

Адаптивный алгоритм глобальной оптимизации на основе взвешенного усреднения координат и нечетко-нейронных сетей

Кошур В.Д.

Красноярский государственный технический университет, koshur@fvt.krasn.ru

Аннотация. Проблема поиска глобального экстремума для негладких, многоэкстремальных целевых функций на ограниченном множестве в пространстве R^n до сих пор не решена в полной мере. Предлагается новая модификация адаптивного нечетко-нейронного алгоритма глобальной оптимизации. Основополагающая идея состоит в применении технологии построения нечетко-нейронных сетей для оценок инверсных регрессий.

Введение

Широкая область компьютерного моделирования прямо или косвенно связана с постановкой и решением задач оптимизации. При этом наибольший интерес представляет определение таких значений параметров моделируемой системы, которые соответствуют глобальному минимуму или максимуму заданной целевой функции.

В настоящее время наиболее распространенным и достаточно эффективным алгоритмом поиска глобального экстремума считается генетический алгоритм (ГА) [1-3] и его модификации (МГА) – гибридные ГА [4,5], адаптивные ГА [6-8]. В программный комплекс MatLab 7, вышедший в декабре 2004 года, включен комплект инструментальных средств на основе ГА и поисковой оптимизации. Дополнительным преимуществом ГА и МГА является возможность организации параллельных вычислений с реализацией задач большой размерности на многопроцессорных вычислительных комплексах и кластерных системах [3,9].

К достаточно универсальным алгоритмам глобальной оптимизации также относятся алгоритмы на основе непараметрических оценок инверсных регрессий [10-12], рандомизированного подхода [13,14] и близких по методологии методов непараметрической поисковой оптимизации [15,16] и взвешенного усреднения координат [17].

Проведённые исследования [18] вычислительной эффективности алгоритмов глобальной оптимизации на основе метода усреднения координат и МГА для набора тестовых примеров негладких многоэкстремальных целевых функций показали, что, при дополнительной эвристической подстройке параметров алгоритма метода усреднения координат, он является конкурентоспособным численным методом. В данной работе рассматривается модифицированный алгоритм поиска глобального экстремума на основе взвешенного усреднения координат, нечеткой логики [19-22] и нейронных сетей [5,23,24].

Идея построения и использования инверсных характеристик чрезвычайно плодотворна как при решении задач технической кибернетики [25,26], так и ре-

шении широкого класса задач вычислительной математики [27-29], даже в тех случаях, когда инверсные связи могут быть неоднозначными. Использование технологий мягких вычислений [30,31] позволяют придать подобным вычислительным методам элементы искусственного интеллекта.

Главная идея излагаемого подхода состоит в том, что сложная задача оптимизации сводится к серии существенно более простых задач минимизации, состоящих в настройке параметров системы нечетко-нейронных сетей с помощью стандартных методов минимизации квадратичной ошибки выходов сети [5,22-24] на заданном множестве пробных точек.

Поясним идею на примере минимизации функции одной переменной $I = f(t)$. Если на рассматриваемом интервале $[t_1, t_2]$ существует обратная функция $t = f^{-1}(I)$, то, вычисляя $t^{[k]} = f^{-1}(I^{[k]})$, при постепенном уменьшении значений целевой функции $I^{[k]}$ и, проверяя условие $t_1 \leq t^{[k]} \leq t_2$, на некотором шаге $k = N$ можно найти приближенное значение $t^{[N]}$ для аргумента целевой функции $t_{\min} = \arg \min f(t)$, $t \in [t_1, t_2]$.

Если обратной функции не существует, что, как правило, имеет место в приложениях, то, на k -том итерационном шаге по заданному множеству пробных точек $t^{(j)[k]}$, $j = 1, 2, \dots, M^{[k]}$ и соответствующих значений целевой функции $I^{(j)[k]} = f(t^{(j)[k]})$, всегда можно найти аппроксимирующую инверсную оценку, которая к обратной зависимости имеет лишь косвенное отношение, и реализовать её в виде нечетко-нейронной сети $FNN^{[k]}$ (Fuzzy-Neural Network) вида $t^{[k]} = FNN^{[k]}(I^{*[k]}, I^{(j)[k]}, t^{(j)[k]}, W^{[k]})$, где $I^{*[k]}$ – входной параметр сети, $W^{[k]}$ – массив настраиваемых параметров, $I^{(j)[k]}, t^{(j)[k]}$ – массивы задаваемых параметров сети. После обучения $FNN^{[k]}$ на заданном множестве значений целевой функции и соответствующих аргументов $\{I^{(j)[k]}, t^{(j)[k]}\}$, вычисляется улучшенное значение аргумента при последовательном уменьшении значений целевой функции: $I^{*[k]} = \min(I^{(1)}, I^{(2)}, \dots, I^{(M^{[k]})}) - \varepsilon^{[k]}$, где $\varepsilon^{[k]} \geq 0$, $\varepsilon^{[k]} \xrightarrow{k \rightarrow \infty} 0$. Для найденного аргумента $t^{[k]}$ вычисляется значение целевой функции $I^{pred[k]} = f(t^{[k]})$ и, если $I^{pred[k]} \leq I^{*[k]}$, то значение аргумента $t^{[k]}$ добавляется к списку пробных точек и осуществляется переход на следующий итерационный шаг $k + 1$, иначе производится уменьшение параметра $\varepsilon^{[k]}$.

1. Теоретические основы метода усреднения координат

1.1. Теорема сходимости [17]. Рассматривается ограниченная, непрерывная почти всюду функция $f(x) : \Omega \rightarrow R$, где $x = (x_1, x_2, \dots, x_n) \in \Omega \subset R^n$. Ограниченное множество Ω представляется как не более чем счетное объединение замкнутых, непересекающихся подобластей, в каждой из которых функция

$f(x)$ непрерывна и ограничена. Необходимо найти точку x_{\min} реализации глобального минимума

$$f(x_{\min}) = \min f(x), \quad x \in \Omega. \quad (1)$$

Пусть $\phi(x)$ – любая непрерывная, ограниченная функция. Вводится последовательность непрерывных функций $P_s(y)$, $s = 1, 2, \dots$, таких что $\forall y \in R$ значения $P_s(y) \geq 0$ и для последовательности вида $P_s(y)/P_s(z)$ выполняется условие монотонного неограниченного возрастания при увеличении номера s и любых фиксированных значениях y, z с условием $y < z$. Примерами функций $P_s(y)$, в частности, являются функции: $\exp(-sy)$, s^{-y} , y^{-s} , а также класс функций вида $(1 - y^r)^s$, для $r = 1, 2, \dots$ при $y \in [0; 1]$.

При выполнении данных условий доказано предельное интегральное соотношение [17]:

$$\phi(x_{\min}) = \lim_{s \rightarrow \infty} \int_{\Omega} \phi(x) p_s(x) dx \quad (2)$$

где $p_s(x)$ – последовательность нормированных ядер:

$$p_s(x) = P_s(f(x)) / \int_{\Omega} P_s(f(x)) dx. \quad (2a)$$

Если функцию $\phi(x)$ определить в виде $\phi_i(x) = x_i$ ($i = 1, 2, \dots, n$), то набор предельных соотношений (2) будет давать координаты точки $x_{\min} = (x_{1,\min}, x_{2,\min}, \dots, x_{n,\min})$ задачи (1). Этот результат может быть модифицирован и для задачи определения глобального максимума.

1.2. Вероятностно-статистическая интерпретация. Смысл предельного соотношения (2) состоит в том, что при фиксированном значении параметра s неотрицательная, нормированная в области Ω весовая функция $p_s(x)$ вида (2a) может рассматриваться как плотность распределения вероятности некоторой случайной величины. При этом значение плотности распределения возрастает при уменьшении значения функции $f(x)$. Тогда соотношение (2) представляет собой предел последовательности математических ожиданий от значений $\phi(x)$. При возрастании параметра селективности s , предел последовательности функций $p_s(x)$ стремится к n -мерной дельта-функции с локализацией в точке x_{\min} [17]. Следует отметить, что если целевая функция $f(x)$ имеет *не единственный* глобальный минимум в области Ω , то, даже если пределы в соотношениях (2)

существуют при $\phi_i(x) = x_i$ ($i = 1, 2, \dots, n$), получаемые предельные значения координат могут не совпадать ни с одним из значений x_{\min} задачи (1).

2. Алгоритм поисковой оптимизации и его модификации

2.1. Вычислительная процедура нахождения приближенного значения координат точки x_{\min} строится на основе организации итерационного процесса [17]

$$x_i^{[k+1]} = \int_{\Omega^{[k]}} x_i p_s^{[k]}(x) dx, \quad (i = 1, 2, \dots, n), \quad (3)$$

$$p_s^{[k]}(x) = P_s^{[k]}(f(x)) / \int_{\Omega^{[k]}} P_s^{[k]}(f(x)) dx. \quad (3a)$$

Здесь k - номер шага вычислительного процесса, $\Omega^{[k]}$ - область усреднения координат на шаге k . По мере роста параметра селективности s крутизна ядер $p_s^{[k]}(x)$ увеличивается, это создает трудности при вычислении интегралов в (3), (3a). Частично эта проблема может быть решена за счет принудительного уменьшения области усреднения. В тоже время следует заметить, что этот прием несет в себе риск потерять искомую точку глобального минимума. Поэтому, представляется, что более эффективным и надежным способом повышения точности вычисления интегралов является последовательное увеличение количества пробных точек $x^{(j)[k]}$, ($j = 1, 2, \dots, M^{[k]}$) на k -том шаге итерации, то есть $M^{[k]} \geq M^{[k-1]}$, при этом область усреднения следует оставлять неизменной. Рациональной модификацией численного алгоритма также является ранжирование (по значениям целевой функции) совокупности пробных точек, которые были получены на предыдущем и следующем итерационном шаге. Или, в простейшей реализации, осуществляется прямое добавление на текущем шаге набора новых пробных точек к имеющемуся набору сгенерированных случайных, равномерно распределенных точек в области Ω .

Для приближенного вычисления интегралов в (3), (3a) используется суммирование значений подынтегральных выражений на заданном наборе пробных точек. Координаты пробных точек генерируются датчиком случайных чисел с равномерным распределением на интервалах $[x_i^{[0]} - d_i, x_i^{[0]} + d_i]$, ($i = 1, 2, \dots, n$). Прямое произведение координатных интервалов образует n -мерную прямоугольную область Π такую, что $\Omega \subseteq \Pi$, с центром в точке $x^{[0]} = (x_1^{[0]}, x_2^{[0]}, \dots, x_n^{[0]})$. В набор пробных точек включаются только точки, которые попадают в область Ω . После аппроксимации соотношения (3), (3a) принимают вид

$$x_i^{[k+1]} = \sum_{j=1}^{M^{[k]}} x_i^{(j)[k]} p_s^{[k]}(x^{(j)[k]}) V^{(j)[k]}, \quad (i = 1, 2, \dots, n), \quad (4)$$

$$p_s^{[k]}(x^{(j)[k]}) = \frac{P_s^{[k]}(g^{[k]}(x^{(j)[k]}))}{\sum_{j=1}^{M^{[k]}} P_s^{[k]}(g^{[k]}(x^{(j)[k]})) V^{(j)[k]}}, \quad (4a)$$

где $V^{(j)[k]}$ соответствует малому n -мерному объему при разбиении области Ω на подобласти, связанные с семейством точек интегрирования $x^{(j)[k]}$, ($j = 1, 2, \dots, M^{[k]}$). Здесь $g^{[k]}(x)$ – вспомогательные функции, масштабирующие целевую функцию $f(x^{(j)[k]})$ к диапазону $[0, 1]$, см. ниже, формулы (6). Учитывая, что при генерации пробных точек используется равномерное случайное распределение, меру малых объемов $V^{(j)[k]}$, ($j = 1, 2, \dots, M^{[k]}$) приближенно можно полагать постоянной величиной, вынести за знаки сумм в выражениях (4), (4a) и сократить. Таким образом, расчетные формулы принимают следующий вид

$$x_i^{[k+1]} = \sum_{j=1}^{M^{[k]}} x_i^{(j)[k]} p_s^{[k]}(x^{(j)[k]}), \quad (i = 1, 2, \dots, n), \quad (5)$$

$$p_s^{[k]}(x^{(j)[k]}) = \frac{P_s^{[k]}(g^{[k]}(x^{(j)[k]}))}{\sum_{j=1}^{M^{[k]}} P_s^{[k]}(g^{[k]}(x^{(j)[k]}))}. \quad (5a)$$

Следует отметить, что для улучшения аппроксимации n -кратных интегралов в виде конечных сумм при генерации последовательностей равномерно распределенных пробных точек $x^{(j)[k]}$, ($j = 1, 2, \dots, M^{[k]}$) целесообразно использовать специализированные датчики псевдослучайных чисел [32].

Для приведения значений целевой функции $f(x^{(j)[k]})$ к диапазону $[0, 1]$ проводится масштабирование и вводятся вспомогательные функции $g^{[k]}(x)$ вида:

$$g^{[k]}(x^{(j)[k]}) = \frac{f(x^{(j)[k]}) - f_{\min}^{[k]}}{f_{\max}^{[k]} - f_{\min}^{[k]}}, \quad (6)$$

где обозначено

$$f_{\min}^{[k]} = \min(f(x^{(1)[k]}), f(x^{(2)[k]}), \dots, f(x^{(M^{[k]})[k]})), \quad (6a)$$

$$f_{\max}^{[k]} = \max(f(x^{(1)[k]}), f(x^{(2)[k]}), \dots, f(x^{(M^{[k]})[k]})) . \quad (6б)$$

Таким образом, соотношения (5), (5а), (6), (6а), (6б) и конкретизация вида ядер $P_s^{[k]}(y)$ полностью определяют вычислительный алгоритм глобальной оптимизации на основе взвешенного усреднения координат.

3. Адаптивный нечетко-нейронный алгоритм поиска глобального минимума

Методологической основой и, в определенной степени, прототипом для предлагаемой нечетко-нейронной адаптивной системы итерационного поиска координат точки x_{\min} глобального минимума задачи (1) является построение непараметрических оценок инверсных регрессий [10,11] и рандомизированный подход [13,14], которые предшествовали детальным исследованиям [15-17]. Взвешенное усреднение пробных координат $x^{(j)[k]}$, ($j = 1, 2, \dots, M^{[k]}$), полученных на k -том итерационном шаге, позволяет найти координаты $(x_1^{[k+1]}, x_2^{[k+1]}, \dots, x_n^{[k+1]})$ точки $x^{[k+1]}$, для которой значение целевой функции меньше, чем ранее найденные: $f(x^{[k+1]}) \approx f_{\min}^{[k]} - \varepsilon^{[k]}$, при этом оценки инверсных регрессий для каждой координаты могут быть построены в виде:

$$x_i^{[k+1]} = \sum_{j=1}^{M^{[k]}} x_i^{(j)[k]} E_i^{(j)[k]} \left(\frac{f^{*[k]} - f(x^{(j)[k]})}{\beta_i^{(j)[k]}} \right) + \Delta x_i^{[k]}, \quad (i = 1, 2, \dots, n), \quad (7)$$

$$f^{*[k]} = f_{\min}^{[k]} - \varepsilon^{[k]}, \quad \varepsilon^{[k]} \geq 0, \quad \varepsilon^{[k]} \xrightarrow{k \rightarrow \infty} 0 . \quad (7a)$$

Здесь $E_i^{(j)[k]}(y)$ – нормированные, симметричные относительно нуля, колоколообразные ядра, $\beta_i^{(j)[k]}$ – коэффициенты размытости ядер, $\Delta x_i^{[k]}$ – малая случайная добавка такая, что $\Delta x_i^{[k]} \xrightarrow{k \rightarrow \infty} 0$. Следует отметить, что приведенный вид соотношений (7) отличается от представленных в работах [10,11]. Отличие заключается в том, что в соотношении (7) допускается возможность независимой индивидуализации вида ядер и параметров размытости не только при переходе к следующей итерации, но и для каждой пробной точки и их координат. По структуре соотношения вида (5) и (7) практически совпадают и соответствуют взвешенному усреднению координат, хотя их методологическая основа различна.

Инверсными соотношениями вида (7) необходимо смоделировать закономерности, которые, как правило, носят неточный и, в ряде случаев, многозначный характер, поэтому целесообразно представить эти инверсные зависимости как нечеткие. Это позволяет придать инверсной регрессии более обобщенный смысл и существенно расширить возможности ее практического применения. Трудности при практической реализации «точных» аналитических соотношений вида (7) связаны с выбором вида ядер $E_i^{(j)[k]}(y)$, коэффициентов размытости $\beta_i^{(j)[k]}$ и параметров пробных движений $\Delta x_i^{[k]}$. При нейросетевой реализации этот выбор осуществляется стандартным образом одним из алгоритмов адаптивного обучения нечетко-нейронных сетей [5,22-24].

В качестве прототипа для построения нечетко-нейронной сети применяется метод Takagi-Sugeno [5,33]. Реализованная четырехслойная сеть состоит из слоя фазификации, слоя применения операторов Т-нормы, слоя нормализации и слоя дефазификации [5]. Для входных переменных z_1, z_2, \dots, z_m и действительной величины y , соответствующей выходному значению, определяется набор нечетких правил вывода в форме:

$$Rule_q : \text{if } z_1 \text{ is } A_{q1} \text{ and } z_2 \text{ is } A_{q2} \dots \text{and } z_m \text{ is } A_m \text{ then } y \text{ is } y_q, \quad (8)$$

где $q = 1, 2, \dots, Q$ (на k -том итерационном шаге $Q = M^{[k]}$), A_{ql} определяют нечеткие множества для q -го правила и l -ой лингвистической переменной, соответствующей z_l ; y_q представляет действительное число – значение требуемого вывода для $Rule_q$. Система нечетко-нейронных сетей, каждая из которых по виду аналогична [34], реализует итерационное улучшение значений координат $x_i^{[k]} = y_i^{[k]}$, где

$$y_i^{[k]} = FNN_i^{[k]}(I^{*[k]}, f(x^{(j)[k]}), x^{(j)[k]}, W_i^{[k]}), \quad (i = 1, 2, \dots, n). \quad (9)$$

Здесь $W_i^{[k]}$ – вектор настраиваемых параметров сети $FNN_i^{[k]}$, входной переменной для семейства нечетко-нейронных сетей (9) является только значение целевой функции $I^{*[k]}$, которое определено ниже. Значения целевой функции $f(x^{(j)[k]})$ и координаты пробных точек $x^{(q)[k]}$ сетью не обрабатываются, а выступают в роли подмножества заданных параметров сети $FNN_i^{[k]}$.

В качестве функций принадлежности используются функции Гаусса:

$$\mu_{(i)(q)(l)}^{[k]}(I) = \exp\left(-\frac{(I - (f(x^{(q)[k]}) + a_{(i)(q)(l)}^{[k]}))^2}{2(b_{(i)(q)(l)}^{[k]})^2}\right), \quad (I = 1, 2, \dots, m), \quad (10)$$

с настраиваемыми параметрами $W_i^{[k]} = \{a_{(i)(q)(l)}^{[k]}\} \cup \{b_{(i)(q)(l)}^{[k]}\}$.

Для реализации логического оператора Т-нормы используется операция произведения:

$$u_q = \mu_{q1}\mu_{q2}\dots\mu_{qm} \quad (q = 1, 2, \dots, Q) \quad , \quad (11)$$

где μ_{qi} – значение функции принадлежности для входной переменной z_i относительно нечеткого множества A_{qi} . Формирование логического вывода по определенной системе нечетких правил (8) и дефаззификация, в форме «центра притяжения» [34], реализуется в виде

$$y_i^{[k]} = \frac{\sum_{q=1}^{M^{[k]}} x_i^{(q)[k]} u_q^{[k]}}{\sum_{q=1}^{M^{[k]}} u_q^{[k]}} \quad . \quad (12)$$

Следует отметить, что при $m = 1$ соотношение (12) соответствует инверсной регрессии (7) и аналогично по форме взвешенному усреднению координат пробных точек (5).

После обучения нечетко-нейронных сетей $FNN_i^{[k]}$, выходы $y_i^{[k]}$ соответствуют улучшенным значениям координат $x_i^{[k]}$, если на вход сетей подается значение $I^{*[k]} = f^{*[k]}$, которое определяется в виде (7а). Обучение проводится на множестве значений $I^{(j)[k]} = f(x^{(j)[k]})$, при которых семейство сетей $FNN_i^{[k]}$ вида (9) должны выдавать приближенные значения соответствующих координат пробных точек $x_i^{(j)[k]}$.

При численной реализации адаптивного нечетко-нейронного алгоритма глобальной оптимизации целесообразно проводить наращивание координатных сетей $FNN_i^{[k]}$ по числу нечетких правил на новом итерационном шаге, если число пробных точек $M^{[k]}$ на новом итерационном шаге включает как уже имеющийся список, так и новый набор пробных точек.

4. Результаты вычислительных экспериментов

С целью апробации разработанного адаптивного нечетко-нейронного алгоритма глобальной оптимизации была проведена серия вычислительных экспериментов. Рассматривались гладкие и негладкие целевые функции различного вида, заданные на ограниченных односвязных областях Ω в пространстве R^n , где $n = 2, 10, 100, 1000$. Учитывая сложность визуализации графических образов в многомерных пространствах, ниже приводятся наглядные графические представления численных результатов решения задачи минимизации (1) для много-

экстремальных целевых функций *двух переменных*, заданных в прямоугольной области $\Pi \subset R^2$.

4.1. Гладкая многоэкстремальная функция. Целевая функция задана аналитическим выражением:

$$f(x, y) = -\frac{1}{\frac{x_1^2 + x_2^2}{200} - \cos(x_1) \cos(x_2 / \sqrt{2}) + 2}. \quad (13)$$

Глобальный минимум, равный -1 , соответствует координатам $x_1 = x_2 = 0$.

На рис. 1 представлен процесс поиска глобального минимума на поверхности целевой функции (13). Шаги итерационного процесса на плоскости x_1, x_2 с изображением изолиний целевой функции приведены на рис. 1а.

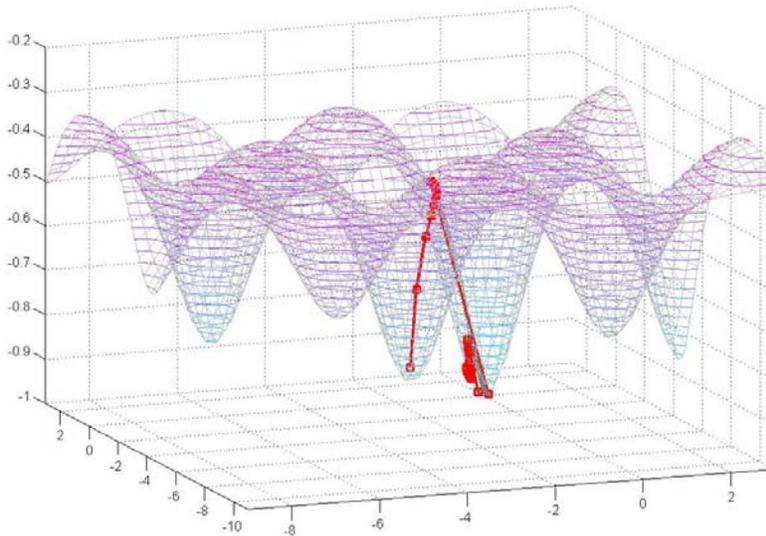


Рис. 1. Поверхность целевой функции (13) и изображение итерационных шагов поиска глобального минимума.

Гистограммы распределения значений целевой функции в пробных точках $x^{(q)[k]}$, ($q = 1, 2, \dots, 100$), использованных на итерационных шагах $k = 1, 2, \dots, 14$, показаны на рис. 1б. Гистограммы иллюстрируют постепенное выделение глобального экстремума в распределении значений целевой функции.

Для гистограмм горизонтальная ось соответствует приведенным к интервалу $[0, 1]$ значениям целевой функции, а вертикальная ось – значениям относи-

тельной частоты появления некоторого фиксированного значения целевой функции. Вычислительны эксперименты были проведены:

- для различного числа итераций ($k = 5, 10, 50, 100, 1000$),
- при различных стартовых условиях – начальных параметрах $W_i^{[k]} (i = 1, 2)$ нечетко-нейронной сети (9),
- для варианта наращивания числа нечетких правил (8) сетей $FNN_i^{[k]}$,
- при фиксации числа нечетких правил (8) и количества используемых пробных точек на каждом итерационном шаге, т.е. $Q = M^{[k]} = 100, \forall k$.

