сети реализуются только вполне определенные виды колебаний и структур. Перед их рассмотрением выведем формулу, позволяющую в установившемся периодическом режиме по известному поведению активностей нейронов найти функциональные зависимости других динамических переменных модели.

Рекуррентные соотношения (13)–(17), определяющие модель, имеют вид

$$A(k+1) = aA(k) + b(k), \quad (27)$$

где  $a$  — вещественная константа, принадлежащая интервалу  $[0, 1]$ ,  $b(k)$  — периодическая функция. Соотношение (27) задает последовательность, члены которой представляют собой суммы конечного степенного ряда

$$A(k+n) = a^n A(k) + \sum_{i=0}^{n-1} a^{n-1-i} b(k+i). \quad (28)$$



Не нарушая общности, можно положить  $n = mT$ , где  $m \in \mathbb{N}$ ,  $T$  — период функции  $b(k)$ . Далее перепишем сумму в (28) в виде двойной суммы

$$\sum_{i=0}^{T-1} \sum_{j=0}^{m-1} a^{mT-1-jT-i} b(k + jT + i).$$

Откуда, учитывая периодичность  $b(k)$  и суммируя по  $j$ , найдем

$$\sum_{i=0}^{T-1} b(k + i) a^{mT-1-i} \frac{1 - a^{-mT}}{1 - a^{-T}}.$$

Подставляя последнее выражение в (28) и переходя к пределу  $i \rightarrow \infty$ , окончательно получим

$$A(k) = \frac{1}{1 - a^T} \sum_{i=0}^{T-1} a^{T-1-i} b(k + i). \quad (29)$$

**Структура сети в установившемся простом периодическом режиме.** Будем рассматривать только простейшие типы колебаний активностей нейронов, имеющие вид (25).

В простом периодическом режиме колебания различных нейронов отличаются только относительными сдвигами фаз. Поэтому нейронная сеть разбивается на кластеры, состоящие из синфазно колеблющихся нейронов. Очевидно, что количество кластеров не может превышать величину  $T$ . Разделение сети на кластеры определяет структуру матрицы  $W_{ij}^0(k)$  в установившемся периодическом режиме.

В качестве примера рассмотрим структуру связей<sup>10</sup>, которая возникает в нейронной сети в простом периодическом режиме при  $T = 6$  и единичной задержке правила Богданова-Хэбба:  $\{m\} = 1$ .

Соотношение (15), определяющее эволюцию связей, будет иметь вид

$$W_{ij}^0(k + 1) = (1 - \mu) W_{ij}^0(k) + \nu N_i(k) N_j(k - 1).$$

Учитывая (29), получим выражение для зависимостей связей от времени в установившемся периодическом режиме

$$W_{ij}^0(k) = \frac{\nu}{1 - (1 - \mu)^T} \sum_{i=0}^{T-1} (1 - \mu)^{T-1-i} N_i(k) N_j(k - 1). \quad (30)$$

Без учета сдвигов фаз количество различных зависимостей произведений  $N_i(k) N_j(k - 1)$  от времени равно

$$\frac{T}{2} + 1 \equiv c_m. \quad (31)$$

<sup>10</sup> Далее под связями будем подразумевать матричные элементы  $W_{ij}^0(k)$ .

Число  $s_m$  определяет количество *типов* связей в сети. Связи одного типа имеют одинаковые формы колебаний (рис. 8) и отличаются только относительными сдвигами фаз, поэтому тип колебаний можно характеризовать средней за период  $T$  величиной связи

$$\frac{1}{T} \sum_{k=1}^T W_{ij}^0(k) \equiv w_m. \quad (32)$$

На рис. 9 построены зависимости от времени некоторых связей нейронной сети при переходе к установившемуся периодическому режиму. Эволюция остальных связей аналогична. Хорошо виден процесс постепенного распределения связей по типам. В нейронной сети реализуются только такие периодические режимы, при которых амплитуды колебаний связей меньше модулей разностей величин (32).

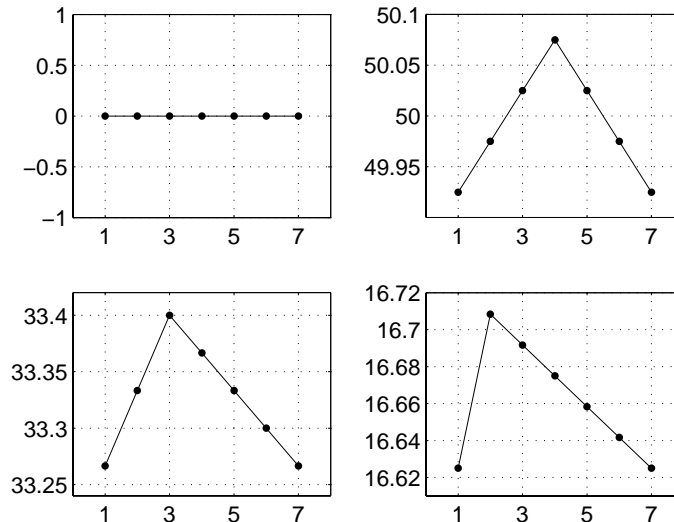


Рис. 8. Типы колебаний связей нейронной сети при  $T = 6$  в установившемся простом периодическом режиме. По горизонтальной оси отложено время, по вертикальной — значения матричных элементов.

Очевидно, что принадлежность связи к тому или иному типу определяется также количеством событий, связанных со срабатыванием правила Богданова-Хэбба, приходящихся на один период колебаний  $T$ . В рассматриваемом примере эти события могут происходить 0, 1, 2 или 3 раза за период. На рис. 8 растущие части графиков соответствуют наличию обучения, нисходящие — только диссипации связей.

Связи, принадлежащие к одним и тем же типам колебаний, образуют в нейронной сети свои собственные структуры. При построении соответствующих *структурных диаграмм* удобно иметь дело не с отдельными нейро-

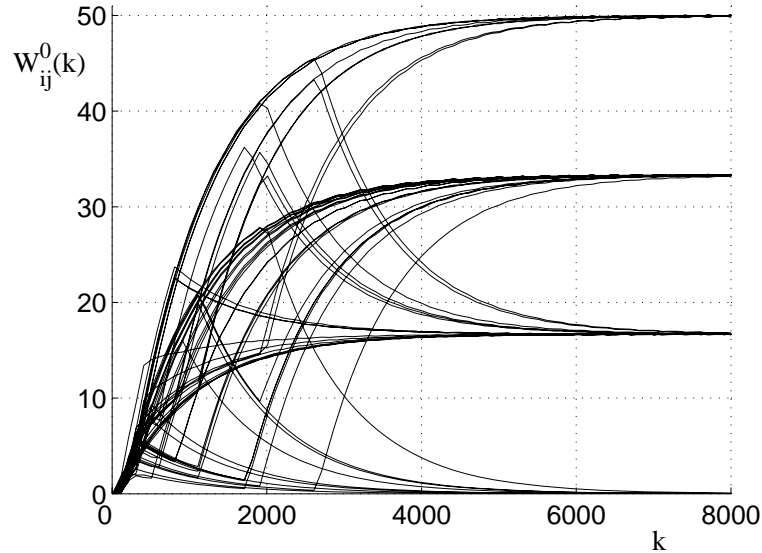


Рис. 9. Эволюция некоторых матричных элементов  $W_{ij}^0(k)$  при переходе к установившемуся периодическому режиму.  $T = 6$ .

нами сети, а с кластерами. Поскольку колебания нейронов, принадлежащих одному кластеру, идентичны, то, следовательно, одинаковы и колебания связей между нейронами одних и тех кластеров (*внутрикластерных связей*), а также связей, направленных от нейронов одного кластера к нейронам другого (*межкластерных связей*).

Будем обозначать кластеры жирными точками, а связи — соединяющими их стрелками. Стрелка символизирует группу связей одного типа, направленных от нейронов одного кластера к нейронам другого. Направление стрелок условимся выбирать, исходя из учета возможности нейронов воздействовать друг на друга. Под связью, *идущей от нейрона  $j$  к нейрону  $i$* , будем понимать матричный элемент  $W_{ij}^0(k)$ , определяющий в соответствии с выражением (13), силу воздействия на нейрон  $i$  со стороны  $j$ -го нейрона.

Расположим точки, символизирующие кластеры, на окружности так, чтобы при обходе по часовой стрелке колебания нейронов каждого следующего кластера отставали по фазе на единицу от колебаний нейронов предыдущего. То есть, если в момент времени  $k$  нейроны кластера 1 (см. рис. 10) становятся активными, то в момент  $k + 1$  активизируется кластер 2 и т. д. Если задержка правила Богданова-Хэбба равна единице, как в нашем примере, то максимальными (в смысле средних значений, взятых за период) будут связи, выходящие из кластеров, опережающих по фазе на единицу кластеры, в которые связи входят. На рис. 10 приведены величины  $w_m$

выходящих связей для одного кластера сети.

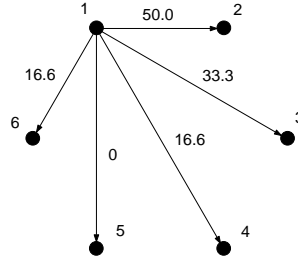


Рис. 10. Типы связей, выходящих из одного кластера нейронной сети. Рядом с каждой стрелкой отмечено среднее за период значение соответствующей связи.

Если для каждого из кластеров на одной диаграмме построить связи только одного типа, то получим искомые структурные диаграммы (рис. 11), где цифрами над каждой диаграммой указаны числа  $w_m$ .

Очевидно, что внутрикластерные связи являются симметричными  $W_{ij}^0(k) = W_{ji}^0(k)$ , в отличие от межкластерных связей, которые в общем случае (при использовании правил Богданова-Хэбба с ненулевыми задержками) несимметричны. В рассматриваемом примере внутрикластерные связи относятся к типу с  $w_m = 33.3$ .

Таким образом, в установившемся простом периодическом режиме вследствие того, что нейронная сеть разбита не более чем на  $T$  кластеров, в ней имеется  $T/2 + 1$  типов структур, сформированных связями, колебания которых различаются только относительными сдвигами фаз.

В рассмотренном выше примере количество кластеров равно величине периода колебаний нейронов  $T$ , то есть в нейронной сети имелись нейроны со всевозможными сдвигами фаз. Экспериментально показано, что реализуются и периодические режимы, где количество кластеров меньше  $T$ , однако только такие, при которых количество различных структур остается равным  $c_m$  (см. формулу (31)). Отметим, что число  $c_m$  определяется количеством *различных относительных* сдвигов фаз колебаний нейронов, поэтому уменьшение числа кластеров в сети может не влиять на количество структур. Для получения соответствующих структурных диаграмм необходимо построить все диаграммы, как если бы в нейронной сети имелись все  $T$  кластеров, а затем исключить из диаграмм отсутствующие кластеры со всеми их связями. На рис. 12 приведен пример, где отсутствует кластер 2 (см. рис. 11).

**Структура сети в установившемся сложном периодическом режиме.** Здесь мы рассмотрим только один тип сложного установившегося периодического режима, при котором формы колебаний нейронов такие же, как и в простых режимах (рис. 4), а количество разных длин периодов

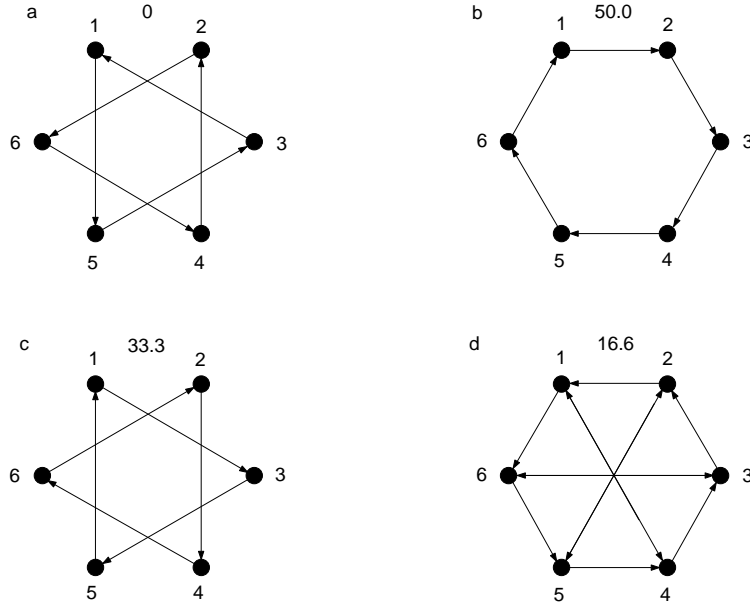


Рис. 11. Виды структур сети в простом установившемся периодическом режиме при  $T = 6$ . Здесь и далее цифры над структурами указывают средние значения связей (см. рис. 10).

равно двум. При этом периоды  $T_1$  и  $T_2$  имеют только соседние значения:  $T_2 = T_1 + 2$ .<sup>11</sup> В неустановившихся сложных режимах число различных периодов колебаний нейронов может быть больше двух, но все подобные режимы являются неустойчивыми, и эволюция сети, как правило, заканчивается обнулением, либо в очень редких случаях — переходом к периодическому режиму с двумя периодами колебаний.

Наличие в нейронной сети двух множеств  $A$  и  $B$ , состоящих из нейронов, периоды колебаний которых равны соответственно  $T_1$  и  $T_2$ , приводит к формированию трех групп структур связей:  $A \leftrightarrow A$ ,  $B \leftrightarrow B$  и  $A \leftrightarrow B$ .

Рассмотрим пример сложного установившегося периодического режима, в котором  $T_1 = 6$  и  $T_2 = 8$ . Группа структур связей между нейронами, имеющими периоды колебаний  $T_1 = 6$  представлена на рис. 11. На рис. 13 показана аналогичная группа структур для  $T_2 = 8$  (как и на рис. 11 вверху каждой диаграммы отмечено среднее за период значение связи  $w_m$ ).

Найдем по формуле (30) зависимости от времени для связей группы  $A \leftrightarrow B$ . Периоды колебаний будут равны  $T_1 T_2 / 2 = 24$ . Имеется два типа связей, отличающихся формами колебаний (см. рис. 14). Но поскольку для всех связей  $w_m = 25$ , и, кроме того, количество событий срабатывания правила

<sup>11</sup>Выполнение данного правила проверено, в частности, для фиксированных в разделе 2 параметрах и интервалов  $\alpha$  и  $\beta$ , использованных при построении фазовой диаграммы.

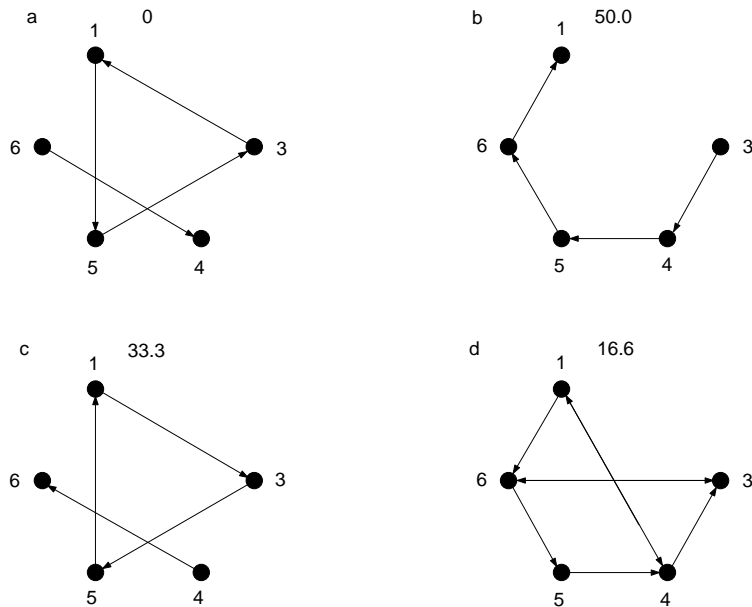


Рис. 12. То же, что и на рис. 11, но отсутствует кластер 2.

Богданова-Хэбба, приходящихся на один период, также одинаково, то все связи из группы  $A \leftrightarrow B$  можно отнести к одному типу.

Таким образом, в рассматриваемом сложном периодическом режиме нейронная сеть разделена на две подсети, внутренние структуры которых эквивалентны структурам сети, находящейся в простых режимах с соответствующими периодами колебаний нейронов. А так как совокупность связей между подсетями неструктурирована, то можно рассматривать эволюцию каждой из подсетей независимо, заменив связи на постоянный фоновый сигнал.

**Вынужденные длиннопериодные колебания.** Эксперименты, в результате которых был найден эффект вынужденных длиннопериодных колебаний, проводились в сотрудничестве с доктором филологических наук заведующей Лабораторией интеллектуальных систем НИИ математики и механики Санкт-Петербургского государственного университета профессором Мариной Абрамовной Красноперовой. Целью экспериментов являлась проверка некоторых гипотез, выдвинутых М. А. Красноперовой в рамках разрабатываемых ею формальных методов в лингвистической поэтике [24]. По результатам первичных исследований была опубликована совместная работа [25].

Эффект состоит в возникновении длиннопериодных колебаний связей

Динамика нейронной сети Кропотова-Пахомова

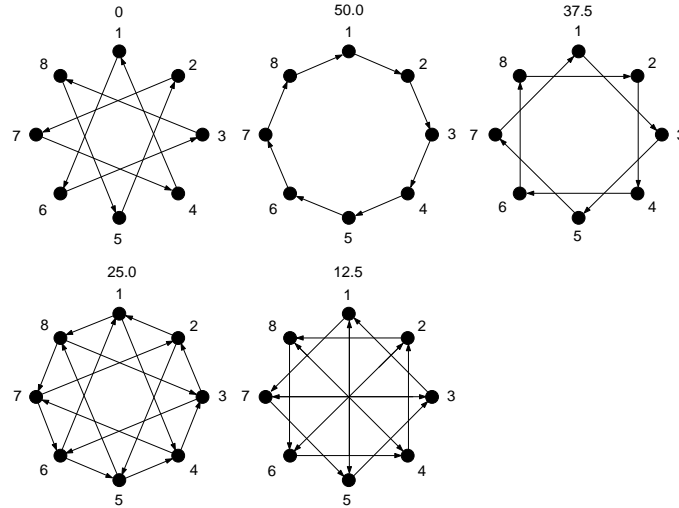


Рис. 13. То же, что и на рис. 11, но при  $T = 8$ .

нейронов, на которые действует периодический внешний сигнал. При этом происходят периодические изменения типов связей воздействуемых нейронов вследствие циклических переходов последних из кластера в кластер.

Длина периода  $T_l$  колебаний удовлетворяет условию

$$T_l = TT_s c, \tag{33}$$

где  $T$  — период собственных колебаний нейронной сети,<sup>12</sup>  $T_s$  — период внешнего сигнала,  $c$  — некоторая целая положительная константа.

Необходимыми условиями возникновения длиннопериодных колебаний являются следующие факторы

- Период внешнего сигнала  $T_s$  не должен быть кратен периоду собственных колебаний сети  $T$ .
- Нейронная сеть должна находиться в установившемся периодическом режиме при параметрах, соответствующих границе фазового перехода между периодическим и непериодическим режимами.

Длиннопериодные колебания связей нейрона устанавливаются через некоторое время после начала действия на этот нейрон внешними периодическими импульсами (рис. 15) и продолжают входить до окончания данного воздействия. После возобновления подачи сигнала длиннопериодные колебания восстанавливаются (рис. 16).

Рассмотрим механизм возникновения колебаний в простом режиме. Напомним, что в периодическом режиме сеть разбита на группы синфазно

<sup>12</sup>Для сложного периодического режима  $T = T_1 T_2$ .

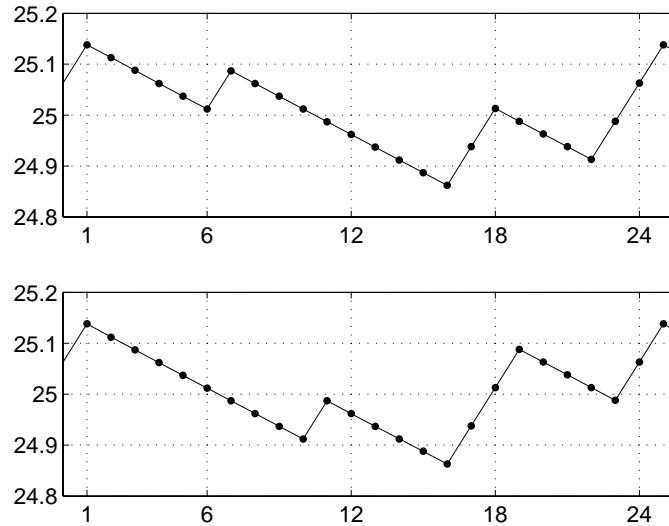


Рис. 14. То же, что и на рис. 8, но при  $T = 8$ .

колеблющихся нейронов — кластеры. В простом периодическом режиме колебания активностей нейронов различаются лишь относительными сдвигами фаз. Было установлено, что тип конкретной связи, характеризующейся величиной  $w_m$  (32), определяется сдвигом фаз колебаний нейронов, между которыми эта связь проходит. Если вследствие, например, внешнего воздействия у какого-то нейрона сети изменяется фаза колебаний, то он переходит из одного кластера в другой. Начинается переходный процесс, в результате которого типы всех связей этого нейрона изменяются.

При длиннопериодных колебаниях нейрон циклически переходит из одного кластера в другой. Один цикл таких переходов образует период длиннопериодных колебаний. Время нахождения нейрона в одном кластере, то есть время между двумя последовательными изменениями фазы колебаний нейрона, называется *сегментом*. Поведение воздействуемого нейрона и сети в целом в рамках одного сегмента можно считать неустановившимся периодическим режимом с характерным стремлением динамических характеристик к своим средним предельным значениям. Если отключить внешний сигнал, то со временем мы получим картину установившегося периодического режима (см. рис. 17).

Особо подчеркнем, что равные длинам сегментов интервалы времени между двумя последовательными сдвигами фаз колебаний нейрона могут значительно превосходить период следования внешних импульсов, воздействующих на нейрон. Откуда следует, что на протяжении одного сегмента нейрон нечувствителен к внешнему воздействию в смысле изменения своей



Динамика нейронной сети Кропотова-Пахомова

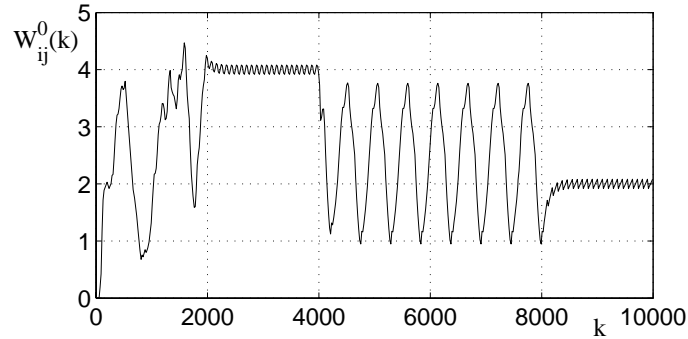


Рис. 15. Длиннопериодные колебания ( $T_l = 540$ ) одной из связей нейрона, на который подается периодический сигнал.  $[0, 2000]$  — интервал «накачки» сети,  $[4000, 8000]$  — интервал воздействия на нейрон внешними импульсами величины 0.5 с периодом  $T_s = 9$ . Период собственных колебаний сети  $T = 10$ .

активности. Однако в точках соединений сегментов, то есть в моменты, когда происходит изменение фазы колебаний активности нейрона, потенциал последнего вследствие действия внешнего импульса на предыдущем шаге получает дополнительное приращение, достаточное для «внеурочного» преодоления порога активации. Сравнивая графики зависимостей потенциала нейрона и его связей от времени (рис. 18), можно увидеть, что в окрестности точки сдвига фазы потенциал испытывает флуктуации. Подобное *нелинейное* поведение потенциала нейрона реализуется в результате совместного действия нескольких параллельно протекающих процессов и становится возможным при соблюдении ряда условий. К примеру, величина амплитуды внешнего сигнала, подбираемая экспериментально, может принимать

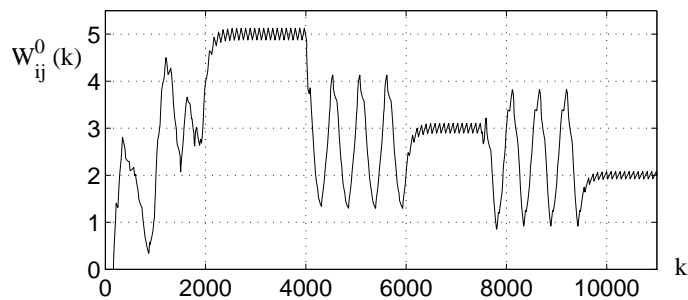


Рис. 16. То же, что и на рис. 15, но внешний периодический сигнал подается на двух интервалах:  $[4000, 6000]$  и  $[7500, 9500]$ .

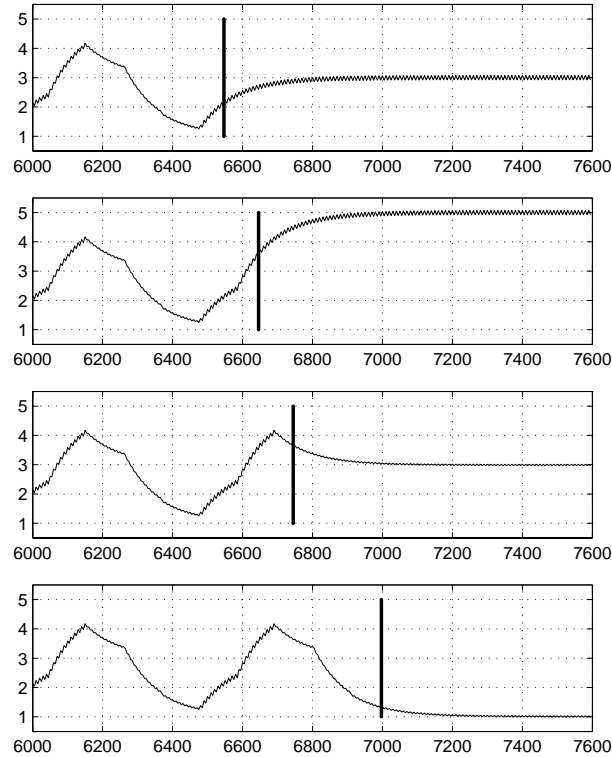


Рис. 17. Эволюция связи нейрона до и после момента (отмечено вертикальным отрезком) прекращения воздействия на него периодическими импульсами.

значения только в достаточно узком интервале. Получить аналитическое описание динамики потенциала при длиннопериодных колебаниях весьма затруднительно, поэтому далее приводится только качественное описание деталей процесса.