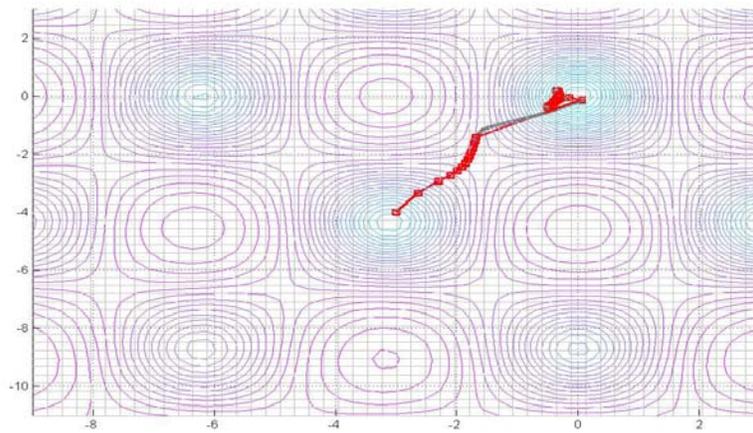


Рис. 1а. Изолинии поверхности целевой функции (13) и последовательные шаги поиска глобального минимума.

Расчеты показали стабильность достижения приближенных значений $x_{1\min}, x_{2\min}$ при работе адаптивно подстраиваемой нечетко-нейронной системы (9). При проведении данных расчетов, как правило, имела место *немонотонная* сходимость значений координат $x_1^{[k]}, x_2^{[k]}$ к координатам $x_{1\min}, x_{2\min}$ точки глобального минимума.

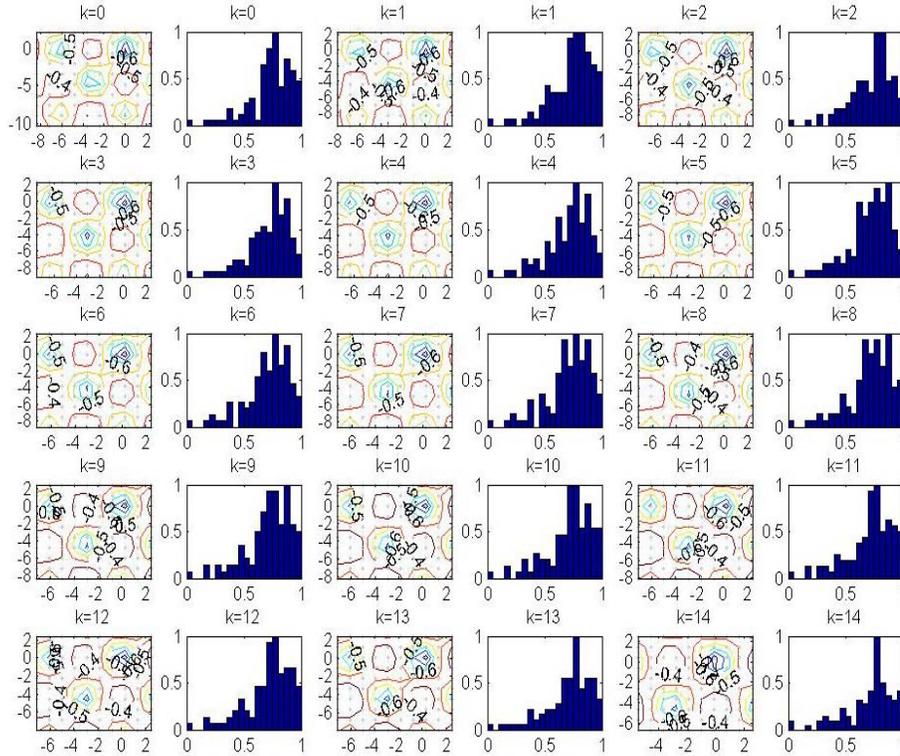


Рис. 16. Гистограммы распределения значений целевой функции (13) в пробных точках для первых 14 шагов итерационного процесса и изолинии, построенные по информации о значениях $f(x^{(q)[k]})$.

4.2. Многоэкстремальная функция Растргина. Целевая функция Растргина задается аналитическим выражением следующего вида [17]:

$$f(x_1, x_2) = 20 + x_1^2 - 10 \cos(2\pi x_1) + x_2^2 - 10 \cos(2\pi x_2). \quad (14)$$

Результаты расчетов по нахождению глобального минимума для функции Растргина показаны на рис. 2, рис. 2а.

Общий характер вычислительного процесса аналогичен представленным результатам пункта 4.1.

Адаптивный алгоритм глобальной оптимизации

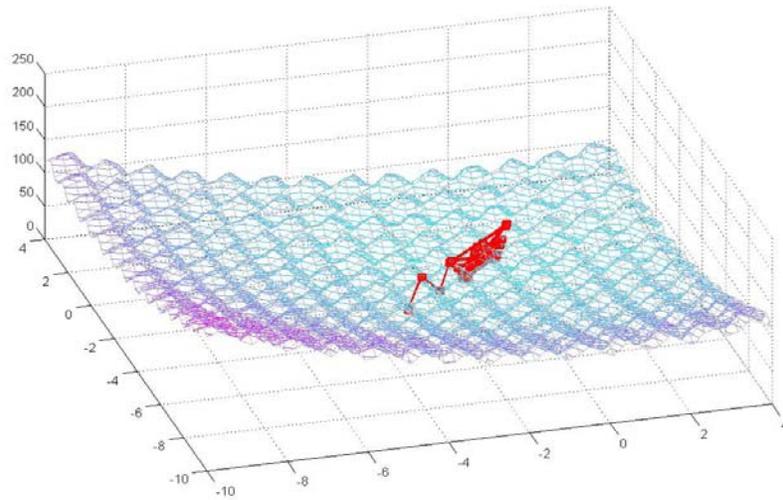


Рис. 2. Поверхность целевой функции (14) и изображение итерационных шагов поиска глобального минимума.

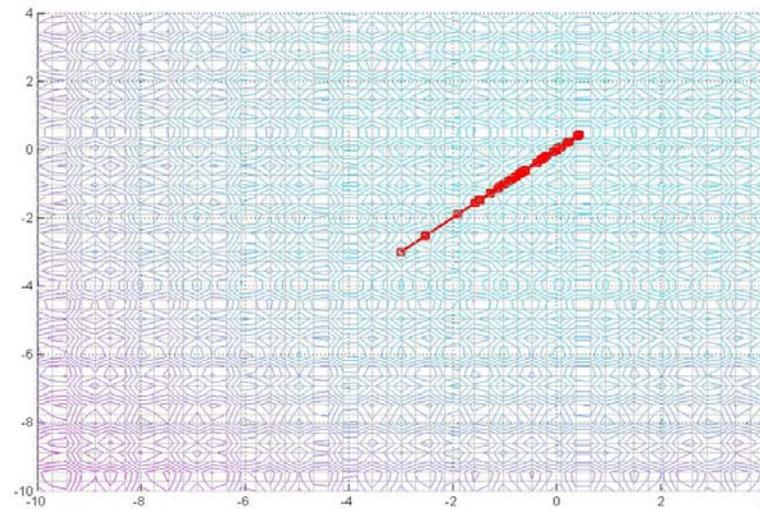


Рис. 2а. Изолинии поверхности целевой функции (14) и последовательные шаги поиска глобального минимума.

4.3. Негладкая 4-х экстремальная функция. Целевая функция задается выражениями вида [17]:

$$\begin{aligned}
 f1 &= -5\exp(-3((\text{abs}(x_1 + 1))^{0.6} + (\text{abs}(x_2 + 1))^{0.6})); \\
 f2 &= -10\exp(-2(\text{abs}(x_1) + \text{abs}(x_2))); \\
 f3 &= -7\exp(-2.5((\text{abs}(x_1 - 1))^{0.8} + (\text{abs}(x_2 - 1))^{0.8})); \\
 f4 &= -3\exp(-((\text{abs}(x_1 - 2))^{0.9} + (\text{abs}(x_2 - 2))^{0.9})); \\
 f(x_1, x_2) &= f1 + f2 + f3 + f4.
 \end{aligned}
 \tag{15}$$

Результаты расчетов по нахождению глобального минимума для 4-х экстремальной функции представлены на рис. 3, рис. 3а.

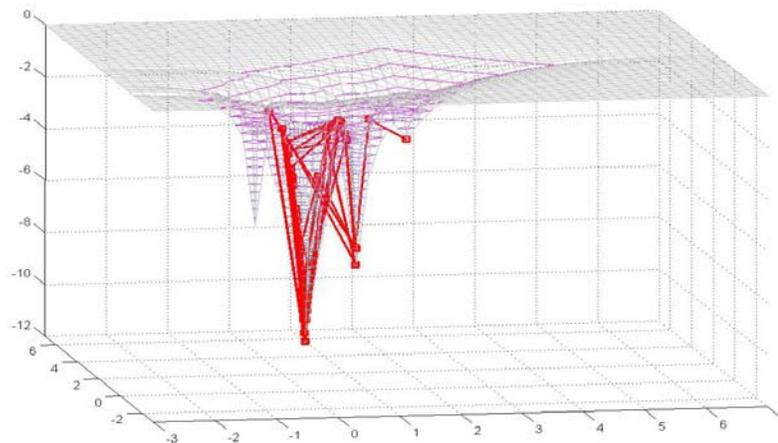


Рис. 3. Поверхность целевой функции (15) и изображение итерационных шагов поиска глобального минимума.

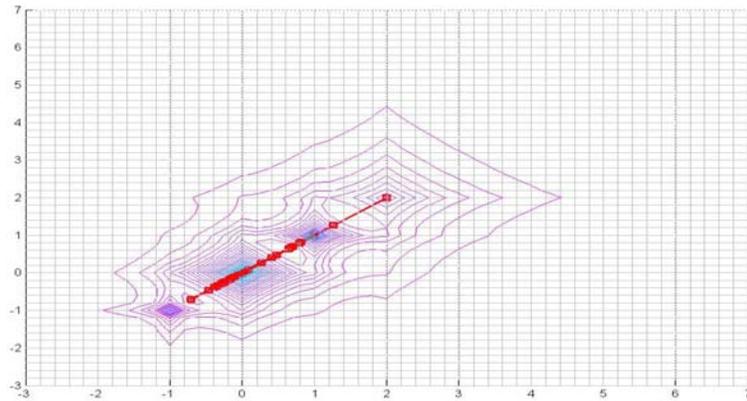


Рис. 3а. Изолинии поверхности целевой функции (15) и последовательные шаги поиска глобального минимума.

4.4. Негладкая 4-х экстремальная функция, имеющая три равных по значению глобальных минимума. Данная функция определяется модифицированными соотношениями (15) в виде:

$$\begin{aligned}
 g1 &= -7\exp(-3((\text{abs}(x_1 + 1))^{0.6} + (\text{abs}(x_2 + 1))^{0.6})); \\
 g2 &= -7\exp(-2(\text{abs}(x_1) + \text{abs}(x_2))); \\
 g3 &= -7\exp(-2.5((\text{abs}(x_1 - 1))^{0.8} + (\text{abs}(x_2 - 1))^{0.8})); \\
 f4 &= -3\exp(-((\text{abs}(x_1 - 2))^{0.9} + (\text{abs}(x_2 - 2))^{0.9})); \\
 f(x_1, x_2) &= g1 + g2 + g3 + f4;
 \end{aligned}
 \tag{16}$$

Результаты расчетов по нахождению глобального минимума для 4-х экстремальной функции с тремя одинаковыми по значению функции точками (1,1), (0,0), (-1,-1) представлены на рис. 4, рис. 4а.

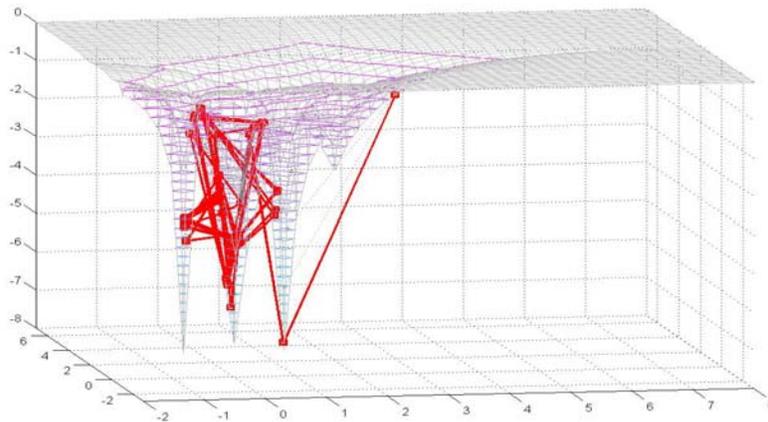


Рис. 4. Поверхность целевой функции (16) и изображение расходящегося процесса поиска неединственного глобального минимума.

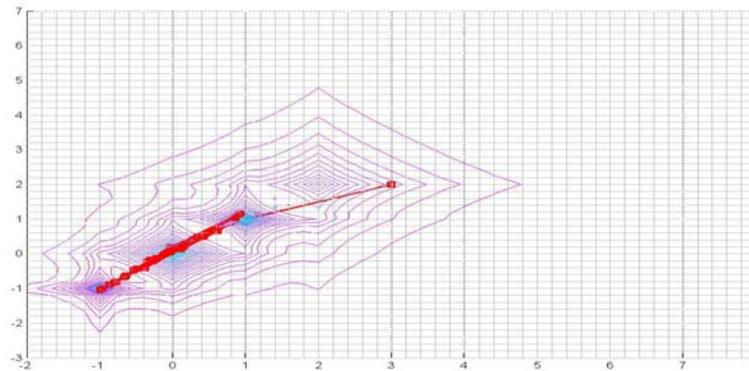


Рис. 4а. Изолинии поверхности целевой функции (16) и последовательные шаги поиска глобального минимума.

Представленные результаты наглядно показывают, что для целевой функции с неединственным глобальным минимумом адаптивный нечетко-нейронных алгоритм (также как и метод взвешенного усреднения координат) не позволяет найти хотя бы один из глобальных минимумов в виду того, что итерационный процесс расходится.

Заключение

Проведенные вычислительные эксперименты свидетельствуют о высокой эффективности предложенного адаптивного нечетко-нейронного алгоритма нахождения глобального экстремума гладких и негладких целевых функций.

Дополнительным и достаточно весомым аргументом в пользу нечеткого представления инверсных характеристик является возможность проведения поиска глобального экстремума для процедурно вычисляемых целевых функций при наличии помех (или погрешностей) и не только аддитивного типа. Исследование этих вопросов, а также анализ применения данного подхода для решения задач многокритериальной оптимизации и решения негладких нелинейных систем уравнений в пространстве R^n предполагается провести в последующих работах.

Литература

1. Holland J.H. *Adaptation in Natural and Artificial Systems*. The University of Michigan Press, Ann Arbor, 1975.
2. Goldberg D.E. *Genetic Algorithms in Search, Optimization, and Machine Learning*. Addison-Wesley, 1989.
3. Курейчик В.М. Генетические алгоритмы. Состояние, Проблемы. Перспективы // Известия Академии наук. Теория и системы управления, 1999, № 1. С. 144-160.
4. Кошур В.Д., Ильин В.А. Нейронная сеть Хопфилда как кроссовер генетического алгоритма // Научная сессия МИФИ-2003. V Всероссийская научно-техническая конференция "Нейроинформатика-2003". Сборник научных трудов. Часть 1. М.: МИФИ, 2003. С. 92-100.
5. Рутковская Д., Пилинский М., Рутковский Л. Нейронные сети, генетические алгоритмы и нечеткие системы. М: Горячая линия – Телеком, 2004. 454 с.
6. Herrera F., Lozano M., Verdegey J.L. Dynamic and heuristic fuzzy connectives-based crossover operators for controlling the diversity and convergence of real coded genetic algorithms // *Int. Journal of Intelligent Systems*, 1996, N 11, pp. 1013-1041.
7. Herrera F., Lozano M. Adaptation of genetic algorithm parameters based on fuzzy logic controllers // *Genetic Algorithms and Soft Computing*, Physica-Verlag, 1996, pp. 95-129.
8. Smith J.L., Fogarty T.C. Operator and parameter adaptation in genetic algorithms // *Soft Computing*, 1997, N 1(2), pp. 81-87.
9. Vishnevsky M.A., Koshur V.D., Legalov A.I., Mirkes E.M. Protective laminar composites design optimisation using genetic algorithm and parallel processing // *Parallel Computing Technologies. Proceedings of 7-th International Conference, PaCT 2003, Nizhni Novgorod, Russia, September 2003*, Ed. by Malyshkin V., Springer, 2003, pp. 394-400.
10. Живоглядов В.П., Медведев А.В. Непараметрические алгоритмы адаптации. Фрунзе: Илим, 1974. 134с.
11. Медведев А.В. Непараметрические системы адаптации. Новосибирск: Наука, 1983. 174 с.
12. Рубан А.И. Метод непараметрической оптимизации стохастических объектов. // *Системы управления: Сборник научных работ. Вып.1. Томск: Изд-во Томского государственного университета, 1975. С. 101-107.*
13. Алексеев В.И., Кориков А.М., Полонников Р.И., Тарасенко В.П. Экстремальная радионавигация. М.: Наука, 1978. 280 с.
14. Алексеев В.И. Субоптимальные рекуррентные алгоритмы оценивания в системах навигации // *Изв. вузов СССР. Радиоэлектроника*. 1987, Том 30, № 3. С. 34-39.

15. Рубан А.И. Метод непараметрической поисковой оптимизации // Изв. вузов. Физика. 1995, Том 38, № 9. С. 65-73.
16. Рубан А.И. Метод глобальной оптимизации функций в непрерывном пространстве // Информатика и системы управления: Межвузовский сборник научных трудов, посвященный памяти профессора Б.П. Соустина, Вып. 8. Красноярск: НИИ ИПУ, 2002. С. 161-173.
17. Рубан А.И. Глобальная оптимизация методом усреднения координат. Красноярск: ИПЦ КГТУ, 2004. 302 с.
18. Власов П.А. Глобальная оптимизация на основе генетического алгоритма и метода усреднения координат. Магистерская диссертация по направлению 552800 - информатика и вычислительная техника / Научный руководитель: д.ф.-м.н., проф. Кошур В.Д. Красноярск, КГТУ, 2005. 68 с.
19. Zadeh L.A. Fuzzy sets // Information and Control, 1965, Vol. 8. pp. 338-353.
20. Заде Л.А. Понятие лингвистической переменной и его применение к принятию приближенных решений. М: Мир, 1976.
21. Nauck D. Neuro-fuzzy systems: review and prospects // Proc. Fifth European Congress on Intelligent Techniques and Soft Computing (EUFIT'97), Aachen, Sep. 8-11, 1997. pp. 1044-1053.
22. Леоненков А.В. Нечеткое моделирование в среде MATLAB и fuzzyTECH. СПб.: БХВ-Петербург, 2003. 736 с.
23. Круглов В.В., Борисов В.В. Искусственные нейронные сети. Теория и практика. – М: Горячая линия – Телеком, 2002. 382 с.
24. Осовский С. Нейронные сети для обработки информации. М: Финансы и статистика, 2002. 344 с.
25. Цыпкин Я.З. Адаптация и обучение в автоматических системах. М.: Наука, 1968. 400 с.
26. Цыпкин Я.З. Сглаженные рандомизированные функционалы и алгоритмы в теории адаптации и обучения // Автоматика и телемеханика, 1971, № 8. С. 29-50.
27. Ogawa T., Kosugi Y., Kanada H. Neural network based solution to inverse problems // Proceedings of 1998 IEEE World Congress on Computational Intelligence, IEEE, 1998, pp. 2411-2416.
28. Бессольцев Д.А., Кошур В.Д. Решение нелинейных уравнений с помощью инверсных нейронных сетей // Научная сессия МИФИ-2004. VI Всероссийская научно-техническая конференция “Нейроинформатика-2004”. Сборник научных трудов. Часть 2. М.: МИФИ, 2004. С. 45-49.
29. Кошур В.Д. Инверсный метод нейросетевых аппроксимаций для решения задач оптимизации и управления // Научная сессия МИФИ-2004. VI Всероссийская научно-техническая конференция “Нейроинформатика-2004”. Сборник научных трудов. Часть 1. М.: МИФИ, 2004. С. 217.
30. Linkens D.A., Nyongesa H.O. Learning system in intelligent control: an appraisal of fuzzy, neural and genetic algorithm control applications // IEE Proc. Control Theory and Applications. 1996, Vol. 143, N 4, pp. 367-386. [Рус. перевод: Обучающиеся системы в интеллектуальном управлении: сопоставление подходов, основанных на размытой логике, искусственных нейронных сетях и генетических алгоритмах. М.: ВИНТИ. Экспресс-информация “Астронавтика и ракетодинамика”. 1997, Вып. 8, реф. 15; 1997, Вып. 9, реф. 16.]
31. Кошур В.Д. Вычислительные технологии на основе мягких вычислений // Вычислительные технологии, 2005, Том 10, Спец. выпуск. С. 77-82.
32. Иванов В.М. Программная реализация алгоритма генерации ЛП_τ последовательностей // Информатика и процессы управления. Межвузовский сборник научных статей. Красноярск: КГТУ, 1995. С. 125-130.

33. Takagi T., Sugeno M. Derivation of fuzzy control rules from human operator's control action // Proceedings of the IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis, July 1983, pp. 55-60.
34. Godjevac J., Steele N. Neuro-fuzzy control of a mobile robot // Neurocomputing, 1999, N 28, pp. 127-143.

Статья поступила 23 января 2006 г.
После доработки 4 сентября 2006 г.

Быстродействующий алгоритм семантической классификации JPEG-изображений

Дорогов А.Ю., Курбанов Р.Г., Разин В.В.

Санкт-Петербургский государственный электротехнический университет
(СПбГЭТУ) "ЛЭТИ", dorogov@lens.spb.ru

Аннотация. В работе обсуждается алгоритм семантической классификации JPEG изображений и результаты его моделирования в программной среде МАТЛАБ. Предлагаемый алгоритм основан на трех основных идеях: 1) для классификации изображений используется спектральное признаковое пространство, формируемое стандартной процедурой блочного кодирования JPEG-формата, что позволяет производить классификацию без восстановления изображения; 2) семантика полного изображения является производной от семантики сегментов изображения, что позволяет реализовать экономную иерархическую процедуру классификации; 3) исключаются какие-либо априорные предположения о конфигурации семантического класса в пространстве признаков, классификация выполняется по достоверным прецедентам базы данных.

В контексте семантической классификации представлены новые алгоритмы адаптивной сегментации изображений, алгоритмы информативной оценки системы первичных признаков и формирования сложных вторичных признаков, алгоритмы нечеткой метрической классификации сегментов изображения, алгоритмы нечеткой иерархической классификации изображений по результатам сегментной классификации.

1. Введение и обзор ключевых работ

1.1. Постановка задачи. По статистическим оценкам в мультимедийных базах данных до 80 процентов изображений представлены в JPEG формате или производных от него (JFIF, SPIFF, JBIG, JPEG-EXIF, MPEG). Формат JPEG представляет собой один из лучших методов сжатия с потерями, в котором обобщен полувековой опыт исследований инженеров и ученых, работающих в компьютерной, телевизионной и других областях, связанных с человеческим зрением и компьютерной графикой. В мультимедиа и Интернет технологиях этот формат широко используется для хранения, обработки и передачи изображений по каналам связи.

Типичной задачей обработки мультимедиа изображений является их классификация. Среди различных видов классификаций по уровню значимости и уровню сложности выделяется задача семантической классификации зрительных образов [1,2], позволяющая получить содержательный для человека ответ на вопрос: «Что изображено на картинке?». Прогресс в распознавании семантики образов, оказывает непосредственное воздействие на развитие систем компьютерного зрения, робототехнических систем, поисковых систем Интернет технологий, специализированных баз данных для мультимедиа систем. Время, затраченное на классификацию образов, имеет решающее значение как для баз данных большого объема, так и для систем, работающих в реальном масштабе времени. Высокое быстродействие систем классификации может быть достиг-

нута за счет сокращения непродуктивных преобразований видеобраза и использования быстрых алгоритмов обработки данных.

1.2. Типичная схема семантической классификации. Анализ семантики является вершиной иерархической процедуры обработки изображений. В основании пирамиды лежат методы формирования первичной системы информативных признаков. Основное требование на данном этапе – обеспечить максимально-возможную инвариантность признаков к топологическим преобразованиям и высокое быстродействие в получении первичной информации. Первичные признаки представлены в пространстве высокой размерности и имеют значительные различия по уровню информативности, эти обстоятельства препятствуют их непосредственному использованию в задаче классификации.

На следующем уровне иерархии формируется система вторичных признаков с примерно одинаковым уровнем значимости. Исходной информацией при этом служит анализ накопленной базы данных, который производится всякий раз при добавлении новых данных. Главное требование к процедуре анализа состоит в том, чтобы подобрать оптимальную систему признаков и сократить размерность признакового пространства. В режиме обучения классифицирующая система, взаимодействуя с оператором, накапливает банк семантических понятий. Временные затраты на анализ базы данных на этом этапе не имеют решающего значения, поскольку не влияют на быстродействие системы в рабочем режиме.

На верхнем уровне пирамиды, параметрическими или непараметрическими методами решается задача семантической классификации образов. Ответ может быть многозначным, поэтому необходимо ранжировать полученные решения, используя ту или иную оценку уровня значимости.

1.3. Обзор существующих методов решения. Известные методы решения задачи семантической классификации изображения, сохраняя в целом типичную схему, имеют большое разнообразие в способах и стратегиях реализации этапов.

Проблема распознавания семантики изображения (в зарубежной литературе используется аббревиатура CBIR – Content-based image retrieval) имеет более чем 20-летнюю историю. Первые работы относятся к 80-м годам прошлого века [3]. Однако наиболее существенное развитие это направление получило в последние несколько лет. Многие из выполненных исследований посвящены развитию стратегий последовательного уточнения запроса и оптимизаций поисковых процедур для изображений в конкретных базах данных большого объема [3-7]. Хорошо известным продуктом является система анализа семантики MARS, разработанная в университете Illinois [4]. К другим средствам относятся системы анализа PicToSeek [6], DrawSearch [7] и Viper [8]. Общая стратегия в разработке таких систем семантического поиска состоит в создании нового запроса, который оптимизируется в процессе диалога с пользователем. Однако для эффективного использования эта стратегия требует сложной трансформации базы данных в лингвистическую модель с взвешенным набором весов для терминологических переменных. Кроме того, в некоторых системах (Viper [8]) генерации подобной модели приводит к очень большой размерности признакового пространства. Процедура рафинирования поискового запроса использует

следующие методы: многомерные индексные структуры [9], набор признаков, выделяемых из примеров, представляющих интерес для пользователя [10], и взвешенные средние для позитивных и негативных примеров [11]. Интерактивный диалог с пользователем в СВIR системах трактуется как парадигма супервизорного обучения, стимулирующая механизм человеческого восприятия образов. В качестве средства формирования весов используются супервизорные нейронные сети [12], методы вероятностной классификации [13] и взвешенные евклидовы расстояния [14] в многомерном пространстве.

В процессе диалога производится назначение весов терминологическим понятиям. В результате строится некоторая функция подобия. Однако разделяющая мощность такой функции сильно ограничена, поскольку из-за проблем вычислительной сложности для построения функций подобия используется, как правило, не более чем квадратичная форма. Нейросетевые модели хорошо подходят для решения данной задачи, но требуют наличия большого объема обучающих данных для каждого нового запроса.

Альтернативный подход не супервизорного обучения, основанный на использовании самоорганизующихся карт Кохонена, рассматривается в работе [15]. Достоинство этого метода в том, что в процессе обучения системы не требуется взаимодействия с человеком-оператором. Но вследствие ограниченности правил формирования карт это, с другой стороны, приводит к снижению разделяющей мощности системы распознавания. Кроме того, для восприятия результата классификации человеком необходимо выполнить трактовку автоматически вырабатываемых концептуальных понятий, что не всегда возможно.

Задача семантического распознавания пересекается с задачей создания искусственного разума [16,17]. Это направление развивается в нашей стране в Институте проблем управления РАН им. В.А. Трапезникова. Сущность подхода состоит в трансформации количественной информации любого вида к «текстовой форме» (точнее к номинальной шкале). В процессе не супервизорного обучения реализуется выделение повторяющихся текстовых цепочек, которые трактуются как концептуальные понятия. Главная проблема использования данного подхода связана с адекватной трансформацией изображения к «текстовому» образу.

Приведенный обзор показывает, что в настоящее время научно-техническое направление семантической классификации изображений активно развивается, но до полного решения еще далеко, каждый из известных методов имеет как достоинства, так и определенные недостатки. Наилучшие результаты следует ожидать при решении конкретных задач.

1.4. Конкретизация целей исследования. В нашей работе предлагается ограничиться задачей семантической классификации JPEG-изображений. Мотивы такого подхода обусловлены тем обстоятельством, что JPEG-сжатие основано на использовании априорных данных о свойствах человеческого зрения, поэтому можно ожидать, что кодированный образ будет представлен в признаковом пространстве с достаточно высокой степенью информативности.

В качестве первичной системы признаков предлагается использовать спектральные коэффициенты блочного кодирования. Размерность пространства первичных признаков существенно ограничена за счет используемой в JPEG схемы

адаптивного сжатия информации. Таким образом, снимается проблема формирования первичной системы признаков и устраняется необходимость их прореживания по уровню информативности, это обеспечивает принципиальную основу для построения быстрых процедур семантической классификации. Главной целью настоящего исследования является разработка методов и алгоритмов семантической классификации JPEG-изображений.

2. Идея исследования

В JPEG формате цветные изображения представляются в виде яркостной (Y) и двух цветоразностных компонент (Cb, Cr). Яркостная и цветоразностные компоненты образуют трехмерное цветовое пространство, подобное цветовому пространству, используемому в телевизионном вещании по системе СЕКАМ. Формулы преобразования цветового пространства RGB («Красный, Зеленый, Синий») к цветоразностному пространству $YCbCr$ приведены в рекомендации JFIF [21].

Алгоритм JPEG-сжатия включает в себя два этапа. На первом этапе выполняются разделение изображения на блоки размеров 8×8 пикселей, которые подвергаются двумерному ортогональному косинусному преобразованию. В результате получаются матрицы спектральных коэффициентов размером 8×8 . Было доказано, что косинусное преобразование является оптимальным для случайных процессов марковского типа. Случайные процессы этого класса в статистическом смысле хорошо аппроксимируют случайное множество изображений. Это означает, что, обеспечивая в среднем достаточно хорошее качество восстановленного изображения, можно ограничиться только несколькими спектральными коэффициентами и отбросить ряд высокочастотных компонент. В результате объем передаваемой информации значительно сокращается, причем оставшиеся коэффициенты несут наиболее существенную информацию об изображении.

Прореживание коэффициентов реализуется за счет процедуры нелинейного квантования при переходе к 8 (или 12) битному представлению данных. Кроме того, при кодировании учитывается также, что человеческий глаз обладает более низким пространственным разрешением для цветовых перепадов по сравнению с разрешением для перепадов яркости. Поэтому цветоразностные компоненты кодируются с вдвое меньшим пространственным разрешением, что в два раза сокращает объем информации для каждой компоненты.

На втором этапе производится сжатие данных без потерь за счет использования кода Хаффмана. В процедуре Хаффмана числа кодируются неравномерным самосинхронизирующимся бинарным кодом. Сжатие достигается за счет того, что статистически наиболее вероятные значения кодируются короткими кодовыми словами, а наименее вероятные - длинными. В стандарте JPEG предусмотрена передача кодирующих таблиц кода Хаффмана вместе с изображением, однако, в большинстве случаев используются рекомендуемые в стандарте [21] примеры кодирующих таблиц, полученные разработчиками стандарта при статистическом исследовании множества изображений. На выходе кодировщика Хаффмана формируется непрерывный битовый поток данных.

При восстановлении изображения, после этапа декодирования Хаффмана, восстанавливаются 8-ми (или 12-ти) битные коды переданных спектральных коэффициентов по яркостной и каждой цветоразностной компонентам. Эти коды можно использовать как первичные информативные признаки, которые уже не требуют какого-либо цензурирования по уровню значимости. Таким образом, для выполнения классификации изображения нет необходимости в его полном восстановлении, поскольку признаковое пространство уже на промежуточном этапе подготовлено процедурой JPEG-кодирования. Исключение этапа полного восстановления изображения обеспечивает значительную экономию времени при решении задачи семантической классификации. Общая схема кодирования/декодирования JPEG изображений на уровне блоков представлена на рис. 1. Вертикальная серая стрелка указывает позицию отбора первичных информативных признаков.

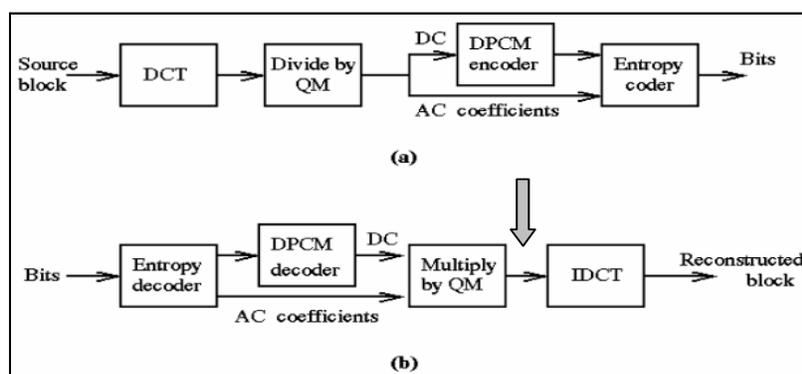


Рис. 1. а) Схема JPEG – кодирования, б) Схема декодирования JPEG-изображений. DCT, IDCT – прямое и инверсное косинусное преобразование, DC – нулевой спектральный коэффициент, AC – множество высокочастотных коэффициентов, QM – процедуры квантования/ деквантования спектральных коэффициентов, Entropy coder/decoder – кодирование и декодирование Хаффмана, Bits –битовый поток в канале связи.

В процедуре JPEG-кодирования для каждого блока изображения вычисляются нулевые спектральные коэффициенты (DC), которые несут информацию о среднем уровне яркостной и цветоразностных компонент. В предлагаемом подходе эта тройка коэффициентов используется для блочной пространственной сегментации изображения. Сегментация реализуется процедурой метрической кластеризации, с метрикой Хэмминга, межклассовое расстояние оценивается по принципу «максимально удаленных соседей». В зависимости от уровня сложности изображение может быть разделено на 1-7 пространственных сегментов. В каждом сегменте в свою очередь выделяются непрерывные компоненты, число которых не ограничивается. Сегменты описываются в 100-мерном признаковом пространстве (принципы формирования вторичных признаков представлены в разделе 3).

В процессе обучения в базе данных накапливаются 100-мерные вектора, каждый из которых соответствует одному сегменту изображения. При формиро-

вании базы данных оператор присваивает сегменту концептуальное понятие (*concept*), которое может быть снабжено определением (*modifier*). Тестовые комбинации *modifier/concept* определяют семантику сегмента. Таким же образом присваивается семантика всему изображению. В базе данных выявляется частотное отношение между семантикой изображения и семантикой сегментов. На основе энтропийной меры вторичные признаки в базе данных оцениваются по уровню информативности. Полученные оценки используются для агрегации вторичных признаков с целью повышения уровня их информативности. В результате на уровне сегментов образуются 20-25 сложных признаков (конкретное значение заведомо не определено и зависит от накопленной базы данных). Все сложные признаки имеют примерно одинаковый уровень информативности. Значение уровня информативности определяет степень доверия к сложному признаку и используется впоследствии для вычисления функций принадлежности к семантическому образу.

Вне зависимости от сегментного деления на уровне образа изображения формируются еще 18 признаков, которые образуют параллельную независимую систему информативных признаков. Эти признаки также адаптивно агрегируются в сложные признаки. В результате образуется 5-7 сложных признаков «уровня образа». Таким образом, семантика образа распознается по двум системам признаков.

На уровне образов это прямое распознавание, которое производится по системе сложных признаков «уровня образа» методами метрической классификации. Результаты накапливаются в базе данных, и их частотность определяет степень доверия к результатам распознавания.

На уровне сегментов распознавание семантики образа выполняется в два этапа. На первом этапе распознается семантика сегментов изображения подобно тому, как это делается для признаков «уровня образа». А на втором этапе на основе найденного ранее по базе данных статистического отношения между семантиками сегментов и образов определяются функции принадлежности для семантики образов.

Результаты по обеим системам признаков объединяются, образуя окончательный результат семантического распознавания. В целом система распознавания неявно реализует нечеткую нейронную сеть, в которой выходными лингвистическими переменными являются семантики образов.

3. Реализация алгоритмов и результаты экспериментов

Алгоритм декодирования JPEG-изображения реализован в соответствии с процедурами установленными рекомендациями T.81, T.83, T.84 [18-20] международного агентства в области коммуникаций ITU-T. Из допустимого множества профилей в рамках проведенного исследования в полном объеме реализован наиболее распространенный профиль *Baseline*, поддержанный рекомендацией JFIF [21]. Процедура сжатия данных основана на блочном кодировании изображений реализуемого на основе дискретного косинусного преобразования (DCT) над блоком размером 8×8 (см. рис. 2).

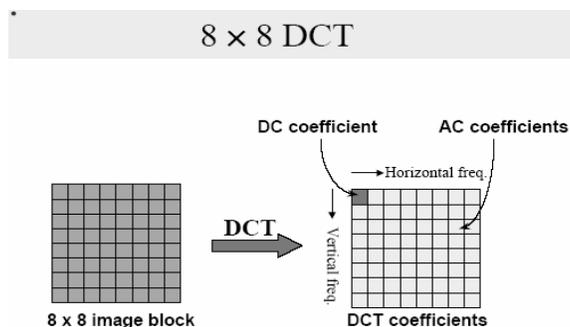


Рис. 2. Блочное кодирование. DCT – дискретное косинусное преобразование, DC – нулевой спектральный коэффициент, AC – множество высокочастотных коэффициентов.

3.1. Формирование первичных информативных признаков. На выходе декодирующей процедуры блок изображения представлен тремя матрицами размером 8×8 . Каждая матрица содержит косинусный двумерный спектр $C_k(i, j)$ ($k=1,2,3$) в цветовом пространстве $YCbCr$. Двумерное косинусное преобразование определяется выражениями:

$$C_k(i, j) = \frac{1}{4} \sum_{x=0}^7 \sum_{y=0}^7 f_k(x, y) \cos \frac{(2x+1)i \cdot \pi}{16} \cos \frac{(2y+1)j \cdot \pi}{16}, \quad i, j > 0,$$

$$C_k(0, 0) = \frac{1}{8} \sum_{x=0}^7 \sum_{y=0}^7 f_k(x, y), \quad i, j = 0,$$

где x, y координаты пикселей компоненты изображения f_k .

Из матриц косинусного спектра формируется четыре векторных признака:

- **DCf[3]** – цветовой фон блока – трехмерный вектор, составленный из значений постоянных составляющих (DC-составляющих) текущего блока.
 $DCf(k) = C_k(0, 0)$.

- **ACf[3]** – вариабельность цветности – трехмерный вектор, составленный из эффективных значений переменных составляющих цветности для трех компонент блока изображения:

$$ACf(k) = \sqrt{\sum_{i=1}^7 \sum_{j=1}^7 C_k(i, j)^2},$$

где $C_k(i, j)$ – спектральные коэффициенты блока для k -й компоненты блока изображения.

- **Cont[3]** – цветовой контраст – трехмерный вектор, координаты которого определяются правилом:

$$Cont(k) = \frac{1}{\sqrt{2}} \frac{ACf(k)}{DCf(k)}.$$

Если все пиксели блока изображения имеют одинаковую цветность, то векторные признаки ACf и $Cont$ равны нулевому вектору.

- **Angle[3]** – доминирующая угловая ориентация градиента цветности. Для вычисления признаковых координат для каждой компоненты вычисляются оценки градиента по трем направлениям:

$$\text{Горизонтальное направление } G_h(k) = \sqrt{\sum_{j=1}^7 C_k(0, j)^2}.$$

$$\text{Вертикальное направление } G_v(k) = \sqrt{\sum_{i=1}^7 C_k(i, 0)^2}.$$

$$\text{Наклонное направление (под углом } \pi/4) \text{ } G_d(k) = \sqrt{\sum_{i=1}^7 C_k(i, i)^2}.$$

Из трех оценок определяется максимальная, и координате вектора признака присваивается положительное значение угла, соответствующее максимальной оценке. Таким образом, для каждой координаты вектора признаков возможны три значения $0, \pi/4, \pi/2$.

3.2. Алгоритм сегментации изображения. DC-коэффициенты блоков изображения по яркостной и двум цветоразностным компонентам образуют множество векторов размерности 3. Множество кластеризуется метрическим методом, расстояние между векторами определяется выражением:

$$L = \sum_{i=1}^3 |x_i - y_i| / D_i,$$

где x_i, y_i – координаты векторов, D_i – диапазон вариаций значений по координате. Расстояние между классами оценивается по принципу «максимально-удаленных соседей».

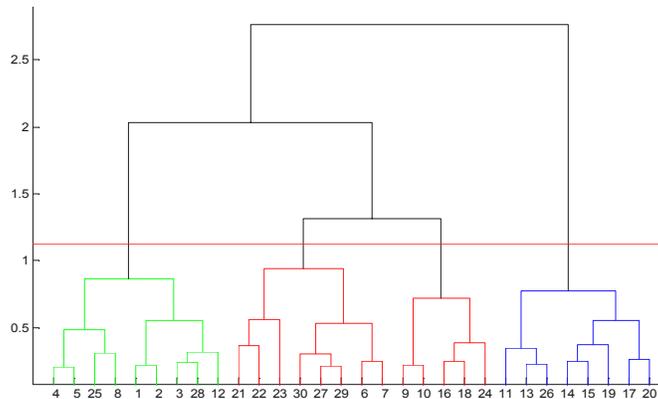


Рис. 3. Дендрограмма сегментации изображения. Горизонтальная линия соответствует адаптивному уровню разделения.

На рис.3 представлена примерная дендрограмма, где ветви дерева отражают кластерную иерархию множества из 30 точек. На горизонтальной оси дендрограммы представлено множество классифицируемых точек, по вертикальной оси отложены расстояния между точками в цветовом пространстве. Горизонтальные дуги объединяют близкие точки в кластеры. В МАТЛАБе для построения дендрограммы используется встроенная функция *dendrogram* из расширения *Statistic Toolbox*. Разделение на классы производится адаптивно, линия кластерного раздела на дендрограмме кластерных расстояний (см. рис. 3) проводится, когда отношение двух смежных уровней дендрограммы падает ниже 85%. По визуальным оценкам этот порог обеспечивает результаты, близкие к субъективным решениям человека.

На рис. 4 показан результат сегментации изображения («оранжевая роза»), соответствующий выше приведенной дендрограмме. Блоки 8×8 помечены точечными маркерами различного вида, каждый из которых отвечает одному из четырех образованных сегментов изображения.

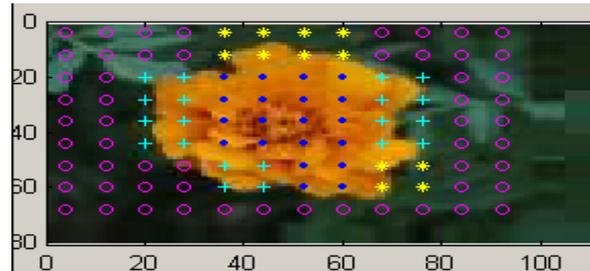


Рис. 4. Сегментированное изображение

3.3. Формирование вторичных информативных признаков. Система вторичных признаков строится на основе первичных информативных признаков. Характерная особенность вторичных признаков состоит в том, что все они имеют определенный физический смысл, и в той или иной мере отражают логическую информацию, которую человек использует при анализе семантики изображения.

3.3.1. Признаковое пространство сегмента. Каждый блок сегмента интерпретируется как точка в дискретной плоскости блоков. Координаты (r,c) точки отсчитываются от левой и верхней грани изображения и выражаются в числе блоков. Площадь сегмента (далее обозначается через S) определяется числом входящих в него точек плоскости блоков. Обозначим через R и C размеры изображения в блоках, по вертикали и горизонтали соответственно. Признаковое пространство для сегмента включает характеристики, представленные в табл. 1.

3.3.2. Признаковое пространство полигона. Известно, что при восприятии изображения человеческий глаз последовательно фиксирует взгляд на ряде характерных точек [22]. Точка фиксации взгляда – это область изображения, ми-

нимально достаточная для выделения человеческим глазом ее отличительных особенностей. При анализе JPEG-изображения естественно принять, что размер точки фиксации взгляда определяется размером одного блока изображения. Число точек фиксации взгляда зависит от уровня сложности изображения, однако в любом случае это число значительно меньше, чем число блоков изображения. Множество точек фиксации взгляда для изображения в целом или его фрагмента далее называется полигоном. При моделировании данной функции человеческого восприятия была выбрана упрощенная схема, закрепляющая число точек полигона в зависимости от представительности объекта. Принято, что полигон всего изображения состоит из 21 точки, полигон сегмента состоит из 14 точек, полигон непрерывной области сегмента состоит из 7 точек. На рис. 5 показаны точки полигона для яркостной компоненты сегмента граничных областей. Точки условно соединены ребрами графа.

Табл. 1. Информативные признаки сегмента

<p>Param.potential = $S/R \cdot C$ – потенциал сегмента равен относительной площади, занимаемой сегментом в поле изображения.</p> <p>Param.YCbCr_DC=CL(k) – массив из трех элементов; содержит яркостную и цветоразностные компоненты Y,Cb,Cr для доминирующего фона сегмента.</p> <p>Param.YCbCr_AC=aYCbCr – вариабельность цветности сегмента, трехмерный вектор, координаты которого равны среднеквадратичным значениям цветовой вариабельности блоков сегмента, для вычисления значений используется следующее правило:</p> $aYCbCr(i) = \frac{1}{S} \sqrt{\sum_s ACf(i)^2} ,$ <p>Param.LPosition=LPosition – структура, содержащая логические позиции сегмента в поле изображения.</p> <p>LPosition.top=1 – сегмент касается верхней границы изображения</p> <p>LPosition.bottom=1 – сегмент касается нижней границы изображения</p> <p>LPosition.left=1 – сегмент касается левой границы изображения</p> <p>LPosition.right=1 – сегмент касается правой границы изображения</p> <p>LPosition.center=1 – сегмент имеет блоки, принадлежащие области центра</p> <p>LPosition.quadrant1=1 – сегмент имеет блоки, принадлежащие первому квадранту</p> <p>LPosition.quadrant2=1 – сегмент имеет блоки, принадлежащие второму квадранту</p> <p>LPosition.quadrant3=1 – сегмент имеет блоки, принадлежащие третьему квадранту</p> <p>LPosition.quadrant4=1 – сегмент имеет блоки, принадлежащие четвертому квадранту</p> <p>Param.Narea – число изолированных областей в сегменте</p>

Координаты (x,y) точек полигона отсчитываются от левой и верхней грани изображения и выражаются в пикселях с точностью до размера блока. Точки полигона выделяются по значению модулей векторов либо цветовой вариабельности $|ACf|$, либо цветового контраста $|Cont|$, либо угловой ориентацией градиента. Выбор критерия в исследовательской моделирующей программе осуществляет оператор. При закреплённом числе точек полигона пороговый уровень выделения точек также является информативным признаком.

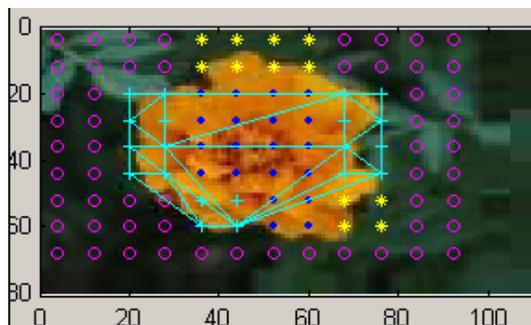


Рис. 5. Полигон сегмента граничных областей.

По полигону строится 12 векторных вторичных признаков, характеризующих его форму. Набор признаков, описывающих полигон, представлен в табл. 2.