Введем понятие *уровня* динамической переменной нейронной сети, находящейся в периодическом режиме. Пусть  $a(k)$  — некоторая динамическая переменная, а  $T$  — период собственных колебаний нейронной сети. Тогда *уровнем* переменной  $a(k)$  назовем последовательность значений, принимаемых  $a(k)$  с интервалом  $T$ . Очевидно, что в дискретном времени количество уровней у одной переменной равно  $T$ , а в установившемся периодическом режиме уровни всех переменных сети суть постоянные функции времени. Отдельные кривые на графиках рис. 3 представляют примеры уровней для средних активности и потенциала.

Рассмотрим динамику потенциала нейрона, на который действует периодический внешний сигнал, на интервале одного сегмента (рис. 19). Воздей-

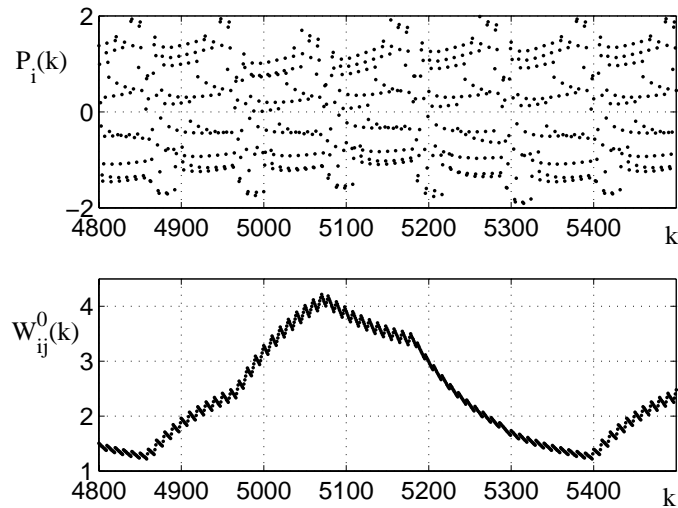


Рис. 18. Один период длиннопериодных колебаний. На верхнем графике показана зависимость потенциала воздействуемого нейрона от времени, на нижнем — одной из его связей.

ствие на нейрон внешним импульсом в некоторый момент времени вызывает в следующий момент приращение его потенциала. Это хорошо видно на представленном графике. Если значение потенциала ниже нуля, а по абсолютной величине меньше амплитуды импульса, то при получении потенциалом приращения его значение станет положительным и вызовет активизацию нейрона, что в свою очередь приведет к сдвигу фазы колебаний последнего.

В рассматриваемом случае это происходит только в моменты времени  $k_1 = 5936$  и  $k_2 = 6044$ . Интервал  $[k_1, k_2]$  определяет длительность текущего сегмента длиннопериодных колебаний. Внешние импульсы, поданные в промежутке  $k_1 < k < k_2$ , не влияют на поведение активности нейрона. Здесь сказываются следующие факторы:

- соотношение периода внешних импульсов и периода собственных колебаний сети;
- соотношение амплитуды импульсов и абсолютного значения разности значений уровней потенциала воздействуемого нейрона;
- скачок уровней потенциала, связанный с событием сдвига фазы активности, и последующее их стремление к предельным значениям.

Основную роль играет расхождение периодов внешних импульсов и собственных колебаний сети. За каждый период колебаний сети импульсы отстают на один шаг, вследствие чего приращения получают разные уровни

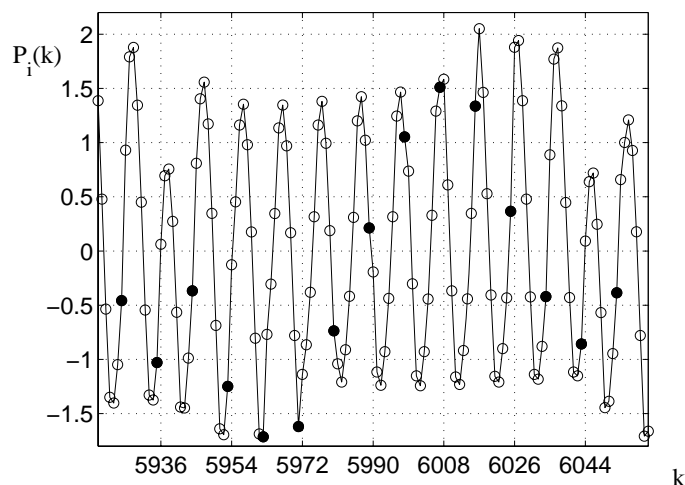


Рис. 19. Динамика потенциала воздействуемого нейрона в промежутке между двумя последовательными изменениями фазы активности. Черные кружки соответствуют моментам подачи внешних импульсов. Период следования внешних импульсов  $T_s = 9$ . Период собственных колебаний сети  $T = 10$ . Величина импульсов равна 0.5. Порог активизации  $h = 0$ .

потенциала (см. рис. 20).

На рис. 20 видна эволюция каждого из уровней потенциала до и после воздействия на них внешними импульсами. Непосредственно после получения потенциалом приращения соответствующий уровень испытывает скачок, а затем стремится к некоторому установившемуся значению. В моменты сдвигов фаз уровни перемешиваются. Фаза колебаний активности нейрона сохраняется до тех пор, пока под воздействием внешнего импульса какой-либо уровень не перейдет из области отрицательных значений потенциала в положительную, где нейрон становится активным.

Представленный пример есть лишь частный случай из многообразия возможных сценариев поведения потенциала. Тем не менее, здесь хорошо прослеживается влияние вышеупомянутых факторов. Отметим, что скачок состояний потенциала в отрицательную область исключил сдвиг фазы на 5954 шаге. Здесь проявилось влияние третьего из вышеперечисленных факторов. Как правило, фактор скачка состояний действует в случаях, когда период следования внешних импульсов сравним с длительностью скачка. Решающую роль в формировании эффекта длиннопериодных колебаний играют специальные амплитуда и период внешнего сигнала. Обязательно должно быть соблюдено условие некратности периодов  $T$  и  $T_s$ .

На рис. 21 показан пример длиннопериодных колебаний связей, где фактор скачка уровней потенциала несущественен, так как период следования

Динамика нейронной сети Кропотова-Пахомова

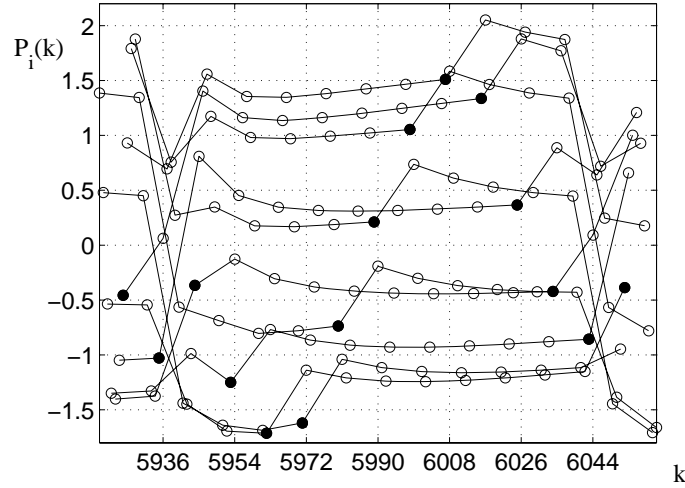


Рис. 20. Динамика уровней потенциала воздействующего нейрона на интервале одного сегмента.

внешних импульсов велик по сравнению с характерным временем *насыщения* уровней.

В отличие от рассмотренного примера длиннопериодные колебания связей могут быть достаточно нетривиальными, а их период весьма длительным даже в случаях, когда частота следования внешних импульсов не намного выше частоты собственных колебаний нейронной сети (см. примеры на рис. 22).

**Собственные длиннопериодные колебания.** В нейронной сети помимо рассмотренных выше вынужденных колебаний могут возникать и собственные колебания, которые в отличие от вынужденных не прекращаются после отключения внешнего воздействия. Кроме того, для появления собственных колебаний стимулирование сети не обязано быть периодическим. Колебания возникают либо «сами собой» через некоторое время (оно может быть достаточно большим) после снятия сигнала накачки, либо как следствие дополнительного воздействия. В обоих случаях факт появления собственных длиннопериодных колебаний — событие достаточно редкое, и можно лишь примерно указать области параметров, где следует их ожидать. Например, на фазовой диаграмме рис. 5 количество пар параметров  $(\alpha, \beta)$ , при которых они возникают, составляет менее 2% от общего числа экспериментально перебранных пар.

Механизм возникновения собственных длиннопериодных колебаний практически аналогичен картине при вынужденных колебаниях. Сеть кластеризована, находясь в установившемся периодическом режиме. Для одного или нескольких нейронов «не находится места», и они начинают периоди-

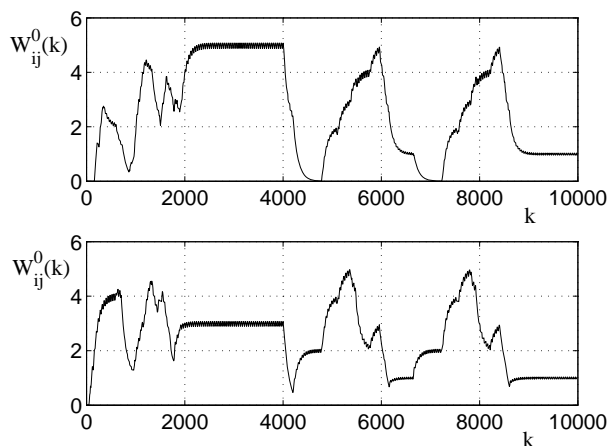


Рис. 21. Длиннопериодные колебания двух связей воздействующего нейрона. Период внешних импульсов  $T_S = 97$ .

чески изменять свое «место жительства», перескакивая из одного кластера в другой. Полный цикл таких перескоков определяет период колебаний. Отличие от вынужденных колебаний состоит в том, что на «бездомные» нейроны не оказывается периодического внешнего воздействия — его заменяет вся сеть.

Наибольший найденный экспериментально период длиннопериодных колебаний для сети с  $n = 64$  равен 39144.

## 7 Непериодический режим

Зависимости от времени активностей  $N_i(k)$  нейронов сети, находящейся в неперриодическом динамическом режиме, можно рассматривать, как последовательности следующих друг за другом отрезков простых периодических колебаний (см. рис. 4) различных периодов и длин. Оказывается, что количество таких периодов сильно ограничено (как правило, не превосходит 3-4-х), и в среднем длины интервалов, на которых не происходит изменения периода колебаний, в несколько раз превышают длины самих периодов. В связи с этим целесообразно ввести понятие частот колебаний нейронов сети для неперриодических режимов.

Рассмотрим зависимость активности одного из нейронов сети от времени  $N_i(k)$  на некотором интервале  $[k_1, k_2]$ . Выберем границы интервала, исходя из условий  $N_i(k_1) \neq N_i(k_1 - 1)$ ,  $N_i(k_2) \neq N_i(k_2 + 1)$ . Зависимость  $N_i(k)$  представляет собой двоичную последовательность. Разделим ее на блоки, каждый из которых будет состоять либо только из нулей, либо из единиц. Обозначим символами  $t_j$  ( $j \in \mathbb{N}$ ) всевозможные длины блоков, имеющиеся на интервале  $[k_1, k_2]$ .

Динамика нейронной сети Кропотова-Пахомова

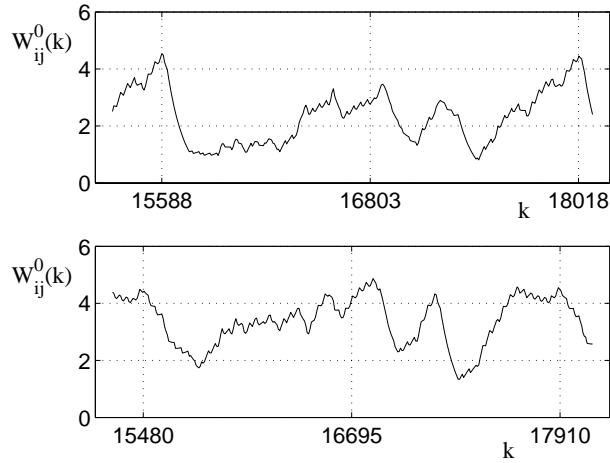


Рис. 22. Один период длиннопериодных колебаний ( $T_l = 2430$ ) двух связей воздействуемого нейрона. Период внешних импульсов  $T_s = 27$ . Период собственных колебаний нейронной сети  $T = 10$ .

Назовем величины

$$\frac{1}{2t_j} \tag{34}$$

частотами колебаний нейрона на интервале  $[k_1, k_2]$ . Очевидно, величины  $2t_j$  имеют смысл длин периодов, соответствующих частотам (34).

В качестве первого примера рассмотрим некоторый интервал *типичной* зависимости  $N_i(k)$  в неперiodическом режиме

```

111111110000000011111111000000001111111100000000111111110000000011111111
000000001111111100000000111111110000000011111111000000001111111100000000
111111110000000011111111000000001111111100000000111111110000000011111111
000000001111111100000000111111110000000011111111000000001111111100000000
111111110000000011111111000000001111111100000000111111110000000011111111
000000001111111100000000111111110000000011111111000000001111111100000000
111111110000000011111111000000001111111100000000111111110000000011111111
000000001111111100000000111111110000000011111111000000001111111100000000
1111111100000000111111110000000011111111000000001111111100000000

```

Здесь есть блоки лишь двух различных длин:  $t_1 = 8$  и  $t_2 = 9$ . Следовательно, имеется две частоты колебаний нейрона

$$\frac{1}{2t_1} = \frac{1}{16}, \quad \frac{1}{2t_2} = \frac{1}{18}.$$

Двоичную последовательность, представляющую зависимость  $N_i(k)$ , можно переписать в виде<sup>13</sup>

$$3t_2 4t_1 22t_2 4t_1 6t_2 6t_1 6t_2 4t_1 6t_2 6t_1 6t_2 4t_1, \quad (35)$$

где, например, величина  $3t_2$  обозначает первые три блока, длина каждого из которых равна 9, а величина  $4t_1$  — четыре последующих блока длины 8.

Подобные последовательности можно представлять в виде<sup>14</sup>

$$\{c_{ij}t_j\}_{i,j=1}^{p,q}, \quad (36)$$

где  $c_{ij} \in \mathbb{N}$ ,  $p$  — количество блоков, а  $q$  — количество различных полуинтервалов  $t_j$ , расположенных на рассматриваемом интервале времени  $[k_1, k_2]$ .

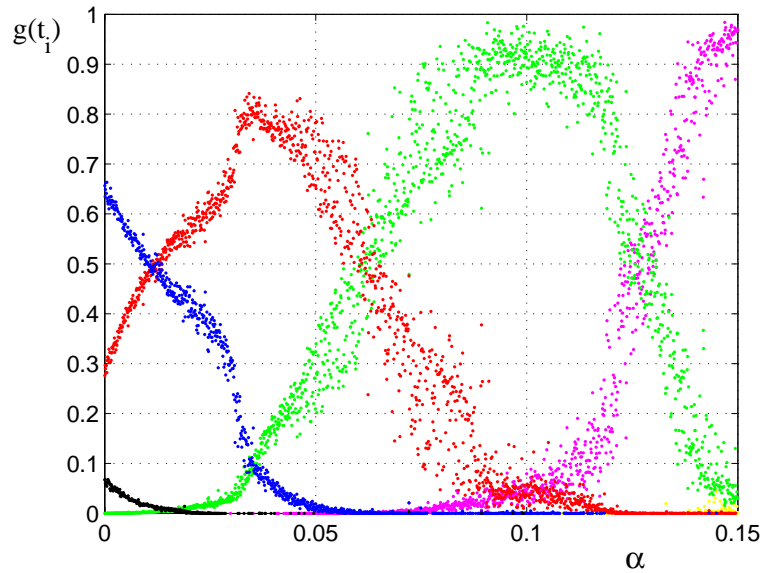


Рис. 23. Зависимости относительных сумм длин интервалов, соответствующих отдельным частотам колебаний нейронов сети, от параметра  $\alpha$  при  $\beta = 0.2$ .  $t_j = 5$  — желтый,  $t_j = 6$  — розовый,  $t_j = 7$  — зеленый,  $t_j = 8$  — красный,  $t_j = 9$  — синий,  $t_j = 10$  — черный.

Поставим эксперимент, состоящий из большого числа опытов над моделью при различных параметрах, соответствующих непериодическому режиму. Рассмотрим одно из сечений областей параметров диаграммы рис. 5,

<sup>13</sup>В записи (35) утерянную информацию о порядке следования единичных и нулевых блоков можно восстановить, указав с какого числа начинается последовательность. Для дальнейшего рассмотрения эта информация излишняя, так как нас будут интересовать только длины блоков.

<sup>14</sup>Здесь и далее, где встречается выражение  $c_{ij}t_j$  суммирование по индексу  $j$  не производится.



относящихся к «сплошной» фазе неперидического режима

$$0 \leq \alpha \leq 0.15, \quad \beta = 0.2. \quad (37)$$

Выполним 1501 опыт для значений параметра  $\alpha = 0.0001k$ ,  $k = 0, \dots, 1500$ . В результате получим, что в 52% опытов количество различных частот  $q$  равно 4, в 43% — 3, в 4% — 2 и на 1% приходятся случаи с  $q = 5$ .

Для каждого из опытов вычислим относительные суммы длин интервалов, соответствующих отдельным частотам колебаний нейронов сети

$$g(t_j) \equiv \frac{\sum_{m=1}^n \sum_i c_{ij}(m) t_j(m)}{nl}, \quad j = 1, \dots, q, \quad \sum_{j=1}^q g(t_j) = 1, \quad (38)$$

где  $c_{ij}(m)$ ,  $t_i(m)$  — коэффициенты и длины блоков в представлении (36) для  $m$ -го нейрона,  $l$  — время наблюдения,  $n$  — количество нейронов сети.

На рис. 23 представлены зависимости величин  $g(t_j)$  от параметра  $\alpha$  для наблюдаемых значений  $t_j = 5, \dots, 10$ .

Отметим также, что численные значения длин полупериодов  $t_j$ , упорядоченных, например, по возрастанию, удовлетворяют условиям  $t_{j+1} = t_j + 1$ ,  $j = 1, \dots, q$ .

Приведенный выше пример отражает общую ситуацию<sup>15</sup>, а именно, динамика модели в неперидическом режиме при условии отсутствия внешнего воздействия такова, что каждый нейрон сети имеет *малое* количество частот колебаний в смысле определения (34), причем еще меньшее количество частот являются *доминирующими* (см. рис. 23).

Вышеупомянутые факты особенно отчетливо проявляются на форме автокорреляционных функций средних по сети динамических характеристик. На рис. 24 показаны графики автокорреляционных функций средних по сети потенциала и матричного элемента матрицы  $W_{ij}^0(k)$ , где хорошо видна периодичность в следовании максимумов. Расстояния между соседними максимумами определяются доминирующими частотами зависимостей активностей нейронов сети от времени  $N_i(k)$ .

Характер изменения среднего времени жизни<sup>16</sup> нейронной сети в области параметров (37) показан на рис. 25.

Если нейронная сеть находится в неперидическом динамическом режиме, то длины  $c_{ij} t_j \equiv \Delta k$  интервалов, соответствующих определенной частоте  $1/2t_j$ , можно рассматривать как значения некоторой случайной величины. Для построения ее плотности распределения  $p(\Delta k, t_j)$  необходимо, чтобы сеть эволюционировала достаточно длительное время.

Рассмотрим эволюцию нейронной сети при параметрах  $\alpha = 0.001$ ,  $\beta = 0.2$  на интервале времени  $l = 10^8$ . По зависимости рис. 25 можно заключить,

<sup>15</sup>Пример показателен еще тем, что большинство опытов (см. рис. 25) относилось к неустойчивому неперидическому режиму, то есть сеть обнулялась за время наблюдения.

<sup>16</sup>Усреднение велось по достаточно узким интервалам  $\alpha$ , содержащим представительное количество точек.

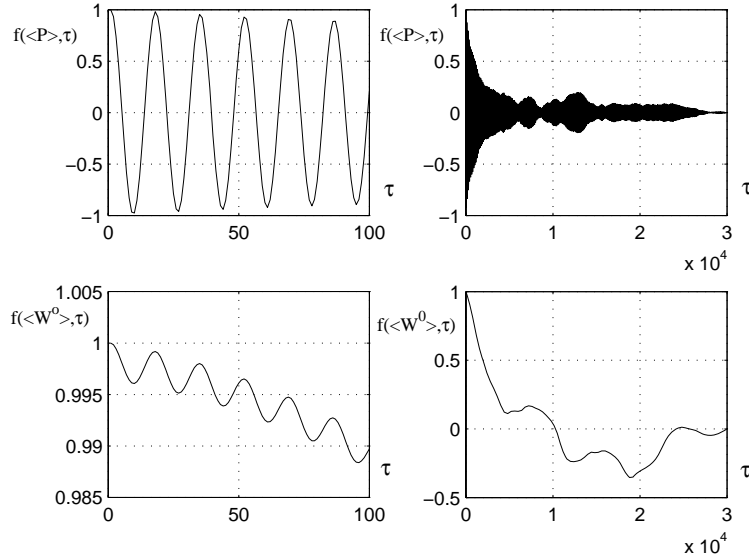


Рис. 24. Автокорреляционные функции средних потенциала (вверху) и матричного элемента матрицы  $W_{ij}^0(k)$  (внизу). Левая пара графиков соответствует малому интервалу сдвигов  $\tau$ , правая — интервалу  $[0, 3 \cdot 10^4]$ .

что среднее время жизни сети при этом значительно превосходит  $10^8$ . Найдем плотность распределения для частоты  $1/18$ , которая согласно рис. 23 является доминирующей (синий график с  $t_j = 9$ ). Плотность распределения будем искать по формуле

$$p(\Delta k, t_j) = \frac{\sum_{m=1}^N \sum_i c_{ij}(m) t_j(m)}{N l g(t_j)}, \quad (39)$$

где  $N$  — количество нейронов,  $l$  — время наблюдения. Здесь знаменатель равен суммарной длине всех интервалов, соответствующих колебаниям с частотой  $1/2t_j$ , а числитель — сумме интервалов того же сорта фиксированной длины  $\Delta k$ , поэтому условие  $\sum_{\Delta k=1}^{\infty} p(\Delta k, t_j) = 1$ , которому должна

удовлетворять плотность распределения, выполнено. Очевидно, что значения величин  $\Delta k$  кратны длине полупериода  $t_j$ , поэтому искомая плотность вероятности будет отлична от нуля только в точках, кратных  $t_j$ .

На рис. 26 представлена зависимость (39) для  $t_j = 9$ . Поскольку по обеим осям отложены логарифмы, взятые от аргумента и значений исследуемой функции, и график имеет явные линейные участки, то приходим к выводу, что плотность распределения имеет кусочно-степенную форму, то

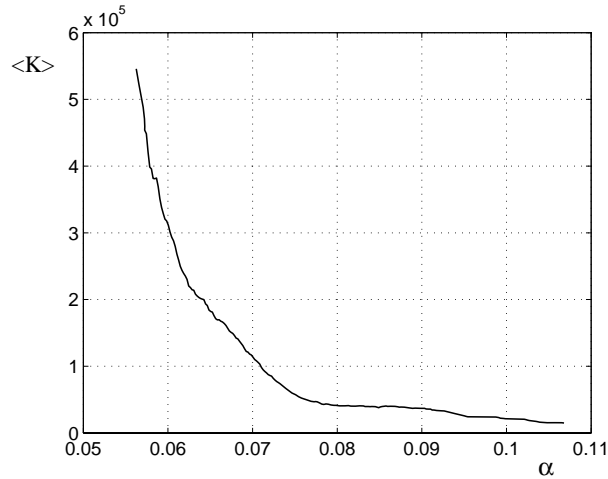


Рис. 25. Среднее время жизни сети при  $0 \leq \alpha \leq 0.15$ ,  $\beta = 0.2$ .

есть на отдельных участках может быть представлена в виде

$$p(\Delta k, t_j) = a \Delta k^{-\varphi}, \quad (40)$$

где  $a$ ,  $\varphi$  — положительные вещественные константы.

Метод наименьших квадратов для случая линейной регрессии дает следующие значения трех коэффициентов  $\varphi$ , относящихся к трем прямолинейным участкам графика рис. 26:

$$\begin{aligned} &1.637 \pm 0.006, \\ &4.76 \pm 0.05, \\ &10.2 \pm 0.3. \end{aligned}$$

### Обсуждение

В заключение приведем некоторые соображения относительно возможных направлений исследования как модифицированной модели Кропотова-Пахомова, так и нейронных сетей, которые могут быть построены с использованием задействованных в ней механизмов.

Оригинальная модель Кропотова-Пахомова строилась авторами с целью изучения процессов, протекающих в коре головного мозга человека. При этом мы должны четко осознавать, что на современном уровне развития науки деятельность, связанная с построением всевозможных компьютерных моделей нейронных сетей, какими бы «реалистическими» они ни были, практически не приближает нас к пониманию механизмов работы мозга,

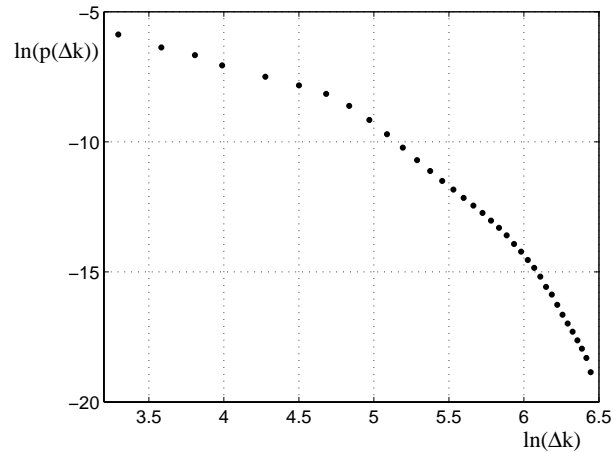


Рис. 26. Плотность распределения для длин временных интервалов, соответствующих частоте  $t_j = 9$ .

связанных с такими феноменами как сознание или разум. Подобная позиция достаточно полно и аргументированно отражена в замечательных книгах Роджера Пенроуза [26].

Во введении отмечалось, что целью настоящей работы не является моделирование каких-либо реальных физических процессов; это относится, в частности, и к процессам, протекающим в коре головного мозга человека. Задача работы более скромная — исследовать саморганизирующуюся нейронную сеть на предмет поиска в ее динамическом поведении свойств, заслуживающих дальнейшего детального изучения уже применительно к реальным физическим объектам. Именно поэтому в качестве основы была взята реалистическая нейронная сеть.