Табл. 2. Информативные признаки полигона

<p>Polygon.Potential=$S/(X*Y)$ – матрица потенциалов полигонов Polygon.Diameter=$D/\sqrt{X^2+Y^2}$ – матрица относительных диаметров полигонов (относительно диагонали отображения) Polygon.FormFactor =S/D^2 – фактор формы Polygon.DAngle – углы ориентаций максимального диаметра для полигонов Polygon.EAngle=EAngle – углы ориентаций вектора главной компоненты полигонов. Polygon.Radius=Radius – максимальные относительные радиусы для полигонов Polygon.EFactor=EFactor – отношение минимального диаметра аппроксимирующего эллипса полигона к максимальному диаметру. Polygon.Mass_center=Mass_center./[X,Y] – матрица относительных координат центра масс точек полигона. Polygon.Polygon_center=Polygon_center./[X,Y] – матрица относительных координат геометрического центра полигона (координаты центра масс выпуклой оболочки полигона). Polygon.Level=Level – пороговые уровни выделения точек полигонов. Polygon.MaxValue – максимальные значения уровней выделения для точек полигона. Polygon.Points=Points – массив ячеек, содержащих координаты точек полигона Polygon.Graph – массив ячеек, содержащий матрицы смежностей графа Делоне для точек полигона</p>
<p>S – площадь полигона в пикселях X*Y – размеры изображения в пикселях D – длина максимальной диагонали изображения</p>

Каждое поле структуры *Polygon* имеет три компоненты. Координаты центра масс и центра масс выпуклой оболочки полигона представляют собой матрицы размером 3×2 . Все координаты представлены в относительных единицах, нормировка выполняется по отношению к фактическим размерам изображения. Линейные размеры нормируются по отношению к максимальной диагонали изображения. Следующие математические выражения используются для формирования признакового пространство полигона:

- *Относительный центр масс полигона* – определяется его координатами:

$$x_c = \frac{1}{X \cdot N} \sum_{i=1}^N x_i, \quad y_c = \frac{1}{Y \cdot N} \sum_{i=1}^N y_i$$

где X, Y размеры изображения, выраженные в пикселях, N – число точек полигона.

- *Относительный диаметр полигона* – равен отношению максимального диаметра полигона к диагонали изображения, где диагональ изображения вычисляется выражением $Diagonal = \sqrt{X^2 + Y^2}$.

- *Доминирующее направление полигона* – выраженный в радианах угол наклона диаметра полигона к положительному направлению горизонтальной оси изображения.

- *Относительный радиус полигона* – максимальное расстояние точки полигона от его центра масс.

- *Коэффициент формы* – отношение площади полигона к квадрату его диаметра.

- *Коэффициент деформации* – отношение длины минимального диаметра эллипса (аппроксимирующего выпуклую оболочку полигона) к максимальному диаметру эллипса. Аппроксимирующий эллипс определяется главными компонентами множества точек выпуклой оболочки полигона. Выпуклая оболочка включает в себя только те точки полигона, которые образуют максимальную выпуклую поверхность.

- *Геометрический центр* – центр масс выпуклой оболочки полигона.

- *Угол наклона аппроксимирующего эллипса* – выраженный в радианах угол наклона максимального диаметра аппроксимирующего эллипса к положительному направлению горизонтальной оси изображения.

Для выделения выпуклой оболочки и формирования параметров полигона используются встроенные МАТЛАБ функции: *convhull*, *pdist*, *mean*, *princomp*. Полигоны строятся для сегмента в целом, каждой непрерывной компоненты сегмента и для всего изображения в целом. На рис. 6 показан полигон уровня изображения для яркостной компоненты.

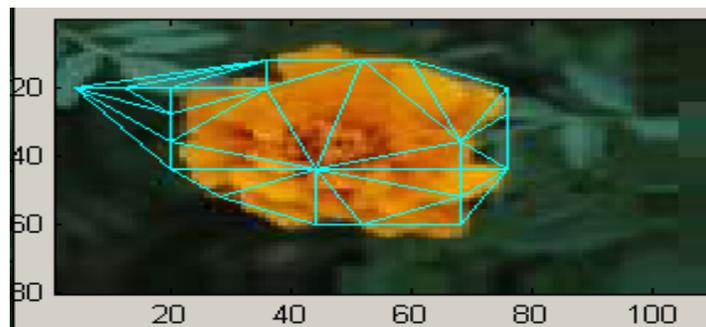


Рис. 6. Полигон уровня изображения

В пространство признаков сегмента включается только максимальная непрерывная компонента, матрица смежностей не используется и зарезервирована для будущих версий. Всего для количественной характеристики сегмента используется 100 вторичных признаков в покоординатном исчислении. По полигону уровня изображений формируется 39 вторичных признаков. Процедура формирования вторичных информативных признаков используется как на этапе накопления базы данных, так и на этапе распознавания.

На этапе накопления базы данных, для каждого сегмента и каждого образа оператор вводит концептуальные понятия (*concept*) и определения к ним (*modifier*) в виде текстовых слов. Эта пара слов считается семантикой изображения или его сегмента. Вместе с вектором признаков эта информация сохраняется в базе данных. По базе данных строится нечеткое отношение *Transit* между семантикой сегмента и семантикой образа. Значения матрицы нечеткого семантического отношения определяется числом совместного проявления в базе данных семантики сегмента с определенной семантикой образа.

3.4. Алгоритм формирования сложных признаков. Сложные признаки – это третий уровень в иерархии пространств информативных признаков. Этот уровень необходим, для того чтобы перед выполнением процедуры классификации сформировать признаковое пространство, анизотропное по информативности.

3.4.1. Оценка информативности. Формирование сложных признаков выполняется на основе анализа информативности признаков в накопленной базе данных. Для анализа информативности признаков используется мера взаимной информации признака [23] с верифицированной семантикой сегмента и/или изображения в целом. Предварительно количественные признаки квантуются с точностью 5% от диапазона изменения, а семантика сегментов и образов кодируется числовыми кодами. В итоге поле признаков приводится к номинальной шкале, где каждое значение можно считать буквой. Для количественных признаков число букв в алфавите равно 20. Для семантики число букв определяется числом различных семантических понятий в базе данных. Оценка информативности признаков выполняется на основе информационной матрицы, по одной координате которой представлены коды признака, а по другой коды семантического образа, пример подобной матрицы показан ниже.

x/y	a	b	c	d	e	f	g	h	
0	m_{00}	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}	m_{06}	m_{07}	m_0
1	m_{10}	m_{11}	m_1
2							m_2
3							m_3
4	m_{40}	m_{41}	m_4
	n_0	n_1	n_2	n_3	n_4	n_5	n_6	n_7	M

Данную матрицу можно рассматривать как накопленную по базе данных статистику работы канала передачи информации. Каждый элемент матрицы указыва-

ет на число переходов i -ой буквы алфавита признака x в j -ю букву алфавита признака y . Числа $m_0 \dots m_4$ – (суммы по строкам) определяют композицию признака x . Числа $n_0 \dots n_7$ – (суммы по столбцам) определяют композицию признака y . Величина $M = \sum_j n_j = \sum_i m_i$ равна числу переданных символов.

Энтропии признаков определяются выражениями:

$$H_x = -\sum_i \frac{m_i}{M} \log \frac{m_i}{M} \quad - \quad \text{энтропия входа,}$$

$$H_y = -\sum_j \frac{n_j}{M} \log \frac{n_j}{M} \quad - \quad \text{энтропия выхода,}$$

$$H_{x \otimes y} = -\sum_i \sum_j \frac{m_{ij}}{M} \log \frac{m_{ij}}{M} \quad - \quad \text{энтропия канала,}$$

$$H_{x;y} = H_x + H_y - H_{x \otimes y} \quad - \quad \text{взаимная энтропия.}$$

Известно [23], что $H_{x;y} \leq H_x$ и $H_{x;y} \leq H_y$. Взаимная энтропия характеризует степень похожести признаков x и y . Удобно использовать нормированную величину $S_{x;y} = H_{x;y} * 2 / (H_x + H_y)$, которая изменяется в диапазоне $[0,1]$. Если $S_{x;y} = 1$, то признаки x и y совпадают (с точностью до замены букв). Если $S_{x;y} = 0$, то признаки x и y сильно различаются. Величина $S_{x;y}$ (когда x является измеряемым информативным признаком, а y – семантикой сегмента) рассматривается как степень надежности признака x и используется далее для вычисления функции принадлежности для лингвистических переменных. Данный метод позволяет выполнить оценку информативности одиночных признаков по отношению к семантике сегмента или образа и отбросить слабые признаки.

3.4.2. Агрегирование признаков. Взаимная энтропия используется также для оценки связи между измеряемыми информативными признаками. Признаки независимы, если их взаимная энтропия равна нулю. Агрегация независимых информативных признаков в сложный признак приводит к усилению разделяющей способности по отношению к семантике образа [23]. Поэтому на этапе классификации все признаки подвергаются агрегации. При равном уровне информативности эффект от агрегации тем выше, чем выше степень взаимной независимости признаков. Агрегация выполняется как адаптивная кластеризующая процедура, реализующая последовательное объединение наиболее независимых признаков, при этом используется тот же механизм адаптации, что и при сегментировании изображения. Мерой расстояния между признаками x и y является величина $1 - S_{x;y}$. В табл. 3 демонстрируется эффект агрегации на примере 5 признаков полигона образа. Числовые значения в таблице равны степени информативности $S_{x;y}$ одиночных и сложных признаков по отношению к семан-

тике образа. В контексте нечеткой классификации информативность трактуется как степень достоверности признака.

Структура сложных признаков переопределяется после каждого изменения базы данных.

Табл. 3. Информативность признаков

Признаки полигона	Одиночный признак	Сложный признак
FormFactor	0.659	0.8719
EFactor	0.677	
Potential	0.673	0.9222
Radius	0.651	
Mass_center (y)	0.637	
Polygon_center (y)	0.643	

3.5. Семантическая классификация. Семантическая классификация выполняется в признаковом пространстве третьего уровня. Предъявляемое изображение разделяется на сегменты, и по каждому сегменту вычисляются значения сложных признаков, которые представляют собой векторы небольшой размерности. Размерность признакового пространства переопределяется заново при пополнении базы данных.

3.5.1. Нечеткая классификация сегментов. Каждый сегмент изображения обрабатывается по той же схеме, что и при формировании базы данных. Текущее значение сложного признака сравнивается со значениями одноименного признака всех верифицированных сегментов базы данных (прецедентами). В результате сравнения определяются стандартные евклидовы расстояния между текущими и верифицированными значениями. Уровень достоверности классификации по данному сложному признаку определяется величиной:

$$\xi_i = 1 - \min_k (d_{ki}) / \max_k (d_{ki}),$$

где d_{ki} – расстояние текущего значения i -го признака до k -го прецедента. Уровень достоверности максимален и равен 1, когда в базе данных существует прецедент, совпадающий с текущим значением признака. Функция принадлежности к прецеденту вычисляется по правилу:

$$\eta_k = \frac{1}{p} \sum_{i=1}^p \min(\xi_i, \chi_i),$$

где p – число сложных признаков на уровне сегментов, χ_i – уровень достоверности сложного признака. Функция принадлежности к семантике s определяется выражением:

$$\mu_s = \max_{\eta_k \in K_s} (\eta_k),$$

где K_s – множество прецедентов с семантикой s . Значения функции μ_s определяют степени принадлежности анализируемого сегмента термам лингвистической переменной «семантика сегмента». На заключительном этапе выделяется терм, которому соответствует максимальное значение функции принадлежности. Семантика этого термина рассматривается как окончательный результат

классификации сегмента, а соответствующее значение функции принадлежности определяет уровень его достоверности.

3.5.2. Нечеткая классификация образов по семантике сегментов. Семантическая классификация образов выполняется на основе матрицы нечеткого отношения *Transit* между семантикой сегментов и семантикой образов. Просматриваются семантики классифицированных сегментов, и выполняется вычисление функции принадлежности к семантике образов, по следующему правилу:

$$\mu_{\text{lm}}(i) = \frac{1}{n \cdot m(i)} \sum_{k=1}^n \mu_s(k) \text{Transit}(i, k),$$

где n – число сегментов в образе, $\mu_s(k)$ – уровень достоверности семантики k -го сегмента, $m(i) = \max_k (\text{Transit}(i, k))$ – нормирующий множитель.

3.5.3. Нечеткая классификация образов по параметрам полигона образа. По каждому сложному признаку полигона образа выполняется метрическая классификация по всем прецедентам базы данных. Текущее значение сложного признака сравнивается со значениями одноименного признака всех прецедентов. В результате сравнения определяются стандартные евклидовы расстояния между текущими и верифицированными значениями. Уровень достоверности классификации по каждому сложному признаку определяется величиной:

$$\xi_i = 1 - \min_k (d_{ki}) / \max_k (d_{ki}),$$

где d_{ki} – расстояние текущего значения i -го признака до k -го прецедента. Достоверность максимальна и равна 1, когда в базе данных существует прецедент, совпадающий с текущим значением признака. Функция принадлежности к прецеденту вычисляется по правилу:

$$\eta_k = \frac{1}{g} \sum_{i=1}^g \min(\xi_i, \chi_i),$$

где g – число сложных признаков на уровне сегментов, χ_i – уровень достоверности сложного признака. Функция принадлежности к семантике образов p определяется выражением

$$\mu_s = \max_{\eta_k \in K_s} (\eta_k),$$

где K_p – множество прецедентов с семантикой s . Значения функции μ_p определяют степени принадлежности анализируемого образа термам лингвистической переменной «семантика образа».

3.5.4. Комплексование результатов. Результаты семантической классификации, полученные по двум параллельным системам признаков, конъюнктивно объединяются. Результирующая функция принадлежности определяется правилом:

$$\mu = \min(\mu_{\text{lm}}, \mu_p).$$

Если множество диагнозов пусто или максимальное значение функции принадлежности меньше заданного порога, то констатируется отказ от классификации.

ции. Данная ситуация возникает, когда база данных недостаточно представительна.

3.6. Результаты эксперимента. Для проведения экспериментальных исследований была разработана программа классификации JPEG-изображений в среде МАТЛАБ. Программа позволяет также формировать базу данных и выполнять оценки информативных признаков. Рабочее окно программы показано на рис. 7. При декодировании изображений в правом информационном поле отображаются обнаруженные маркеры JPEG-формата. Основным управляющими элементами интерфейса являются управляющие кнопки *Decode JPEG-file* – запуск процедуры декодирования файла, *Dialog* – вызов окна управления базой данных, *Classify* – запуск процедуры классификации. Диалоговое окно, представленное в нижнем правом углу скриншота, реализует ввод семантических понятий и управление базой данных.

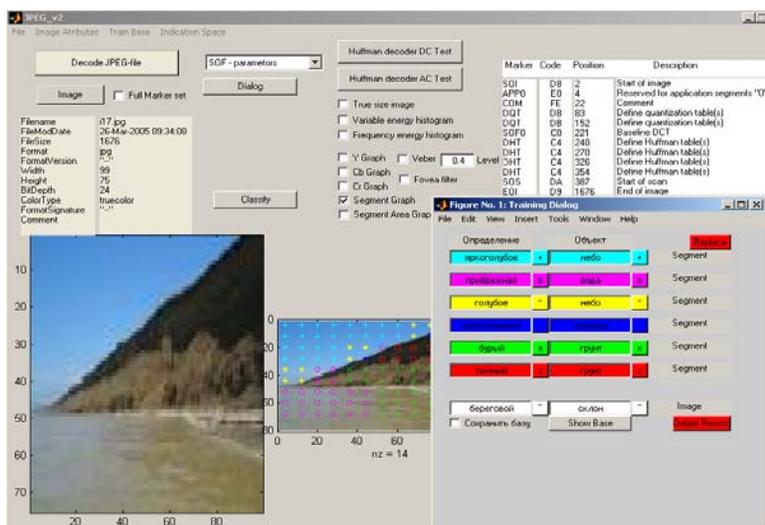


Рис. 7. Скриншот рабочего окна

Для проведения экспериментов с помощью данной программы была создана база данных, содержащая информацию по 100 изображениям. Характеристики базы данных представлены в табл. 4.

Из таблицы видно, что представительность базы данных в среднем не высокая; на каждое семантическое понятие *Concept/Modifier* на уровне образов приходится 1.5 примера, а для уровня сегментов 2.5 примера. Однако некоторые концепты представлены достаточно полно. В табл. 5 показана частотность наиболее представленных концептов на уровне сегментов. Результаты экспериментов показали, что изображения, которые содержат наиболее представительные концепты базы данных, достаточно уверенно распознаются по семантике.

Табл. 4. Характеристика базы данных

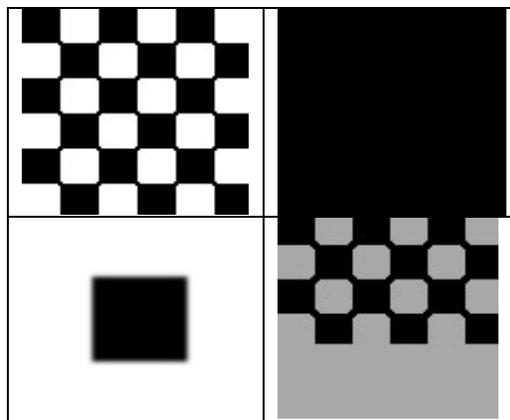
Число примеров в базе данных=100
Число сегментов в базе данных=535
Число концептуальных понятий в базе данных=86
Число модификаторов понятий в базе данных=154
Число сочетаний <i>Concept/Modifier</i> на уровне сегментов=214
Число сочетаний <i>Concept/Modifier</i> на уровне образов=67
Средний размер изображений 120*120 пикселей

Табл. 5. Представительность базы данных

Частотность <i>Concept/Modifier</i> на уровне сегментов	Частотность <i>Concept</i> на уровне сегментов
цветотеневая граница 54	небо 111
текстурный фрагмент 20	граница 81
голубое небо 20	листва 43
белесое небо 17	вода 24
облачное небо 16	грунт 23
светло-голубое небо 16	берег 21
зеленая листва 12	фрагмент 20
темно-зеленая листва 12	пятно 17

Ниже приведены два характерных примера работы алгоритма семантической классификации.

Табл. 6. Текстурный фрагмент



1. *Текстурный фрагмент.* В первом, втором и третьем квадранте табл. 6 приведены образы, входящие в базу данных, в четвертом квадранте (правый нижний угол) представлен тестовый образ.

На рис. 8 показан скриншот результата классификации. Левое окно – результат классификации по признакам полигона, среднее окно – результат классификации по признакам сегментов, правое окно – комплексный результат.

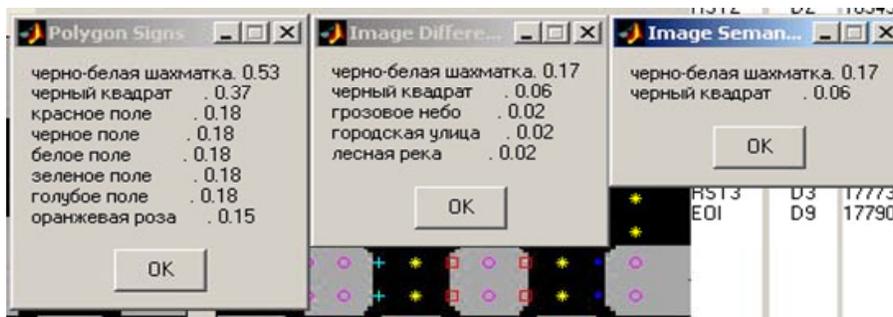


Рис. 8. Результаты классификации

Числа определяют значения функции принадлежности лингвистическому термину. При выводе на экран результаты семантической классификации образов упорядочиваются по убыванию функции принадлежности.

2. *Лесная река.* В первом, втором и третьем квадрантах табл. 7 выборочно показаны примеры, входящие в базу данных, в четвертом квадранте представлено тестовое изображение.

Табл. 7. Лесная река



Результаты семантической классификации тестового изображения показаны на рис. 9. Правое окно соответствует комплексному результату. Первую позицию с наибольшим значением достоверности занимает семантика «лесная река», что соответствует фактическому содержанию тестируемого изображения. Интересно отметить, что остальные результаты оказались по смыслу близкими к достоверному семантическому значению.

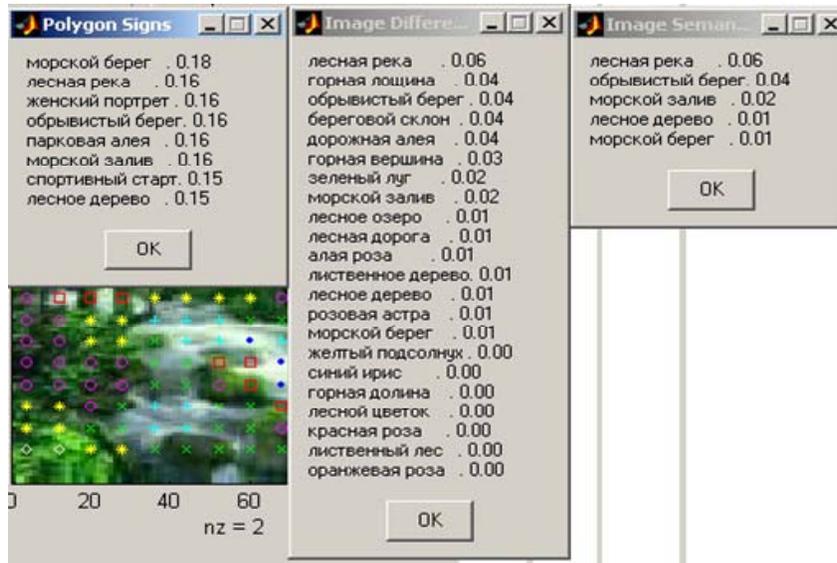


Рис. 9. Результаты классификации образа «Лесная река»

4. Выводы и обсуждение результатов

Проведенные исследования доказали принципиальную возможность реализации процедуры семантической классификации изображений в признаковом пространстве JPEG-формата. Исключение процедуры полного восстановления изображения является принципиальным решением, которое дает возможность обеспечить высокое быстродействие классифицирующих подсистем. Для экспериментальных исследований в проекте использовалась программная среда MATLAB, что было обусловлено богатством ее функциональных возможностей в программном и пользовательском интерфейсе и высоким уровнем оперативности при отладке программ. Однако данная среда является интерпретирующей, поэтому скорость выполнения программ существенно ниже, чем в любой компилирующей среде.

Примерные оценки быстродействия дают следующие цифры. Программная процедура, написанная на языке MATLAB, выполняла полное декодирование и сегментацию JPEG-изображения размером 100*150 пикселей за 10 сек (на процессоре Intel 2,4 ГГц). Семантическая классификация изображения с отображением результата выполнялась за 7 сек. Встроенная функция MATLAB, также реализующая полное JPEG-декодирование, выполняется за 0.1 сек. Поэтому можно ожидать, что время выполнения процедуры семантической классификации в компилирующей программной среде для данного размера изображения будет также на уровне 0.1 сек.

Представительность базы данных должна быть существенно выше, чем в проведенных экспериментах. Удовлетворительная классификация может быть обеспечена при среднем уровне 15-20 образов на семантический класс. Накопление базы данных – процесс достаточно длительный, поскольку связан с действиями оператора. На описание одного изображения (в сегментном представлении) оператором тратится 2-4 минуты. Кроме того, семантическое описание подвержено субъективным взглядам оператора. Субъективность накопленных данных можно уменьшить двумя путями: 1) объединением баз данных созданных различными операторами, 2) добавлением уровня онтологий, определяющего отношения между семантиками. Уровень онтологий может быть выстроен в автоматическом режиме по совпадениям концептов или модификаторов. Этот путь является более перспективным, поскольку позволяет без расширения базы данных добавить еще один уровень в иерархию семантической классификации, увеличив тем самым достоверность ее работы.