Среди примечательных свойств предложенной модели, безусловно, стоит упомянуть явления, связанные с кластеризацией, кусочно-степенной вид распределений, а также особенности динамики нейронной сети в переходных процессах и на границах фазовых переходов.

Одной из ярковыраженных отличительных черт модифицированной модели является наличие высокочастотной составляющей в зависимостях динамических переменных от времени. Причина состоит в использовании быстроизменяемых связей, что на практике реализовано применением видоизмененного правила Богданова-Хэбба (8). Вопрос о реальном существовании быстромодифицирующихся синапсов можно не ставить, поскольку все, что относится к короткопериодическим осцилляциям, целесообразно рассматривать в качестве фона, на котором развивается интересная для исследователя динамика. При таком подходе периодические режимы, рассмотренные в разделе 6, соответствуют различным *стационарным* состояниям кластеризованной нейронной сети. При длиннопериодических коле-

баниях уже сама стационарная сеть играет роль фона для периодической эволюции кластера нейронов. Таким образом, именно режимы, при которых наблюдаются длиннопериодические колебания, следует считать «истинно» периодическими. При проведенных экспериментах наличие длиннопериодических колебаний встречалось довольно редко, что побуждает исследовать это явление более основательно, в особенности, с целью выяснения уровня универсальности подобного поведения самоорганизующихся нейронных сетей.

Идеологию исключения высокочастотной составляющей следует распространить и на непериодические режимы. Делать это можно разными способами: от построения огибающих для временных зависимостей динамических переменных с последующей интерполяцией до применения частотных фильтров. Неявно один из подобных способов использовался при построении распределения, показанного на рис. 26.

Поведение нейронов сети в непериодическом режиме можно интерпретировать как переключение между несколькими стационарными состояниями. Тогда каждый из цветных графиков на рис. 23 будет соответствовать «весу» того или иного состояния, а зависимость, представленная на рис. 26, — распределению вероятностей времён жизни нейронов сети в состоянии, имеющем наибольший «вес». Изломанность графика рис. 26 может служить отправной точкой еще для одного направления исследования. Было бы интересно выявить именно те внутренние механизмы модели, которые приводят к подобным кусочно-степенным распределениям, а также осуществить поиск аналогов этих механизмов в природных процессах.

Динамика модели в окрестностях фазовых переходов, особенно на границах, где с одной стороны расположена область устойчивого непериодического режима, по-видимому, является одной из наиболее заманчивых тем. Начать было бы полезно, например, с построения диаграмм наподобие рис. 5, но уже для зависимостей энтропий от ключевых параметров модели.

Интересно проследить влияние различных задержек модифицированного правила Богданова-Хэбба на динамические и статистические характеристики нейронной сети. Как уже отмечалось, применение задержек позволяет получить несимметричную матрицу связей, что помимо простого обобщения исходной модели дает возможность записывать в сеть и затем воспроизводить примитивные временные последовательности образов.

Если говорить о проведении каких-либо параллелей между динамикой модели и, к примеру, поведением групп нейронов головного мозга человека, то здесь, по-видимому, следует начинать с вычисления известных или построения новых численных характеристик, которые позволили бы сравнивать результаты на формальном уровне. Определенный оптимизм, основанный пока только на качественных результатах, не вошедших в настоящую работу, на этот счет есть. Отметим, в частности, определенное сходство Фурье-образов, построенных для некоторых зависимостей динамических переменных сети от времени, и реальных энцефаллограмм.

Нельзя не упомянуть о совершенно ином подходе, использованном ав-

торами настоящей статьи совместно с Ю.А. Купериным и И.В. Пранявичуте (Санкт-Петербургский государственный университет), при работе с модифицированной моделью Кропотова-Пахомова. Речь идет о построении музыкальных текстов на основе временных рядов, генерируемых нейронной сетью. В процессе создания музыкального образа использовалось несколько глобальных параметров, таких как темп, положение октав и минимальный квант времени; однако очередность следования отдельных нот и их относительные высоты определялись исключительно внутренней динамикой сети. Интересно, что испытуемые, в числе которых были и профессиональные музыканты, отмечали в предложенных для прослушивания примерах наличие четко выраженной музыкальной составляющей. Материал был представлен на Второй международной конференции по когнитивной науке (2006, Санкт-Петербург) и IX Всероссийской научно-технической конференции «Нейроинформатика-2007» (Москва). Примеры музыкальных образов можно загрузить из интернета, воспользовавшись ссылкой [hep.niif.spbu.ru/files/musik/patterns.rar](http://hep.niif.spbu.ru/files/musik/patterns.rar). Тот факт, что созданные столь витиеватым образом мелодии субъективно воспринимаются человеком как нечто интересное, несомненно заслуживает внимания.

Очевидно, что модифицированная модель Кропотова-Пахомова является очень богатым объектом с точки зрения разнообразия динамического поведения, и изучать ее достаточно интересно и полезно, даже учитывая пока еще призрачный характер ее связи с реальными физическими процессами. Но в естественно-научной области деятельности ценность модели определяется степенью ее скоррелированности с явлением, для описания которого она строилась. Поэтому, безусловно, дальнейшая работа с представленной в настоящей статье моделью нейронной сети, большей частью должна быть направлена именно на поиск точек соприкосновения с природными процессами.

## Приложение 1

По мере того как складывается центральная нервная система и обособляются ее высшие центры, возрастает роль специально-психического подбора; под его контролем, сначала из рефлексов, а затем рядом с ними вырабатываются привычные реакции. Однако в момент своего возникновения «привычная» реакция не заслуживает этого имени. Она представляет тогда еще «непривычное» приспособление, приспособление к условиям, еще только впервые переживаемым, не пережитым раньше. Следовательно, она возникает, как пластичная реакция, и потом уже, в ряду поколений, вызываемых повторением тех же условий, приобретает характер привычной. Когда человек еще только начинает плавать, его движения имеют характер пластичных; после ряда повторений они являются привычными.

При этом возрастает, с одной стороны, определенность, отчетливость и целесообразность реакции, с другой стороны — ее консерватизм. Первое происходит в силу того, что в ряду повторений приспособление ста-

новится совершеннее: продолжительное действие подбора со стороны повторяющихся условий устраняет наименее подходящие к этим условиям элементы реакции, тогда как наиболее подходящие сохраняются. Второе изменение, стоящее в неразрывной связи с первым, зависит, главным образом, от накапливающегося действия психического подбора. Каждый раз, как психический подбор оказывается благоприятным для реакции, т.е. каждый раз, как сама она оказывается непосредственно целесообразной, ее элементы и их взаимная связь становятся прочнее, возможность повторения реакции возрастает. Следовательно, можно сказать следующее

*Возможно, что сопротивление нервных проводников уменьшается по мере их функционирования в силу чисто физических условий самого проведения. Это, может быть, послужит исходной точкой самого простого и непосредственного объяснения процессов «привыкания»: при таких условиях каждое повторение реакции должно облегчать следующее ее повторение.*

Так и развивается привычная реакция: она не могла бы часто повторяться и стать «привычной», если бы не оказывалась в своих повторениях непосредственно целесообразной.

Повторяясь чаще и чаще, привычная реакция становится все в большей степени устойчивой и автоматичной, приближаясь во всех отношениях к рефлексам (даже, как увидим, по степени сознательности). Таковы, например, движения при ходьбе или профессиональные движения работника, выполняющего какую-нибудь простую техническую операцию. Очень вероятно, что многие инстинкты, например, строительные инстинкты птиц, возникли из привычных действий, повторявшихся при сходных условиях в ряду поколений, и приобретающих все более консервативный и стихийный характер.

## Приложение 2

Назовем *образом* группу активных в фиксированный момент времени нейронов, то есть таких, для которых  $N_i(k) = 1$ , а *последовательностью образов* — цепочку друг за другом поочередно активизирующихся групп нейронов. Запоминание сетью цепочек образов проявляется в виде формирования специальных структур связей в результате соответствующего внешнего воздействия, а воспроизведение связано с распространением сигналов вдоль этих структур.

Рассмотрим подробнее, как образуются связи и распространяется сигнал. Пусть имеется два изолированных нейрона с номерами  $m_1$  и  $m_2$ , то есть все соответствующие матричные элементы матрицы связей равны нулю:

$$W_{im_1}(k) = W_{m_1i}(k) = W_{m_2i}(k) = W_{im_2}(k) = 0, \quad i = 1, \dots, n.$$

Если в момент времени  $k$  подействовать внешним импульсом на нейрон с номером  $m_1$ , то (при достаточной величине импульса) в последующий момент  $k + 1$  за счет появления в правой части (13) дополнительного слагаемого  $S_{m_1}(k)$  потенциал данного нейрона превысит порог активности  $h_{m_1}$ , и нейрон станет активен ( $N_{m_1} = 1$ ). Заметим, что посредством подбора параметров  $\alpha$  и  $\beta$  можно добиться того, что еще через один шаг по времени потенциал  $P_{m_1}$  снова окажется по величине меньше  $h_{m_1}$ , и нейрон станет неактивен. Далее, подадим такой же по величине импульс на второй нейрон, но в момент  $k + 1$ . Активность этого нейрона наступит при  $k + 2$  и в (15) появится положительное слагаемое  $\nu N_{m_2}(k + 2)N_{m_1}(k + 1)$ . В результате в матрице связей возникнет ненулевой матричный элемент  $W_{m_2 m_1}$ , при этом симметричный ему останется нулевым. Если теперь через некоторое время нейрон  $m_1$  снова станет активен, например, опять после внешнего воздействия, то в (13) под знаком суммы появится слагаемое  $W_{m_2 m_1} N_{m_1}$ , что может привести к последующей активизации нейрона  $m_2$ . Активность же нейрона  $m_2$  не окажет никакого влияния на нейрон  $m_1$ , поскольку матричный элемент  $W_{m_1 m_2}$  равен нулю. Поэтому здесь можно говорить о появлении связи, направленной от нейрона  $m_1$  к нейрону  $m_2$ , и о распространении сигнала вдоль этой связи.

Аналогичная картина будет и при последовательной подаче импульсов на группы нейронов (рис. 27), причем связи между нейронами внутри групп возникать не будут, так как нейроны остаются активны только в течение одного шага по времени.

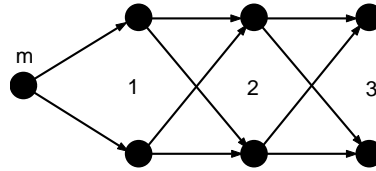


Рис. 27. Структура связей нейронной сети, соответствующая цепочке из 4-х образов ( $m, 1, 2$  и  $3$ ).

Повторное однократное воздействие, например, на нейроны из группы 1 вызовет последовательную однократную активизацию групп 2 и 3, при этом новые связи возникать не будут, а матричные элементы существующих связей увеличатся по величине, то есть связи укрепятся. Следовательно, даже при наличии диссипации связей возможно многократное воспроизведение одной и той же последовательности образов.

Очевидно, что для корректного воспроизведения разные образы последовательности не должны пересекаться, то есть один нейрон может быть задействован только в одном образе. При использовании структур специального типа (рис. 28), где есть выделенная цепочка управляющих нейронов, можно записывать и последовательности с пересекающимися образами. Од-



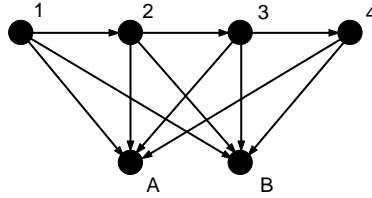


Рис. 28. Пример структуры, допускающей пересечение образов.  $1 \rightarrow 2 \rightarrow 3 \rightarrow 4$  — цепочка управляющих нейронов.  $A, B$  — нейроны образов. Воздействие на нейрон с номером 1 вызовет последовательную активизацию групп нейронов:  $(1), (2, A, B), (3, A, B), (4, A, B), (A, B)$ .

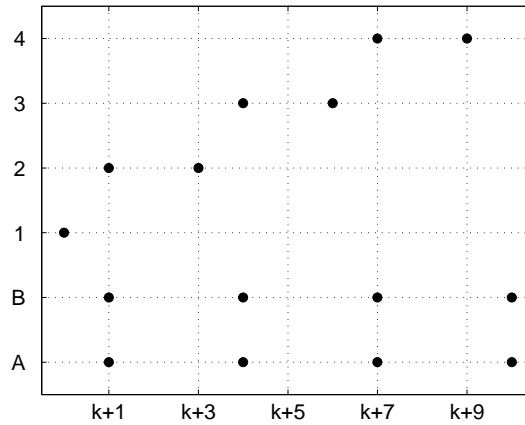


Рис. 29. Внешние импульсы, формирующие структуру на рис. 28. По горизонтальной оси отложено время, по вертикальной — номера стимулируемых нейронов.

нако для формирования подобных структур необходимо либо особым образом воздействовать на сеть (см. рис. 29), либо видоизменять модель. Во втором случае отбирать управляющие нейроны можно, например, по принципу, используемому в сетях Кохонена [27].

На рис. 28 приведен один из простейших примеров, где последовательность образов состоит из двух постоянно активных нейронов  $A$  и  $B$ . Но в общем случае здесь может быть гораздо большее количество нейронов, образующих своеобразный экран, на который управляющие нейроны будут «проецировать» свои образы.

Конечно описанный выше механизм записи последовательностей образов следует считать скорее дополнительной иллюстрацией, позволяющей лучше понять принципы работы модели, нежели руководством к практическому применению, поскольку здесь практически теряется одно из осново-

ных преимуществ нейросетей — работа с большим количеством информации при минимуме задействованных объектов.

### Список литературы

- [1] А. А. Богданов. *Познание с исторической точки зрения*. СПб. 1901.
- [2] В. Miller, J. Reinhardt. *Neural networks. An introduction*. Springer-Verlag. 1991.
- [3] М. Б. Беркинблит. *Нейронные сети*. МИРОС. 1993.
- [4] И. С. Суровцев, В. И. Ключкин, Р. П. Пивоварова. *Нейронные сети*. Воронеж, ВГУ. 1994.
- [5] А. К. Jain, J. Mao, K. M. Mohiuddin. Artificial neural networks: a tutorial. *Computer*. 1996. V. 29. P. 31-44.
- [6] H.-J. Chang, W. J. Freeman, B. C. Burke. Biologically modeled noise stabilizing neurodynamics for pattern recognition. *Int. J. Bifurcation and Chaos*. 1998. V. 8. No. 2. P. 321-345.
- [7] K. Tateno, H. Hayashi, S. Ishizuka. Complexity of spatiotemporal activity of a neural network model which depends on the degree of synchronization. *Neur. Netw.* 1998. V. 11. P. 5-23.
- [8] M. Yoshida, H. Hayashi, K. Tateno, S. Ishizuka. Stochastic resonance in the hippocampal CA3-CA1 model: a possible memory recall model. *Neur. Netw.* 2002. V. 15. P. 1171-1183.
- [9] Y. Chen, Y. H. Wang, K. Q. Yang. The macroscopic dynamics in separable neural networks. *Phys. Rev. E*. 2001. V. 63, 041901.
- [10] R. Albert, A-L. Barabási. Statistical mechanics of complex networks. *Rew. Mod. Phys.* 2002. V. 74. P. 47-97.
- [11] S. Bornholdt, T. Röhl. Self-organized critical neural networks. *Phys. Rev. E*. 2003. V. 67, 066118.
- [12] W. K. Theumann. Mean-field dynamics of sequence processing neural networks with finite connectivity. *Physica A*. 2003. V. 328. P. 1-12.
- [13] M. Kawato. Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*. 1999. V. 9. P. 718-727
- [14] M. Heerema, W. A. Leeuwen. A recurrent neural network with ever changing synapses. *J. Phys. A: Math. Gen.* 2000. V. 33. P. 1781-1795.
- [15] S. Miyoshi, H.-F. Yanaib, M. Okada. Associative memory by recurrent neural networks with delay elements. *Neur. Netw.* 2004. V. 17. P. 55-63.

- [16] Ю. Д. Кропотов, С. В. Пахомов. Математическое моделирование механизмов обработки сигналов нейронными популяциями в головном мозге. Сообщение I. Постановка задачи и основные свойства модели. *Физиология человека*. 1981. Т. 7. № 1. С. 152-162.
- [17] Ю. Д. Кропотов, С. В. Пахомов. Математическое моделирование механизмов обработки сигналов нейронными популяциями в головном мозге. Сообщение II. Влияние синаптической пластичности на свойства нейронной сети с локальными связями в стационарном режиме. *Физиология человека*. 1983. Т. 9. № 5. С. 787-795.
- [18] Ю. Д. Кропотов, С. В. Пахомов. Математическое моделирование механизмов обработки сигналов нейронными популяциями в головном мозге. Сообщение III. Изучение вызванных реакций нейронных ансамблей. *Физиология человека*. 1984. Т. 10. № 5. С. 813-821.
- [19] Дж. Г. Николас, А. Р. Мартин, Б. Дж. Валлас, П. А. Фукс. *От нейрона к мозгу*. УРСС. 2003.
- [20] К. А. Mardanov, Y. M. Pismak, Y. M. Potyagailo. General properties of realistic neural network dynamics. *Comp. Math. Appl.* 1997, V. 34, N. 7/8, P. 675-685.
- [21] Г. А. Черных, Ю. М. Письмак. Модифицированный принцип Хэбба в модели нейронной сети Кропотова. *Вестник С.-Петербург. ун-та*. 2001. Сер. 4. Вып. 3. №20. С. 104-107.
- [22] К. Kaneko, I. Tsuda. *Complex systems: chaos and beyond. A constructive approach with applications in life sciences*. Springer-Verlag. 2001.
- [23] D. O. Hebb. *The organization of behavior. A neuropsychological theory*. N.Y.: Wiley and Sons. 1949. P. 355.
- [24] М. А. Красноперова. *Основы реконструктивного моделирования стихосложения*. С.-Петербург. ун. 2000.
- [25] М. А. Красноперова, Ю. М. Письмак, Г. А. Черных. О возможности применения моделей нейронных сетей в исследовании ритмики текста. *Формальные методы в лингвистической поэтике*. Сборник научных трудов, посвященный 60-летию профессора Санкт-Петербургского государственного университета М. А. Красноперовой. С.-Петербург. ун. 2001.
- [26] Р. Пенроуз. *Тени разума. В поисках науки о сознании*. Ч. 1,2. Москва-Ижевск: Институт компьютерных исследований. 2005.
- [27] Т. Kohonen. Self-organized formation of topologically correct feature maps. *Bio. Cybern.* 1982. V. 43. N. 1. P. 56-69.

Статья поступила 16 декабря 2005 г.

После доработки 19 февраля 2007 г.

# Кластерный принцип формирования доменов в задачах бинарной минимизации

Литинский Л.Б.

НИИ системных исследований РАН, Москва

[litin@mail.ru](mailto:litin@mail.ru)

**Аннотация.** При минимизации квадратичного функционала, зависящего от большого числа бинарных переменных, имеет смысл агрегировать переменные, объединяя независимые переменные в более крупные блоки – *домены*. Этот прием позволяет находить более глубокие локальные минимумы. В настоящей работе обосновывается метод формирования доменов, основанный на кластеризации матрицы связей. Результаты компьютерных экспериментов свидетельствуют о том, что кластерные домены позволяют получить более глубокие локальные минимумы, чем домены, сформированные случайным образом.

## 1. Введение

Обсуждается задача минимизации квадратичного функционала от  $N$  бинарных переменных:

$$E(\mathbf{s}) = -(\mathbf{J}\mathbf{s}, \mathbf{s}) = - \sum_{i,j=1}^N J_{ij} s_i s_j \longrightarrow \min_{\mathbf{s}} \text{ где } s_i = \{\pm 1\}. \quad (1)$$

Без ограничения общности матрицу связей  $\mathbf{J} = (J_{ij})_1^N$  можно считать симметричной, с нулевыми элементами на главной диагонали:  $J_{ij} = J_{ji}$ ,  $J_{ii} = 0$ . Используя физическую терминологию [1] бинарные переменные  $s_i = \{\pm 1\}$  будем называть *спинами*,  $N$ -мерные векторы  $\mathbf{s} = (s_1, s_2, \dots, s_N)$  – *конфигурационными векторами* или просто *конфигурациями*, а минимизируемую характеристику  $E(\mathbf{s})$  – *энергией* состояния  $\mathbf{s}$ .

Общепотребительная процедура локальной минимизации функционала (1) состоит в следующем: стартовав с какой-то начальной конфигурации, случайным образом перебирают спины, и каждому спину присваивают значение, равное знаку действующего на него локального поля

$$h_i(t) = \sum_{j=1}^N J_{ij} s_j(t). \quad (2)$$

Иными словами, если текущее состояние спина  $s_i(t)$  совпадает со знаком  $h_i(t)$  – если спин *удовлетворен*:  $s_i(t)h_i(t) \geq 0$ , – то его значение остается неизменным:  $s_i(t+1) = s_i(t)$ ; если же  $s_i(t)$  и  $h_i(t)$  противоположных знаков – если спин

неудовлетворен:  $s_i(t)h_i(t) < 0$ , – то в следующий момент времени спин переворачивают:  $s_i(t+1) = -s_i(t)$ :

$$s_i(t+1) = \begin{cases} s_i(t), & \text{если } s_i(t)h_i(t) \geq 0; \\ -s_i(t), & \text{если } s_i(t)h_i(t) < 0. \end{cases} \quad (3)$$

Легко видеть, что переворот неудовлетворенного спина понижает энергию состояния:  $E(\mathbf{s}(t+1)) = E(\mathbf{s}(t)) - 4 |h_i(t)|$ . Если позволить системе эволюционировать достаточно долго, рано или поздно она окажется в состоянии, являющемся локальным минимумом по энергии. В этом состоянии все спины будут удовлетворены и эволюция системы прекратится. Данный динамический сценарий будем называть *случайной динамикой*; в теории нейронных сетей его называют асинхронной динамикой [2].

Другой широко известный тип динамики – *синхронная динамика* [2], когда одновременно переворачивают *все неудовлетворенные* спины. Легко показать, что одновременный переворот сразу нескольких неудовлетворенных спинов не гарантирует, что после переворота они станут удовлетворенными. Синхронную динамику редко используют для минимизации функционала (1): для нее невозможно гарантировать монотонность убывания функционала  $E(\mathbf{s})$ . Кроме того, для синхронной динамики характерно наличие предельных циклов длины 2, когда система “заикливается” между двумя состояниями, последовательно перескакивая из одного в другое и обратно. Предельные циклы не имеют никакого отношения к локальным минимумам и только уменьшают их области притяжения.

В работах [3],[4] предложено обобщение случайной динамики – *доменная динамика*. Суть обобщения состоит в том, что спины случайным образом объединяют в более крупные блоки, или (как называли их авторы) – в *домены*. Затем, при определенных условиях, весь домен переворачивают целиком: все спины, принадлежащие к одному домену, одновременно меняют свои состояния на противоположные. Доменная динамика напоминает синхронную динамику, однако отличается от нее условиями, при которых домен переворачивают. Эти условия естественным образом обобщают случайную динамику на случай агрегированных переменных. Доменная динамика гарантирует монотонное понижение энергии  $E(\mathbf{s})$  и не приводит к предельным циклам. Кроме того, доменная динамика в  $k^2$  раз быстрее случайной динамики, где  $k$  – средняя длина домена (число объединенных в один блок спинов).

В вычислительных экспериментах со случайными Хеббовскими матрицами<sup>1</sup> было установлено [5], что доменная динамика позволяет отыскать в среднем более глубокие локальные минимумы, чем случайная динамика. Однако остается открытым вопрос – каким образом следует формировать домены? Должны ли домены быть случайными, или имеет смысл формировать их как-то

<sup>1</sup> Хеббовскими называются корреляционные матрицы, элементы которых суть скалярные произведения векторов с бинарными координатами  $\pm 1$ ; эти матрицы играют большую роль в теории нейронных сетей [2].

иначе? Зависит ли результат минимизации функционала (1) от способа формирования доменов? Настоящая работа посвящена обсуждению именно этих вопросов.

В следующем разделе дается формальное определение доменной динамики. В третьем разделе обосновывается метод формирования доменов, основанный на идее кластеризации матрицы связей  $\mathbf{J}$ . Четвертый раздел посвящен изложению результатов компьютерного моделирования. Несколько общих замечаний вынесено в заключительный раздел, а технические доказательства – в Приложение.

## 2. Доменная динамика

Для простоты изложения все определения будут даны для случая, когда в первый домен объединены  $k$  первых спинов, во второй домен –  $k$  следующих спинов, и так далее, в последний,  $n$ -й домен, объединены  $k$  последних спинов:

$$\mathbf{s} = (\underbrace{s_1, \dots, s_k}_{1\text{-й домен}}, \underbrace{s_{k+1}, \dots, s_{2k}}_{2\text{-й домен}}, \dots, \underbrace{s_{(n-1)k+1}, \dots, s_N}_{n\text{-й домен}}).$$

Здесь  $N = kn$ ,  $\mathbf{s}$  – произвольная конфигурация.

Суммарное воздействие на первый домен со стороны всех остальных доменов есть суперпозиция воздействий, которые спины этих доменов оказывают на спины первого домена. Тогда *доменное локальное поле*, действующее на  $i$ -й спин первого домена, равно

$$h_i^{(d)}(t) = \sum_{j=k+1}^N J_{ij} s_j(t) = h_i(t) - \sum_{j=1}^k J_{ij} s_j(t), \quad \forall i \leq k, \quad (4)$$

где  $h_i(t)$  есть определенное выше локальное поле (2). Иными словами, доменное локальное поле  $h_i^{(d)}(t)$  получается из локального поля  $h_i(t)$  устранением влияния тех спинов, которые принадлежат вместе с  $i$ -м спином к одному и тому же домену. Тогда энергия взаимодействия первого домена с остальными доменами равна

$$E_1(t) = -F_1(t) = -\sum_{i=1}^k s_i(t) h_i^{(d)}(t).$$

Аналогично определяются доменное локальное поле, действующее на спины  $l$ -го домена,

$$h_i^{(d)}(t) = h_i(t) - \sum_{j=(l-1)k+1}^{lk} J_{ij} s_j(t), \quad \forall (l-1)k+1 \leq i \leq lk, \quad l \in [2, n],$$

и энергия взаимодействия  $l$ -го домена с остальными доменами:

$$E_l(t) = -F_l(t) = - \sum_{i=(l-1)k+1}^{lk} s_i(t)h_i^{(d)}(t), \quad l = 2, \dots, n.$$

Доменная энергия состояния  $\mathbf{s}(t)$  есть

$$E^{(d)}(t) = \sum_{l=1}^n E_l(t) = - \sum_{l=1}^n F_l(t). \quad (5)$$

Минимизируемая характеристика  $E(t)$  (1) отличается от доменной энергии (5) на сумму величин  $E_l^{(in)}$ , которые характеризуют внутри-доменные взаимодействия между спинами:

$$E(t) = E^{(d)}(t) + \sum_{l=1}^n E_l^{(in)}(t), \quad \text{где } E_l^{(in)}(t) = - \sum_{i=(l-1)k+1}^{lk} \sum_{j=(l-1)k+1}^{lk} J_{ij} s_i(t) s_j(t). \quad (6)$$

Доменная динамика определяется следующим образом [3],[4]: в случайном порядке перебирают  $n$  доменов, и если для  $l$ -го домена выполняется неравенство

$$F_l(t) = \sum_{i=(l-1)k+1}^{lk} s_i(t)h_i^{(d)}(t) \geq 0,$$

то домен оставляют неизменным; если же  $F_l(t) < 0$ , то в следующий момент времени  $l$ -й домен переворачивают – всем входящим в него спином одновременно присваивают противоположные значения:

$$s_i(t+1) = \begin{cases} s_i(t), & \text{если } F_l(t) \geq 0, \\ -s_i(t), & \text{если } F_l(t) < 0, \end{cases} \quad \forall i \in [1+(l-1)k, lk].$$

Сравнивая последнее выражение с аналогичным выражением (3) для случайной динамики, убеждаемся в том, что доменная динамика превращается в стандартную случайную динамику, когда длина доменов равна 1:  $k=1$ .

Из выражения (5) следует, что доменная динамика сопровождается монотонным понижением доменной энергии состояния: если в момент времени  $t$  переворачивают  $l$ -й домен, то доменная энергия понижается на  $4|F_l(t)|$ :

$$E^{(d)}(t+1) = E^{(d)}(t) - 4|F_l(t)|.$$

Заметим – и это очень важно, – что одновременно ровно на такую же величину понижается энергия  $E(\mathbf{s}(t))$ . Это следует из выражения (6) и того очевидного факта, что одновременный переворот всех спинов  $l$ -го домена не меняет значения  $l$ -й внутри-доменной энергии  $E_l^{(in)}$ .

Поскольку доменная энергия состояния не может понижаться бесконечно, рано или поздно система окажется в *доменном* локальном минимуме. В этом состоянии для всех доменов будет выполняться  $F_i(t) \geq 0$ , и эволюция доменной динамики прекратится. Однако доменный локальный минимум совсем не обязательно будет минимумом функционала (1). Поэтому домены здесь «размораживаются» и переходят к стандартной случайной динамике. Система получает возможность опуститься по энергии еще глубже, в минимум функционала (1).

Основные надежды, которые возлагают на последовательное использование сначала доменной динамики, а затем – случайной, связаны со следующими соображениями. Энергетическую поверхность функционала (1) можно представить как своего рода овражистый ландшафт в многомерном пространстве, склоны которого изрыты огромным числом мелких локальных минимумов. Минимизация (1) осуществляется в результате скатывания по этому склону динамической системы, запущенной из некоторой начальной точки. При использовании случайной динамики два последовательных состояния системы отличаются друг от друга противоположным значением всего одной бинарной координаты – той, которую переворачивают на данном такте эволюции. Понятно, что система при этом имеет много шансов застрять в первом попавшемся локальном минимуме, даже и неглубоком. Напротив, при использовании доменной динамики два последовательных состояния системы отличаются друг от друга противоположными значениями сразу нескольких спиновых координат. Доменную динамику следует уподобить скатыванию вниз по склону более «крупным шагом». Можно ожидать, что такой более крупный шаг позволит системе миновать множество мелких локальных минимумов, в которых она застряла бы при случайной динамике.

Для простоты все определения давались выше в предположении, что домены имеют одинаковую длину  $k$ , а объединяемые в домен спины имеют последовательные номера. Понятно, что в общем случае домены могут иметь разные длины  $k_i$ , и объединять в домен можно любые спины. В связи с этим возникают вопросы: влияет ли тот или иной состав доменов на результаты минимизации? Каким образом следует формировать домены? Должны ли они быть постоянными для данной матрицы  $\mathbf{J}$ , или их следует формировать случайно на каждом такте эволюции? Можно ли здесь опереться на какие-либо общие соображения? В следующих разделах мы постараемся ответить на эти вопросы.

**Замечание.** По-видимому, идея доменной динамики не раз выдвигалась различными авторами. При подготовке настоящей публикации мы установили, что доменная динамика эквивалентна введенной в [6] *block-sequential dynamics*. Отметим, что *block-sequential dynamics* изучалась в [6] применительно к задаче увеличения емкости памяти модели Хопфилда. Насколько нам известно, использование доменной динамики для минимизации функционала (1) впервые изучалось именно в работах [3]-[5].



### 3. Кластерный принцип формирования доменов

**1. Хеббовская матрица специального вида.** Для модели Хопфилда известна ситуация, когда домены естественно порождаются особенностями устройства матрицы связей  $\mathbf{J}$  [7]. Приведем соответствующие результаты, расставляя нужные нам акценты. В конце раздела будет сформулирован общий принцип формирования доменов.

Пусть имеется  $N$   $M$ -мерных векторов-столбцов  $\mathbf{x}_i$  с бинарными координатами:  $x_i^{(\mu)} = \pm 1$ ,  $i = 1, \dots, N$ ,  $\mu = 1, \dots, M$ . Векторы  $\mathbf{x}_i$  нумеруются нижним индексом  $i \in [1, N]$ , а их координаты – верхним индексом  $\mu \in [1, M]$ ; соотношение между размерностью векторов ( $M$ ) и их числом ( $N$ ) для наших целей значения не имеет.

Образуем  $(M \times N)$ -матрицу  $\mathbf{X}$ , столбцами которой являются  $M$ -мерные векторы  $\mathbf{x}_i$ :

$$\mathbf{X} = \begin{pmatrix} x_1^{(1)} & x_2^{(1)} & \dots & x_N^{(1)} \\ x_1^{(2)} & x_2^{(2)} & \dots & x_N^{(2)} \\ \dots & \dots & \dots & \dots \\ x_1^{(M)} & x_2^{(M)} & \dots & x_N^{(M)} \end{pmatrix}.$$

В теории нейронных сетей  $\mathbf{X}$  называют *матрицей паттернов* [2]. С ее помощью строится Хеббовская матрица связей –  $(N \times N)$ -матрица скалярных произведений векторов-столбцов  $\mathbf{x}_i$ :

$$J_{ij} = \frac{(1 - \delta_{ij})}{M} (\mathbf{x}_i, \mathbf{x}_j) = \frac{(1 - \delta_{ij})}{M} \sum_{\mu=1}^M x_i^{(\mu)} x_j^{(\mu)}, \quad 1 \leq i, j \leq N.$$

Поскольку длины векторов-столбцов  $\mathbf{x}_i$  равны  $\sqrt{M}$ , матричные элементы  $J_{ij}$  суть косинусы углов между соответствующими векторами:  $|J_{ij}| \leq 1$ ; на главной диагонали матрицы стоят нули.

Мы рассмотрим тот специальный случай, когда число  $n$  различных векторов-столбцов у матрицы  $\mathbf{X}$  меньше, чем  $N$ :  $\mathbf{x}_1 \dots \neq \dots \mathbf{x}_l \dots \neq \dots \mathbf{x}_n$ ,  $n < N$ . Иными словами, каждый вектор-столбец  $\mathbf{x}_i$  матрицы  $\mathbf{X}$  может по несколько раз повторяться. Если  $k_l$  – число повторений вектора-столбца  $\mathbf{x}_l$ , то  $\sum_1^n k_l = N$ . Без ограничения общности можно считать, что первые  $k_1$  векторов-столбцов матрицы  $\mathbf{X}$  совпадают между собой, следующие  $k_2$  векторов-столбцов совпадают между собой, и так далее – последние  $k_n$  векторов-столбцов тоже совпадают:

$$\mathbf{X} = \left( \begin{array}{ccc|ccc|ccc} x_1^{(1)} & \dots & x_1^{(1)} & x_2^{(1)} & \dots & x_2^{(1)} & \dots & x_n^{(1)} & \dots & x_n^{(1)} \\ x_1^{(2)} & \dots & x_1^{(2)} & x_2^{(2)} & \dots & x_2^{(2)} & \dots & x_n^{(2)} & \dots & x_n^{(2)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ x_1^{(M)} & \dots & x_1^{(M)} & x_2^{(M)} & \dots & x_2^{(M)} & \dots & x_n^{(M)} & \dots & x_n^{(M)} \end{array} \right), \quad k_l \geq 1. \quad (7)$$

Ясно, что соответствующая Хеббовская матрица будет состоять из  $n \times n$  блоков. Размерность  $(lm)$ -го блока равна  $(k_l \times k_m)$ , а все его элементы равны одному и тому же числу  $J_{lm}$  – косинусу угла между векторами  $\mathbf{x}_l$  и  $\mathbf{x}_m$  ( $l, m = 1, \dots, n$ ). На главной диагонали матрицы  $\mathbf{J}$  стоят квадратные  $(k_l \times k_l)$ -блоки, состоящие из сплошных 1 (за исключением диагональных матричных элементов, которые по определению равны 0):

$$\mathbf{J} = \left( \begin{array}{ccc|ccc|ccc} \overbrace{0 \ 1 \ 1}^{k_1} & & & \overbrace{J_{12} \ \dots \ J_{12}}^{k_2} & & & \overbrace{J_{1n} \ \dots \ J_{1n}}^{k_n} & & & \\ 1 & \ddots & 1 & \vdots & J_{12} & \vdots & \dots & \vdots & J_{1n} & \vdots \\ 1 & 1 & 0 & J_{12} & \dots & J_{12} & & J_{1n} & \dots & J_{1n} \\ J_{21} & \dots & J_{21} & 0 & 1 & 1 & & J_{2n} & \dots & J_{2n} \\ \vdots & J_{21} & \vdots & 1 & \ddots & 1 & \dots & \vdots & J_{2n} & \vdots \\ J_{21} & \dots & J_{21} & 1 & 1 & 0 & & J_{2n} & \dots & J_{2n} \\ \vdots & & \vdots & & & \ddots & & \vdots & & \vdots \\ J_{n1} & \dots & J_{n1} & J_{n2} & \dots & J_{n2} & & 0 & 1 & 1 \\ \vdots & J_{n1} & \vdots & \vdots & J_{n2} & \vdots & \dots & 1 & \ddots & 1 \\ J_{n1} & \dots & J_{n1} & J_{n2} & \dots & J_{n2} & & 1 & 1 & 0 \end{array} \right), \quad |J_{lm}| < 1, \quad l, m \in [1, n]. \quad (8)$$

Оказывается [7], что подобное устройство матрицы связей накладывает достаточно жесткие ограничения на вид конфигураций, доставляющих локальные минимумы функционалу (1). А именно, локальный минимум непременно должен иметь «кусочно-постоянный» вид:

$$\mathbf{s} = (\underbrace{s_1 \ \dots \ s_1}_{k_1}, \underbrace{s_2 \ \dots \ s_2}_{k_2}, \dots, \underbrace{s_n \ \dots \ s_n}_{k_n}), \quad s_l = \pm 1, \quad l = 1, \dots, n. \quad (9)$$

Иными словами, первые  $k_1$  координат локального минимума должны быть одинаковыми, следующие  $k_2$  координат должны быть равны друг другу, и так далее – последние  $k_n$  координат должны быть равны друг другу. Блоки «знакопостоянства» в выражении (9) определяются блочной структурой Хеббовской матрицы (8); или (что есть то же самое) – структурой одинаковых столбцов матрицы паттернов  $\mathbf{X}$  (7). Доказательство того, что локальные минимумы (1) непременно имеют вид (9), основано на том, что спины из одного блока имеют одинаковые

связи с остальными спинами, и под действием локального поля ведут себя одинаково (доказательство приведено в п.1 Приложения).

Не все конфигурации вида (9) будут локальными минимумами функционала (1), но искать минимумы надо среди таких конфигураций. Иначе говоря, с точки зрения минимизации  $E(\mathbf{s})$  бессмысленно рассматривать конфигурации с координатами разных знаков *внутри* блоков знакопостоянства. Имеет смысл рассматривать только конфигурации вида (9). Все спины из одного блока знакопостоянства либо одновременно удовлетворены, либо нет. В последнем случае не имеет никакого смысла переворачивать неудовлетворенные спины по отдельности – ведь при этом возникнут конфигурации с координатами разных знаков внутри блока знакопостоянства, а рассматривать такие конфигурации просто бессмысленно. И оказывается, что при матрице связей вида (8) можно одновременно переворачивать целый блок неудовлетворенных спинов – энергия (1) будет при этом понижаться. Иными словами, блоки знакопостоянства в (9) выступают здесь как *естественные* домены.

**2. Кластерный принцип формирования доменов.** Вообще говоря, и для блочно-постоянной матрицы связей (8) можно сформировать домены случайным образом, а затем использовать их для доменной минимизации функционала (1). Экспериментальная часть настоящей работы как раз и посвящена исследованию того, какой из способов формирования доменов лучше. Но прежде чем переходить к описанию экспериментов сформулируем в двух пунктах общий принцип формирования «правильных доменов»:

**а) в один домен объединяют спины, связанные друг с другом сильнее, чем с другими спинами;**

**б) значения спинов, объединенных в один домен, устанавливаются одинаковыми.**

Этот рецепт естественно назвать *кластерным принципом* формирования доменов из-за его очевидной связи с задачей кластеризации симметричной матрицы [8]. В этой задаче требуется привести матрицу к блочно-диагональному виду таким образом, чтобы элементы внутри диагональных блоков превосходили матричные элементы вне диагональных блоков. Более подробно мы обсудим эту задачу в конце статьи.

Мотивировка кластерного принципа вполне очевидна. Матричный элемент  $J_{ij}$  является мерой связи между  $i$ -м и  $j$ -м спинами. Кластерный принцип формирования доменов исходит из того, что два сильно связанных спина будут, скорее всего, сходным образом взаимодействовать и с остальными спинами. Тогда локальные поля, действующие на  $i$ -й и  $j$ -й спины, будут иметь близкие значения. Естественно предположить, что под действием одинаковых локальных полей спины будут вести себя одинаково. Вопрос о том, какие спины можно считать «сильно связанными», обсуждается в следующем разделе.

В заключение заметим, что Хеббовская матрица (8) отвечает предельно идеализированной ситуации – ее диагональные блоки состоят из 1. Формируя «правильные домены» согласно (9), мы будем объединять в один домен спины, связанные друг с другом предельно сильно. Такие матрицы редко встречается на практике. В компьютерных экспериментах мы исследовали не только идеализи-

рованный случай предельно сильных внутри-доменных связей, но и более реалистичные матрицы.

#### 4. Компьютерное моделирование

**1. Сильно связанные группы спинов.** Первая серия экспериментов проводилась для Хеббовских матриц (8) с предельно сильными внутри-групповыми связями – см. п.1 предыдущего раздела. Внешние параметры задачи были равны следующим значениям: число векторов-столбцов матрицы паттернов  $N=1000$ , их размерность  $M=60$ , число доменов  $n=40$ , размеры доменов  $k_i$  – случайные числа из интервала  $[1, 45]$ :  $\sum_{i=1}^{40} k_i = 1000$ . Координаты  $x_i^{(\mu)}$  в матрице паттернов (7) принимали значения  $\pm 1$  с равными вероятностями. Как результат, на главной диагонали Хеббовской матрицы располагались 40  $(k_i \times k_i)$ -блоков, состоящих из одних единиц:  $J_{ii}^{(in)} = 1$ . Элементы вне-диагональных блоков суть случайные величины со средним значением 0 и дисперсией  $1/M$ :  $\langle J_{lm}^{(out)} \rangle = 0$ ,  $\sigma(J_{lm}^{(out)}) = 1/\sqrt{M}$ . Кластерный принцип дает здесь 40 доменов кусочно-постоянного вида (9), образованных сорока группами сильно связанных спинов.

Всего генерировалось 200 подобных матриц, и для каждой матрицы функционал (1) минимизировали с использованием 3 динамических сценариев:

**а) RANDOM:** случайную динамику запускали с 1000 случайных стартовых конфигураций;

**б) DM-RND** (случайные домены): с тех же самых случайных конфигураций запускали доменную динамику, сформировав  $n=40$  случайных доменов одинаковой длины  $k=25$ ; в домены объединяли по 25 спинов со случайными номерами, а спинам внутри домена случайным образом присваивали значения  $\pm 1$ ;

**в) DM-CLS** (кластерные домены): с 1000 случайных стартовых конфигураций блочно-постоянного вида (9) запускали доменную динамику.

По достижении доменного минимума домены «размораживали», и система (уже в случайной динамике) опускалась в более глубокий минимум функционала (1)<sup>2</sup>. Таким образом, для каждой матрицы получали 3000 локальных минимумов функционала (1). Затем определяли самый глубокий из них, и вычисляли частоту его отыскания в каждом из трех динамических сценариев: RANDOM, DM-RND и DM-CLS. Тот тип динамики, для которого частота отыскания самого глубокого минимума будет наибольшей, следует, очевидно, считать наилучшим.

На рис.1 для всех трех динамических сценариев приведены результаты, усредненные по 200 случайным испытаниям. По оси абсцисс даны 3 типа дина-

<sup>2</sup> В случае кластерных доменов вида (9) доменный локальный минимум является и локальным минимумом функционала (1) – см. п.2 Приложения; в этом случае размораживание доменов не дает дополнительного выигрыша по энергии.

мики, по оси ординат – средняя частота отыскания самого глубокого минимума (выраженная в процентах). Мы видим, что доменная динамика с кластерными доменами (DM-CLS) приводит в самый глубокий минимум в среднем **в 30 раз чаще**, чем случайная доменная динамика (DM-RND) или случайная динамика (RANDOM). Преимущества кластерного принципа формирования доменов здесь очевидны.

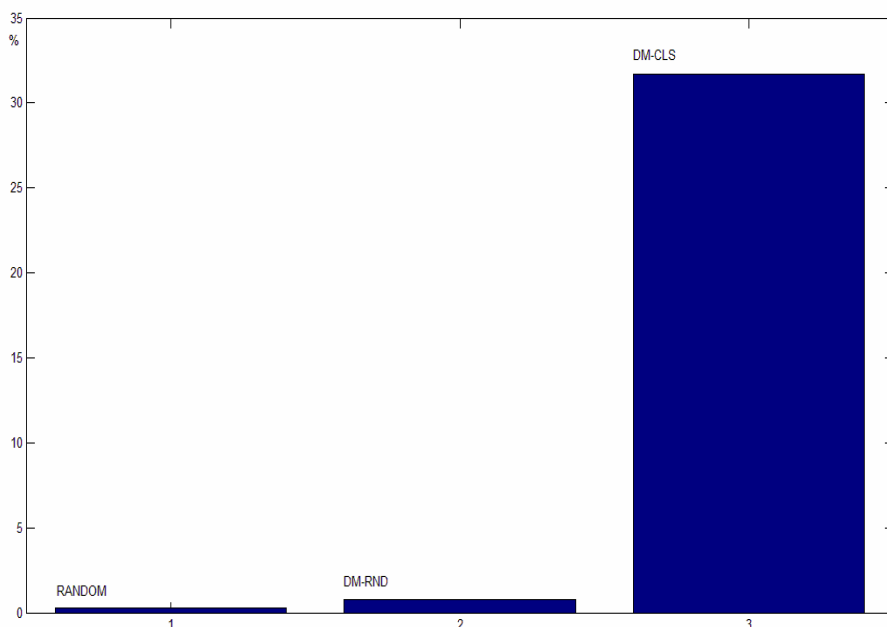


Рис.1. Средняя частота отыскания самого глубокого локального минимума для всех трех динамик.

**2. Ослабление связей внутри сильно связанных спиновых групп.** В описанных экспериментах связи между спинами внутри «правильных» групп были равны максимально возможному значению  $J_{ll}^{(in)} = 1$ , поскольку у матрицы паттернов (7) все  $k_l$  векторов-столбцов  $l$ -й группы были одинаковыми. Что произойдет, если эти группы векторов слегка «размыть»? Если векторы-столбцы внутри «правильных» групп будут не одинаковыми, а слегка отличающимися друг от друга? Как это отразится на результатах минимизации?

Описанные эксперименты были повторены с «размытыми» группами векторов-столбцов  $x_l$ . Теперь первые  $k_l$  векторов-столбцов матрицы (7) брались не идентичными друг другу, а возникали в результате случайного мультипликативного искажения  $M$ -мерного вектора  $x_l$ , а именно: координаты вектора

$x_1$  с вероятностью  $b$  независимым и случайным образом умножались на «-1». Следующие  $k_2$  векторов-столбцов матрицы  $X$  получались в результате аналогичного искажения  $M$ -мерного случайного вектора  $x_2$ , и так далее – последние  $k_n$  векторов-столбцов матрицы  $X$  являлись искажениями  $M$ -мерного случайного вектора  $x_n$ .

Теперь связи между спинами внутри «правильных» групп уже не равны тождественно 1, а являются случайными величинами со средним значением  $\langle J_{ij}^{(in)} \rangle = (1 - 2b)^2$ . Параметр искажения  $b$  характеризует степень неоднородности «правильных» групп спинов: чем больше  $b$ , тем меньше среднее значение внутри-групповой связи, тем менее однородными являются «правильные» группы спинов. Связи между спинами из двух различных групп по-прежнему оцениваются как  $J_{ij}^{(out)} \sim \pm 1/\sqrt{M}$ .

Величины  $b$  и отвечающие им средние значения внутри-групповой связи приведены в табл.1. Только для  $b=0.02$  и  $b=0.05$  «правильные» группы спинов можно считать сильно связанными. Действительно, среднее значение внутри-групповой связи в этих случаях равно  $\langle J_{ij}^{(in)} \rangle \approx 0.9$  и  $\langle J_{ij}^{(in)} \rangle \approx 0.8$  соответственно. Иными словами, углы между  $M$ -мерными векторами  $x_i$  из «правильной» группы меньше  $45^\circ$ ; такие группы векторов еще можно считать плотными, а соответствующие группы спинов – сильно связанными. Но уже для  $b=0.1$  имеем  $\langle J_{ij}^{(in)} \rangle = 0.64$ , а для  $b=0.2$  среднее значение внутри-групповой совсем невелико:  $\langle J_{ij}^{(in)} \rangle = 0.36$ . При таком среднем значении матричного элемента то, что мы по инерции продолжаем считать «правильной» группой спинов, является объединением нескольких слабо связанных подгрупп. Объединяя в один домен несколько слабо связанных спиновых групп, мы, фактически, формируем случайные домены. Можно ожидать, что с ростом  $b$  результаты для DM-CLS-динамики будут все больше походить на результаты для DM-RND-динамики.

**Таблица 1.** Среднее значение внутри-групповой связи между спинами  $\langle J_{ij}^{(in)} \rangle$  в зависимости от  $b$ .

$b$	0.02	0.05	0.1	0.2
$\langle J_{ij}^{(in)} \rangle$	0.92	0.81	0.64	0.36

На рис.2 для трех типов динамики показано, как от параметра  $b$  зависит средняя частота отыскания самого глубокого минимума. Мы видим, что с ростом  $b$  результаты для «правильной» доменной динамики (DM-CLS) все меньше отличаются от результатов для случайной доменной динамики (DM-RND) - в особенности, начиная с  $b=0.1$ . Что и следовало ожидать, поскольку с ростом  $b$  «правильные» группы спинов все меньше отличаются от случайных доменов.

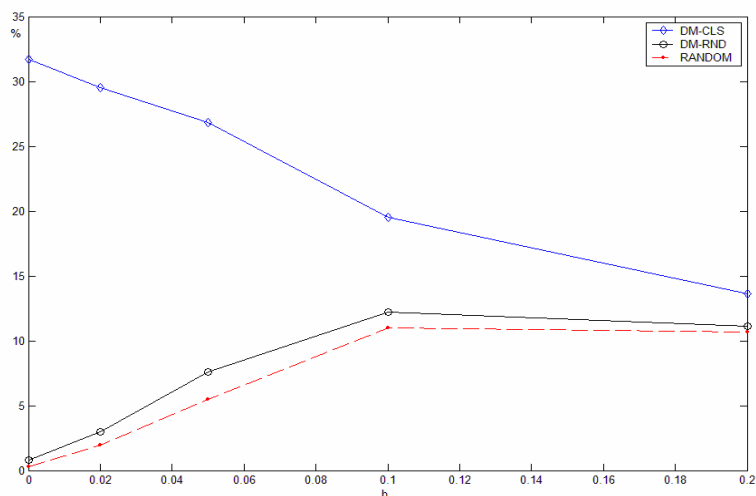


Рис.2. Средняя частота отыскания самого глубокого локального минимума для всех трех динамик как функция параметра  $b$ .

Отметим, что результаты для случайной доменной динамики лишь незначительно лучше результатов для случайной динамики: DM-RND-кривая лишь слегка превосходит RANDOM-кривую. Возможно, это объясняется тем, что длина случайных доменов в наших экспериментах ( $k = 25$ ) далека от оптимальной длины. В работе [5] было установлено, что наилучшие результаты для доменной динамики получаются при  $k=2$ .

**3. Вклады различных динамик в глубину локального минимума.** После остановки в доменном локальном минимуме переходят к случайной динамике, и система опускается еще глубже, в локальный минимум функционала (1). Условимся под *глубиной* локального минимума понимать абсолютное значение соответствующей энергии. Пусть  $D$  – глубина доменного локального минимума, а  $E \geq D$  – глубина окончательного локального минимума функционала (1), в котором система оказывается после «размораживания» доменов. Введем величины  $d=D/E$  и  $r=(E-D)/E$ . Они неотрицательны, и их сумма равна 1:

$$0 \leq d, r \leq 1, d + r = 1.$$

Величины  $d$  и  $r$  характеризуют в относительных единицах то, какая часть глубины минимума определяется доменной динамикой, а какая – случайной динамикой. Например, если  $r \approx 0$ , то можно считать, что практически вся глубина локального минимума определяется доменной динамикой, а увеличение глубины за счет случайной динамики незначительно. Если же  $r \approx 1$ , то имеет место об-

ратная ситуация: доменная динамика почти не приводит к понижению энергии, а вся глубина локального минимума определяется случайной динамикой. Фактически,  $d$  и  $r$  характеризуют относительный вклад в глубину локального минимума доменной и случайной динамик. Интересно сопоставить друг с другом значения  $d$ - и  $r$ -характеристик. Зависит ли соотношение между ними от типа доменной динамики или нет?