5. Литература

1. Smith J.R., S.F.Chang Visualseek: A fully automated content based image query system. In: Proc. ACM Multimedia, Boston, MA, Nov, 1996.
2. Milind R.N., T.S.Huang Extracting Semantics from Audiovisual Content: The Final Frontier in Multimedia Retrieval.
3. Salton G., McGill M.J. Introduction to Modern Information Retrieval. New York: McGraw-Hill, 1983.
4. Rui Y., Hyang T.S. and Mehrotra S. Content-based image retrieval with relevance feedback in MARS. In: Proc.IEEE Int. Conf. Image Processing, Santa Barbara, CA, 1997, pp. 815-818.
5. Celentano A., Sciasico E.D. Feature integration and relevance feedback analysis in image similarity evaluation. J.Electron. Imaging, vol. 7, no.2, pp. 308-317, 1998.
6. Gevers T., Smeulders A.W.M. PieToSeek: Combining color and shape invariant features for image retrieval. IEEE Trans. Image Processing, vol. 9, pp. 102-119, 2000.
7. Sciascio E.Di., Mongiello. DrawSearch: A tool for interactive content-based image retrieval over the net. Proc. SPIE, vol. 3656, pp. 561-572, 1999.
8. Müller H., Müller S., Marcand-Maillet, and Squire D. McG. Strategies for positive and negative relevance feedback in image retrieval. In: Proc. Int. Conf. Pattern Recognition, Barcelona, Spain, 2000.
9. Porkaew S., Mehrotra S., Ortega M. and Chakrabarti K. Similarity search using multiple examples in MARS. In: Visual Information and Information Systems. New York: Springer-Verlag, 1999, pp. 68-75.
10. Ciocca G., Schettini R. Using a relevance feedback mechanism to improve content-based image retrieval. In: Visual Information and Information Systems, New York: Springer-Verlag, 1999, pp. 105-114.
11. Muneesawang P., Guan L. Anonliner RBF model for interactive content-based image retrieval. In: Proc. 1st IEEE Pasific-Rim Conf. Multimedia, Syney, Australia, Dec. 2000, pp. 188-191.
12. Ei-Nada, Wernick M.N., Yang. Y., Galatsonos N.P. Image retrieval based on similarity learning. In: Proc. IEEE. Conf. Image Processing, vol. 3. Vancouver, BC, Canada, 2000, pp.772-775.
13. Peng J., Bhanu B., Oing S. Probabilistic feature relevance learning for content-based image retrieval. Comput. Vision Image Understanding, vol. 75, no. ½ pp. 150-164, 1999.

14. Sclaroff S., Taycher L., Cascia M.L. ImageRover: A content-based image browser for the world wide web. In: Proc. IEEE Workshop Content-Based Access Image Video Libraries, Puerto Rico, June 1997, pp. 2-9.
15. Muneesawang P., Guan L. Automatic machine interactions for content-based image retrieval using a self-organizing tree map architecture. IEEE Trans. on Neural Networks, vol. 13, no 4, July 2002.
16. Бодякин В.И. "Куда идешь, Человек? Основы эволюциологии. (информационный подход), М.: СИНТЕГ, 1998, 332с.
17. Бодякин В.И. "Исследование структурных моделей открытых динамических систем", специальность: 05.13.01 (Управление в технических системах), автореферат диссертации и диссертация на соискание ученой степени к.ф.-м.н., Москва – 1999г.
18. T.81 - Information technology – Digital compression and coding of continuous-tone still images – Requirements and guidelines
19. T.83 - Information technology – Digital compression and coding of continuous-tone still images: Compliance testing.
20. T.84 - Information technology – Digital compression and coding of continuous-tone still images: Extensions.
21. JFIF 1.02 (JPEG File Interchange Format).
22. Ю.К. Гаврилей, А.И. Самарин, М.А. Шевченко Активный анализ изображений в системах с фовеальным восприятием. – Нейрокомпьютеры в системах обработки изображений, №7-8, 2002.- С.34-46.
23. Гоппа В.Д. Введение в алгебраическую теорию информации. М.: Наука. Физматлит, 1995. 112с.

Статья поступила 31 мая 2006 г.
После доработки 29 августа 2006 г.

Корректная селекция алгоритмов классификации

Хомич А.В.

BaseGroup Labs

ho76@mail.ru; supremum76@rambler.ru

Аннотация

Рассматривается проблема определения доверительных интервалов надежности классифицирующих алгоритмов. Показано, что при селекции классификаторов по результатам тестирования необходимо учитывать количество тестируемых классификаторов. Продемонстрировано смещение оценок надежности при селекции классификаторов по результатам теста. Предлагается отказаться от селекции только по результатам тестирования и ввести ограничение на количество параметров в алгоритмах классификации. Предложен метод поиска классификатора, не приводящий к увеличению длины доверительных интервалов и смещению оценок надежности. Часть полученных результатов может быть применена к задаче регрессии.

1. Введение

В настоящей работе рассматривается следующая задача. Пусть имеется конечный набор векторов. Каждому вектору неким «учителем» сопоставлен его класс. В качестве «учителя» может выступать эксперт, измерительный прибор и т. п. Вектор с указанным классом называют примером. Набор имеющихся примеров является выборкой из генеральной совокупности всех возможных в данной предметной области примеров. Требуется создать классификатор, правильно классифицирующий векторы, не участвующие в обучении, основываясь только на анализе имеющихся примеров. Под надежностью классификатора будем понимать вероятность правильной классификации на генеральной совокупности примеров. Все подходы к решению задачи получения надежного классификатора условно можно условно разделить на два случая:

- Подход, основанный на явном использовании сведений о распределении классов.
- Подход, основанный на представлении классификатора в виде «черного ящика» [1].

Первый подход позволяет провести статистически обоснованный анализ надежности классификатора. Но он применим только в относительно простых задачах, в которых наблюдаются хорошо изученные распределения классов и зависимости восстанавливаются в достаточно узком классе функций. Второй подход, напротив, позволяет решать сложные нелинейные задачи, восстанавливая зависимости широкого класса, но вызывает затруднение оценка надежности классификатора. В методе «черного ящика» при общепринятом подходе все примеры разделяются на непересекающиеся наборы обучающих и тестовых примеров. Параметры «черного ящика» настраиваются на правильную классификацию обучающих примеров. На тестовых примерах оценивается его надежность.

Дальнейшие рассуждения основываются на методе определения доверительного интервала надежности классификатора, предложенного в работе [2]. Новым результатом является то, что в данной работе продемонстрирована зависимость доверительного интервала от количества тестируемых вариантов классификатора (используется метод «черного ящика»).

2. Случай тестирования одного алгоритма классификации

Введем обозначения:

- P – значение некоторой вероятности;
- $\Pr(A)$ – безусловная вероятность наступления события A ;
- $\Pr(A|B)$ – вероятность наступления события A при условии B ;
- P_{est} – оценка на тестовых примерах вероятности правильной классификации (оценка надежности);
- $\forall P_{est}$ – читается, как «Все P_{est} , вычисленные в серии тестов, удовлетворяют условию...»;
- $\exists P_{est}$ – читается, как «Среди вычисленных оценок надежности найдется оценка равная P_{est} » или «Среди вычисленных оценок надежности найдется оценка P_{est} , удовлетворяющая условию...».

Пусть m – число бинарных позиций, принимающих значение 0 или 1; k – число позиций, принявших единичное значение ($m \geq k$). Число возможных вариантов размещения k единиц среди m бинарных позиций [3] равно

$$N(m, k) = C_m^k = \frac{m!}{(m-k)!k!}. \quad (1)$$

Вероятность k из m правильных ответов при вероятности правильного ответа P_{true} определяется выражением [2]

$$\Pr(P_{est} | P_{true}, m) = N(m, k) P_{true}^k (1 - P_{true})^{m-k}, \quad (2)$$

где $P_{est} = \frac{k}{m}$.

Оценка P_{est} может оказаться как больше, так и меньше действительной надежности классификации P_{true} . Вероятность надежности классификации P_{true} при полученной оценке P_{est} по результатам m тестов можно выразить, используя формулу Байеса [4,5], как

$$\Pr(P_{true} | P_{est}, m) = \frac{\Pr(P_{est} | P_{true}, m)}{\sum_{P_{true} \in \Omega} \Pr(P_{est} | P_{true}, m)}, \quad (3)$$

где Ω – множество всех возможных значений P_{true} . Если количество возможных значений позиций вектора примера (факторов) конечно, то Ω конечно. В противном случае Ω должна быть заменена конечным множеством. Этого можно достичь, применив следующий прием. Предположим наличие некоторого большого, но конечного количества примеров и примем, что вероятность правильной классификации P_{true} приближенно равна отношению количества правильно решенных примеров к общему количеству примеров. Пусть количество примеров равно некоторому достаточно большому числу M . Тогда вместо $P_{true} \in [\alpha, \beta]$ можно записать

$$P_{true} \in \left\{ \alpha, \frac{\alpha M + 1}{M}, \frac{\alpha M + 2}{M}, \dots, \frac{\beta M - 1}{M}, \beta \right\}.$$

Значение M подбирают, соблюдая компромисс между точностью и сложностью вычислений. Будем обозначать множество, полученное таким способом, как $\Omega[\alpha, \beta]$.

Для более полного анализа надежности классификатора необходимо определить доверительный интервал P_{true} . Обозначим нижнюю границу доверительного интервала P_{true} как P_{below} , а верхнюю границу как P_{above} или $P_{below} \leq P_{true} \leq P_{above}$. Доверительные интервалы P_{true} можно определить по формуле

$$\Pr(P_{true} \in [P_{below}, P_{above}] | P_{est}, m) = \sum_{P_{true} \in \Omega[P_{below}, P_{above}]} \Pr(P_{true} | P_{est}, m) = \alpha, \quad (4)$$

где α – доверительная вероятность. При необходимости определить P_{above} и P_{below} можно воспользоваться формулой (4). Например, задавшись определенными значениями P_{est} , m , α , и перебрав различные варианты $[P_{below}, P_{above}]$ можно выбрать доверительный интервал наименьшей длины (интервал с наименьшим из рассмотренных вариантов значением $P_{above} - P_{below}$).

Наибольший интерес в большинстве приложений вызывает нижняя граница доверительного интервала. Она показывает, какую минимальную надежность классификации можно гарантировать. В таблицах 1 и 2 приведены значения P_{below} , вычисленные по формуле (4) (при расчетах использовалось значение $M = 100$).

Таблица 1 – Значения нижней границы доверительного интервала при $\alpha = 0.95$

P_{est} \ M	10	20	30	40	50
0.6	0.33	0.40	0.43	0.45	0.46
0.7	0.43	0.50	0.54	0.56	0.57
0.8	0.53	0.61	0.65	0.67	0.68
0.9	0.63	0.72	0.76	0.79	0.80
1.0	0.76	0.86	0.90	0.93	0.94

Таблица 2 – Значения нижней границы доверительного интервала при $\alpha = 0.9$

P_{est} \ M	10	20	30	40	50
0.6	0.37	0.43	0.45	0.47	0.48
0.7	0.48	0.53	0.56	0.58	0.59
0.8	0.58	0.65	0.68	0.69	0.70
0.9	0.69	0.76	0.79	0.81	0.82
1.0	0.81	0.89	0.92	0.94	0.95

3. Случай тестирования множества алгоритмов классификации

Для получения классификатора с удовлетворительной надежностью классификации чаще всего приходится тестировать набор различных вариантов классификаторов. Перебор вариантов и их оценка может осуществляться автоматически. Исследователь тестирует набор различных классификаторов и, основываясь на результатах тестирования, выбирает самый надежный. Как и выше, считаем, что оценка надежности при тестировании определенного классификатора равна отношению числа правильных результатов классификации к общему числу тестов для данного классификатора $P_{est} = \frac{k}{m}$. Покажем, что в этом случае проблема расхождения оценки и истинной надежности классификации становится особенно острой.

Пусть n различных классификаторов тестируются на m примерах. Обычно значения надежности классификаторов различны. Под ошибочным выбором классификатора будем понимать выбор классификатора с наименьшей ошибкой тестирования, но не наименьшей реальной ошибкой обобщения. Вероятность ошибочного выбора классификатора зависит от сочетания значений надежности тестируемых классификаторов. Наилучшим случаем является ситуация, когда часть тестируемых классификаторов реально обладают нулевой вероятностью ошибки классификации, а оставшаяся часть обладает единичной вероятностью ошибки классификации. В этом случае ошибка тестирования будет полностью соответствовать реальной ошибке классификации. Следовательно, ошибка тестирования позволит однозначно выявить лучшие и худшие классификаторы.

Наихудшим случаем является ситуация, когда истинная надежность всех тестируемых классификаторов одинакова и недостаточна для практического приложения. В этом случае разброс значений ошибок тестирования отражает лишь наличие неопределенности в оценке ошибки классификации. Случайно полученное малое значение ошибки тестирования может быть ошибочно расценено как признак достаточно малой для практического приложения реальной ошибки классификации. Трудно оценить сочетание реальных значений надежности тестируемых классификаторов без привлечения дополнительных сведений. В отсутствии дополнительных сведений приходится исходить из наихудшего случая.

Допустим, истинная надежность всех классификаторов одинакова и равна P_{true} . Это предположение соответствует наихудшему случаю и используется во всех дальнейших рассуждениях. Если заменить его более слабым ограничением на равенство средней надежности классификаторов некоторому значению, то очевидно с максимальной вероятностью будет выбран классификатор с максимальной надежностью, и его надежность будет не ниже средней надежности по всем классификаторам. Только при условии равенства реальной надежности классификаторов максимальное и среднее значения надежностей совпадают.

Пусть все классификаторы тестируются независимо друг от друга и пусть исследователем задано значение удовлетворительной оценки надежности P_{well} . Вероятность случайно получить, в ходе исследования, удовлетворительную оценку надежности P_{well} определяется по формуле (2), то есть

$$\Pr(P_{est} = P_{well} | P_{true}, m) = \Pr(P_{well} | P_{true}, m). \quad (5)$$

Но надо учитывать и случаи, когда получена оценка больше P_{well} . То есть надо учитывать все случаи, когда может быть получено правильных ответов $mP_{well}, mP_{well} + 1, mP_{well} + 2, \dots, m$. Согласно формуле, описывающей наступление любого несовместного события [4,5], вероятность получения оценки, не меньшей P_{well} , определяется выражением:

$$\Pr(P_{est} \geq P_{well} | P_{true}, m) = \sum_{i=0}^{m(1-P_{well})} \Pr\left(P_{est} = \frac{mP_{well} + i}{m} | P_{true}, m\right). \quad (6)$$

Вероятность того, что такая оценка при n испытаниях ни разу не будет получена, подчиняется закону совместного наступления независимых событий

$$\Pr(\forall P_{est} < P_{well} | P_{true}, n, m) = (1 - \Pr(P_{est} \geq P_{well} | P_{true}, m))^n. \quad (7)$$

Вероятность того, что хотя бы одному классификатору удастся при тестировании показать оценку надежности не меньше P_{well} , равна

$$\begin{aligned} P_{\text{deception}} &= \Pr(\exists P_{\text{est}} \geq P_{\text{well}} \mid P_{\text{true}}, n, m) = \\ &= 1 - (1 - \Pr(P_{\text{est}} \geq P_{\text{well}} \mid P_{\text{true}}, m))^n \end{aligned} \quad (8)$$

В случае, если $P_{\text{true}} < P_{\text{well}}$ величина $P_{\text{deception}}$ характеризует вероятность получить завышенную оценку надежности.

Для демонстрации значимости проблемы завышения оценки надежности классификаторов приведем следующий пример. Пусть тестируется 1000 классификаторов на 100 тестовых примерах. Пусть все классификаторы обладают надежностью 0.5. Вероятность получить оценку надежности, не меньше 0.65, при однократном тестировании приближенно равна 0.0015. Вероятность получить оценку надежности не меньше 0.65, хотя бы при одном испытании из 1000, приближенно составляет 0.79. Как видно, вероятность получения существенно разнящейся оценки и истинной надежности при многократных испытаниях может быть большой даже при малой вероятности при однократном испытании.

Напомним, было принято ограничение, что реальная надежность всех тестируемых классификаторов равна P_{true} . Вероятность того, что в n тестах на m примерах только одна оценка надежности классификации будет равна P_{est} , определяется как

$$n \Pr(P_{\text{est}} \mid P_{\text{true}}, m) (1 - \Pr(P_{\text{est}} \mid P_{\text{true}}, m))^{n-1}.$$

Вероятность того, что в n тестах на m примерах ровно две оценки надежности классификации будут равны P_{est} , определяется, как

$$N(n, 2) \Pr(P_{\text{est}} \mid P_{\text{true}}, m)^2 (1 - \Pr(P_{\text{est}} \mid P_{\text{true}}, m))^{n-2}.$$

Рассуждая далее аналогично, получим ряд несовместных событий. Для того чтобы получить вероятность того, что в n тестах на m примерах хотя бы одна оценка надежности классификации будет равна P_{est} , достаточно сложить вероятности этих событий или

$$\begin{aligned} \Pr(\exists P_{\text{est}} \mid P_{\text{true}}, n, m) &= \\ &= \sum_{i=1}^n N(n, i) \Pr(P_{\text{est}} \mid P_{\text{true}}, m)^i (1 - \Pr(P_{\text{est}} \mid P_{\text{true}}, m))^{n-i}. \end{aligned} \quad (9)$$

Имея формулу вероятности получения конкретного значения оценки надежности, можем определить вероятность истинной надежности. Вероятность надежности классификации P_{true} при полученной оценке P_{est} по результатам тестирования n классификаторов на m тестах можно выразить, используя формулу Байеса [4,5]

$$\Pr(B_i | A) = \frac{P(B_i)P(A | B_i)}{\sum_{j \in [1,i) \cup (i,k]} P(B_j)P(A | B_j)},$$

где k – количество учитываемых несовместных событий, влияющих на вероятность реализации события A . Задача состоит в получении формулы для определения $\Pr(P_{true} | P_{est}, n, m)$. Формально для этого необходимо знать все условные вероятности $\Pr(P_{est} | P_{true}, n, m)$ и безусловные вероятности $\Pr(P_{est})$. Условие несовместности событий выполняется, так как, очевидно, что в одном тесте невозможно одновременно получить две разные ошибки тестирования. $\Pr(P_{est} | P_{true}, n, m)$ интерпретируется, как условная вероятность получить при тестировании ошибку тестирования, равную P_{est} . Исходя из этой интерпретации, $\Pr(P_{est} | P_{true}, n, m)$ можно заменить на $\Pr(\exists P_{est} | P_{true}, n, m)$. $\Pr(P_{est})$ интерпретируется, как вероятность получения при произвольном тестировании произвольного классификатора ошибку тестирования, равную P_{est} . Сложность определения $\Pr(P_{est})$ заключается в том, что ни тестовый набор примеров, ни тестируемый классификатор явно ничем не ограничен. Конечно, можно рассмотреть все множество возможных пар <тестовых примеры, тестируемый классификатор>. Но не факт, что на практике каждая такая пара имеет одинаковую вероятность реализации. Примем как гипотезу, что $\Pr(P_{est})$ равномерно распределено на отрезке $[0,1]$. В этом случае $\Pr(P_{est})$ в формуле Байеса сокращается.

$$\Pr(P_{true} | P_{est}, n, m) = \frac{\Pr(\exists P_{est} | P_{true}, n, m)}{\sum_{P_{true} \in \Omega[0,1]} \Pr(\exists P_{est} | P_{true}, n, m)}. \quad (10)$$

Формула (10) выводится аналогично формуле (3) [2], но с учетом множественности проводимых тестов и применения процедуры выбора классификатора с максимальной оценкой P_{est} .

Доверительные интервалы P_{true} в случае тестирования n классификаторов на m тестах можно определить по формуле

$$\begin{aligned} \Pr(P_{true} \in [P_{below}, P_{above}] | P_{est}, n, m) &= \\ &= \sum_{P_{true} \in \Omega[P_{below}, P_{above}]} \Pr(P_{true} | P_{est}, n, m) = \alpha \quad , \end{aligned} \quad (11)$$

где α – доверительная вероятность, $\Omega[P_{below}, P_{above}]$ – конечное перечислимое множество, определяемое так же, как и в формуле (4).

На практике формулу (11) можно использовать для вычисления довери-

тельных интервалов P_{true} . Например, задавшись определенными значениями P_{est} , m , n , α , и перебрав различные варианты $[P_{below}, P_{above}]$, можно выбрать доверительный интервал наименьшей длины (интервал с наименьшим из рассмотренных вариантов значением $P_{above} - P_{below}$).

Представленная методика позволяет определить доверительный интервал надежности классификатора и обоснованно выбирать количество тестовых примеров. Как видно из таблиц 1 и 2, для оценки надежности классификатора рекомендуется использовать не менее 50 тестовых примеров. При использовании 50 или менее тестовых примеров следует наряду с полученной оценкой указывать и нижнюю границу доверительного интервала надежности классификатора. При тестировании набора конкурирующих классификаторов следует увеличить число тестовых примеров. Также можно дать рекомендацию по возможности избегать многократных тестов. В следующем разделе приводится алгоритм, позволяющий выбрать наиболее надежный классификатор без необходимости большого количества повторений процедуры тестирования.

4. Эффект структурного переобучения

В качестве цели оптимизации структуры и параметров классификатора часто рассматривается [6,13,15] увеличение надежности классификации. На практике используются и другие цели оптимизации структур и параметров, но в данной статье внимание сконцентрировано именно на надежности классификации. При решении задачи регрессии критерием оптимальности может служить среднеквадратичное отклонение по всей области возможных входных значений. Часто эти критерии объединяют общим термином – ошибка обобщения. Непосредственно измерить ошибку обобщения невозможно. Прибегают к оценкам, выводимым из ошибок на тестовых примерах (ошибка тестирования). Параметры алгоритма настраиваются с целью минимизации ошибки на обучающих примерах (ошибка обучения). Надежность алгоритма оценивают по ошибке тестирования. Только при неограниченном увеличении количества тестовых примеров разница между ошибкой тестирования и ошибкой обобщения стремится к нулю. При оптимизации структуры алгоритма приходится перебирать множество вариантов структур [6]. В качестве критерия селекции структур может служить ошибка тестирования.

Практика показывает [7,8,9], что селекция только по ошибке тестирования, в случае перебора большого количества алгоритмов, не гарантирует получение алгоритма с удовлетворительной ошибкой обобщения. Одной из причин этого служит возможность случайного получения малой ошибки тестирования. Если при оценивании 1-5 вариантов структур вероятность такого мала, то при переборе 100 и более структур вероятность случайного результата резко возрастает. Для устранения этого недостатка необходимо увеличивать количество тестов.

Другим источником проблем служит эффект “структурного переобучения”. Эффект “переобучения” при настройке параметров хорошо известен [10]. Эффект переобучения проявляется в том, что, начиная с некоторого порога,

уменьшение ошибки обучения сопровождается увеличением ошибки тестирования. Такое поведение обусловлено тем, что параметры классификатора начинают настраиваться на шумы и закономерности, присущие только обучающему набору примеров. При оптимизации структуры классификаторов также возникает подобный эффект. Алгоритм с выбранной структурой показывает малую ошибку тестирования, но ошибка обобщения остается высокой. Особенно явно этот эффект проявляется в условиях малого количества доступных примеров (50-200). Подобный эффект наблюдался и другими исследователями [7,8,9].

При малом количестве доступных примеров применяются алгоритмы перекрестного тестирования с множеством вариантов разбиений на обучающую и тестовую группы примеров. В исследованиях автора статьи, в которых наблюдался эффект “структурного переобучения”, использовалась bootstrap-оценка [11,12] (один из вариантов метода скользящего контроля). При вычислении bootstrap-оценки генерируется множество случайных разбиений примеров на тестовую и обучающую группы. Тестовая группа содержит фиксированное количество примеров. В данной работе использовались тестовые группы из 10 примеров. Итоговая ошибка тестирования определяется суммированием ошибок тестирования по всем тестовым группам. Эффект “структурного переобучения” уже нельзя устранить одним увеличением количества тестов. В данной работе общее количество тестов доводилось до 500, но эффект продолжал проявляться. По всей видимости, эффект “структурного переобучения” является следствием наличия в выборках ограниченного размера зависимостей, свойственных только этим выборкам (не наблюдаемых на генеральной совокупности примеров). Если при обучении одного алгоритма происходит подстройка под обучающие примеры, то при селекции по результатам тестирования множества алгоритмов происходит подстройка под тестовые примеры. На первый взгляд можно из множества перебираемых алгоритмов выбирать тот, который после обучения демонстрирует наименьшую ошибку тестирования. Но вероятность ошибки на тестовых примерах вовсе не обязательно равна вероятности ошибки на произвольных примерах. При отборе может быть выбран алгоритм, который демонстрирует малую ошибку тестирования не в силу высоких способностей к обобщению, а в силу «удачного» сочетания структуры, начального приближения параметров и свойств алгоритма оптимизации параметров. Если для поиска алгоритма используются генетические алгоритмы, известные своей способностью решать сложные задачи, а поиск такого «удачного» сочетания очевидно сложная задача, то такая ситуация весьма вероятна.