Для ответа на эти вопросы мы усредняли  $r$ -характеристику сначала – по 1000 случайным стартам для каждой матрицы, а затем – по 200 случайным матрицам. Такое усреднение проделывалось для двух вариантов доменной динамики: со случайными доменами (DM-RND) и с кластерными доменами (DM-CLS). Графики усредненных  $r$ -характеристик приведены на рис.3. По оси абсцисс отложен уровень искажений  $b$ , по оси ординат – усредненная по 200000 стартов  $r$ -характеристика для кластерной доменной динамики (DM-CLS) и для случайной доменной динамики (DM-RND).

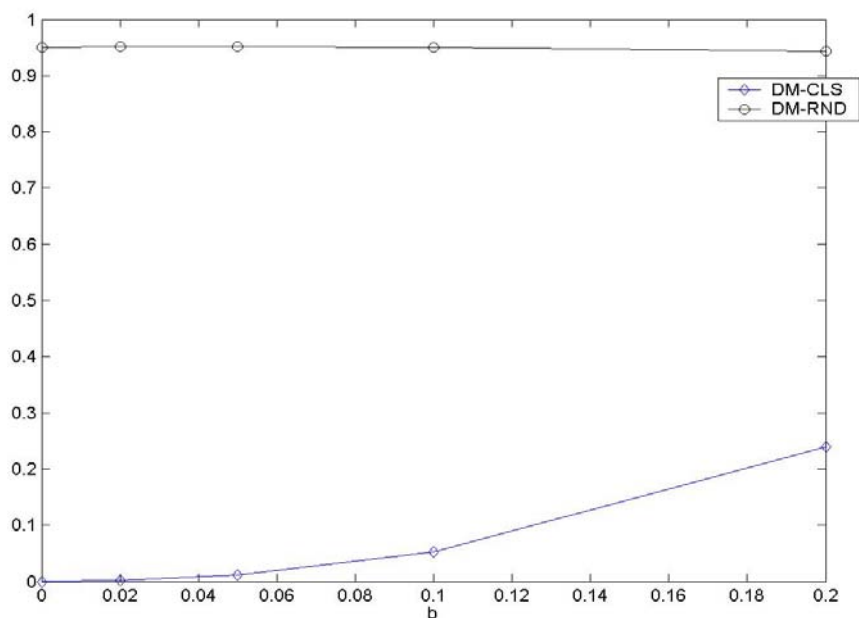


Рис. 3. Усредненные значения  $r$ -характеристики для кластерных доменов (DM-CLS) и случайных доменов (DM-RND) – см. текст.

Мы видим, что для случайной доменной динамики выполняется приближенное равенство  $r \approx 1$ . Иначе говоря, использование случайных доменов длины  $k=25$  характеризуется тем, что в среднем 95% глубины локального минимума определяется случайной минимизацией. Доменная динамика дает весьма незна-



чительный вклад в глубину локального минимума (около 5%). Этот результат практически не зависит от уровня искажений  $b$  – см. ход (DM-RND)-кривой. Можно сказать, что использование случайных доменов длины  $k=25$  лишь незначительно улучшает результаты стандартной случайной минимизации.

Напротив, при использовании кластерных доменов имеем:  $r \ll 1$ . Большая часть глубины локального минимума определяется здесь именно доменной динамикой, а вклад случайной динамики сравнительно мал – см. ход (DM-CLS)-кривой. При нулевом уровне искажений ( $b=0$ ) доменные локальные минимумы вообще являются минимумами функционала (1) (см. п.2 Приложения); в этом случае выполняется точное равенство  $r = 0$ . По мере увеличения параметра  $b$  растет и значение  $r$ -характеристики, однако происходит это очень медленно.

## 5. Заключение

Мы видели, что когда в эмпирических данных можно обнаружить четко выраженные группы сильно связанных спинов, их использование в качестве доменов позволяет на порядок чаще находить глобальный минимум, чем с помощью случайных доменов. Ослабление связей внутри таких групп сильно связанных спинов, «размытие» этих групп, стирает разницу между результатами, характерными для двух способов формирования доменов. Иными словами, результаты компьютерных экспериментов свидетельствуют о продуктивности кластерного принципа формирования доменов. Число и состав доменов определяются при этом в результате кластеризации матрицы  $\mathbf{J}$ .

Под кластеризацией матрицы понимаются приведение ее к такому блочно-диагональному виду, когда матричные элементы внутри диагональных блоков превосходят матричные элементы вне этих блоков. Конкретная постановка задачи кластеризации определяется тем, в каком смысле понимать превосходство внутри-диагональных матричных элементов над вне-диагональными. Широкий диапазон различных подходов к решению каждой конкретной задачи. Обзоры методов кластеризации можно найти в [8].

Для матрицы связей корреляционного типа, когда матричные элементы являются скалярными произведениями некоторого набора векторов, хорошо зарекомендовала себя процедура кластеризации, предложенная в [9]. Эта процедура позволяет не только выделить плотные группы векторов, но и разобраться в том, какие из этих групп расположены ближе друг к другу, а какие дальше. Именно эта информация и требуется для объединения сильно связанных спинов в домены. Вычислительная сложность процедуры кластеризации  $\sim O(N^2)$ . Однако обобщить ее на матрицы связей общего вида – необязательно корреляционные – пока не удается.

Автор признателен А.А.Мурашкину, выполнившему для настоящей работы весь вычислительный эксперимент. Работа выполнялась в рамках программы "Фундаментальные основы информационных технологий и систем" (проект 1.8) при финансовой поддержке грантом РФФИ 06-01-00109.

## Приложение

Докажем здесь два утверждения относительно локальных минимумов функционала (1) для случая, когда матрица связей имеет блочную структуру (8); эти утверждения были использованы нами выше.

**1. Кусочно-постоянный вид локальных минимумов.** Покажем, что первые  $k_1$  координат локального минимума должны быть одинаковыми – см. выражение (9). Рассуждение, устанавливающее справедливость этого утверждения, проведем для первых двух координат. Пусть конфигурация  $\mathbf{s} = (s_1, s_2, s_3, \dots, s_N)$  является локальным минимумом. Тогда каждая координата

$s_i$  должна совпадать со знаком локального поля  $h_i$ :  $s_1 h_1 = s_1 \sum_{j=1}^N J_{1j} s_j \geq 0$  и

$s_2 h_2 = s_2 \sum_{j=1}^N J_{2j} s_j \geq 0$ . Поскольку матрица  $\mathbf{J}$  имеет вид (8), локальные поля можно

записать следующим образом:  $h_1 = s_2 + H$ ,  $h_2 = s_1 + H$ . Второе слагаемое

$H = \sum_{j=3}^N J_{ij} s_j$  в обоих случаях одинаково, поскольку все матричные элементы

первой и второй строк с номерами  $j > 2$  одинаковы:  $J_{1j} = J_{2j}$ ,  $j > 2$  – см. выражение (8). Итак, одновременно должны выполняться два неравенства:

$s_1 s_2 + s_1 H \geq 0$  и  $s_2 s_1 + s_2 H \geq 0$ . Если предположить, что координаты  $s_1$  и  $s_2$  различны, то  $s_1 s_2 < 0$  и одновременно должно выполняться:  $1 \leq H$  и  $1 \leq -H$  – что невозможно. Следовательно, наше предположение неверно и координаты  $s_1$  и  $s_2$  совпадают. Доказательство того, что локальные минимумы функционала (1) с матрицей  $\mathbf{J}$  (8) имеют кусочно-постоянный вид (9), на этом закончено.

**2. Доменные локальные минимумы кусочно-постоянного вида.** Легко сообразить, что любая конфигурация вида (9), являющаяся доменным локальным минимумом, будет и локальным минимумом функционала (1). Действительно, пусть для первого домена выполняется неравенство:  $F_1 = \sum_{i=1}^k s_i h_i^{(d)} \geq 0$ .

Поскольку первые  $k_1$  координат  $s_i$  (которые как раз и объединены в первый домен) равны друг другу, сумма в правой части последнего выражения есть сумма  $k_1$  одинаковых слагаемых:  $F_1 = k_1 s_i h_i^{(d)}$ ; следовательно,

$s_1 h_1^{(d)} = s_1 \sum_{j=k+1}^N J_{1j} s_j \geq 0$ . Используя выражение (4), связывающее локальное поле

$h_i$  с доменным локальным полем  $h_i^{(d)}$ , получаем, что положительным будет и следующее выражение:

$$s_1 h_1 = s_1 h_1^{(d)} + s_1 \sum_{j=1}^{k_1} J_{1j} s_j = s_1 h_1^{(d)} + (k_1 - 1) > 0. \quad (10)$$

Это означает, что координата  $s_1$  совпадает со знаком локального поля  $h_1$ , действующего на нее со стороны всей сети. Иначе говоря, доменный локальный минимум будет и локальным минимумом функционала (1). Следовательно, «размораживание» здесь доменов не приведет к дальнейшему понижению энергии состояния. В терминах п.4.3 это означает, что  $r=0$ .

Обратное неверно: из того, что кусочно-постоянная конфигурация вида (9) является локальным минимумом функционала (1), вовсе не следует, что она будет и доменным локальным минимумом. Иначе говоря, из того, что  $s_1 h_1 > 0$ , совсем не следует, что  $s_1 h_1^{(d)} > 0$ . Это легко увидеть, обратившись к выражению (10):  $s_1 h_1^{(d)}$  может быть отрицательным, но если число спинов  $k_1$  в домене достаточно велико, сумма в правой части выражения (10) будет положительной. Иначе говоря, одновременно будут выполняться два неравенства:  $s_1 h_1 > 0$ ,  $s_1 h_1^{(d)} < 0$ . Это означает, в частности, что доменная динамика позволяет иногда выбираться из не очень глубоких локальных минимумов функционала (1).

## Литература

- [1] A.K. Hartmann, H. Rieger. Optimization Algorithms in Physics. Wiley-VCH, Berlin, 2001.
- [2] J. Hertz, A. Krogh, R. Palmer. Introduction to the Theory of Neural Computation. Addison-Wesley, 1991.
- [3] Б.В. Крыжановский, Б.М. Магомедов, А.Л. Микаэлян. Доменная модель нейронной сети. Доклады Академии наук. 2005. Т. 401. №4. С. 462-466.
- [4] Б.В. Крыжановский, Б.М. Магомедов. Доменный алгоритм релаксации нейронной сети и оптимизация квадратичного функционала. Труды "МСО-2005". М.: МГУ, 2005. С. 607-613.
- [5] V. Kryzhanovsky, B. Magomedov. On the probability of finding local minima in optimization problems. *Proceedings of IJCNN'2006*, Vancouver: 2006. PP. 5882–5887.
- [6] A.V.M. Herz, C.M. Marcus. Distributed dynamics in neural networks. *Phys. Rev. E*47. 1993. PP.2155-2161.
- [7] Л.Б. Литинский. О прямом вычислении неподвижных точек нейронной сети. Теоретическая и математическая физика. 1994. Т.101(3). С. 467-479.
- [8] Н.Г. Загоруйко. Прикладные методы анализа данных и знаний. Новосибирск: Изд-во Ин-та математики, 1999.
- [9] L. Litinskii, D. Romanov. Neural network clustering based on distances between objects. St. Kollias, A. Stafylopatis, W Duch, E.Oja (Eds.). Artificial Neural Networks – ICANN 2006. 16h International Conference, Athens, Greece, September 10-14, 2006, Proceedings, Part II. PP.437-443.

Статья поступила 06 февраля 2007 г.  
После доработки 29 августа 2007 г.

# Перспективы исследований на стыке информатики и биологии\*

Редько В.Г.  
НИИ системных исследований РАН  
E-mail: vgrecko@gmail.com

**Аннотация.** В статье подчеркивается актуальность исследований на стыке информатики и биологии. Выделены уже сложившиеся научные направления: нейронные сети, эволюционное моделирование, адаптивное поведение, модели происхождения молекулярно-генетических систем. Указана ключевая проблема – проблема происхождения правил логического вывода. Аргументируется, что необходима серьезная академическая программа «Эволюция, мозг, интеллект». В качестве первого приближения предлагается вести исследования на феноменологическом уровне, путем анализа когнитивной эволюции – эволюции познавательных свойств биологических организмов. Предлагаются контуры программы исследований когнитивной эволюции. Характеризуются определенные заделы исследований по предлагаемой программе.

## 1. Введение. Направления исследований на стыке информатики и биологии

Информатика и биология – активно развивающиеся науки 21 века. На стыке этих наук можно ожидать интересные и перспективные направления будущих исследований. Для того чтобы наметить такие перспективы, целесообразно охарактеризовать уже складывающиеся в этой области направления, выделить связанную с развитием научного миропонимания нетривиальную проблему и наметить подходы к ее исследованию. Этому и посвящена настоящая статья.

Далее кратко характеризуются уже сложившиеся научные направления, выделяется ключевая проблема – проблема обоснования применимости логического мышления в научном познании – и указываются подходы к ее исследованию путем построения моделей когнитивной эволюции.

На стыке информатики и биологии развивается несколько интересных научных направлений: *нейронные сети, эволюционное моделирование, адаптивное поведение, модели происхождения молекулярно-генетических систем*. Работы по нейронным сетям и эволюционному моделированию хорошо известны, и мы не будем на них сейчас останавливаться. Отметим только книги и лек-

---

\* Работа выполнена при финансовой поддержке программы Президиума РАН «Интеллектуальные компьютерные системы» (проект 2-45) и РФФИ (проект № 07-01-00180).

ции, характеризующие эти направления [1-5]. Модели адаптивного поведения и происхождения самовоспроизводящихся молекулярно-генетических систем управления простейших организмов недостаточно известны специалистам и кратко характеризуются в данном разделе. Подчеркнем, что именно путем моделирования эволюции адаптивного поведения, путем моделирования эволюции познавательных свойств биологических организмов, исследования когнитивной эволюции целесообразно искать подходы к изучению проблемы обоснования применимости логического мышления в научном познании. Также важно, что в работах по анализу происхождения самовоспроизводящихся молекулярно-генетических систем сложились четкие математические и компьютерные модели, для которых уже сейчас можно провести определенное упорядочивание.

Направление «Адаптивное поведение» (АП) активно развивается с начала 1990-х годов [6-8]. Основной подход этого направления – конструирование и исследование искусственных (в виде компьютерной программы или робота) «организмов», способных приспосабливаться к внешней среде. Эти организмы называются «аниматами» (от англ. animal + robot = animat). Также часто используются термины «агент», «автономный агент».

Программа-минимум направления «Адаптивное поведение» – исследовать архитектуры и принципы функционирования, которые позволяют животным или роботам жить и действовать в переменной внешней среде.

Программа-максимум этого направления – попытаться проанализировать эволюцию когнитивных способностей животных и эволюционное происхождение человеческого интеллекта [9].

Для исследований АП характерен *синтетический подход*: здесь конструируются архитектуры, обеспечивающие «интеллектуальное» поведение аниматов. Причем это конструирование проводится как бы с точки зрения инженера: исследователь сам «изобретает» архитектуры, подразумевая конечно, что какие-то подобные структуры, обеспечивающие адаптивное поведение, должны быть у реальных животных.

Часто системы управления АП рассматривают как автономные системы управления [10].

Хотя «официально» направление АП было провозглашено в 1990 году, были явные провозвестники этого направления. Например, в нашей стране модели поведения животных в 1960-70-х годах строились и исследовались М.Л. Цетлиным, М.М. Бонгардом, Д.А. Поспеловым [11-13].

В исследованиях АП используется ряд нетривиальных компьютерных методов:

- нейронные сети [1-3],
- генетический алгоритм и другие методы эволюционной оптимизации [4,5],
- классифицирующие системы (Classifier Systems) [14],
- обучение с подкреплением (Reinforcement Learning) [15].

Достоинство анимат-подхода – он привязан к конкретным задачам: во-первых, к задаче моделирования поведения искусственных, модельных организмов, в частности, роботов, и, во-вторых, к задаче моделирования адаптивного поведения биологических организмов. Исследователи направления АП ста-

раются строить именно такие модели, которые применимы к описанию поведения *как реального животного, так и искусственного анимата* [16,17].

Отметим, что в АП в основном используется *феноменологический подход* к исследованиям систем управления адаптивным поведением. Предполагается, что существуют формальные правила адаптивного поведения, и эти правила не обязательно связаны с конкретными микроскопическими нейронными или молекулярными структурами, которые есть у живых организмов. Скорее всего, такой феноменологический подход для исследований АП вполне имеет право на существование. В пользу этого тезиса приведем аналогию из физики. Есть термодинамика, и есть статистическая физика. Термодинамика описывает явления на феноменологическом уровне, статистическая физика характеризует те же явления на микроскопическом уровне. В физике термодинамическое и статистическое описания относительно независимы друг от друга и, вместе с тем, взаимодополнительны. По-видимому, и для описания живых организмов может быть аналогичное соотношение между феноменологическим (на уровне поведения) и микроскопическим (на уровне нейронов и молекул) подходами. При этом естественно ожидать, что для исследования систем управления адаптивным поведением феноменологический подход должен быть более эффективен (по крайней мере, на начальных этапах работ), так как очень трудно сформировать целостную картину поведения на основе анализа всего сложного многообразия функционирования нейронов, синапсов, молекул.

Подчеркнем, что «Адаптивное поведение» – активно развивающееся направление исследований. Есть научное общество «The International Society for Adaptive Behavior» (<http://www.isab.org.uk>). Регулярно проводятся международные конференции «Simulation of Adaptive Behavior (From Animal to Animat)». Издается журнал «Adaptive Behavior» (<http://www.isab.org.uk/journal>).

Обзор отечественных работ в области адаптивного поведения содержится в книге [7].

В 1960-80 годы проблема возникновения самовоспроизводящихся **молекулярно-генетических информационных систем** в процессе происхождения жизни заинтриговала многих ученых. Целая плеяда Нобелевских лауреатов (М. Эйген, Ф. Крик, Ф. Дайсон, Ф. Андерсон) предприняла попытки представить и промоделировать сценарии возникновения предбиологических информационных систем. В результате этих попыток был разработан ряд интересных математических и компьютерных моделей (квазивидов, гиперциклов и сайзеров), позволяющих представить возможные схемы возникновения молекулярно-генетических самовоспроизводящихся систем простейших организмов [18-22].

Модель *квазивидов* [18,19,23,24] – модель эволюции информационных последовательностей – описывает возможный процесс того, как могли бы эволюционно возникнуть самые простейшие макромолекулы, кодирующие наследственную информацию. В модели рассматривается эволюция популяции цепочек РНК. Эволюция такой популяции приводит к отбору квазивида – такого распределения цепочек РНК, в которое входит как «наилучшая РНК» (размножающаяся с максимальной скоростью), так и близкие к ней цепочки, отличающиеся от этой наилучшей мутационными заменами. Подчеркнем, что эта мо-

дель представляет собой «каноническую» схему эволюции информационных последовательностей, отражающую общие черты эволюционных процессов.

В конце 1970-х годов М. Эйген и П. Шустер предложили следующую по отношению к квазивидам модель – модель *гиперциклов* [19]. В гиперцикле к цепочкам РНК добавляются цепочки аминокислот – ферменты, которые выполняют определенные каталитические функции и вместе с цепочками РНК формируют целостную систему кооперативно взаимодействующих макромолекул. Ферменты циклически катализируют репликацию РНК и способствуют повышению точности копирования, циклическая организация гиперцикла обеспечивает его структурную стабильность.

Еще более близкие «конструкции» к самовоспроизводящейся молекулярно-генетической системе живой клетки – «сайзеры» – были предложены В.А. Ратнером и В.В. Шаминам [25]. Термин «сайзер» был предложен В.А. Ратнером и происходит от *System of self-reproduction (syser)*. В сайзерах ферменты репликации цепочек РНК и трансляции (синтеза ферментов) действуют примерно так же, как и в живой клетке. Модель сайзеров ближе к простейшим биологическим организмам, чем гиперциклы и позволяет анализировать возможные этапы эволюции от простейших мини-сайзеров до протоклеток. Например, модель адаптивного сайзера [26] характеризует процесс эволюционного возникновения системы автономного управления адаптивным поведением в переменной внешней среде.

Очерченные модели – только фрагменты и только возможного сценария эволюции предбиологических информационно-кибернетических систем. Тем не менее, эти модели позволяют четко и на серьезном математическом уровне представить гипотетические этапы возникновения самовоспроизводящейся молекулярно-генетической системы живой клетки (рис.1).

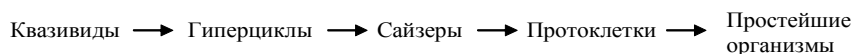


Рис. 1. Схема гипотетической эволюции самовоспроизводящихся макромолекулярных систем.

Подробное описание математических моделей предбиологических молекулярно-генетических информационных систем содержится в [27,28].

## 2. Ключевая проблема – эволюционное происхождение правил доказательства теорем

### 2.1. Гносеологическая проблема

Для серьезной постановки исследовательских работ целесообразно выделить ключевую проблему, связанную с фундаментальной наукой, такую проблему, изучение которой может внести радикальный вклад в научное миропонимание. В данном разделе характеризуется такая проблема.

Существует глубокая гносеологическая проблема: почему логическое мышление *человека* применимо к познанию *природы*?

Действительно, есть фундаментальная естественнонаучная дисциплина – физика. Мощь физики связана с эффективным применением математики. Но математик строит свои теории чисто логическим путем, совсем независимо от внешнего мира, используя свое мышление (в тиши кабинета, лежа на диване, в изолированной камере...). Почему же результаты, получаемые математиком, применимы к реальной природе?

Рассмотрим вопрос более подробно. Допустим, физик, изучая динамику некоторого объекта, сумел в определенном приближении свести описание объекта к дифференциальному уравнению (рис. 2). Далее он интегрирует полученное уравнение согласно известным из математики правилам и получает характеристики движения объекта (в рамках используемого приближения). Переход от дифференциального уравнения к характеристикам движения происходит в сознании физика, а используемые при этом правила интегрирования получены логическим путем математиками. Переход этот чисто дедуктивный, но, если быть предельно строгим, применимость данного перехода к физическому объекту надо обосновывать: поведение физического объекта совершенно необязательно должно соответствовать правилам логики человеческого мышления.<sup>1</sup>

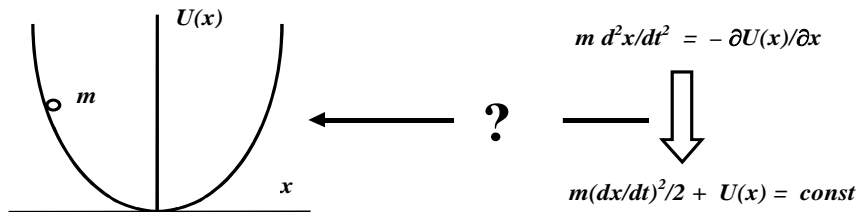


Рис. 2. Почему логика человеческого мышления применима к познанию природы? Левая часть рисунка иллюстрирует движение материальной частицы в потенциальном рельефе, правая часть – дедуктивный логический вывод закона сохранения энергии из второго закона Ньютона для рассматриваемого объекта. Вопрос состоит в том, почему логический вывод, сделанный человеком, применим к реальному объекту в природе.

В более общей формулировке проблему можно поставить так: *почему логика человеческого мышления применима к познанию природы?* Действительно,

<sup>1</sup> На сомнения в применимости математики к познанию природы могут навести и размышления о неожиданных поворотах в науке. Например, многие ученые задумывались над тем, почему общей теории относительности и квантовой механике вдруг потребовались достижения с переднего фронта математики – неевклидова геометрия и теория эрмитовых операторов над векторами в гильбертовом пространстве. Могло сложиться впечатление, что путь развития математики в какой-то степени предопределен развитием содержательного познания природы, хотя математики в своих построениях стремились к максимальной обоснованности и общности своих теорий.



с одной стороны, логические процессы вывода происходят в нашем, человеческом мышлении, с другой стороны, процессы, которые мы познаем посредством логики, относятся к изучаемой нами природе. Эти два типа процессов различны. Поэтому далеко не очевидно, что мы можем использовать процессы первого типа для познания процессов второго типа.

Можно ли конструктивно подойти к решению этих вопросов? Чтобы продемонстрировать такую возможность, будем рассуждать следующим образом.

Рассмотрим одно из элементарных правил, которое использует математик в логических заключениях, правило *modus ponens*: «если имеет место  $A$ , и из  $A$  следует  $B$ , то имеет место  $B$ », или  $\{A, A \rightarrow B\} \Rightarrow B$ .

А теперь перейдем от математика к собаке И.П. Павлова. Пусть у собаки вырабатывают условный рефлекс, в результате в памяти собаки формируется связь «за УС должен последовать БС» (УС – условный стимул, БС – безусловный стимул). И когда после выработки рефлекса собаке предъявляют УС, то она, «помня» о хранящейся в ее памяти «записи»: УС  $\rightarrow$  БС, делает элементарный «вывод»:  $\{УС, УС \rightarrow БС\} \Rightarrow БС$ . И у собаки, ожидающей БС (скажем, кусок мяса), начинают течь слюнки.

Конечно, применение правила *modus ponens* (чисто дедуктивное) математиком и индуктивный «вывод», который делает собака, явно различаются. Но можем мы ли думать об эволюционных корнях логических правил, используемых в математике? Да, вполне можем – умозаключение математика и «индуктивный вывод» собаки качественно аналогичны.

Итак, мы можем думать над эволюционными корнями логического мышления, над происхождением тех правил, которые используются при математических доказательствах. Для решения проблемы применимости нашего мышления в научном познании целесообразно разобраться, как произошли познавательные способности человека в процессе биологической эволюции и как эволюционно появилась их применимость к познанию явлений природы. Наиболее четкий путь подобного исследования – построение математических и компьютерных моделей эволюции познавательных способностей животных и эволюционного происхождения мышления человека.

Рассматриваемая гносеологическая проблема – фундаментальная проблема. Она касается принципиальных возможностей человеческого познания, возможностей научного познания природы. Естественно ожидать, что ее изучение позволит укрепить основания всего величественного здания науки. Следовательно, эта проблема заслуживает тщательного исследования.

## 2.2. Почему проблема происхождения правил логического вывода

Каковы же правила логического вывода, происхождение которых необходимо исследовать? Для ответа на этот вопрос обратимся к математической теории логического вывода. Методы логического вывода в контексте математики 20-го века были формализованы в 1930 годы немецким математиком Г. Генценом, который в качестве базового примера рассматривал евклидово доказательство теоремы: «*Существует бесконечно много простых чисел*».

Генцен [29,30] выделил правила, используемые при математических доказательствах, и проанализировал связь математической теории логического вывода с разработанными в то время исследованиями логики. Правила математических доказательств были более общими, чем указанное выше правило *modus ponens*, но, тем не менее, именно «формулы», подобные *modus ponens*, были в их основе.

Таким образом, правила, происхождение которых необходимо исследовать, в определенном приближении, известны.