Для устранения эффекта “структурного переобучения” в работе [8] предлагается выбирать среди множества конкурирующих алгоритмов, алгоритм не с самой лучшей оценкой точности. В данной статье предлагается использовать иной подход, заключающийся во введении явных ограничений на максимальное количество параметров (МКП) в алгоритме. Зависимости, свойственные только обучающей выборке, сформированы случайно, следовательно, сложнее общих зависимостей (наблюдаемых на генеральной совокупности примеров). Поэтому структура малой сложности не позволит их выявить и будет вынуждена настраивать свои параметры на общие зависимости. Аналогично устраняют эффект “параметрического переобучения”, когда ограничивают количество параметров и разброс их значений [13]. В обоснование введения МКП можно также

сослаться на принцип минимальной длины описания [14].

Оптимальное значение МКП можно приближенно определять методом, предложенным в работе [15]. В этой работе предложено выбирать количество настроечных параметров равным $\sqrt{K * N}$, где K – размерность вектора примера, N – количество обучающих примеров. При необходимости, МКП можно уточнять методами одномерной оптимизации (например, методом золотого сечения). На минимальную сложность алгоритма ограничения не накладываются. Селекция алгоритмов, удовлетворяющих требованию МКП, выполняется по ошибке обучения. После завершения оптимизации структуры и параметров алгоритма выполняется проверка на тестовом наборе примеров. Запоминается алгоритм с наименьшей ошибкой тестирования. Затем можно скорректировать МКП и повторить поиск. При оптимизации МКП количество пробующихся вариантов следует ограничивать 3-10 значениями. Во-первых, вычисление критерия оптимальности требует много вычислительных затрат. Во-вторых, при увеличении числа проб растет вероятность случайно получить хороший результат.

Опишем использованный в данной работе алгоритм поиска надежных алгоритмов классификации или регрессии. В процессе поиска ищется алгоритм ALG с наименьшей ошибкой обобщения. Количество итераций алгоритма ограничено величиной T , задаваемой пользователем.

1. $t := 1$;
2. Выбирается МКП_{*t*};
3. Создается множество A_t алгоритмов с количеством свободных параметров в каждом, не превышающим МКП_{*t*};
4. Каждому алгоритму из A_t назначаются значения параметров, доставляющие минимум ошибки обучения (эмпирической ошибки);
5. Из A_t выбирается алгоритм ALG_t с наименьшей ошибкой обучения (эмпирической ошибкой);
6. Для алгоритма ALG_t вычисляется ошибка тестирования (оценка ошибки обобщения) $E(ALG_t)$;
7. $t := t + 1$;
8. Если $t > T$, то переход на шаг 9, иначе переход на шаг 2;
9. $ALG = \arg \min_{ALG_t} E(ALG_t), t = 1, \dots, T$.

Описанный алгоритм не ограничивает методы создания множества алгоритмов, настройки параметров алгоритмов, разделения примеров на обучающие и тестовые наборы, выбора МКП. Алгоритм служит для поиска других алгоритмов с заданными характеристиками. С учетом всего этого представленный алгоритм может считаться метаалгоритмом.

Для проверки гипотезы наличия эффекта структурного переобучения был выполнен ряд экспериментов по восстановлению модельных функций (таб. 3). Результаты модельных экспериментов приведены в таблице 4. Эксперименты проводились с помощью симулятора искусственных нейронных сетей Neuro-Genesis [16]. В каждом эксперименте создавалось 1000 нейронных сетей с различной структурой. На структуры нейронных сетей накладывались только общие ограничения. Ограничивалось количество скрытых слоев. Разрешались структуры нейронных сетей содержащие от 1 до 3 скрытых слоев. В части экс-

периментов ограничивалось максимальное суммарное количество синапсов в нейронной сети. Количество нейронов явно не ограничивалось. Каждый вариант структуры нейронной сети оценивался перекрестным тестированием (bootstrap-оценка). В качестве решения выбиралась нейронная сеть с наименьшей ошибкой тестирования. Затем нейронная сеть тестировалась на дополнительном наборе из 1000 примеров. Вероятность ошибки на дополнительных примерах принималась за приближенное значение вероятности ошибки на произвольных примерах.

Таблица 3. Модельные функции

Обозначение	Определение
F ₁	$F_1(x_1, x_2, x_3, x_4) = \begin{cases} 1, & \text{если } x_1x_2 > x_3x_4 \\ 0, & \text{в противном случае} \end{cases}$
F ₂	$F_2(x_1, x_2, x_3, x_4) = \begin{cases} 1, & \text{если } \max\{x_1, x_2\} > \max\{x_3, x_4\} \\ 0, & \text{в противном случае} \end{cases}$
F ₃	$F_3(x_1, x_2, x_3, x_4, x_5, x_6) = \begin{cases} \text{sign}(x_5), & \text{если} \\ \min\{x_1, x_2, x_1x_2\} > \min\{x_3, x_4, x_3x_4\} \\ \text{sign}(x_6), & \text{в противном случае} \end{cases}$ $\text{sign}(x) = \begin{cases} 1, & \text{если } x > 0 \\ 0, & \text{если } x \leq 0 \end{cases}$
F ₄	$F_4(x_1, x_2, x_3, x_4) = \begin{cases} \text{sign}(x_1x_2), & \text{если } x_1x_2x_3x_4 > 0 \\ \text{sign}(x_3x_4), & \text{в противном случае} \end{cases}$

Таблица 4. Результаты экспериментов по восстановлению модельных функций

Восстанавливаемая функция	Кол-во примеров	Отсутствие ограничений на кол-во синапсов		Наличие ограничений на максимальное кол-во синапсов	
		Минимальная ошибка тестирования	Ошибка на произвольных примерах	Минимальная ошибка тестирования	Ошибка на произвольных примерах
F ₁	200	19%	60%	30%	32%
F ₂	300	18%	43%	19%	19%
F ₃	500	22%	53%	32%	33%
F ₄	500	21%	54%	22%	26%

Из таблицы 4 видно, что оптимизация структур нейронных сетей в отсутствии ограничений на максимальное количество синапсов приводит к большому

отличию вероятности ошибки на тестовых примерах от вероятности ошибки на произвольных примерах. При введении ограничений на максимальное количество параметров вероятность ошибки на тестовых примерах намного точнее приближает вероятность ошибки на произвольных примерах. Эффект “структурного переобучения” ярко выражен во всех 4-х модельных экспериментах.

5. Заключение

Выбор классификатора из множества вариантов по результатам тестирования ведет к увеличению длины доверительных интервалов надежности. В статье показаны комбинаторно-вероятностные причины данного явления. Также установлено наличие эффекта подстройки под тестовые примеры, что приводит к занижению оценок ошибки обобщения. Все это затрудняет селекцию алгоритмов классификации или регрессии по результатам тестирования. В качестве альтернативы предлагается ввести дополнительный уровень настройки алгоритма поиска классификатора или регрессора. Предлагается метод максимизации надежности алгоритмов классификации и регрессии, использующий ограничения на количество настроечных параметров в алгоритмах. В предложенном методе поиск ведется среди алгоритмов, удовлетворяющих ограничению числа параметров, селекция выполняется по ошибке обучения, по окончании поиска выполняется тестирование выбранного алгоритма. При необходимости ограничение на максимальное количество параметров меняется, и поиск повторяется снова. При таком подходе можно ограничиться вычислением 3-10 оценок ошибки обобщения по ошибке тестирования, что практически не приведет к увеличению длин доверительных интервалов или смещению оценок.

Литература

1. Эшби У.Р. Введение в кибернетику — М.: УРСС, 2005. — 432 с. — ISBN: 5-484-00031-9
2. Highlayman W. H. The design and analysis of pattern recognition experiments // *Bell System Technical Journal*. — 1962. — vol. 41. — p. 723—744.
3. Нефедов В. Н., Осипова В. А. Курс дискретной математики — М.: МАИ, 1992. — 263 с.
4. Гмурман В. Е. Теория вероятностей и математическая статистика — М.: Высш. шк., 2003. — 479 с.
5. Кендалл М., Стьюарт А. Статистические выводы и связи — М.: Наука, 1973. — 900 с.
6. Хомич, А. В., Жуков Л. А. Метод эволюционной оптимизации и его приложение к задаче синтеза искусственных нейронных сетей // *Нейрокомпьютеры: разработка, применение*. — 2004. — № 12. — С. 3—15.
7. Tukey J. W. Comparing individual means in the analysis of variance // *Biometrics*. — 1949. — № 9. — p. 99—114.
8. Andrew Y. Ng. Preventing overfitting of cross-validation data // *Proc. 14th International Conference on Machine Learning*. — San Mateo, CA, USA: Morgan Kaufman, 1997. — p. 245—253.
9. Klockars A. J., G. Sax G. Multiple Comparisons — Sage Publications, 1986. — 88 p.
10. Lawrence S., Giles C. L. Overfitting and neural networks: Conjugate gradient and backpropagation // *Proceedings of the IEEE International Conference on Neural Net-*

А.В. Хомич

- works (IJCNN'2000). — IEEE Press, 2000. — p. 114—119.
11. Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection // IJCAI. — San Francisco, California, USA: Morgan Kaufmann, 1995. — p. 1137—1145.
 12. Efron B. The Jackknife, the Bootstrap, and other resampling plans — Philadelphia, USA: SIAM, 1982. — 92 p. — ISBN 0-89871-179-7.
 13. Bartlett P. L. For valid generalization, the size of the weights is more important than the size of the network // Advances in Neural Information Processing Systems 9 (1996). — USA: MIT Press, 1997. — p. 134—140.
 14. Rissanen J. Modeling by shortest data description // Automatica. — 1978. — №14. — p. 465—471.
 15. Ежов А. А., Шумский С. А. Нейрокомпьютинг и его применения в экономике и бизнесе. Сер. учебники экономико-аналитического института МИФИ. Под ред. проф. В. В. Харитонов. М.: МИФИ, 1998. — 220 с.
 16. Хомич А. В. А. с. 2005611168 РФ. Программа для ЭВМ “Neurogenesis” №2005611168; Опубл. 24.02.2005.

Статья поступила 14 мая 2006 г.
После доработки 20 сентября 2006 г.

Самоорганизация в семантическом пространстве глаголов естественного языка

Коршаков А. В.

ВНЦ «Курчатовский институт», Москва, parano@mail.ru

Аннотация. Объектом анализа в статье являются множества английских глаголов. Глаголы рассматриваются как объёмные тела в некоем «семантическом» пространстве, на котором введена определённая метрика. В основе определения расстояний в упомянутом пространстве положено число существующих адекватных взаимнообратных переводов у обрабатываемых слов-глаголов на ряд (13) других индоевропейских языков. Целью такого исследования было подтвердить или опровергнуть существование нерегулярных структур в заполнении объёмными телами (глаголами) или кластеров, а также прояснить особенности строения семантического пространства, влияющие на расположение в нём слов. Для этих целей использовалась программная эмуляция нейронной сети Хопфилда, в которой каждый нейрон соответствовал одному из рассматриваемых глаголов. Веса нейронов задавались в соответствии с «расстоянием» между понятиями в смысле выбранной метрики. Рассмотрено большое число слов, среди которых выделяется группа «многозначных» и группа с «конкретным» значением или смыслом. Причём последняя группа не образует очевидных кластеров в пространстве глаголов. Глаголы естественных языков, символизирующие набор базовых, имеющих особый смысл для повседневной жизни действий, на которые способен человек, распределены равномерно по семантическому пространству.

1. Введение

Компьютерная обработка текста и человеческой речи, в прикладном смысле, сталкивается со многими проблемами, связанными с многозначностью слов. Одним из аспектов этой проблемы является (возможно, наиболее острый её аспект) работа с глаголами, так как многие из них и, особенно группа наиболее часто встречающихся в речи, имеет множество переводов с одного языка на другой. Например, английский глагол «Take» («*Брать*») может иметь более чем 20 различных выражений в других языках.

Всё множество глаголов может быть упорядочено в соответствии с убыванием числа возможных переводов на иностранные языки, и только около 300-й позиции этого списка число различных переводов внутри семейства индоевропейских языков падает до 2-4. Возможно, что такая картина сохраняется и для языков других семейств и структур. Такое явление легко прослеживается при анализе многоязыковых online-словарей, существующих в Интернете. Выбор необходимого и адекватного перевода из набора, который предоставляет словарь (в случае перевода с одного языка на другой), – в некоторых случаях достаточно трудная задача для человека, а компьютер здесь сталкивается и вовсе с непреодолимыми трудностями. Человек при распознавании речи использует, например, контекст (а ещё интонации голоса, мимику и многое другое), в котором произнесено слово, чтобы исключить двусмысленность. Под контекстом здесь следует понимать некий блок окружающего слово текста или состояние окру-

жающей среды (например, при анализе смысла одиночного, короткого высказывания – слова). Компьютер также должен быть способен решать подобные задачи. Это требует формализации процедуры обработки нечёткой информации, содержащейся в такой категории как контекст.

Обычно слова рассматриваются только как состоящие в определённом словаре, и никаких иных свойств, кроме того, что они собственно означают, и ещё, быть может, их порядкового номера в последовательности, упорядоченной в соответствии с алфавитом, им не приписывается. Такие словари, составленные по принятым в современном мире правилам, представляют собой список состоящих в них слов и статей, относящихся к ним. Такие пары «слово-статья» более или менее независимы друг от друга. В лучшем случае их связь выражается в указании синонимов друг друга. Более тонкие взаимоотношения слов и представленных ими понятий рассматриваются как явления, проявляющие себя только в огромных корпусах текста, содержащих миллионы слов, и нередко только в статистическом смысле [1]. Устроенная в соответствии с этими принципами процедура работы с языковой информацией кажется трудоёмкой и не слишком удобной, так как каждый раз приходится обращаться к большим объёмам текстовых данных, имеющих лишь опосредованное отношение к выполняемой прикладной задаче, – например, переводу (межъязыковому) конкретного текста с конкретным смыслом. Трудоёмкость становится очевидной при рассмотрении задач, требующих решения «на лету», например, в задачах интерпретации речи. Необходим более лаконичный способ представления лингвистической информации. Есть основания полагать, что такой способ существует, так как кажется сомнительным, чтобы человеческий мозг, вполне успешно справляющийся с оговоренным кругом задач, каждый раз при работе с речью прибегал бы к статистическому анализу всей, хранящейся в памяти лингвистической информации.

Можно выделить два пути, ведущих к получению данных относительно вида представления лингвистической информации в сознании человека. Среди существующего множества методов изучения нервных процессов наиболее эффективными можно назвать техники трехмерной визуализации активности нервной ткани [2, 3]. Это первый путь. Он состоит в изучении нервной системы, например, при помощи томографов ядерного магнитного резонанса. Он даёт определённые знания о процессах, протекающих в мозге человека во время обработки лингвистической информации, но сопряжён с целым рядом трудностей различного характера. Другой путь заключается в том, чтобы исследовать непосредственно продукт этих (языковых) процессов, то есть речь, учитывая фактор представляемой в ней информации.

Преыдушие исследования [4] показывают, что слова и отражаемые ими понятия, связаны сильнее, а, главное, глубже, чем это может показаться при рассмотрении структур «словарного» типа. Нужно скорее говорить не о жёсткой связи слов и отражаемых ими понятий, а об их «нечёткой» связи, не оперируя при этом только терминами «одно и тоже», для синонимов, и, «означают существенно разное», для слов, не имеющих общих значений даже в переносных смыслах.

В качестве подтверждения существования сильной связи слов, их смыслов и даже составляющих их частей, можно привести тот факт, что частота использо-

вания символов языка, то есть составных частей речи (звуки речи, фонемы, собственно слова) в различных языках (в [5] это показано на примере слов повседневного лексикона индоевропейской группы языков), подчиняется одинаковой и строгой функциональной зависимости [5, 6, 7]. Кривые частот встречаемости символов монотонно падают от максимального до почти нулевого значения, соответствующего наиболее редкому символу. Кривые, полученные, более чем для 1000 слов упомянутой группы языков, совпадают с высокой точностью, и рассогласования не превышают 5%, что укладывается в коридор статистической погрешности [8]. Для каждого из рассматриваемых языков последовательность символов специфична, что является следствием различия языков.

Употребление определённых фонем в речи является подписью каждого языка. Некоторые языки избегают употребления звуков, свойственных оставшемуся большинству.

Такие зависимости демонстрируют строгую математическую основу процесса формирования слов. Вероятнее всего, алгоритмы, применяемые мозгом для распознавания и генерации речи, независимы от конкретного языка, и могут настраиваться на определённую кодировку фонетических символов с различной степенью эффективности.

2. Многообразие глаголов в тезаурусах естественных языков

Человеческий язык или естественный язык – это открытая система, готовая принимать или «изобретать» новые слова или же отбрасывать старые слова. Множество существующих слов, имеющих специальное значение, могут быть просто неизвестны среднестатистическому пользователю лексической и семантической основ языка. То есть, язык слишком мощное средство, чтобы охватить его целиком. При исследовании такой сложной и динамичной системы, следует каким-то образом ограничить область анализа, которая, тем не менее, должна содержать, основные свойства языка. Одним из путей такого ограничения является подход, при котором ограничиваются работой с малой частью языка, соответствующей какой-то конкретной прикладной проблеме. Например, если иметь в виду задачу распознавания речи, то такой малой «областью интересов» семантики языка могут быть имена числительные и географические названия. Область применения таких систем распознавания – автоматические информационные бюро в аэропортах и голосовой набор номеров в аппаратах сотовой связи. К сожалению, на сегодняшний день, технические решения, относящиеся к этим областям применения распознавания, используют некие механистические принципы, а не внутренние закономерности языка, и, как правило, работают не идеально.

Возможно, что это явление есть следствие неверного выбора области семантики, корректно представляющей язык и отношение между его частями – словами. Для успешного решения задачи предлагается использовать критерием выбора подгрупп семантик для анализа не критерий «области практического применения», а критерий «исчерпывающего описания законов взаимодействия элементов языка внутри системы». То есть, необходимо для начала сориентировать русло решения от прикладных проблем в сторону теоретического обоснования решения.

Кажется разумным, в этом контексте, обратиться к такому словарному множеству языка как глаголы.

Глаголы составляют от 20% до 25% всех слов (при рассмотрении выборки из 10000 наиболее часто используемых слов английского, немецкого, голландского и французского языков) [9]. В противовес именам числительным глаголы чрезвычайно полисемичны (многозначны), то есть, их употребление возможно в языке во множестве значений и контекстов. Это является одним из обстоятельств (а может и основным обстоятельством), существенно усложняющим разработку технических средств и программного обеспечения обработки текста и речи. Сюда относятся все виды устройств и программного обеспечения, работающие только с текстом, только с речью, их комбинации, программы по генерации текста по распознанной речи и генерации речи по имеющемуся текстовому файлу. С другой стороны, если посмотреть на явление полисемии (многозначности), как на характеризующее язык или мышление вообще, то можно использовать его для построения системы отношений между глаголами, а в последствии и между словами вообще, что позволит построить более прогрессивные алгоритмы для работы с естественным языком и алгоритмы понимания человеческой речи.

3. Пространство глаголов

Рассмотрим множество наиболее распространённых в речи глаголов (например, для английского языка), таких как: *TAKE, GET, SET, GO, PUT, RUN, MAKE, MOVE, HOLD, COME, DO, TURN...* Эта последовательность может быть переведена на русский язык как: *БРАТЬ, ПОЛУЧАТЬ, УСТАНАВЛИВАТЬ, ИДТИ*, и т. д. При попытке продолжить эту цепочку для русского языка, используя любой англо-русский словарь, можно получить представление о степени полисемии этих слов.

Переведём эти слова на родственные индоевропейские языки, используя обычную процедуру перевода и классические словари, составленные представителями лингвистической науки. Подобная процедура может быть выполнена в полуавтоматическом режиме [10]. После чего, аналогичным образом, с использованием словарей родственных языков полученные слова переведём обратно на английский язык. В результате таких действий мы получим множество английских глаголов, мощностью значительно большей, чем исходное. После нескольких шагов такой процедуры процесс расходится и уже на втором шаге мы имеем около 300 слов, при начальных 15. При сопоставлении массивов «взаимобратных переводов» становятся очевидными тонкие и чрезвычайно густые связи между глаголами внутри языка. Число переводов для каждого слова можно рассматривать как меру в некотором пространстве. Вообще говоря, для лучшего представления пространственных отношений между словами, как смысловыми понятиями в некоем абстрактном «смысловом» пространстве, удобно представлять каждое понятие как объёмное тело, в общем случае произвольной формы. Тогда всё смысловое пространство можно представить заполненным этими телами, которые по некоторым признакам, как выяснится в дальнейшем, могут быть сгруппированы в кластеры. Отдельные тела могут смыкаться или даже взаимно проникать. Можно представить и некоторые незапол-

ненные части пространства понятий, как области, которые не соответствуют никакой фонетической конструкции естественного языка. В первом, обсуждаемом здесь, приближении предполагается примерно сферическая форма тел понятий. Таким образом, можно ввести в рассмотрение понятие расстояния между словами-понятиями. Расстояние, по определению, это геометрическое понятие, содержание которого зависит от объектов, для которых оно определяется. В соответствии с этим, расстояние между двумя точками есть длина соединяющих их прямой, в общем случае, многомерной [11]. Интуитивно расстояние между словами-понятиями можно рассматривать как длину отрезка, соединяющего точки центров сфер, как виртуальных образов слов-понятий в семантическом пространстве.

Однако, если существуют расстояния, необходимо ввести также метрику, посредством которой это расстояние может быть измерено. Метрика вводится следующим образом. Число взаимобратных переводов $\rho_{i,j}$ для двух разных глаголов i, j исходного списка можно рассматривать как характеристику расстояния между ними.

Пользуясь такой посылкой, как введение концепции «расстояния», можно сформировать «матрицу расстояний» для достаточно большого множества слов-понятий, которые могут попасть в поле рассмотрения.

Матрица расстояний, как основная характеристика обсуждаемого множества глаголов, и, как следствие, пространства, в котором они располагаются, строилась для каждой доступной для анализа пары глаголов для 13 родственных языков (болгарского, русского, чешского, польского, латвийского, шведского, немецкого, голландского, французского, итальянского, румынского, албанского и греческого). В качестве базового языка для рассмотрения, был выбран английский язык, но мог быть выбран и любой другой. Результат представляет собой квадратную, симметричную, положительно определённую матрицу, с длиной грани, равной числу рассматриваемых слов.

Для удобства работы с такими расстояниями, данными только в «дискретном» представлении, их желательно перевести в континуальное пространство. Это также способствует снижению чёткости определений значений слов и связей между ними, что более соответствует действительности («границы» понятия «размываются» и перестают быть резкими).

Для дальнейшего необходимо ввести ещё одну дополнительную величину – «объём» слова N_i . Объём равен общему числу переводов на все родственные языки для одного конкретного глагола. После расчёта этих величин для каждого из слов определим расстояние $D_{i,j}$ между глаголами-понятиями по следующей формуле:

$$D_{i,j} = -\ln \left(\frac{\rho_{i,j}^2}{N_i N_j} \right), \quad (1)$$

где $\rho_{i,j}$ – число взаимобратных переводов для глаголов i, j , N_i, N_j – объёмы этих слов, $D_{i,j}$ – расстояние между словами-понятиями. Деление на объёмы слов

производит нормировку расстояния, а функция логарифма выполняет здесь роль «сглаживателя» границ значений.

Таким образом, в нашем распоряжении имеется массив расстояний между точками в пространстве, которые это пространство и определяют. Для целей поиска возможных интересных топологий внутри этого пространства можно поставить задачу определения координат этих точек, каждая из которых соответствует понятию. Из сказанного следует, что мы имеем набор, вообще говоря, n -мерных объёмных тел, расположенных в n -мерном пространстве, о которых известны только расстояния между их «геометрическими» центрами. Необходимо найти координаты центров этих тел. Для решения введем соответствующую систему координат. Воспользуемся декартовой 4-мерной системой координат, причём предполагаем, что 4-я координата будет принимать значения, малые по сравнению с первыми 3-мя (а в отдельных «стабильных» случаях вообще принимать только нулевые значения), неся, таким образом, информацию о флуктуациях системы. Результаты расчётов, приведённые ниже, показали, что такой выбор был оправдан в том смысле, что «понятия» в большинстве случаев размещались и в более жёстких рамках пространств с меньшей размерностью.

Далее, по определению, в декартовых координатах расстояние между i -й и j -й точками равно:

$$D_{i,j} = \sqrt{(x0_i - x0_j)^2 + (x1_i - x1_j)^2 + (x2_i - x2_j)^2 + (x3_i - x3_j)^2}, \quad (2)$$

где $x\{k_i\}$ – k -я координата i -й точки, $D_{i,j}$ – расстояние между i -й и j -й точкой (словом), определяемое заданным элементом «логарифмированной матрицы расстояний» по формуле (1).

Выражение (2) даёт систему нелинейных уравнений, которая с увеличением числа глаголов, начиная с некоторого момента, становится переопределённой. Её решением и будут искомые координаты.

Система решалась методом простой итерации:

$$\mathbf{x}^{(k+1)} = \mathbf{F}(\mathbf{x}^{(k)}). \quad (3)$$

Для поиска решения переопределённой системы использовались дополнительные соображения, связанные со способом выбора уравнений для решения в каждом конкретном шаге процесса, связанным с добавлением в рассмотрение нового слова. В решаемой на каждом этапе процесса системе, всегда присутствовало уравнение, отвечающее расстоянию от слова (из множества базовых слов-понятий), называемого «главным» («центром кластеризации», «кластеро-образователем»), это слово выбиралось экспериментатором, и от него начинались все «построения». Таким образом, при решении системы первым рассматривалось уравнение относительно расстояния от центра кластеризации. «Вторым» типом уравнений системы были соотношения для расстояний между добавляемым понятием, и словами, уже принадлежащими кластеру, безотносительно к центру.

Уравнения, соответствующие словам, «бесконечно» удалённым от центра данного смыслового кластера (или центра тела главного слова) давали в ходе

расчётов недопустимые значения координат или были значительно удалены от центра кластера. «Бесконечно» удалённые слова – слова, расстояния между которыми, согласно матрице расстояний, равны величине, принятой в расчётах за бесконечность. Фактическое значение «условной бесконечности» выбиралось как много большее, чем максимальное расстояние, присутствующее в матрице расстояний.

Таким образом, слова, не связанные с данным «кластеризующим» центром, «не подходили кластеру» и исключались из рассмотрения. Однако исключались не абсолютно, и при добавлении в кластер нового слова, из еще не обчисленного списка, предъявлялись системе вновь, в надежде, что только что добавленное слово стабилизирует решение. И иногда действительно встречались слова, которые могли присутствовать в кластере (то есть иметь значения координат близкие к таковым для центра кластеризации) только в случае присутствия своего «соседа» – например слова «*Drag*» и «*Haul*». («*Тащить*» и «*Тянуть*») Уравнения, соответствующие словам, которые отвергались всё время, вплоть до окончания процесса всегда оказывались чуждыми по смыслу данной выборке или кластеру, – например, из выборки «*Destroy*» («*Уничтожать*») на бесконечность всегда уходило слово «*Connect*» («*Соединять*»).

После выбора уравнений для решения запускался итерационный процесс, который продолжался при постоянном контроле сходимости к стабильному решению.

После того, как очередное решение системы уравнений было получено, то есть, было получено очередное приближение кластера с выбранным «ключевым словом», к нему применялся ряд критериев, по которым определялось, соответствует ли полученному приближению определение понятия кластера как совокупности точек, отвечающих словам, группируемым по принципу похожего смысла. Приведём некоторые из них.

Решаемая система уравнений переопределена, и, следовательно, получить точное решение принципиально невозможно. Поэтому, прежде всего, проверялось, не превышает ли погрешность полученного решения некоторого интуитивного порога, выставяемого оператором. Таким порогом чаще всего выбиралась погрешность в 100% в расчёте на одно кластеризуемое слово-понятие, содержащееся в кластере. Это означает, что точка, соответствующая слову, проверяемому на совместимость с кластером, может отклониться от своего среднего значения решения на его удвоенную величину и всё равно остаться членом кластера. Такой выбор «порога принадлежности» является эмпирическим и позволяет получать смысловые агрегации, состав которых не противоречит истинному значению глаголов.

К словам в кластере также применялся критерий отсутствия доминирующей роли осцилляций в кластере. Последнее означает, что осциллирует не более чем определённый процент от всех слов кластера. Под осцилляциями понимается поочерёдное принятие словом (центром объёмного тела) 2-х наборов координат от итерации к итерации. Этот процесс купировался добавлением нового слова в кластер, обладающего для осциллирующего, стабилизирующими свойствами. Такое, однако, происходило далеко не всегда.

При не выполнении хотя бы одного из этих критериев вновь добавляемое к кластеру слово считалось непригодным и отбрасывалось, в соответствии с методом, описанным выше.

Если такую вычислительную процедуру применять «локально», то есть, не выходя за определённые границы по количеству одновременно рассматриваемых уравнений, то она сходится к разумному решению в пределах допустимых величин погрешности. Разрешение такой ограниченной системы уравнений даёт величины координат точек для не слишком многочисленной группы глаголов. Стоит отметить, что при попытке выйти за границу, обусловленную, возможно, некоторыми внутренними свойствами рассматриваемой группы, наблюдается катастрофическая потеря точности вычислений и итеративный процесс расходится. В ходе экспериментов было рассмотрено большое количество таких образований или групп глаголов (кластеров). Оказалось, что все они обладают двумерной геометрией, или, (иногда, в зависимости от выбора центра кластеризации) различаются взаимным расположением и координатами точек внутри каждой группы и достаточно точно вписываются в плоскость. Из-за явления расходимости вычислительной процедуры при рассмотрении больших групп понятий расчёты носили локальный характер, и, следовательно, об ориентации полученных плоскостей в пространстве, без дополнительных допущений относительно привязки к плоскостям системы координат, нельзя говорить даже в относительном смысле.

В качестве примеров описанных плоскостных топологий в пространстве понятий можно привести, например, кластер «Уничтожать-Ломать», где наблюдалось, как в ходе итерационного процесса точка «*BREAK*» («*Ломать*») притягивала в свою группу точки пространства, отвечающие следующим понятиям: «*CRUSH*», «*SHATter*», «*WRECK*», «*deSTROY*», «*SMASH*»¹, которые потом и вошли в результирующий «кластер». Также характерна и стабильность кластера «Тянуть-Ташить» («*PULL*»),² с членами своего семейства: «*TUG*», «*DRAW*», «*DRAG*», «*HAUL*», «*TRAIL*»². Хорошо видно, что приведённые группы слов, внутри каждой группы, несут вполне определённо общее значение, и не являются при этом синонимами, в словарном смысле этого слова, по крайней мере в некоторых языках из группы рассмотренных. Это кажется особенно замечательным, если учесть, что такие образования были получены из стандартных словарей, путём применения к ним чисто математических процедур. Структуры кластеров «Уничтожать-Ломать» и «Тянуть-Ташить» представлены на рис. 1,2.

Можно думать, что подобные группы слов просто «покрывают» некоторую область пространства понятий, которая отвечает определённой смысловой нагрузке, передать каковую посредством языка от одного человека к другому, можно используя одно из слов, входящих в группу. Соответствующие слова другого языка «покрывают» ту же самую область смыслового пространства.

¹ Соответственно: «*Давить*», «*Раздробить*», «*Крушить*», «*Уничтожать*», «*Разбивать*».

² Соответственно: «*Дёргать*», «*Волочить*», «*Тащить*», «*Тянуть*», «*Тащиться (идти сзади)*».

Самоорганизация в семантическом пространстве глаголов

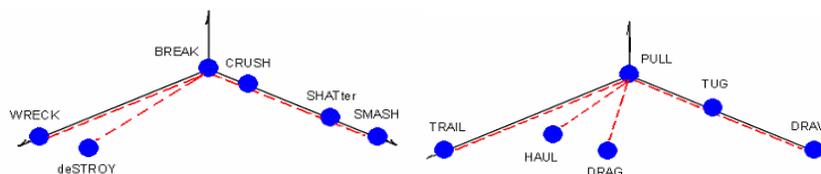


Рис. 1. Кластер «Уничтожать-Ломать»

Рис. 2. Кластер «Тянуть-Ташить»

Согласно исследованиям методами магнитно-резонансной томографии (МРТ) высшей нервной деятельности людей, разговаривающих на двух разных языках с раннего детства, корковые зоны, активные при речи на первом или втором языке в значительной степени совпадают, что может говорить об обращении к одному «полно» памяти при формировании высказываний. И хотя этот факт всё ещё является предметом оживлённых дискуссий, любопытно обратить внимание на следующее. Так в [12], при помощи контрастной функциональной МРТ по определению уровня кислорода в крови исследовалась нервная деятельность людей бегло говорящих на двух сильно различных языках – английском и мандарине (диалект китайского). Картина активности коры головного мозга регистрировалась в то время, как субъекты называли или договаривали слова на обоих языках, по приведённому визуальному стимулу или высказанному началу слова. При работе на каждом из двух языков наблюдалась активность в префронтальной, височной, теменной и добавочной моторной областях. Во всех случаях местоположение пиковой активности совпадало. По крайней мере на уровне работы с единичными изолированными словами общая степень регистрируемой активности, или количество активных вокселей¹, в опытах была характерна для вызванной активности. Разница в степени активности для разных языков имела, однако, существенной разницы в местоположении очага активности между двумя языками выявлено не было (рис. 3). Особенно замечательно то, что разницы не было выявлено, в том числе и в префронтальных «лингвистических» областях, которые традиционно считаются ответственными за генерацию смысловой составляющей речи. Такая картина активности не зависит от возраста и времени начала полноценного использования второго языка.

То, что в каждой смысловой группе или кластере содержится несколько слов, возможно, является причиной того, что часто бывает трудно подобрать подходящий перевод для конкретного слова. Слова, не имеющие «общего» смысла со словами рассматриваемого кластера, расположены «бесконечно» далеко от кластеризующего центра (расстояние между такими словами равно «условной бесконечности») и в ходе итерационного процесса исторгались из текущей смысловой структуры. С другой стороны, доступность множества слов, состоящих в «соседстве», делают возможными несколько версий перевода определённого предложения.

¹ Воксель – объёмный эквивалент пикселя (элементарного представимого в ЭВМ элемента изображения)

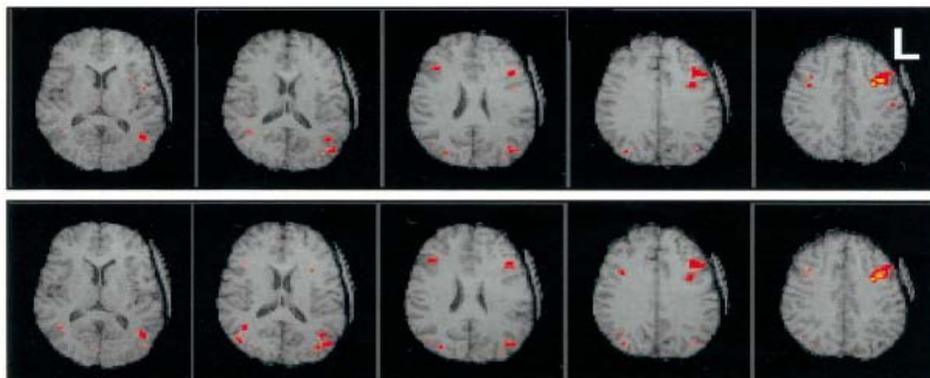


Рис. 3. Типичные картины активности при использовании английского и мандарина (Соответственно верхняя и нижняя линейки). Левое полушарие с правой стороны [12].

Как бы то ни было, при возможности локального картирования пространства понятий, используя метод решения системы нелинейных алгебраических уравнений, полное картирование или, что, то же самое, нахождение связей между различными кластерами, возможно либо путём сопоставления локальных данных, либо с применением некоторой другой вычислительной процедуры. Причины локальности геометрической модели, возможно, связаны с тем, что пространство понятий, само по себе подчиняется иной, чем евклидова, геометрии. Быть может, более уместным тут было бы рассмотрение точек, соответствующих словам-понятиям, на сферической поверхности.

Для картирования всего пространства понятий был выбран путь, связанный с использованием нейронной сети Хопфилда.

4. Сетевая система глаголов

Метод подразумевал следующие действия с имеющейся в распоряжении нормированной матрицей расстояний. Для рассмотрения пространства понятий, в котором состоят глаголы и предоставленного для анализа списка, строилась полносвязная нейронная сеть Хопфилда, каждый нейрон в которой соответствовал конкретному глаголу списка, то есть число нейронов в сети (или размерность сети) было равно числу рассматриваемых слов. Активное или пассивное состояние нейронов говорило о присутствии глагола в конкретной группировке (кластере) или исключении его из неё. Величины или веса связей между нейронами выставлялись в соответствии с обучением, которое, в свою очередь проводилось на основе нормированной логарифмической матрицы расстояний между глаголами. Процедура обучения будет описана ниже. Из рассмотрения были исключены первые несколько английских глаголов рассматриваемого списка. Взаимосвязи с прочими словами, для этой группы глаголов слишком сложны для анализа в этом приближении, что было причиной чрезмерной «загруженности» сети значащими образами, и, вело к преобладанию «эпилептической активности» (все нейроны сети находились в активном состоянии) при функцио-

нировании сети среди всех возможных картин активности. Значения в речи глаголов, входящих в исключённую группу, часто модулируются приставками и предлогами, анализ которых для прочих слов не рассматривался в этом исследовании. Кроме того, из рассмотрения преимущественно исключались глаголы, участвующие в кластерах, полученных при исследовании пространства понятий при помощи системы нелинейных алгебраических уравнений.

Таким образом, все рассмотренное поле английских глаголов можно метафорически уподобить сетчатке глаза. В ней, как известно, существует область «острого зрения» и область «нормального зрения». При нейросетевом анализе основной упор делался на исследование свойств «нормальной области», то есть таких частей пространства, которые заполнены словами, не вошедшими ни в один из смысловых кластеров. Топологические структуры, полученные методом решения нелинейных алгебраических уравнений, можно уподобить области «острого зрения», где взаимосвязи между отдельными частями «изображения» хорошо видны.

Начальная выборка для английского языка составляла 256 глаголов, после исключения, оговоренных выше подмножеств, остаток составил около 60 слов, которые использовались как центры кластеров при попытках их конструирования нейросетевым методом. Конкретные слова выбирались случайным образом в каждом из экспериментов, причём предпочтение отдавалось словам, не участвовавшим в кластерах, полученных вышеописанным способом.

С сетью проводилось две серии опытов для каждой выборки слов. В первой серии опытов количество обучающих паттернов ограничивалось эмпирической формулой для максимальной ёмкости памяти сети Хопфилда:

$$L = \frac{N}{4 \cdot \ln(N)}, \quad (4)$$

где L – ёмкость памяти, N – число нейронов сети [13], $N = 60$.

Во второй серии сеть обучалась N обучающим паттернам для N нейронов. Каждый обучающий паттерн представлял собой «гипотетическое» множество слов с родственным смыслом. Тренировочные паттерны синтезировались в изоляции для каждого слова как «центра кластеризации». Не входящими в данную группу словами считались слова, расположенные бесконечно удалённо от центра кластеризации согласно нормированной логарифмической матрице расстояний (такие слова не имели общих переводов на другие рассматриваемые языки). Технически процедура формирования обучающих паттернов для нейронной сети выглядит следующим образом. Сначала выбиралось слово – центр кластеризации. В этой роли, как уже говорилось, по очереди выступают все слова рассматриваемой выборки. После чего формировалась «гипотеза» кластера, первоначально состоящая из всех остальных слов рассматриваемой выборки, кроме текущего центра кластеризации. После этого из гипотезы кластера изымались те слова, расстояния между которыми и словом-центром кластеризации, согласно нормированной логарифмической матрице расстояний было равно бесконечности.

Процесс обучения формирует зоны притяжения (аттракции) некоторых точек равновесия, соответствующих обучающим данным. При использовании ассоциативной памяти мы имеем дело с множеством обучающих векторов \bar{X} , которые, в данной работе, формировались вышеописанным способом, и, которые в результате проводимого обучения определяют расположение конкретных аттракторов, то есть конкретных картин активности с характерным набором активных нейронов, выражающих принадлежность соответствующих им слов рассматриваемому кластеру, символическим выражением которого является наблюдаемый аттрактор. Каждый нейрон имел биполярную пороговую функцию активации.

Для обучения сети использовалась классическая процедура обучения нейронных сетей Хопфилда – обобщённое правило Хэбба, в соответствии с которым [14]:

$$w_{i,j} = \frac{1}{N} \sum_{k=1}^p x_i^{(k)} \cdot x_j^{(k)}, \quad (5a)$$

$$T_i = - \sum_{k=1}^p x_i^{(k)}, \quad (5b)$$

где p – число обучающих векторов в матрице \bar{X} , $w_{i,j}$ – веса синапсов нейронов, T_i – пороги нейронов.

Фаза обучения сети Хопфилда ориентирована на формирование таких значений весов, при которых в режиме функционирования задание начального состояния нейронов, близких к одному из обучающих векторов, приводит в результате функционирования сети к стабильному состоянию, в котором реакция нейронов остаётся неизменной в любой момент времени.

В экспериментах исследовались сети с параллельной динамикой, то есть характеризующиеся синхронным функционированием нейронных элементов сети. При этом за один такт работы сети все нейроны одновременно изменяют своё состояние согласно формуле:

$$y_i(t+1) = F \left(\sum_{\substack{j=1 \\ j \neq i}}^N w_{j,i} \cdot y_j(t) - T \right), \quad (6)$$

где y – наблюдаемая в момент времени t активность сети, F – биполярная пороговая функция активации, T – текущее значение порога нейронов, а i пробегает все значения от 1 до N .

При функционировании сети можно наблюдать только устойчивые стационарные точки и циклы длиной два. Это свойство естественным образом испол-

зовалось при отождествлении картин активности сети с искомыми группами слов – кластерами.

В первых сериях опытов упор делался на ассоциативные функции нейронной сети Хопфилда. Так как ассоциативная память демонстрирует способность к коррекции, то при представлении тестовой выборки, например, отличающейся некоторым количеством битов на отдельных позициях вектора, нейронная сеть может скорректировать эти биты и завершить процесс классификации на нужном аттракторе. Следовательно, согласно этому режиму функционирования предполагалось, что, сеть по предоставленным примерам будет способна сформировать ассоциативные связи между словами согласно их смыслу, отражённо-му в матрице расстояний, по которой и формировались обучающие паттерны.

Вторая серия опытов, кроме ассоциативной функции, предполагала использование некой «креативности» сети. При использовании сети Хопфилда как ассоциативной памяти, в ней неизбежно формируется «паразитная память» - набор паттернов, которым сеть изначально не обучалась. Такое явление особенно характерно для сети, в которую записывают объём информации, превышающий номинальную ёмкость её памяти L . В этой серии опытов, наборы картин активности, относящихся к паразитной памяти, рассматривались как полноценные смысловые кластеры.

В обоих случаях после обучения анализировались полученные для каждого значения порога сети, одинакового для всех нейронов, картины активности. Для этого применялась специальная автоматическая процедура просмотра всех значений порогов сети и записи всех зарегистрированных паттернов активности сети.

Процедура тотального картирования памяти нейронной сети состояла в том, что для данного эксперимента при обучении, кроме значений массива весов, рассчитывался массив порогов нейронных элементов по формуле (5б). При минимальной величине порога T_{min} , содержащейся в этом массиве, сеть гарантировано переходит в «эпилептический» режим, когда активны все нейроны. При максимальной величине порога T_{max} , сеть находится в перманентной «коме», т.е. ни один нейрон не получает достаточно возбуждающих воздействий, чтобы превзойти порог и перейти в активное состояние. Далее, на каждом шаге процедуры картирования последовательно выбиралось значение порога из диапазона $[T_{min}, T_{max}]$, и присваивалось значению текущего порога каждого нейрона сети.

При использовании такой динамики сети результат работы обсуждаемых алгоритмов оказался одинаковым, независимо от серии опытов. Тотальный просмотр всей возможной активности сети, для всех серий опытов, не выявил смысловых агрегаций слов, содержащих более чем 2 слова. Следует напомнить, что здесь рассматривались слова, не вошедшие в смысловые кластеры. Типичные картины активности сети представлены на рис. 4. Показаны только информативные области паттернов. Красным отмечены активные нейроны, символизирующие присутствие соответствующего слова в кластере. Разумеется, картины активности сети, с количеством активных нейронов, больше 2-х, существовали, но они включали в себя «слова», никак не связанные между собой по смыслу, и содержали очень большое число составляющих, охватывающих практически всю выборку слов. Данное явление можно интерпретировать как следствие гомогенности распределения глаголов в смысловом пространстве.



Рис. 4. Паттерны активности кластеризующей сети.

Таким образом, за исключением области «острого» восприятия смысла, содержащего ограниченное количество, очевидно, наиболее применимых в повседневной речи, либо более значимых по какой-то другой причине, слов-глаголов, и стягивающих к себе слова с близким значением, остальное пространство понятий не содержит слов, несущих близкие значения и способных к группировке, по крайней мере, в рамках рассматриваемых моделей. Смысловые связи в этой области семантического пространства существуют только с «ближайшими» соседями, причём такие связи слабее, чем связи внутри смыслового кластера.

5. Выводы

В работе приведены два взаимодополняющих метода, работающих с трудно формализуемой лингвистической информацией. Цели функционирования этих

алгоритмов несколько различны. Формирование смысловых кластеров через анализ значений их координат в семантическом пространстве, полученных путём решения систем нелинейных уравнений, может продемонстрировать сильное сродство значений некоторых слов друг к другу. В то время как «нейросетевой» метод даёт представление о заселённости областей смыслового пространства, не занятого кластерами, словами, в эти кластеры не вошедшими.

Использование обоих методов может показать, что содержащиеся в мультязыковых словарях лингвистические данные могут быть систематизированы с применением к ним категории «смысла», который кодирует слово внутри человеческого языка, а не простого статистического упорядочения.

Слова-понятия могут быть размещены в некотором пространстве, к которому применим достаточно эффективный математический аппарат. При локальном рассмотрении некоторые области семантического пространства выглядят как плоскости в евклидовой геометрии.

Подобная процедура смыслового упорядочения кажется полезной для её применения при организации словарей электронных систем перевода и интерпретации речи. Если заметить, что человек, занимающийся переводом, переводит смысл текста, а не слова текста, то такое усовершенствование может улучшить качество предоставляемых электронными переводчиками услуг.

В отличие от, например, [15] в данной работе фонетическая форма слова вообще не рассматривалась. Тем не менее, в упомянутой статье показано, что слова могут быть эффективно упорядочены и по признаку «фонетической формы» корня слова, причём эффективнее, чем простая сортировка по алфавиту. Это поднимает вопрос о взаимосвязи между этими двумя пространствами – фонетическим (пространством комбинаций фонетических символов) и смысловым, или об отображении одного пространства на другое.

Литература

1. McEnery, T. and Wilson, A.: *Corpus Linguistics*, Edinburgh: Edinburgh University Press. (1996)
2. Price, C., Indefrey, P., van Turenhout M.: The Neural Architecture, Underlying the Processing of Written and Spoken Word Forms. In: Brown, C.M., Hagoort P. (eds.): *The Neurocognition of Language*. Oxford Univ. Press, NY (1999) 211–240
3. Simos, P.G., Breier, J.I., Fletcher, J.M., Bergman, E., Papanicolaou, A.C.: Cerebral Mechanisms Involved in Word Reading in Dyslexic Children: a Magnetic Source Imaging Approach. *Cerebral Cortex* 10(8) (2000) 809–816
4. Коршаков А. В.: Процедура построения пространства понятий как часть системы машинного перевода. *Нейроинформатика 2006, VIII Всероссийская Научно-Техническая Конференция*, Сборник научных трудов, часть 3, стр. 155–162.
5. Vvedensky V.L., Korshakov A.V., (2004). Visualization of the Basic Language Thesaurus. *Proc. VII International Conf. «Cognitive Modelling in Linguistics»*, Varna, pp. 308–313.
6. Введенский В. Л., Коршаков А. В.: Естественно упорядоченный алфавит индоевропейских языков. *Нейроинформатика 2004, VI Всероссийская Научно-Техническая Конференция*, Сборник научных трудов, часть 2, стр. 18–24.
7. Shannon C. E. A Mathematical Theory of Communication. *The Bell System Technical Journal*, Vol. 27, pp. 379–423, 623–656, July, October, 1948.