Подчеркнем, что пока мы не останавливаемся на вопросе о том, как может возникнуть доказательство той или иной теоремы, решение той или иной проблемы. Примером схемы возникновения научного решения может служить известное описание А. Пуанкаре, который рассказывает о том, как однажды он долго не смог заснуть и обдумывал свою математическую задачу [31]. В конце концов, путем случайной комбинации частичных решений у него сложилось общее решение той задачи, которую он обдумывал. В настоящей работе основное внимание концентрируется на ключевой проблеме – проблеме анализа эволюционного происхождения логических правил математических доказательств.

Среди подходов, подобных развиваемым здесь, отметим также яркую работу Ю. Вигнера [32], заострявшего внимание на проблеме «непостижимой эффективности математики в естественных науках». Однако в отличие от Ю. Вигнера, который подчеркивал многочисленные трудности, здесь мы намечаем подходы к конструктивному анализу указанной ключевой проблемы путем исследования эволюции познавательных процессов, приведших к логическим правилам, используемым в математике.

Естественно, что мышление человека не сводится только к правилам доказательства теорем, но именно эти правила четко формализованы, лежат в основе математического знания и на них имеет смысл сконцентрироваться на начальном этапе исследований.

### 2.3. Взгляды философов. Становление эволюционной эпистемологии

Над проблемой применимости мышления в научном познании думали многие философы, работавшие в области теории познания. Отметим некоторые точки зрения.

В «Исследовании о человеческом познании» (1748) Давид Юм подверг сомнению понятие причинной связи [33]. Он задался вопросом: почему когда мы видим, что за одним явлением *A* постоянно следует другое *B*, то мы приходим к выводу, что *A* является причиной *B*? Например, когда мы наблюдаем, что Солнце освещает камень, и камень нагревается, то мы говорим, что солнечный свет есть причина нагревания камня.

Юм взглянул на наш познавательный процесс со стороны, извне. Он как бы вышел на некий мета-уровень рассмотрения наших собственных познавательных процессов и задался вопросом о том, откуда взялись эти познавательные процессы и почему они работают.

Острота сомнений Юма была в том, что, развивая его идеи, можно ставить вопрос о принципиальной способности человека познавать мир.

Остроту сомнений Юма хорошо почувствовал Иммануил Кант. Но Кант также видел мощь и силу современной ему науки. Тогда уже была глубокая, серьезная и развитая математика, мощная ньютоновская физика, дающая картину мира, позволяющая объяснить множество явлений на основе немногих четких предположений, использующая многозвенные и сильные математические дедуктивные выводы. Что было делать Канту? Подвергнуть сомнению все эти познавательные процессы? И развивая сомнения Юма дальше, отвергнуть всю науку? Ведь на самом деле – драма!

Конечно же, Кант, как научно образованный человек, не стал отвергать современную ему науку, а постарался разобраться, как же работают познавательные процессы. В результате появились знаменитая «Критика чистого разума» [34] и популярное изложение «Критики...» «Пролегомены ко всякой будущей метафизике, могущей появиться, как наука» [35]. Кант провел исследование познавательных процессов в определенном приближении – приближении фиксированного мышления взрослого человека. Он не задавался вопросом, *откуда* берутся познавательные способности, он просто констатировал факт, что они существуют, и исследовал, *как* они работают. В результате этого анализа Кант пришел к выводу, что существует система категорий, концепций, логических правил и методов вывода (таких как заключения относительно причинных связей между событиями), которые используются в познании природы. Эта система «чистого разума» имеет априорный характер, – она существует в нашем сознании прежде всякого опыта – и является основой научного познания природы.

Приближение фиксированного мышления человека наложило свой отпечаток: Кант утверждает – и вполне логично (!) – что так как «чистый разум» априорен, то наш рассудок в познавательном процессе предписывает свои законы природе:

*« ... хотя в начале это звучит странно, но, тем не менее, верно, если я скажу: рассудок не черпает свои законы (a priori) из природы, а предписывает их ей»* [35].

Наверно, во времена Канта было разумно ограничиться приближением фиксированного мышления взрослого человека – все сразу не охватишь. Кроме того, не было еще теории Чарльза Дарвина. По-видимому, если бы Кант знал теорию происхождения видов, то он явно бы задумался об эволюционном происхождении «чистого разума». Тем более что эволюционные идеи ему были не чужды – вспомним его теорию происхождения Солнечной системы. Но приближение фиксированного мышления взрослого человека накладывает свои ограничения – оно не позволяет ответить на вопросы: откуда же взялись познавательные способности, познаем ли мы истинные законы природы или наш рассудок «предписывает их ей». Фактически Кант ушел от наиболее острой части вопроса, поставленного Юмом, – он не задавался вопросом, откуда взялся «чистый разум», а только тщательно и детально исследовал свойства «чистого разума» и применение его в научном познании.

Естественно, что после появления теории происхождения видов Дарвина должна была произойти ревизия концепции априорного «чистого разума». И она произошла. Очень четко ее выразил Конрад Лоренц в знаменитой статье

«Кантовская доктрина априорного в свете современной биологии» [36]. Согласно Лоренцу, кантовские априорные категории и другие формы «чистого разума» произошли в результате естественного отбора: «Наши категории и формы восприятия, данные до индивидуального опыта, адаптированы к внешнему миру точно по той же причине, по какой копыто лошади адаптировано к почве степи и плавник рыбы адаптирован к воде до того, как рыба вылупится из икринки» [36].<sup>2)</sup>

Составляющие «чистого разума» возникали постепенно в процессе эволюции, в результате многочисленных взаимодействий с внешним миром. В эволюционном контексте «чистый разум» совсем *не априорен*, а имеет явные эволюционные *эмпирические* корни.

Такое представление дало начало эволюционной эпистемологии, современной эволюционной теории познания, развитой в философско-методологических работах К. Лоренца, К. Поппера, Д. Кэмпбелла, Г. Фоллмера. Краткий обзор эволюционной эпистемологии дан в представленной в Интернет-журнале работе [39].

Однако в работах философов как-то не ощущается понимания остроты проблемы о принципиальной способности человека познавать мир. Проблемы: почему с помощью нашей *человеческой* логики, нашего человеческого мышления, нашего интеллекта, нашего «чистого разума» мы способны познавать *природу*. Но ведь есть драма. Примерно такая же, какую, возможно, ощутил И. Кант перед «Критикой чистого разума». Надо либо подвергнуть сомнению все научные знания, – которые получены с помощью человеческого мышления, в котором можно сомневаться, – либо заняться обоснованием самого мышления. И как уже отмечалось, наиболее естественный подход к решению проблемы – исследовать биологические корни наших познавательных способностей и постараться разобраться, как и почему эти способности возникли, и почему в процессе их эволюционного возникновения появилась возможность познания природы. Наиболее четкий путь такого исследования – построение математических и компьютерных моделей когнитивной эволюции, эволюции познавательных свойств биологических организмов. Причем здесь, как это ни удивительно, возможен симбиоз философской науки, эпистемологии, с естественнонаучными исследованиями.

Как же конкретно возникали составляющие «чистого разума» в процессе эволюции, за счет чего они обеспечивали познание закономерностей природы животными, как эволюция познавательных способностей животных привела к возникновению интеллекта человека?

Есть ли задел исследований, направленных на решение этих интригующих вопросов? Оказывается, что да, есть. Такой задел развивается в направлении исследований «Адаптивное поведение» (см. Введение), дальняя цель которого близка к задаче исследования когнитивной эволюции. В следующих разделах мы обсудим методы построения моделей эволюции познавательных способностей на базе исследований эволюции адаптивного поведения.

---

<sup>2)</sup> Отметим также работы И.М. Сеченова [37,38], который один из первых занимался проблемой эволюционного происхождения логического мышления человека.

### **3. Необходима серьезная академическая программа «Эволюция, мозг, интеллект». Контур программы**

Исследования когнитивной эволюции будут вестись эффективно, если они будут идти по добротной академической программе. В качестве рабочего названия программы исследования эволюционного происхождения мышления предлагаем следующее: «Эволюция, мозг, интеллект». В рамках этой программы предлагается провести исследования когнитивной эволюции на базе методов, сложившихся в направлении «Адаптивное поведение».

Исследования целесообразно вести методом последовательных приближений. В качестве первого приближения разумно использовать сложившийся в «Адаптивном поведении» феноменологический подход.

Предложим контуры программы исследований эволюционного происхождения мышления человека.

#### **А. Разработка схем и моделей поведения аниматов, имеющих естественные потребности**

Разумно начать с моделирования целостного адаптивного поведения простых аниматов, имеющих естественные потребности: питания, размножения, безопасности. Работы [40,41] представляют собой задел таких исследований.

#### **Б. Исследование перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов**

Такой переход можно рассматривать, как появление в «сознании» животного свойства «понятие». Обобщенные образы можно представить как мысленные аналоги наших слов, не произносимых животными, но реально используемых ими. Например, у собаки явно есть понятия «хозяин», «свой», «чужой», «пища». И важно осмыслить, как такой весьма нетривиальный переход мог произойти в процессе эволюции.

#### **В. Исследование процессов формирования причинной связи в памяти животных**

По-видимому, запоминание причинно-следственных связей между событиями во внешней среде и адекватное использование этих связей в поведении – одно из основных свойств активного познания животным закономерностей внешнего мира. Такая связь формируется, например, при выработке условного рефлекса: животное запоминает связь между условным стимулом (УС) и следующим за ним безусловным стимулом (БС), что позволяет ему предвидеть события в окружающем мире и адекватно использовать это предвидение.

Естественный следующий шаг – переход от отдельных причинных связей к «базам знаний», к логическим выводам на основе уже сформировавшихся знаний.

#### **Г. Исследование процессов формирования логических выводов в «сознании» животных**

Фактически, уже на базе классического условного рефлекса животные способны делать «логический вывод» вида:  $\{УС, УС \rightarrow БС\} \Rightarrow БС$  или «Если имеет место условный стимул, и за условным стимулом следует безусловный, то нужно ожидать появления безусловного стимула». Такие выводы подобны выводам математика, доказывающего теоремы (см. п. 2.1). И целесообразно разобраться в системах подобных выводов, понять, насколько адаптивна логика поведения животных и насколько она подобна нашей, человеческой логике. Возможно, здесь были бы полезны семантические сети, предложенные разработчиками искусственного интеллекта [42], и сопоставление процессов выводов на семантических сетях с «выводами» поведенческой логики животных.

#### **Д. Исследование коммуникаций, возникновения языка**

Наше мышление тесно связано с языком, с языковым общением между людьми. Поэтому целесообразно проанализировать: как в процессе биологической эволюции возникал язык общения животных, как развитие коммуникаций привело к современному языку человека, как развитие коммуникаций и языка способствовало развитию логики, мышления, интеллекта человека.

Конечно же, перечисленные пункты формируют только контуры плана будущих исследований по предлагаемой программе. Тем не менее, уже сейчас видно, сколь широк фронт исследований, и как много нетривиальной, интересной и важной работы предстоит сделать.

### **4. Заделы работы по программе**

Для того чтобы представить, как вести исследования по программе, необходим хороший аналитический обзор уже имеющего задела. Такой обзор необходимо сделать в будущем. В данном разделе будут охарактеризованы только отдельные фрагменты проведенных работ.

#### **4.1. Почему «Эволюция»**

Познавательные способности биологических организмов сложились в процессе эволюции.

Кроме того, сейчас активно развивается эволюционное моделирование, работы которого могут быть использованы при исследовании когнитивной эволюции. Например, исследования модели квазивидов (об этой модели см. Введение) позволяют сделать оценку эффективности эволюционного поиска оптимального решения [23,24,43]. За счет того, что эволюционный процесс идет путем последовательного уточнения найденных решений посредством отбора и

мутаций, оптимальное решение находится достаточно быстро. В простейшем случае единственного оптимума эволюционный процесс сходится за порядка  $N$  поколений, при этом общее число особей, участвующих в эволюционном процессе при поиске оптимального решения, составляет порядка  $N^2$ . Здесь  $N$  – длина гена особи, число двоичных символов, составляющих ген особи. Эволюционный процесс в данном случае можно назвать «субоптимальным»: он не обеспечивает максимальную скорость поиска, тем не менее, он намного быстрее случайного перебора.

В определенных случаях эволюционный поиск нетривиального решения может быть более эффективен, чем обучение. Так, например, в [44,45] методом компьютерного моделирования исследовалась модель эволюции популяции самообучающихся автономных агентов и анализировалось взаимодействие между обучением и эволюцией. Система управления отдельного агента была основана на нейросетевых адаптивных критиках [46,47]. Модель отработывалась на примере агента-брокера, который имеет ресурсы двух типов: деньги и акции; сумма этих ресурсов составляет капитал агента. Агент стремился увеличить свой капитал, покупая и продавая акции.

Сравнивались три случая адаптации агента: случай L (чистое обучение), случай E (чистая эволюция), случай LE (эволюция + обучение). Обучение нейронных сетей агента проводилось методом обучения с подкреплением [15], при этом использовалось  $\varepsilon$ -жадное правило, в котором с малой вероятностью  $\varepsilon$  действия агента выбирались случайными, а с вероятностью  $1-\varepsilon$  – оптимальными в соответствии с полученным обучением.

Было проведено сравнение ресурса, приобретаемого агентами за  $T = 200$  временных тактов для этих трех способов адаптации. Для случаев E и LE считалось, что  $T$  – продолжительность поколения и регистрировалось максимальное значение ресурса в популяции  $R_{max}(n_g)$  в конце каждого поколения, в этих случаях  $n_g$  – номер поколения. В случае L (чистое обучение) рассматривался только один агент, ресурс которого для удобства сравнения со случаями E и LE обнулялся каждые  $T = 200$  тактов времени. В этом случае индекс  $n_g$  увеличивался на единицу после каждых  $T$  временных тактов.

Зависимости  $R_{max}(n_g)$  показаны на рис. 3. Чтобы исключить уменьшение значения  $R_{max}(n_g)$  из-за случайного выбора действий при применении  $\varepsilon$ -жадного правила [15] для случаев LE и L полагалось  $\varepsilon = 0$  после  $n_g = 100$  для случая LE и после  $n_g = 2000$  для случая L (на рис. 3 видно резкое увеличение  $R_{max}(n_g)$  после  $n_g = 100$  и  $n_g = 2000$  для соответствующих случаев).

Рис. 3 показывает, что обучение, объединенное с эволюцией (случай LE), и чистая эволюция (случай E) дают одно и то же значение конечного ресурса  $R_{max}(500) = 6,5$ , соответствующее оптимальной стратегии поведения (покупки и продажи акций). Однако эволюция и обучение вместе обеспечивают нахождение больших значений  $R_{max}$  быстрее, чем эволюция отдельно – существует симбиотическое взаимодействие между обучением и эволюцией.

В случае L (чистое обучение) асимптотическое значение ресурса ( $R_{max}(2500) = 5,4$ ) было существенно меньше, чем для случаев E и LE. Анализ компьютерных экспериментов показал, что одно обучение обеспечивает нахождение только «субоптимальной» стратегии поведения: агент держит капитал в

акциях при росте и при слабом падении курса и переводит капитал в деньги при сильном падении курса.

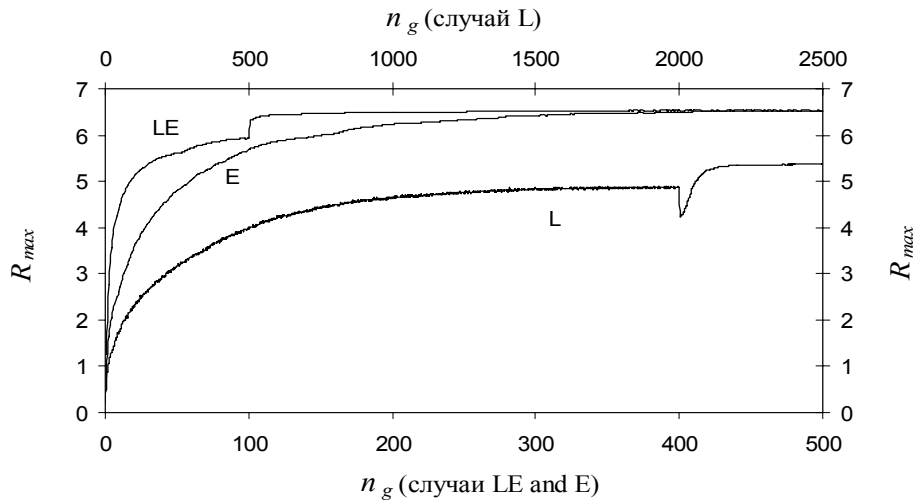


Рис. 3. Зависимости  $R_{max}(n_g)$ . Кривая LE соответствует случаю эволюции, объединенной с обучением, кривая E – случаю чистой эволюции, кривая L – случаю чистого обучения. Временная шкала для случаев LE и E (номер поколения  $n_g$ ) представлена снизу, для случая L (индекс  $n_g$ ) – сверху. Кривые усреднены по 1000 компьютерным экспериментам.

Итак, приведенные результаты демонстрируют, что эволюционная адаптация в рассмотренной модели более эффективна, чем обучение. Это связано с тем, что обучение в нейронных сетях агентов имеет ряд промежуточных шагов, а эволюция идет к поиску оптимума более прямым путем. Тем не менее, хотя обучение и несовершенно, оно способствует более быстрому нахождению оптимальной стратегии поведения по сравнению со случаем чистой эволюции (см. графики LE и E на рис. 3).

Таким образом, эволюционный подход важен как с точки зрения развития исследования когнитивной эволюции в целом, так и с учетом того, что эволюционные процессы представляют собой эффективный метод адаптации.

#### 4.2. Почему «Мозг»

Познавательные способности животных обеспечиваются их мозгом. Исследования нейронных сетей идут широким фронтом. И хотя эти исследования в основном прикладные, тем не менее, есть большое количество работ, которые анализируют эволюционные аспекты становления нейронных структур мозга, обеспечивающих познавательные процессы. Например, в [48] проанализирована эволюция нейронных структур мозга животных, играющих важную роль при выборе действий, обеспечивающих адаптивное поведение.



Многие работы по нейронным сетям посвящены распознаванию образов, т.е. анализу того, как происходит сокращение сенсорной информации – от набора признаков к классам образов. См. например, теоретические работы [49,50]. Это связано с эволюционным процессом перехода от физического уровня обработки информации в нервной системе животных к уровню обобщенных образов, изучение которого представляет собой один из пунктов предлагаемой программы исследований (см. п. Б в разделе 3).

Работа [51] непосредственно посвящена изучению процессов в нейронных моделирующих полях, процессов формирования иерархической системы обобщенных понятий с помощью нейроподобной структуры.

Начата работа по конструированию схем адаптивного поведения анимата с естественными потребностями (проект «Мозг анимата») на основе теории функциональных систем [52-55]. Этот проект может быть положен в основу базовых моделей когнитивной эволюции.

#### **4.3. Почему «Интеллект»**

Конечной целью предлагаемых исследований должен быть процесс эволюционного формирования логического мышления человека. Эта цель, несомненно, актуальна, так как сейчас идут активные работы по попыткам моделирования естественного и искусственного интеллекта. В частности, такие исследования непосредственно связаны с разработкой роботов, обладающих элементами интеллекта, см. например, работы [56].

Интересные работы связаны с методами формализации правил принятия решения. Например, в [57] предложены обобщенные, соответствующие разным уровням биологической эволюции правила принятия решений. Правила учитывают ассоциативные модели, модели классического и инструментального условного рефлексов, модели предвидения результатов действия. Проработаны схемы обучения и принятия решения на основе этих правил, проведено соответствующее компьютерное моделирование, подтвердившее эффективность предложенных правил принятия решений.

Выделим также работу [58], в которой разработаны схемы моделирования мышления и организации интеллектуального поведения и подчеркивается, что интеллект необходим для организации адаптивного поведения.

В данном разделе отмечены только некоторые научные работы, которые связаны с предлагаемой программой. Однако число таких работ велико и анализ задела исследований по программе должен быть продолжен.

### **5. Заключение**

Итак, предложена программа исследований на стыке информатики и биологии, таких исследований, которые могут внести радикальный вклад в научное миропонимание. Работы по программе связаны с современными исследованиями по теории познания и когнитивным наукам. Выделена ключевая проблема – проблема применимости логического мышления человека в научном познании природы. В качестве первого приближения целесообразно изучать

эволюцию познавательных способностей животных на феноменологическом уровне, используя подходы математического и компьютерного моделирования, сложившиеся в современном направлении «Адаптивное поведение».

### Литература

1. Хайкин С. Нейронные сети: полный курс, 2-е издание. М.: Издательский дом «Вильямс», 2006. 1104 с.
2. Осовский С. Нейронные сети для обработки информации. М.: Финансы и статистика, 2002.
3. Лекции по нейронным сетям. Подборка по конференциям «Нейроинформатика-1999» – «Нейроинформатика-2007». Под ред. Ю.В. Тюменцева. Представлены на сайте: <http://www.niisi.ru/iont/ni/Library/>
4. Holland J.H. Adaptation in Natural and Artificial Systems. Ann Arbor, MI: The University of Michigan Press, 1975 (1st edn). Boston, MA: MIT Press., 1992 (2nd edn).
5. Емельянов В.В., Курейчик В.М., Курейчик В.В. Теория и практика эволюционного моделирования. М.: Физматлит, 2003.
6. Meyer J.-A., Wilson S. W. (Eds) From animals to animats. Proceedings of the First International Conference on Simulation of Adaptive Behavior. The MIT Press: Cambridge, Massachusetts, London, England. 1991.
7. От моделей поведения к искусственному интеллекту. Серия «Науки об искусственном» (под ред. Редько В.Г.). М.: УРСС, 2006.
8. Nolfi S., Baldassarre G., Calabretta R., Hallam J., Marocco D., Miglino O., Meyer J-A, Parisi D. (Eds). From animals to animats 9: Proceedings of the Ninth International Conference on Simulation of Adaptive Behaviour. LNAI. 2006. Volume 4095. Berlin, Germany: Springer Verlag.
9. Donnart J.Y., Meyer J.A. Learning reactive and planning rules in a motivationally autonomous animat // IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, 1996. V. 26. N. 3. PP. 381-395. See also: <http://animatlab.lip6.fr/index.en.html>
10. Жданов А.А. Метод автономного адаптивного управления // Изв. РАН. Теория и системы управления. 1999. N. 5. С. 127-134.
11. Цетлин М.Л. Исследования по теории автоматов и моделирование биологических систем. М.: Наука, 1969. 316 с.
12. Бонгард М.М., Лосев И.С., Смирнов М.С. Проект модели организации поведения – «Животное» // Моделирование обучения и поведения. М.: Наука, 1975. С.152-171.
13. Гаазе-Рапопорт М.Г., Поспелов Д.А. От амебы до робота: модели поведения. М.: Наука, 1987, М.: УРСС, 2004.
14. Holland J.H., Holyoak K.J., Nisbett R.E., Thagard P. Induction: Processes of Inference, Learning, and Discovery. Cambridge, MA: MIT Press, 1986.
15. Sutton R., Barto A. Reinforcement Learning: An Introduction. Cambridge: MIT Press, 1998. See also: <http://www.cs.ualberta.ca/~sutton/book/the-book.html>
16. Непомнящих В.А. Аниматы как модель поведения животных // IV Всероссийская научно-техническая конференция «Нейроинформатика-2002». Материалы дискуссии «Проблемы интеллектуального управления – общесистемные, эволюционные и нейросетевые аспекты». М.: МИФИ, 2003. С. 58-76. См. также <http://www.keldysh.ru/pages/BioCyber/RT/Nepomn.htm>
17. Непомнящих В.А. Поиск общих принципов адаптивного поведения живых организмов и аниматов // Новости искусственного интеллекта. 2002. N. 2. С. 48-53.
18. Эйген М. Самоорганизация материи и эволюция биологических макромолекул. М.: Мир, 1973. 216 с.
19. Эйген М., Шустер П. Гиперцикл. Принципы самоорганизации макромолекул. М.: Мир, 1982. 270 с.

20. Crick F.H.C. The origin of the genetic code // *Journal of Molecular Biology*. 1968. Vol. 38. No.3. PP. 367-379.
21. Dyson F.J. A model for the origin of life // *J. Mol. Evol.* 1982. Vol.18. No.5. PP.344-350.
22. Anderson P.W. Suggested model for prebiotic evolution: the use of chaos // *Proc. Natl. Acad. Sci. USA*. 1983. V.80. No.11. PP. 3386-3390.
23. Редько В.Г. Оценка скорости эволюции в моделях Эйгена и Куна // *Биофизика*. 1986. Т. 31. N.3. С. 511-516.
24. Редько В.Г., Цой Ю.Р. Оценка эффективности эволюционных алгоритмов // *Доклады АН*, 2005. Т. 404. N. 3. С. 312-315.
25. Ратнер В.А., Шамин В.В. Сайзеры: моделирование фундаментальных особенностей молекулярно-биологической организации. Соответствие общих свойств и конструктивных особенностей коллективов макромолекул // *Журн. общ. биологии*. 1983. Т.44. N.1. С. 51-61.
26. Редько В.Г. Адаптивный сайзер // *Биофизика*. 1990. Т. 35. Вып. 6. С. 1007-1011.
27. Редько В.Г. Эволюционная кибернетика. М.: Наука, 2001. 156 с.
28. Редько В.Г. Эволюция, нейронные сети, интеллект: Модели и концепции эволюционной кибернетики. Серия «Синергетика: от прошлого к будущему». М.: УРСС, 2005. 224 с.
29. Генцен Г. Исследования логических выводов // *Математическая теория логического вывода*. М.: Наука. 1967. С.9-76.
30. Генцен Г. Непротиворечивость чистой теории чисел // *Математическая теория логического вывода*. М.: Наука. 1967. С.77-153.
31. Пуанкаре А. О науке. М.: Наука. 1990, 736 с. (стр. 404-405).
32. Вигнер Е. Непостижимая эффективность математики в естественных науках // *Успехи физических наук*. 1968. Т. 94. Вып. 3. С. 535-546.
33. Юм Д. Исследование о человеческом познании. Соч. в 2-х томах. Т.2. М.: Мысль, 1966. С. 5-169.
34. Кант И. Критика чистого разума. Соч. в 6-ти томах. Т.3. М.: Мысль, 1964. С. 69-695.
35. Кант И. Прологомены ко всякой будущей метафизике, могущей появиться как наука. Соч. в 6-ти томах. Т.4, часть 1. М.: Мысль, 1965. С. 67-210.
36. Lorenz K. Kant's doctrine of the a priori in the light of contemporary biology // *Learning, Development and Culture* (ed. Plotkin H.). N.Y. 1982.
37. Сеченов И.М. Рефлексы головного мозга. Серия «Из наследия мировой психологии». Изд.2. УРСС, 2007. 128 с.
38. Сеченов И.М. Элементы мысли. Сб. избр. статей. М. — Л., 1943. СПб.: Питер, 2001. 416 с.
39. Кузнецов Н.А., Баксанский О.Е., Гречишкина Н.А. Происхождение знания: истоки и основания // *Информационные процессы*. 2007. Том 7. №1. С. 72-92. <http://www.jip.ru/2007/72-92-2007.pdf>
40. Бурцев М.С., Гусарев Р.В., Редько В.Г. Исследование механизмов целенаправленного адаптивного управления // *Изв. РАН. Теория и системы управления*. 2002. No 6. С. 55-62.
41. Бурцев М.С. Модель эволюционного возникновения целенаправленного адаптивного поведения. 2. Исследование развития иерархии целей // *Препринт ИПМ РАН*, 2002, N. 69.
42. *Semantic Networks in Artificial Intelligence* (F. Lehmann, E.Y. Rodin, eds.), Pergamon Press, Oxford, 1992.
43. Редько В.Г. Спиновые стекла и эволюция // *Биофизика*. 1990. Т. 35. Вып. 5. С. 831-834.
44. Red'ko V.G., Mosalov O.P., Prokhorov D.V. A model of evolution and learning // *Neural Networks*, 2005. Vol. 18, No 5-6. PP. 738-745.

45. Мосалов О.П., Прохоров Д.В., Редько В.Г. Сравнение эволюции и обучения как методов адаптации агентов // Научная сессия МИФИ - 2006. VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006»: Сборник научных трудов. В 3-х частях. Ч.1. М.: МИФИ, 2006. С. 201-208.
46. Редько В.Г., Прохоров Д.В. Нейросетевые адаптивные критики // Научная сессия МИФИ-2004. VI Всероссийская научно-техническая конференция «Нейроинформатика-2004». Сборник научных трудов. Часть 2. М.: МИФИ, 2004. С. 77-84.
47. Prokhorov D.V., Wunsch D.C. Adaptive critic designs // IEEE Trans. Neural Networks, 1997. Vol. 8. No.5. PP. 997-1007.
48. Prescott T.J. Forced moves or good tricks in design space? Landmarks in the evolution of neural mechanisms for action selection // Adaptive Behavior. 2007. Vol. 15. No. 1. PP. 9-31.
49. Крыжановский Б.В., Крыжановский В.М., Микаэлян А.Л. Доменная модель нейронной сети // Доклады АН. 2005. Т. 401. №4. С. 462-466.
50. Крыжановский Б.В., Крыжановский В.М., Микаэлян А.Л. Применение процедуры клиппирования в задачах бинарной минимизации квадратичного функционала // Доклады АН. 2007. Т. 413. №6. С. 730-733.
51. Перловский Л.И. К физической теории мышления: теория нейронных моделирующих полей // Нейроинформатика, 2006. Том 1, № 2. С. 175-196.  
<http://www.niisi.ru/iont/ni/Journal/N2/>
52. Анохин П.К. Принципиальные вопросы общей теории функциональных систем // Принципы системной организации функций. М.: Наука, 1973. С. 5-61.
53. Анохин К.В., Бурцев М.С., Зарайская И.Ю., Лукашев А.О., Редько В.Г. Проект «Мозг анимата»: разработка модели адаптивного поведения на основе теории функциональных систем // Восьмая национальная конференция по искусственному интеллекту с международным участием. Труды конференции. М.: Физматлит, 2002. Т.2. С.781-789.
54. Red'ko V.G., Prokhorov D.V., Burtsev M.S. Theory of functional systems, adaptive critics and neural networks // International Joint Conference on Neural Networks, Budapest, 2004. PP. 1787-1792.
55. Red'ko V.G., Anokhin K.V., Burtsev M.S., Manolov A.I., Mosalov O.P., Nepomnyashchikh V.A., Prokhorov D.V. Project "Animat Brain": designing the animat control system on the basis of the functional systems theory // In Butz, M.V., Sigaud, O., Pezzulo, G., Baldassarre, G. (Eds.), Anticipatory Behavior in Adaptive Learning Systems: From Brains to Individual and Social Behavior. LNAI 4520, Berlin Heidelberg: Springer Verlag, 2007. PP. 94-107.
56. Станкевич Л.А. Нейрологические средства систем управления интеллектуальных роботов // VI Всероссийская научно-техническая конференция «Нейроинформатика-2004». Лекции по нейроинформатике. Часть 2. М.: МИФИ, 2004. С. 57-110.
57. Witkowski M. An action-selection calculus // Adaptive Behavior. 2007. Vol. 15. No. 1. PP. 73-97.
58. Вайнцвайг М.Н., Полякова М.П. О моделировании мышления // От моделей поведения к искусственному интеллекту. Серия «Науки об искусственном» (под ред. Редько В.Г.). М.: УРСС, 2006. С. 280-286.

Статья поступила 23 августа 2007 г.

# Применение вейвлет-анализа для исследования импульсной активности нейронов головного мозга человека

Седов А.С., Раева С.Н.

Институт химической физики им. Н.Н. Семенова РАН, Москва

[sedovas@pochta.ru](mailto:sedovas@pochta.ru), [raeva@photonics.ru](mailto:raeva@photonics.ru)

**Аннотация.** Исследуется возможность использования метода вейвлет-анализа для изучения импульсной активности нейронов головного мозга человека. Предложен подход, позволяющий изучать с помощью вейвлет-преобразований динамику межклеточных взаимодействий в виде быстропротекающих (0,5-1,5 сек) явлений локальной частотной и фазовой синхронизации (когерентности) со стабилизацией осцилляторной (3-5 Гц) активности. Показано, что применение этого подхода позволяет количественно исследовать явления локальной синхронизации и осцилляторной импульсной активности, возникающие в нейронах некоторых подкорковых структур мозга человека в функционально-значимые этапы реализации сознательной опосредованной речью деятельности. Показана эффективность использования вейвлет-анализа для исследования частотных характеристик различных паттернов импульсной активности, а также для устранения шума и выделения полезного сигнала из натуральных записей нейронной активности, регистрируемой с помощью микроэлектродной техники при проведении стереотаксических нейрохирургических операций у больных с дискинезиями.

## Введение

Изучение различных классов колебательных процессов является одной из актуальных проблем современной биологии и смежных с ней нейронаук. Эти процессы происходят на всех уровнях организации живой материи – от биохимических реакций до крупных экосистем. Особый интерес для понимания механизмов нервной деятельности человека в условиях нормы и патологии имеют осцилляторные синхронизированные электрические процессы, регистрируемые как на уровне ЭЭГ биопотенциалов в коре больших полушарий [1-9], так и импульсной активности нейронов в некоторых подкорковых структурах мозга [10-15]. Применение микроэлектродной техники с лечебно-диагностическими целями во время нейрохирургических операций сделало возможным прямое экспериментальное изучение этих явлений в клеточных элементах человеческого мозга и роли их в передаче информации при осуществлении целенаправленной опосредованной речью двигательной и умственной деятельности [12, 16-21].

Применение традиционно используемых в нейрофизиологии методов статистики и Фурье-анализа оказалось малоэффективным для исследования быстропротекающих импульсных реакций, так как не позволяло достоверно выявлять изменения частотных (до 10 Гц) характеристик, развивающихся на временных масштабах в доли секунды. Именно эти особенности динамики кратковременно протекающих (0,5-1,5 сек) нейронных перестроек обусловили

перспективность применения вейвлет-анализа. Основным полем применения вейвлет-преобразований является анализ нестационарных (во времени) или неоднородных (в пространстве) сигналов. Он имеет широкий спектр применения для исследования частотно-временных характеристик нестационарных процессов, для фильтрации сигналов, сжатия данных, распознавания образов и решения ряда других задач [22-29]. В электрофизиологии метод вейвлет-анализа нашел применение при исследовании частотной синхронизации ЭЭГ ритмов в различных корковых областях на коротких временных интервалах (до 100 мс) при поиске вербальных ассоциаций [7] и когнитивной деятельности человека [4,5], а также для выявления эпизодов кратковременной фазовой ЭЭГ синхронизации [8,9] и дискриминации нейронной активности [30]. Метод вейвлет-анализа нашел применение для исследования динамики нейронных реакций в мозге крыс при многократной стимуляции, а также для анализа динамики взаимодействия внутриклеточных процессов [31]. В литературе мы не нашли работ по использованию вейвлет-анализа для изучения явлений синхронизации и осцилляторной активности в нейронах головного мозга человека.

Настоящая работа посвящена описанию применения метода вейвлет-анализа для количественного изучения быстропротекающих явлений локальной синхронизации и осцилляторной импульсной активности, а также частотно-временных характеристик паттернов реакций нейронов головного мозга человека. Апробация разработанного метода была реализована на экспериментальном материале ранее полученных данных исследования динамики импульсных реакций в 144 нейронах неспецифического парафасцикулярного комплекса (ядра CM-Pf) таламуса человека при целенаправленной словесно направляемой деятельности [20,21,32].

### **Методика исследований**

Экспериментальные данные, используемые в работе, были получены коллективом Лаборатории клеточной нейрофизиологии человека ИХФ РАН в ходе стереотаксических нейрохирургических операций, проводимых с применением микроэлектродной техники в лечебных целях на больных спастической кривошеей в Институте нейрохирургии РАМН (хирурги Н.Я. Васин, М.Р. Меджидов). С помощью вольфрамовых микроэлектродов (диаметр кончика 1 мкм, сопротивление 1-5 МОм) внеклеточно отводилась импульсная активность нейронов ненаркотизированного мозга человека. Параллельно с нейрограммой (НГ) регистрировалась фонограмма (ФГ) речевых (или звуковых) сигналов, электроэнцефалограмма в стандартных отведениях (ЭЭГ), электромиограмма (ЭМГ) конечностей и шейных мышц. Исследуемые биоэлектрические процессы регистрировались на 14-ти канальном магнитографе ЕАМ-500, 8-канальном электроэнцефалографе фирмы "Медикор" и фоторегистраторе. Данные с магнитографа с помощью аналого-цифрового преобразователя вводились в ПК.

Использовался ранее разработанный методологический подход [12, 17-19, 21]. Он позволял исследовать на клеточном (НГ), корковом (ЭЭГ) и

периферическом (ЭМГ) уровнях динамику процессов, возникающих в разные функционально-значимые этапы осуществления целенаправленного словесно направляемого поведенческого акта. Анализировались изменения, возникающие при предъявлении двух основных видов стимулов: 1) функционально-значимых словесных стимулов, вызывающих активацию внимания и/или стимулов, побуждающих человека к выполнению какого-либо произвольного движения (типа “Левую руку...сожмите!” и др.); 2) функционально незначимых словесных (или звуковых) раздражителей, неинформативных для больного (например, “Кадр 36”, “Глубина 40”). Характеристики реакций нейронов и их латентные периоды (ЛП) анализировались в определенной последовательности – от момента предъявления стимула-команды (по данным регистрации ФГ), от момента начала реализации движения (по данным регистрации ЭМГ). Подробнее методические особенности проведения исследований описаны ранее [12, 17, 18, 20, 21].

Первичная обработка данных включала фильтрацию наводок и выделение активности отдельных клеток методами оконной дискриминации и дискриминации по форме сигнала. Дальнейшая обработка включала построение гистограмм текущей частоты спайков и межимпульсных интервалов, анализ авто- и кросскорреляционных функций и др. Динамику межклеточных взаимодействий исследовали с помощью методов корреляционного анализа путем построения кросскорреляционной матрицы и перистимульной диаграммы рассеяния по данным текущей частоты разрядов. Исследование динамики частотных характеристик импульсных процессов осуществляли с помощью вейвлет-анализа. Статистический анализ полученных данных, а также достоверность полученных результатов контролировали с помощью критерия Стьюдента с применением программы Statistica.

### **Описание метода вейвлет-анализа**

Для исследования импульсной активности нейронов мозга человека использовался метод дискретного и непрерывного вейвлет-анализа. В отличие от Фурье-анализа, в котором анализирующая функция (синус, косинус) покрывает всю временную ось, вейвлеты представляют собой быстротатахующие функции с компактным носителем (рис. 1). Двухпараметрическая анализирующая функция хорошо локализована как во временном, так и в частотном пространствах. При этом частота и время рассматриваются как независимые переменные. Это свойство вейвлетов позволяет выявлять не только определенные частотные характеристики сигналов, но и локализовать их пространственные и временные координаты. Основной принцип построения базиса вейвлет-анализа состоит в использовании масштабирующей функции и фазовых смещений. За счет масштабных преобразований этой функции вейвлеты способны выявить различие в частотных характеристиках сигнала, а путем смещения анализировать свойства сигнала в разных точках временного интервала [22, 24, 28, 29].

Для устранения шума из исходной записи нейронной активности использовалось дискретное вейвлет-преобразование. Исходная запись

нейрограммы раскладывалась на аппроксимирующие коэффициенты, которые представляют сглаженный сигнал, и детализирующие коэффициенты, описывающие колебания. Шумовая компонента больше отражается в детализирующих коэффициентах, поэтому именно они подвергались обработке. Процедура удаления шума состояла из трех шагов: 1) вычисление вейвлет-разложения исходного сигнала до уровня  $N$ ; 2) обнуление детализирующих коэффициентов ниже определенного порогового значения (пороговая обработка); 3) вейвлет-реконструкция, основанная на первоначальных аппроксимирующих коэффициентах уровня  $N$  и модифицированных детализирующих коэффициентах уровней от 1 до  $N$ . В качестве материнского вейвлета использовали Симлет, либо Койфлет с 4 и выше нулевыми моментами и разложением до 8 уровня ( $N=8$ ). Пороговая обработка осуществлялась с помощью метода жесткого многоуровневого трешолдинга. При жесткой пороговой обработке сохраняются все коэффициенты, большие или равные по абсолютной величине порога  $\tau$ , а меньшие коэффициенты обращаются в нуль. Оптимальное для шумоподавления нейрограммы значение порога  $\tau$  выбиралось по минимаксному критерию и изменялось от уровня к уровню.

Более подробно особенности применения вейвлет-анализа для фильтрации сигналов описаны в работах Смоленцева [29] и Laubach [30].

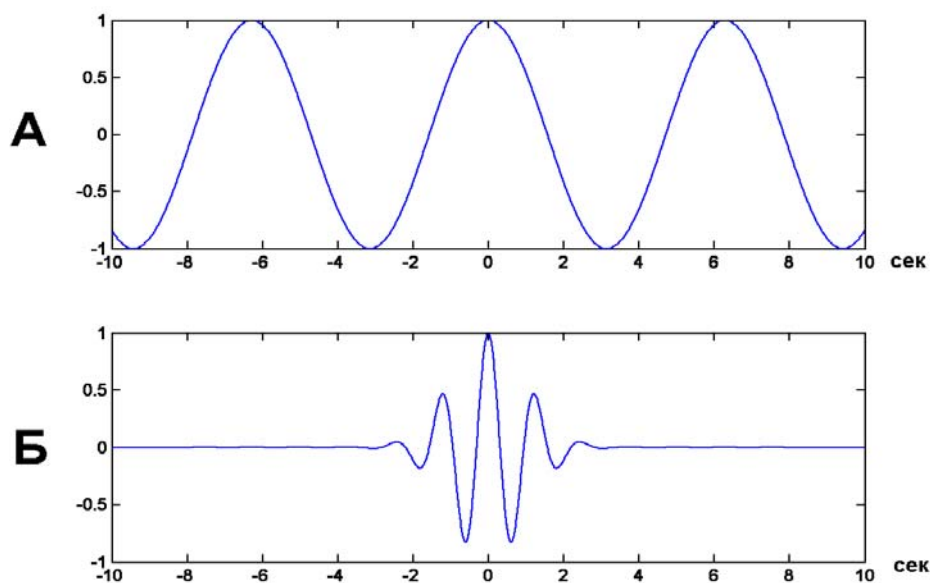


Рис 1. Система базисных функций для Фурье-анализа (А) и для вейвлет-анализа (на рисунке представлена действительная часть вейвлета Морле) (Б). По абсциссе – время, сек; по ординате – значение функции.



Для исследования динамики частотных характеристик импульсной активности нейронов использовалось непрерывное вейвлет-преобразование. Вейвлет-коэффициенты  $C_{a,b}$  сигнала  $x(t)$ , соответствующие масштабному коэффициенту  $a$  и положению  $b$ , определяются формулой:

$$C_{a,b} = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \overline{\psi\left(\frac{t-b}{a}\right)} dt. \quad (1)$$

Практически текущая частота разрядов  $x(t)$  задается выборкой значений  $\{x_k\}$ . Тогда в качестве  $x(t)$  выбирается кусочно-постоянная функция  $x(t) = x(k\Delta t) = x_k$  при  $t \in [k\Delta t, (k+1)\Delta t]$ , где  $\Delta t$  – бин текущей частоты нейронов. Параметр  $b$  пробегает значения с шагом  $\Delta t$ , масштабный параметр также меняется, пробегая дискретный ряд значений.

В качестве материнского вейвлета использовался вейвлет Морле, поскольку он хорошо локализован как во временной, так и в частотной области [23, 27, 29]. В общем случае вейвлет Морле комплекснозначный. В настоящей работе использовалась его действительная часть, определяемая следующим аналитическим выражением:

$$\psi(t) = \cos(k_0 t) \cdot e^{-t^2/2} \quad (k_0=6), \quad (2)$$

где  $\psi$  – вейвлет Морле,  $t$  – переменная (время),  $k_0$  – константа.

Значения текущей частоты разрядов  $x_k$  одиночных нейронов или ансамбля близлежащих клеток, сглаживались с помощью Гауссовского фильтра. Бин текущей частоты выбирался, исходя из особенностей паттерна импульсаций и исследуемых частотных характеристик (50-200 мс). Вычисленные по формулам (1) и (2) вейвлет-коэффициенты сглаживались с помощью фильтра скользящего среднего и брались по модулю.

Для исследования характеристик частотной и фазовой синхронизации активности нескольких близлежащих нейронов был разработан подход, основанный на вычислении матрицы произведения вейвлет-коэффициентов двух близлежащих нейронов. Он состоял из нескольких этапов:

- 1) вычисление вейвлет-коэффициентов ( $C_{a,b}(x)$ ,  $C_{a,b}(y)$ ) по формулам (1),(2) для текущей частоты ( $x_k$ ) одного и ( $y_k$ ) другого нейрона;
- 2) вычисление матрицы произведения вейвлет-коэффициентов

$$K_{a,b} = C_{a,b}(x) \cdot C_{a,b}(y); \quad (3)$$

- 3) нормировка и сглаживание полученной матрицы.

Согласно (3) матрица произведения вейвлет-коэффициентов  $K_{a,b}$  считается поэлементно: величина  $K_{a,b}$  есть произведение величин  $C_{a,b}(x)$  и  $C_{a,b}(y)$ . Полученная матрица позволяла выявить и локализовать моменты возникновения частотной (при  $K_{a,b} < 0$ ) и фазовой (при  $K_{a,b} > 0$ ) синхронизации, возникающей в ансамбле близлежащих нейронов.

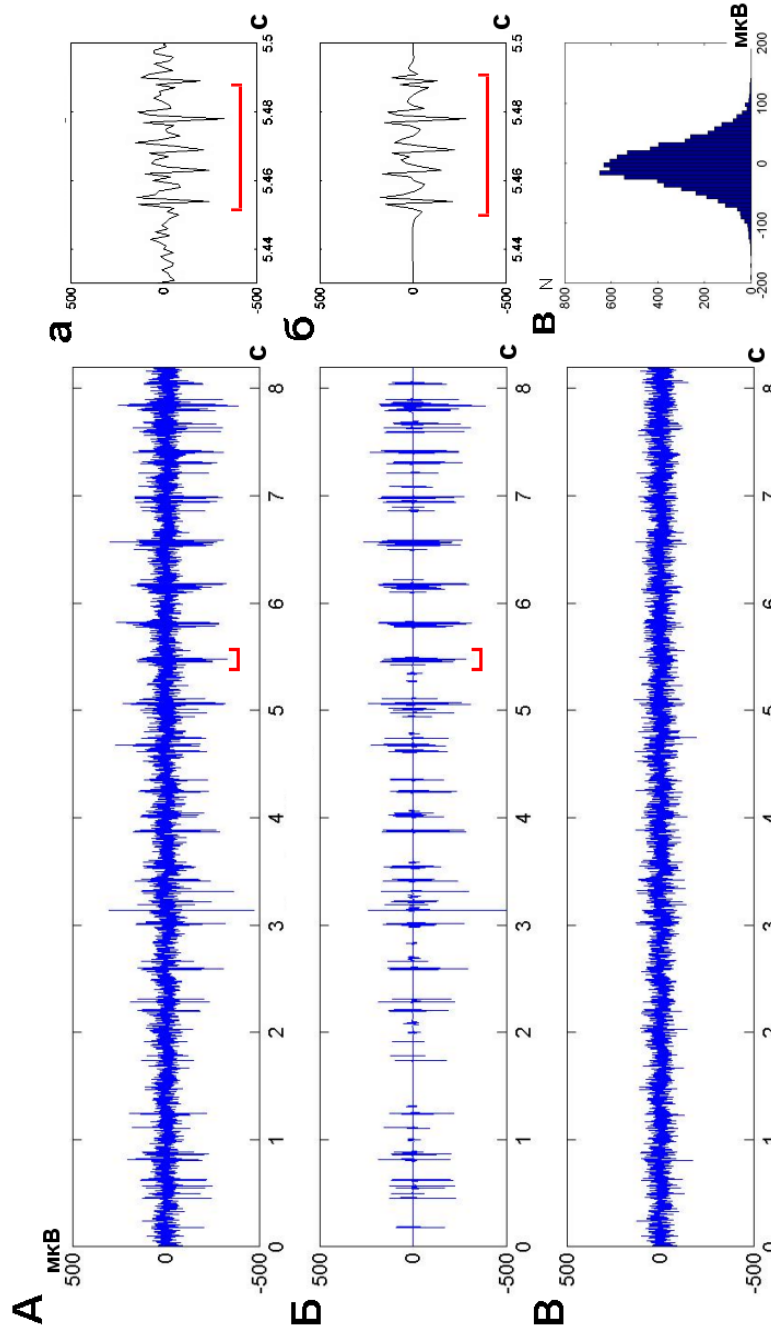


Рис 2. Результаты применения дискретного вейвлет-преобразования для устранения шума из нейрограммы. А – исходная запись нейрограммы, а – фрагмент нейрограммы на большой развертке, соответствующий отрезку записи, обозначенному сплошной линией, Б – отфильтрованная исходная запись нейрограммы, б – то же на большой развертке, В – устраненный шум, в – гистограмма распределения плотности шумового сигнала: по абсциссе – амплитуда мкВ, по ординате – число отсчетов (количество бинов – 50)

Разработанный подход и метод вейвлет-преобразований использовался в настоящей работе для исследования импульсной активности нейронов одного из важных подкорковых центров головного мозга человека – неспецифического парафасцикулярного комплекса (ядра СМ-Pf) таламуса. Апробация метода включала анализ активности 28 нейронов парафасцикулярного комплекса у 5 больных спастической кривошеей.

## Результаты исследований

Использование дискретного вейвлет-преобразования для анализа натуральной записи импульсной активности позволяло осуществлять фильтрацию нейронного шума без значимой потери полезной информации. Об этом свидетельствует гауссовское распределение шумового сигнала, а также сопоставление на большой развертке исходной и отфильтрованной записи нейронной активности (рис. 2). Фильтрация с применением дискретного вейвлет-анализа значительно упрощала дальнейший анализ – дискриминацию спайков и разделение ансамблевой активности на отдельные нейроны.

Использование непрерывного вейвлет-преобразования (Морле) по текущей частоте разрядов нескольких близлежащих СМ-Pf нейронов давало возможность выявлять на коротких интервалах времени (0.5-1.5 сек) появление ритмической (3-5 Гц) импульсной активности. Из рис. 3 видно, что предъявление функционально значимого словесного стимула-команды “Приготовились!”, мобилизующего селективное внимание человека, сопровождается возникновением в близлежащих нейронах неспецифического СМ-Pf комплекса таламуса быстротекающей локальной синхронизации и осцилляторной (3-5 Гц) активности. Сходные кратковременно протекающие явления выявляются также на высоте реализации произвольного движения в момент наибольшего напряжения мышц сгибателей пальцев руки больного.

Матрица вейвлет-коэффициентов ( $C_{a,b}$ ) исследуемого клеточного ансамбля количественно отражает моменты возникновения осцилляторной (5 Гц) активности нескольких элементов, составляющих клеточный ансамбль. Так, на 4-й секунде появление синхронизированной осцилляторной (5 Гц) активности коррелирует во времени с предъявлением релевантного вербального стимула “Приготовились!”, на 9-й секунде – с осуществлением произвольного движения человека.

Применение непрерывного вейвлет-преобразования (Морле) позволило также исследовать динамику возникновения локальной синхронизации и ритмической (3-5 Гц) активности в клеточных ансамблях при многократном повторении однотипных (двигательных или умственных) тестов. Как видно из рис. 4, выполнение двигательного теста (сжатие пальцев руки в кулак) сопровождалось торможением активности двух близлежащих пачечных ритмических (3,5 Гц) нейронов в СМ-Pf комплексе таламуса. Оба нейрона синергично реагировали тоническим торможением в момент предъявления функционально-значимой побудительной команды – “Сожмите руку в кулак!” с латентным периодом 100-200 мс длительностью 1,5 сек. Реакции этих нейронов опережали начало ЭМГ изменений на 0,8 сек. На высоте реализации движения и

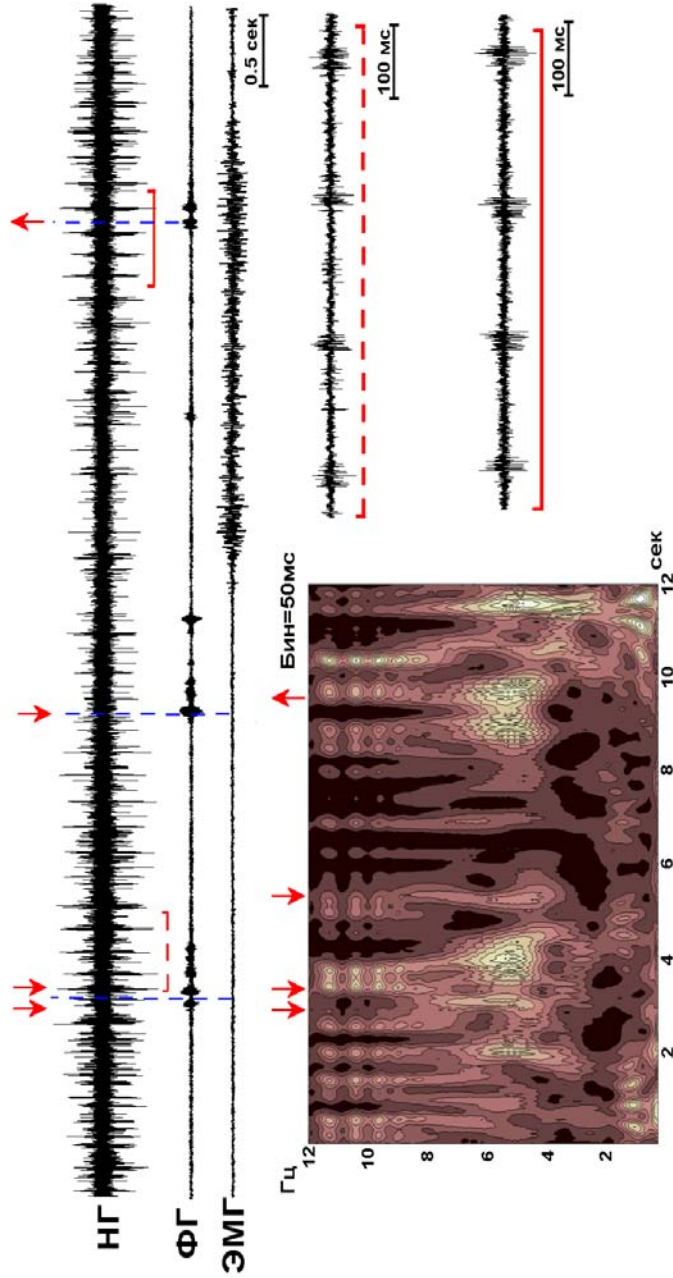


Рис 3. Применение вейвлет-анализа для выявления быстропротекающей (0.5-1.5 сек) синхронизированной и осциллирующей (5 Гц) активности в ансамбле близлежащих нейронов SM-Pf комплекса таламуса головного мозга человека. Последовательно представлены: НГ – запись нейрограммы, ФГ – фонограмма, ЭМГ – электромиограмма мышц сгибателей левой руки большого. Слева внизу – вейвлет-преобразование текущей частоты разрядов нейронов: по абсциссе – время, сек; по ординате – частота, Гц; цветом указано распределение вейвлет-коэффициентов (белый-максимум, черный-минимум). Справа – фрагмент нейрограммы на большой развертке, соответствующий интервалу времени, обозначенному пунктирной и сплошной линиями. Две стрелки вниз – момент предъявления команды “Приготовились!”, одна стрелка вниз – момент предъявления команды “Сожмите палец в кулак!”, стрелка вверх – “Разожмите!”

в период последствия эти же нейроны с пачечным ритмическим режимом работы реагировали синергичными активационными изменениями. Повторное предъявление двигательного теста приводило к уменьшению периода инициации произвольного движения по данным ЭМГ потенциалов (с 0,8 сек до 0,2 сек) и к уменьшению длительности реакции (с 1,5 сек до 0,9 сек), т.е. к ускорению ответной двигательной реакции человека.

Использование предложенного подхода для анализа межнейронного взаимодействия позволило выявить и локализовать моменты возникновения отрицательных значений матрицы произведения вейвлет-коэффициентов ( $K_{a,b} < 0$ ), которые появлялись на высоте реализации движения при первом (I) предъявлении двигательного теста. Это свидетельствует о возникновении кратковременной частотной синхронизации между клетками, разряжающимися в противофазе. Повторное предъявление двигательного теста приводило к изменению характера межнейронного взаимодействия. Так, при четвертом (IV) выполнении пробы в определенные моменты времени значения матрицы произведения вейвлет-коэффициентов принимают положительные значения ( $K_{a,b} > 0$ ), что свидетельствует о возникновении, наряду с частотной, также и фазовой синхронизации (когерентности) со стабилизацией ритма на частоте 3.5-4 Гц (рис. 4).

Анализ полученной матрицы ( $K_{a,b}$ ) выявляет особенности динамики межклеточных корреляций: укорочение латентного периода от момента возникновения синхронизации (на 1 сек,  $p < 0.01$ ), увеличение длительности локальной синхронизации с 1-1,5 сек до 2-2,5 сек ( $p < 0.01$ ) и усиление её выраженности на максимуме реализации движения, а также в период последствия от него.

В ранее проведенных исследованиях [20, 32] впервые было показано, что осуществление различных двигательных тестов, особенно с вовлечением пораженных мышц шеи, сопровождалось быстропротекающими перестройками или трансформацией исходной импульсной активности СМ-Pf нейронов у больных спастической кривошеей. Эти нейрональные перестройки в 88% случаев коррелировали с усилением тонического напряжения пораженных мышц шеи, что дало основание предположить об участии неспецифического парафасцикулярного комплекса таламуса в механизмах двигательной патологии спастической кривошеи.

На рис. 5 приведены данные, демонстрирующие возможность применения метода вейвлет-преобразований для количественного исследования частотной характеристики возникновения явления группирования разрядов в СМ-Pf нейронах, коррелирующего со значительным усилением тонического напряжения пораженных мышц шеи больного. Метод позволяет анализировать особенности нейронных перестроек, характеризующихся периодическими увеличениями частоты импульсаций (от 10-15 имп/с до 40-50 имп/с,  $p < 0.01$ ) с вариабельной (0,7-1,2 Гц) компонентой частоты следования групп. Эти частотные составляющие появлялись, как правило, на коротких промежутках времени и исчезали после завершения движения и/или уменьшения напряжения мышц шеи.

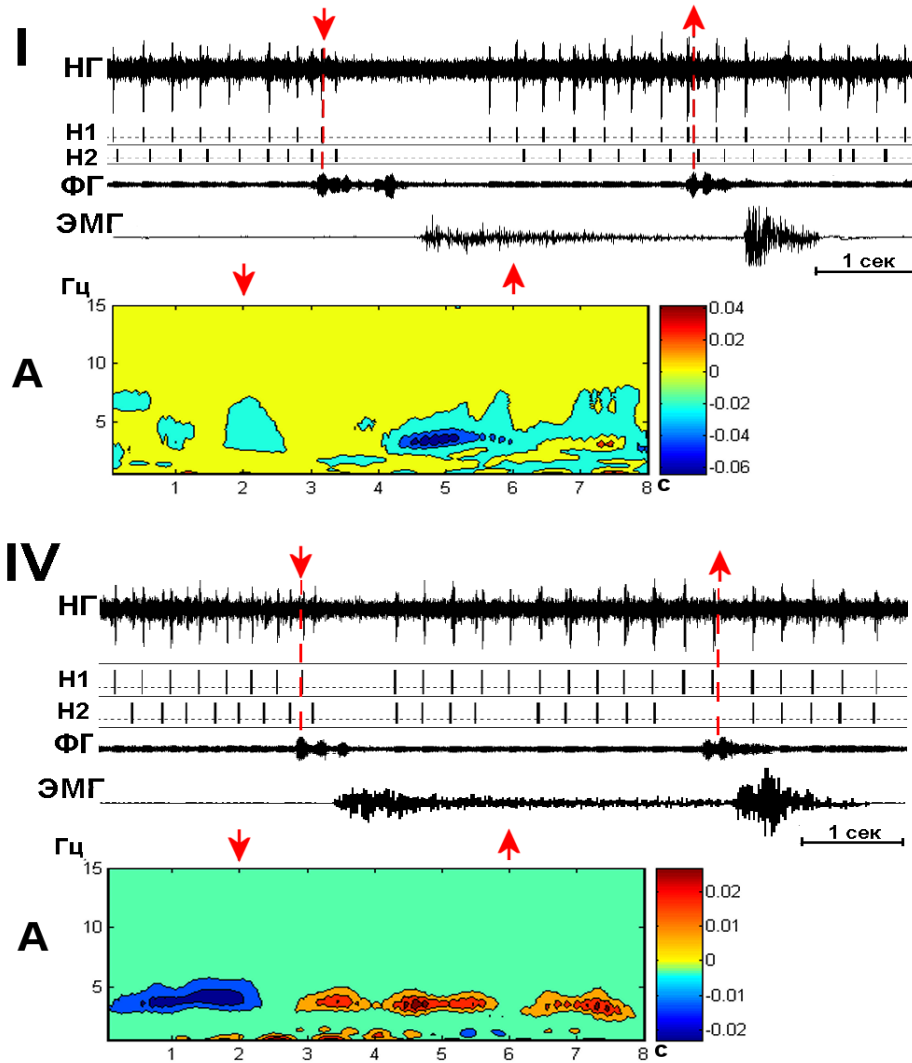


Рис 4. Динамика локальной синхронизации и осцилляторной активности СМ-Pf нейронов таламуса человека при предъявлении четырех (I-IV) однотипных двигательных тестов – сжимании пальцев руки в кулак.

Последовательно представлены: фрагмент нейрограммы (НГ), выделенные спайки двух нейронов (Н1, Н2), фонограмма (ФГ), электромиограмма (ЭМГ) сгибателей пальцев руки. А – нормированная матрица произведения вейвлет-коэффициентов для Н1 и Н2: по абсциссе – время, сек; по ординате – частота, Гц; Стрелки те же, что на рис 3.

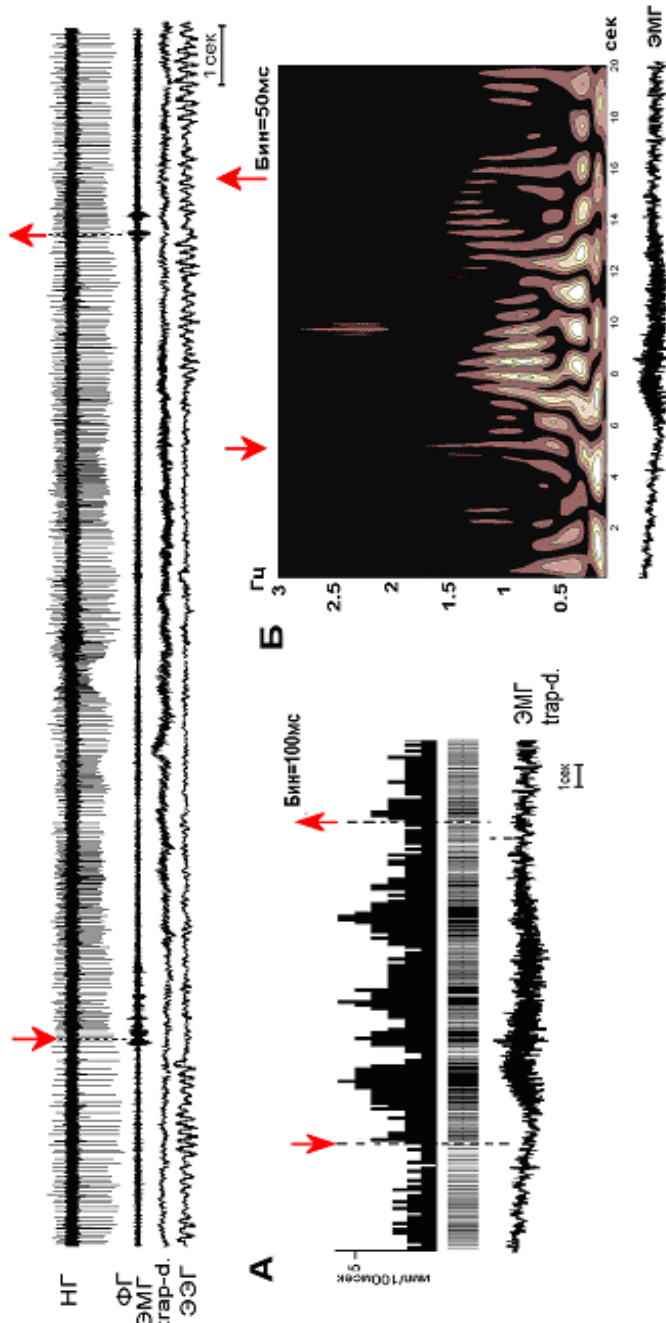


Рис 5. Динамика частотных характеристик нейронов при произвольном напряжении пораженных мышц шеи у больного спастической кривошеей.

Последовательно представлены: нейтрограмма (НГ) фонограмма (ФГ), электромиограмма (ЭМГ) мышц шеи, электроэнцефалограмма (ЭЭГ). А – текущая частота импульсных разрядов: по абсциссе – время, сек; по ординате – число импульсов за 100 мс; расстроенная диаграмма, ЭМГ мышц шеи, Б – вейвлет-преобразование текущей частоты разрядов нейронов: по абсциссе – время, сек; по ординате – частота, Гц; цветом указано распределение вейвлет-коэффициентов (белый-максимум, черный-минимум), ЭМГ мышц шеи. Стрелка вниз – момент предъявления команды “Напрягите шею!”, вверх – “Расслабьте!”.

## Обсуждение

В многолетних исследованиях М.Н. Ливанова с сотрудниками [1, 2] было показано важное значение динамики пространственной ЭЭГ синхронизации и осцилляторных (2-5 Гц) процессов в различных областях коры больших полушарий головного мозга при выработке условных рефлексов у животных, а также при различных функциональных проявлениях нормальной и патологической деятельности человека. Было установлено, что пространственная синхронизация в определенном частотном диапазоне ЭЭГ биопотенциалов и когерентность этих процессов имеет определяющее значение для проведения возбуждения между различными отделами коры больших полушарий, а также между корой и подкорковыми структурами. Позднее были получены новые данные, углубляющие понимание этих процессов как основы кортико-таламо-кортикального взаимодействия между различными структурами мозга в механизмах осознания, внимания, ассоциативного обучения и других когнитивных функций [3, 4, 7, 33, 34]. Согласно литературным данным в основе этих процессов лежит усиление синаптической взаимосвязи между нейронами, обусловленное усилением общего входа клеток [33, 35, 36]. В модельном эксперименте было показано важное значение параметров межнейронных коррелятов и связей между клеточными осцилляторами при селективном внимании, памяти и других когнитивных процессах [35, 37-41]. Учитывая кратковременность динамики протекания перестроек, развивающихся в мозге при обработке информации на временных масштабах в доли секунды, ряд авторов использовал аппарат вейвлет-анализа для количественного исследования этих процессов, в частности, явлений синхронизации и осцилляций [4, 7-9]. Так, в работе Николаева с соавт. [7] на основе предложенного метода корреляции вейвлет-коэффициентов были выявлены моменты синхронизации ритмов ЭЭГ на коротких временных интервалах при поиске словесных ассоциаций и принятии решения. Этот подход, позволяя исследовать временную и частотную синхронизацию, однако не давал возможность исследовать динамические параметры когерентности двух и более сигналов. В ряде других исследований метод вейвлет-преобразований использовался в модельном эксперименте для выявления кратковременной фазовой синхронизации ритмической активности нейронов при изменении коэффициента взаимосвязи между этими клетками [8,9].

Ранее нами [12, 17, 18, 21] было показано, что в функционально значимые этапы осуществления осознанной словесно-опосредованной двигательной и когнитивной деятельности в конвергентных нейронах неспецифического парафасцикулярного (СМ-Pf) комплекса и ряда других подкорковых структур мозга человека (Rt, MD, Cd) возникали быстропротекающие явления локальной синхронизации и ритмической (3-4 Гц) активности. Эти явления носили динамический кратковременный (0,5-1,5 сек) характер и коррелировали во времени и функционально с моментом фокусировки селективного внимания при предъявлении релевантного вербального (или сенсорного) стимула, а также с реализацией произвольного движения или психологического теста.

Предложенный в настоящей работе подход с использованием метода вейвлет-преобразований показал эффективность его применения для



количественного исследования динамики этих импульсных процессов, развивающихся при реализации сложных форм ВНД, связанных с речевым мышлением. Полученные данные с использованием этого подхода на экспериментальном нейронографическом материале позволили детально анализировать динамику частотной и фазовой быстропротекающей синхронизации и осцилляторной активности, возникающих в ансамбле близлежащих полимодальных нейронов в неспецифическом СМ-Pf комплексе таламуса человека при осуществлении целенаправленной опосредованной речью деятельности. Показано, что применение непрерывного вейвлет-разложения текущей частоты нескольких нейронов с построением матрицы произведения их вейвлет-коэффициентов позволяет детально исследовать динамику и тип синхронизации в клеточных ансамблях. Это дает возможность достоверно выявлять временную взаимосвязь межклеточных взаимодействий с активацией селективного внимания при предъявлении релевантного вербального стимула, с осуществлением произвольного движения, а также с проявлением двигательной патологии у больных спастической кривошеей.

Дискретное вейвлет-преобразование нашло широкое применение для фильтрации одномерных и двумерных сигналов, восстановления изображений, сжатия данных и во многих других задачах [24, 25, 28, 29]. В настоящей работе этот подход успешно использовался для удаления шума из натуральной записи нейрограммы. Учитывая, что импульсная активность отводилась экстраклеточно в условиях операционного блока в ходе стереотаксических нейрохирургических операций, активность нейронов регистрировалась на фоне шума. Использование фильтрации с помощью вейвлетов позволяло эффективно улучшать соотношение сигнал/шума без значимого искажения полезного сигнала. Это значительно облегчало дальнейший анализ и упрощало проведение дискриминации спайков от отдельных нейронов.

В целом, результаты настоящей работы свидетельствуют об эффективности и перспективности применения предложенного подхода на основе метода вейвлет-преобразований для количественного анализа динамики импульсных процессов головного мозга человека. На экспериментальном материале нейронографических записей обосновано его использование для выявления быстропротекающей локальной синхронизации и осцилляторной импульсной активности, динамики межнейронных взаимодействий, частотно-временных характеристик нейрональных реакций, фильтрации нейрограммы от шума с выделением полезного сигнала. Представленные материалы, подтверждая ранее полученные данные, углубляют понимание роли осцилляторных и синхронизированных импульсных процессов в обеспечении условий для передачи информации при целенаправленной деятельности, реализуемой человеком при участии речевых систем.

Авторы выражают глубокую благодарность сотруднику Лаборатории клеточной нейрофизиологии человека Института химической физики РАН Кудинову И.В., внесшему большой вклад в разработку метода вейвлет-анализа для исследования импульсных процессов головного мозга человека.

Работа выполнена при поддержке РФФИ (Грант №06-04-48251а).

## Литература

1. Ливанов М.Н. Нейронные механизмы памяти // Успехи физиол. наук., 1975, т.6 с. 66-89
2. Ливанов М.Н. Пространственная организация головного мозга // М., Наука, 1972, 462 с.
3. Gray C., Konig P., Engel A.K., Singer W. Oscillatory responses in cat visual cortex exhibit intercolumnar synchronization which reflects global stimulus properties // Nature, 1989, v 338(6213), pp. 334-337
4. Basar E., Demiralp T., Schürmann M., Basar-Eroglu C. and Ademoglu A. Oscillatory brain dynamics, wavelet analysis, and cognition // Brain and language, 1999, v. 66, pp. 146-183
5. Basar E., Schürmann M., Demiralp T., Basar-Eroglu C., Ademoglu A. Event-related oscillations are 'real brain responses' – wavelet analysis and new strategies // Int. J. Psychophysiol., 2001, v. 39, pp. 91-127
6. Stinivasan R., Russel D.P., Edelman G.M., Tononi G. Increased synchronization of neuromagnetic responses during conscious perception // J. Neurosci., 1999, v 19, p. 5435-5448
7. Николаев А.Р., Иваницкий Г.А., Иваницкий А.М. Исследование корковых взаимодействий в коротких интервалах времени при поиске вербальных ассоциаций // Журнал высшей нервной деятельности, 2000, т. 50, №1, с. 44-60
8. Varela F., Lachaux J-P., Rodriguez E. and Martinerie J. The brainweb: phase synchronization and large-scale integration // Nature reviews, 2001, v. 2, pp. 229-239
9. Quyen M.L.V., Foucher J., Lachaux J-P., Rodriguez E., Lutz A., Martinerie J., Varela F.J. Comparison of Hilbert transform and wavelet methods for the analysis of neuronal synchrony // J. Neurosci. Methods, 2001, v. 111, pp. 83-98
10. Albe-Fessard D., Arfel G., Guiot G., Hardy J., Vourch G., Hertzog E., Aleonard P., Derome P.: Derivations d'activites spontanees et evoquees dans les structures cerebrales profondes de l'homme // Rev. Neurol., 1962, pp. 89-105
11. Jasper H., Bertrand G.: Stereotaxic microelectrode studies of single thalamic cells and fibres in patients with dyskinesia // Trans. Am. Neurol. Assoc., 1964, v. 89, pp.79-82
12. Раева С.Н. Микроэлектродное исследование активности нейронов головного мозга человека // М.: Наука, 1977, 208 с.
13. Jeanmonod D., Magnin M., Morel A. Low-threshold calcium spike bursts in the human thalamus. Common physiopathology for sensory, motor and limbic positive symptoms // Brain, 1996, v. 119, pp. 363-375
14. Lenz F.A., Jaeger C.J., Seike M.S., Lin Y.C., Reich S.G., DeLong M.R., Vitek J.L.: Thalamic single neuron activity in patients with dystonia: dystonia-related activity and somatic sensory reorganization // J. Neurophysiol., 1999, v. 82, pp.2372-2392.
15. Magnin M., Morel A., Jeanmonod D. Single-unit analysis of the pallidum, thalamus and subthalamic nucleus in parkinsonian patients // Neuroscience 2000, v. 96, pp. 549-564.
16. Raeva S. Localization in human thalamus of units triggered during 'verbal commands,' voluntary movements and tremor // Electroencephalogr. Clin. Neurophysiol., 1986, v.63, pp.160-173.
17. Раева С.Н. Нейронные механизмы словесно направляемого движения // В кн.: Механизмы деятельности мозга человека. Ч.1: Нейрофизиология человека. Л.: Наука, 1988, с. 245-299
18. Раева С.Н., Лукашев А.О., Кадин А.Л., Васин Н., Шабалов В.А., Гроховский Н.П.: Динамика нейронных взаимодействий в ретикулярном ядре таламуса мозга человека при речевых стимулах разной сигнальной значимости // Нейрофизиология, 1990, 22:451-459

19. Raeva S., Lukashev A. Unit activity in human thalamic reticularis neurons. II. Activity evoked by significant and non-significant verbal or sensory stimuli // *Electroencephalogr. Clin. Neurophysiol.*, 1993, v. 86, pp. 110-122
20. Раева С.Н. Особенности фоновой активности нейронов парафасцикулярного комплекса (СМ-Pf) таламуса человека при изменении функционального состояния мозга // *Рос. физиол. журн. им. И.М. Сеченова.*, 2004, т. 90, с. 756-768.
21. Раева С.Н. Роль парафасцикулярного комплекса (СМ-Pf) таламуса человека в нейронных механизмах произвольного внимания // *Рос. физиол. журн. им. И.М. Сеченова.*, 2005, т. 91 с.225-238
22. Астафьева Н.М. Вейвлет-анализ: основы теории и примеры применения // *Успехи физических наук*, 1996, т. 166, №11, с. 1145-1170
23. Новиков Л.В. Основы вейвлет-анализа сигналов // *Учебное пособие. Санкт-Петербург 1999*, 153 с.
24. Дремин И.М., Иванов О.В., Нечитайло В.А. Вейвлеты и их использование // *Успехи физических наук*, 2001, т. 171, №5, с. 564-601
25. Переберин А.В. О систематизации вейвлет-преобразований // *Вычислительные методы и программирование*, 2001, т.2, с 15-40
26. Hasegawa H. Stochastic resonance of ensemble neurons for transient spike train: Wavelet analysis // *Physical review E.*, 2002, v. 66, p 021902
27. Shyu H.C. Construction of a Morlet wavelet power spectrum // *Multidimensional systems and signal processing*, 2002, v.13, pp.101-111
28. Liò P. Wavelets in bioinformatics and computational biology: state of art and perspectives // *Bioinformatics*, 2003, v. 19, pp. 2-9
29. Смоленцев Н.К. Основы теории вейвлетов. Вейвлеты в Matlab // *Москва*, 2005, 304 с.
30. Laubach M. Wavelet-based processing of neuronal spike trains prior to discriminant analysis // *J. Neurosci. Methods*, 2004, v. 134, pp. 159-168
31. Pavlov A.N., Makarov V.A., Mosekilde E. and Sosnovtseva O.V. Applications Of wavelet-based tools to study the dynamics of biological processes // *Briefings in bioinformatics*, 2006, v. 7(4), pp. 375-389
32. Раева С.Н., Седов А.С. Нейронные механизмы обработки моторного сигнала в парафасцикулярном комплексе (СМ-Pf) таламуса человека при произвольном движении в норме и патологии // *Рос. физиол. журн. им. И.М. Сеченова*, 2004, т. 90, с. 398-399
33. Steriade M. Corticothalamic resonance, states of vigilance and mentation // *Neuroscience*, 2000, v. 101, pp. 243-276
34. Киров В.Н., Белова Е.И. Механизмы формирования и роль осцилляторной активности нейронных популяций в системной деятельности мозга // *Журн. высш. нервн. деят.*, 2000, т. 50, с. 179-191
35. Абарбанель Г., Рабинович М., Сельверстон А., Баженов М., Хуэрта Р., Сущик М., Рубчинский Л. Синхронизация в нейронных ансамблях // *Успехи физических наук*, 1996, т. 166, с. 363-390
36. Contreras D., Steriade M. Synchronization of low-frequency rhythms in corticothalamic networks // *Neuroscience*, 1997, v.76, pp.11-24
37. Kryukov V.I. An attention model based on the principle of dominant // In: *Neurocomputers and attention. Vol. I: Neurobiology, synchronization and chaos. Manchester Univ. Press*, 1991, pp. 319-351
38. Borisyuk R.M. Interacting neuronal oscillators can imitate selective attention // In: *Neurocomputers and attention. Vol. I: Neurobiology, synchronization and chaos. Manchester Univ. Press*, 1991, pp. 189-200
39. Борисюк Г.Н., Борисюк Р.М., Казанович Я.Б., Иваницкий Г.Р. Модели динамики нейронной активности при обработке информации мозгом – итоги “десятилетия” // *Успехи физических наук*, 2002, т.172, с.1189-1214

А.С. Седов, С.Н. Раева

40. Borisyuk R.M., Kazanovich Y.B. Oscillatory neural network model of attention focus formation and control // *BioSystems*, 2003, v.71, pp. 29-38
41. Игумен Феофан (Крюков В.И.) Принципы сенсорной интеграции: иерархичность и синхронизация // *Журнал высшей нервной деятельности*, 2005, т.55, с. 163-169

Статья поступила 20 ноября 2006 г.  
После доработки 02 октября 2007 г.