В.А. Коршаков

8. Введенский В. Л.: Математические закономерности словообразования в европейских языках. *Нейроинформатика 2005, VII Всероссийская Научно-Техническая Конференция*, Сборник научных трудов, часть 2, стр. 263-270.

9. Der Dokumenten-Server der Universitat Leipzig, <http://wortschatz.uni-leipzig.de/index.html>

10. Foreignword.com, The language site <http://www.foreignword.com>

11. *Математический энциклопедический словарь*. Москва «Советская энциклопедия» 1988.

12. Michael W. L. Chee, Edsel W. L. Tan, and Thorsten Thiel. Mandarin and English Single Word Processing Studied with Functional Magnetic Resonance Imaging. *The Journal of Neuroscience*, April 15, 1999, 19(8):3050–3056

13. Головкин В.А. *Нейронные сети: обучение, организация и применение*. Издательское предприятие журнала «Радиотехника». Москва 2001.

14. Осовский С. *Нейронные сети для обработки информации*. Москва «Финансы и статистика» 2002.

15. Введенский В. Л.: Сеть корней глаголов русского языка. *Нейроинформатика 2006, VIII Всероссийская Научно-Техническая Конференция*, Сборник научных трудов, часть 3, стр. 236-243.

Статья поступила 3 июля 2006 г.

После доработки 11 октября 2006 г.

К физической теории мышления: теория нейронных моделирующих полей

Перловский Л.И.

Гарвардский университет, leonid@deas.harvard.edu

Аннотация. В статье сделан обзор подходов к моделированию мышления в таких направлениях исследований как искусственный интеллект, распознавание образов, нейронные сети. Дается математическая формулировка нейронных моделирующих полей, объединяющая механизмы мышления в систему, соответствующую данным психологии и нейробиологии. Подчеркивается роль инстинкта к знанию в активном процессе познания окружающего мира.

Ключевые слова: мышление, физика, эмоции, концепции, сознание, инстинкт к знанию.

1. Физика и мышление

Методы нейронных сетей в настоящее время активно используются для решения разнообразных практических задач. Однако известные нейронные сети слишком просты для объяснения человеческого мышления. В этой статье мы попытаемся не просто улучшить методы для решения практических задач, но также сделать шаг к тому, чтобы объяснить мышление, как физический процесс, так же как Ньютон объяснил движение планет. Но возможна ли физическая теория мышления и в чём могла бы быть её суть? Предполагается, что, говоря о мышлении и мозге, мы обсуждаем одну и ту же физическую систему на разных уровнях описания. Эта ситуация для физики не нова, например, термодинамика и статистическая физика связаны подобным образом. Физическое описание на микроскопическом уровне (атомы, молекулы, элементарные частицы, струны) предпочтительно не всегда, часто феноменологическое описание более эффективно. Эйнштейн однажды заметил, что термодинамика его привлекает именно тем, что определена на феноменологическом уровне. Физическая теория отличается поиском основных законов, нескольких универсальных принципов, из которых математически следует описание экспериментальных данных в широкой области.

Проблема мышления обсуждалась в течение тысячелетий, от философов Древней Греции до современных математиков и исследователей в области когнитивных наук. И всё же терминология в этой области ещё не установилась. Слова: *мышление, мысль, воображение, эмоции, концепции, концепты* используются различным образом в повседневной жизни, также они не были чётко определены в когнитивной науке и являются предметом дискуссий [1]. Мы считаем, что стандартизированные определения уместны лишь после того, как теория создана (например, «сила» была определена ньютоновскими законами после тысяч лет менее точного использования). Поскольку теория мышления ещё соз-

даётся, в данной статье придерживаемся следующего порядка словоупотребления. Предлагаемые термины должны: (1) соответствовать современным дискуссиям в научной и математической среде, (2) соответствовать тысячелетним философским обсуждениям и принятому употреблению в культуре, (3) быть чёткими и описываемыми математически, и (4) отклонения от этих принципов и противоречия в употреблении терминов должны отмечаться и обсуждаться. В соответствии с определением в словаре [2], которое мы принимаем лишь как отправную точку, мышление включает сознательные и бессознательные процессы, основными среди которых являются не только мысль как концептуализация, но также восприятие, эмоция, воля, память, воображение; мышление формируется в мозге. Мнения о том, какие математические методы подходят для описания мышления, сильно различаются. Основатели искусственного интеллекта, включая Аллана Ньюэлла и Марвина Минского, считали, что для этой цели адекватна формальная логика [3,4] и никакие специальные новые математические методы не требуются. Противоположная точка зрения развивалась Брайаном Джозефсоном и Роджером Пенроузом, считавшими, что мышление не может быть понято на основе современной физики; для объяснения мышления потребуются новые неизвестные ещё физические законы и явления [5,6]. Рассматривались квантовые процессы, которые могут происходить в нейронах [5-8]. Некоторые авторы, рассматривая процессы классической физики в нейронах, считают, что небольшое число основных принципов нейронной организации играют фундаментальную роль, определяя основные законы мышления, и необходимо найти математическое описание этих основных законов. Этому мнению придерживаются Стефан Гроссберг, предложивший, что основные законы включают резонансное совпадение афферентных и эфферентных сигналов (идущих сверху-вниз и снизу-вверх) [9] и эмоциональную оценку концептуального содержания [10], Лотфи Заде, предложивший теорию гранул [11], Алекс Мейстель, разрабатывающий иерархическую многомасштабную организацию мышления [12]; Джералд Эдельман, предложивший отбор нейронных групп [13]; и автор данной статьи, предложивший инстинкт к знанию, эстетические эмоции и динамическую логику [8,14-16].

2. Логика и мышление

Длительное время считалось, что интеллект эквивалентен концептуальному пониманию и рассуждению. Часть этой «веры» состояла в том, что мышление основано на логике. Основатели искусственного интеллекта в 1950-х и 1960-х годах верили, что, основываясь на правилах логики, они вскоре создадут компьютеры, чей интеллект намного превзойдёт человеческий.

Начало этой истории обычно относят к Аристотелю, сформулировавшему правила логики [17]. Однако Аристотель не считал, что мышление работает логически; он открыл логику, как лучший способ аргументации, а не как теорию мышления. Например, в книге «Риторика для Александра» Аристотель перечисляет десятки тем, на которые Александру Македонскому приходится произносить политические речи [18]. На каждую тему Аристотель приводит два противоположных мнения (например, заключать мир или объявлять войну, использовать пытки при допросе или нет...). Для каждого из противоположных

мнений Аристотель даёт свои логические аргументы. Ясно, что логика, по Аристотелю, – это инструмент для выражения ранее принятых решений, а не механизм мышления. Логика может лишь быть использована для дедукции из основных принципов, но не способна указать, в чём эти принципы должны состоять. В последующие столетия тонкости мысли Аристотеля не всегда верно понимались. С развитием науки идея, что интеллект эквивалентен логике, стала захватывать умы. В XIX веке математики обратились к логике. Джордж Буль заметил противоречие в теории Аристотеля. Начиная от Аристотеля, логика основывалась на законе исключённого среднего (или исключённого третьего): каждое утверждение либо истинно, либо ложно, любая средняя альтернатива исключается [19]. Но Аристотель также подчёркивал, что логические утверждения не должны формулироваться слишком точно (например, не следует определять меру пшеницы – мешок – с точностью до одного зерна), что язык подразумевает необходимую неопределённость.

Буль считал, что следует исправить противоречие между точностью закона исключённого среднего и нечёткостью языка. Родилась новая область математики, *формальная логика*, которой занимались Джордж Буль, Готтлоб Фреге, Георг Кантор, Бертран Рассел, Давид Гильберт, Курт Гёдель. Логика «выбросили» нечёткость языка и основали формальную математическую логику на законе исключённого среднего. Гильберт разрабатывал направление названное формализмом, отрицающее интуицию, и хотел определить объекты науки формально, в терминах логических аксиом. Он был уверен, что его логическая теория также описывает механизмы мышления: «Основная идея моей теории доказательств – это не что иное как описание нашей деятельности понимания, это протокол правил, согласно которым наше мышление в действительности происходит» [20]. В 1900 г. он сформулировал знаменитую Entscheidungsproblem: определить систему логических правил, достаточную для доказательства всех прошлых и будущих математических теорем. Это подразумевало формализацию научного творчества и всего человеческого мышления.

Почти сразу, после того как Гильберт сформулировал свою программу формализации, обнаружилась первая дыра. В 1902 г. Рассел показал несостоятельность формальной процедуры, определив множество R следующим образом: *R есть множество всех множеств, которые не включают самих себя*. Включено ли R в R ? Если нет, то R должно быть включено в R по определению, но если R включено в R , это противоречит определению. В любом случае получаем противоречие. Этот пример получил название парадокс Рассела. В последующие 25 лет математики пытались построить самосогласованную математическую логику, свободную от подобных парадоксов. Но в 1931 г. Гёдель доказал, что это невозможно [21], формальная логика оказалась противоречивой, не самосогласованной [22].

Пенроуз посвятил книгу [5] вопросу о мышлении в свете теории Гёделя. Он пришёл к выводу, что мышление не может быть объяснено как вычислительная процедура (то есть не может быть описано существующей математикой и не может быть реализовано ни в компьютере, ни в какой системе классической физики). Разумеется, Пенроуз встретил много возражений среди коллег. Книга [5] разбирает эти возражения, и Пенроуз подчёркивает, что рассмотрел все возражения. Настоящая статья, однако, даёт новое возражение: Логика не

является ни фундаментальным механизмом мышления, ни его основной. Способность к логическому мышлению возникает из нелогических механизмов мышления.

3. Обзор математических теорий интеллекта. Логика и сложность

Простое восприятие вовлекает сигналы из сенсорных органов и внутренние представления мышления (память) об объектах. В процессе восприятия мышление ассоциирует подмножества сигналов, соответствующих объектам, с представлениями об этих объектах. Далее активируются нейронные сигналы, ведущие к мысленным и поведенческим реакциям, являющимся частью понимания объекта или ситуации.

Математическое описание этого *первого шага*, казалось бы, простого процесса ассоциация-распознавание-понимание оказалось далеко не простым делом, и это связано с понятием комбинаторной сложности (КС) [23]. Это понятие подчёркивает многочисленность комбинаций элементов в сложной системе; например, восприятие ситуации обычно требует одновременного распознавания её многочисленных элементов, которые могут встретиться в различных комбинациях. Например, рассмотрим 100 элементов (не слишком большое число); число комбинаций из 100 элементов, 100^{100} превышает число всех элементарных частиц во Вселенной за всё время её существования. Никакой компьютер не будет в состоянии провести столько вычислений.

Эта проблема впервые была замечена в задачах распознавания образов в 1960-х годах и получила название «проклятие размерности» [24]. Казалось, что адаптивные самообучающиеся алгоритмы и нейронные сети могли бы обучиться решению любой задачи «сами по себе», они лишь нуждались в достаточном числе обучающих примеров. Последовавшие тридцать лет создания адаптивных самообучающихся алгоритмов и нейронных сетей привели к выводу, что требуемое число обучающих примеров часто оказывалось комбинаторно большим. Таким образом, самообучающиеся методы встретились с *КС обучения*.

В 1970-х годах для решения проблемы КС были предложены *системы правил* [25,26]. Первоначальная идея состояла в том, что знания, необходимые для решения задач определённого круга, будут сформулированы в виде логических правил, и необходимость в обучении отпадёт. Однако в непрерывно изменяющемся мире количество необходимых правил росло; правила оказались зависимы от других правил и необходимо рассматривать комбинации правил. Системы правил встретились с *КС правил*.

Начиная с 1980-х годов, были предложены *системы моделей*. Они использовали модели, зависящие от адаптивных параметров. Идея состояла в том, чтобы соединить преимущества систем правил с адаптивностью и обучаемостью. Знания предполагалось вложить в модели, а изменяющимся аспектам конкретных ситуаций система должна была обучаться сама, подгоняя параметры моделей [27,28]. Подгонка моделей к данным требовала идентифицировать подмножества данных, соответствующих различным моделям. Количество подмножеств, однако, комбинаторно велико. Популярный алгоритм для одновре-

менной подгонки многих моделей к данным, «проверка многих гипотез» [29], известен тем, что сталкивается с КС вычислений. Системы моделей встретились с *КС вычислений*.

Последовавшие исследования обнаружили связь КС с логикой в различных алгоритмах [23]. Формальная логика использует «закон исключённого среднего», который исключает возможность ввести меру близости логических высказываний. Поэтому, алгоритмы, использующие логику, рассматривают каждое даже небольшое изменение в данных или моделях, как новое высказывание (гипотезу); более того, комбинации данных рассматриваются как новые высказывания. Это приводит к КС. Более того, КС алгоритмов, основанных на логике, связана с гёделевской теорией: проявляется несогласованность логики в конечных системах [30]. Многозначные логики и нечеткая логика были предложены для того, чтобы преодолеть ограничения, связанные с законом исключённого третьего [31]. Однако, математика многозначных логик не отличается в принципе от формальной логики, «исключенное третье» заменяется на «исключенное $n+1$ ». Нечеткая логика встретила сложность степени нечеткости, если берётся слишком большая нечеткость, решение не достигает необходимой точности: если нечеткость слишком мала, метод приближается к формальной логике. Сложные системы требуют различных степеней нечеткости в различных частях системы, изменяющихся в процессе работы системы. Поиск подходящих степеней нечеткости для различных комбинаций элементов системы вновь приводит к КС.

Возможна ли логика после Гёделя? Недавний обзор этой области [32], показывает, что логика после Гёделя усложнилась, стала «менее логичной», чем предполагалось основателями искусственного интеллекта, и оказывается напрямую связанной с проблемой КС. Повторю, Пенроуз считал, что из теории Гёделя следует невычислимость мыслительных процессов и необходимость новой физической теории [5]. Противоположная точка зрения, высказываемая в данной статье, состоит в том, что из невычислимости логики не следует невычислимость мышления. Логика не является основным механизмом мышления.

4. Структуры мышления

В 1950-60-х годах исследователи в области искусственного интеллекта наивно верили, что вскоре они создадут компьютеры, превосходящие человеческий интеллект, и что математика логики окажется достаточной для этой цели. Однако, как мы обсудили, логика не объясняет мышления, но мышление работает. Обратимся к механизмам мышления, основываясь на представлениях, сложившихся в психологии, философии, когнитивной науке и нейробиологии. Основные механизмы мышления включают инстинкты, концепции, эмоции, поведение.

Инстинкты – это врождённые способности, не обучаемые, обычно адаптивные. В человеке и высших животных инстинкты связаны с эмоциями.

Слово концепты или концепции мы используем для обозначения общего смысла таких понятий как концепция, идея, понимание, мысль, представление, понятие. Смысл этих слов состоит в обозначении категории, группы, или класса предметов, явлений, событий или отношений. Концепции абстрактны в

том смысле, что они не различают индивидуальные явления. Подчёркивая это свойство, средневековые философы использовали слово «универсалии». Платон и Аристотель называли их идеями или формами и считали их основой для способности к пониманию мира. Подобным образом Кант считал, что они составляют содержание чистого разума, механизма способности к пониманию. По Юнгу обучение сознательным концепциям основано на врождённых бессознательных структурах психики, архетипах. Современная наука обычно связывает механизмы концепций с внутренними представлениями-понятиями об объектах и об отношениях между ними.

Эмоциями называют внутренние состояния, связанные с чувствами, а также преувеличенно-экспрессивное поведение, используемое для коммуникаций чувств. Многие исследователи (например, Левайн и Ливен [33]) подчёркивают, что в каждом решении обычно присутствуют в той или иной мере и эмоции и логика.

Важная роль эмоций в мышлении разбиралась многими авторами с различных точек зрения: в философии (Ренэ Декарт [34], Иммануил Кант [35], Жанн Поль Сартр [36]); в аналитической психологии (Карл Юнг [37]); в психологии и нейронных механизмах (Стефан Гроссберг и Даниэл Левайн [10], Эндрю Ортони [30], Джозеф ЛеДу [39]); в философской лингвистике (П. Гриффис [40]); в нейрофизиологии (Антонио Дамазио [41]); в адаптации, обучении и познании (Л.И. Перловский [16,42]).

Декарт попытался дать научное описание эмоций («страстей»). Он рационализировал эмоции, объяснил их как объекты и соотнёс с физиологическими процессами. По Канту эмоции тесно связаны с суждениями о том, какие конкретные восприятия и ситуации соответствуют каким общим концепциям-понятиям (и наоборот). Эмоции оказываются неотъемлемой частью мышления. Способность к эмоциям-суждениям – основание для всех высших познавательных способностей, включая способность к прекрасному и возвышенному. Сартр приравнивал эмоции в значительной мере к бессознательному содержанию психики; сегодня это не кажется адекватным. Юнг проанализировал сознательные и бессознательные аспекты эмоций. Он указал на недифференцированный статус примитивных психологических состояний, в которых эмоции, концепции и поведение слиты, и на их роль в повседневной жизни и в психозах. Он также подчеркнул рациональный аспект сознательных дифференцированных эмоций. Ортони объяснял эмоции в терминах концептуально-понятийных представлений и подчёркивал роль абдуктивной логики, как механизма для определения эмоций других людей. ЛеДу анализировал нейронные структуры и пути распространения нейронных сигналов, участвующих в эмоциональных процессах, в особенности, при страхе. Гриффис рассматривает основные эмоции и их эволюционное развитие в социальном взаимодействии. Дамазио считает, что эмоции – это в основном телесные ощущения и восприятие эмоций в мозгу вовлекает «телесные точки». Гроссберг и Левайн считают, что эмоции – это нейронные сигналы, связывающие инстинктивные и концептуальные мозговые центры. В процессах восприятия и познания эмоции оценивают насколько концепции-модели объектов и ситуаций удовлетворяют или не удовлетворяют инстинктивные потребности. Математическое описание в следующей секции близко связано с идеями Канта, Юнга, Гроссберга и Левайна.

Объяснение основных механизмов мышления, повторю, не требует мистических предположений; каждый механизм поддается математическому описанию [43]. Из механизмов мышления концепции-модели наиболее доступны сознанию. Рэй Джакендофф [44] также считает, что наиболее адекватный термин для механизма концепций – это модель, или внутренняя модель мышления. Концепции – это модели в буквальном смысле. Они моделируют в нашем мышлении объекты и ситуации в мире. Например, при зрительном восприятии объекта, концепция-модель в нашей памяти проецирует изображение объекта на зрительную кору мозга. Там это изображение совмещается с изображением, проецируемым из сетчатки глаза.

Концепции используются при удовлетворении основных инстинктов, которые возникли как механизмы выживания задолго до концепций. Гроссберг и Левайн [10] выделили инстинкты, как внутренние сенсоры, измеряющие основные потребности, и отделили их от «инстинктивного поведения», которое надлежит описывать подходящими механизмами. Соответственно, в настоящей статье «инстинкт» – это механизм внутренних сенсоров. Например, когда уровень сахара в крови опускается ниже определённого уровня, инстинкт «сообщает» мозгу, что нужна еда.

Как мы узнаём об инстинктивных нуждах? Мы не слышим инстинктивных повелений и не читаем циферблатов инстинктивных сенсоров. Инстинкты-сенсоры связаны с пониманием и поведением эмоциями. Фундаментальная роль эмоций в системе мышления состоит в том, что эмоциональные нейронные сигналы оценивают концепции для целей удовлетворения инстинктов [10]. Этот эмоциональный механизм оказался принципиально важен для того, чтобы «разорвать замкнутый круг» комбинаторной сложности.

Наш анализ приводит к неотвратимому выводу: человек и высшие животные обладают особым инстинктом, ответственным за познание. Очевидно, человек и высшие животные способны к исследовательскому поведению, даже когда основные телесные потребности, такие как голод, удовлетворены. Биологи и психологи обсуждали различные аспекты этого поведения. Харри Харлоу открыл, что обезьяны и люди стремятся к положительным стимулам, независимо от удовлетворения голода [45]; Дэвид Берлин в этой связи обсуждал любопытство [46]; Леон Фестинджер ввёл понятие когнитивного диссонанса и описал многочисленные эксперименты, раскрывающие потребность человека в уменьшении диссонанса [47]. Фундаментальная природа инстинкта к знанию связана с тем, что наши знания всегда должны быть подогнаны под конкретную окружающую ситуацию. Без адаптации концепций-моделей мы не смогли бы воспринимать непрерывно изменяющийся мир. Мы бы не могли ни ориентироваться, ни удовлетворять наши телесные потребности. Поэтому у нас есть врождённая потребность, инстинкт к улучшению наших знаний. Я назвал его *инстинктом к знанию*. Математически он описывается максимизацией меры близости между концепциями-моделями и миром (как он измеряется сенсорными органами; надо также иметь в виду, что сенсорные органы тоже адаптируются в процессе восприятия).

Эмоции, оценивающие удовлетворение или неудовлетворение инстинкта к знанию, не связаны непосредственно с телесными потребностями. Это, следовательно, духовные или эстетические эмоции. Концептуально-эмоциональное

восприятие мира ведёт к действиям в окружающем мире и к действиям внутри мышления, приводит к увеличению знаний. Математическая теория концептуально-эмоционального восприятия и понимания описывается в следующем разделе. Мышление включает иерархию многих уровней концепций-моделей, от простейших элементов восприятия (линии, движущиеся точки) до концепций-моделей объектов, отношений между объектами, до сложных ситуаций.

5. Теория нейронных моделирующих полей (НМП)

Теория нейронных моделирующих полей (НМП) математически описывает механизмы мышления, обсуждавшиеся выше. Это многоуровневая гетеро-иерархическая система [8], это не строгая иерархия, в ней множество обратных связей между соседними уровнями, поэтому это именно гетеро-иерархия. На каждом уровне НМП действуют механизмы концепций-моделей, содержащие знание; они генерируют эфферентные сигналы, распространяющиеся «сверху-вниз» по иерархии и взаимодействующие с афферентными сигналами, поступающими из предыдущих уровней иерархии, «снизу-вверх». Взаимодействие между этими сигналами управляется инстинктом к знанию, который «направляет» концепции-модели к улучшению, то есть направляет обучение, адаптацию, и возникновение новых концепций-моделей для улучшения их соответствия входным сигналам, идущим снизу-вверх.

В этом разделе излагается основной механизм взаимодействия между двумя соседними уровнями иерархии, между афферентными (идущими вверх) и эфферентными (идущими вниз) сигналами. Эти сигналы в нейронной архитектуре представляют собой поля нейронных активностей (этот аспект НМП подобен нейронным активностям в теории адаптивного резонанса [48]), отсюда название этой архитектуры, НМП. Иногда удобнее говорить об этих двух уровнях сигналов, как о входе и выходе одного уровня обработки сигналов. На каждом уровне выходные сигналы – это концепции-модели, воспринятые или образованные из входных сигналов. Входные сигналы ассоциируются с концепциями (или группируются в концепции) в соответствии с моделями и инстинктом к знанию. Эта общая структура НМП соответствует нашим знаниям о нейронной организации мозга; в этой статье, однако, мы не проводим детальное сопоставление математических механизмов с конкретными нейронами и синапсами.

5.1. Инстинкт к знанию

Инстинкт к знанию описывается математически как максимизация меры схожести моделей и входных сигналов. В процессе обучения и понимания входных сигналов, модели адаптируются так, чтобы лучше представлять входные сигналы, и чтобы схожесть между ними увеличивалась. Это увеличение схожести удовлетворяет инстинкт к знанию и ощущается как эстетическая эмоция.

На каждом иерархическом уровне нейроны нумеруются индексом $n = 1, \dots, N$. Эти нейроны получают входные сигналы $X(n)$, идущие снизу-вверх, из нижних уровней иерархии. $X(n)$ – это поле активности нейронных входных сигналов, поступающих от нейронов на предыдущем (нижнем) уровне. У каждого

нейрона есть несколько синапсов; в общем случае, мы описываем нейрон набором чисел $\mathbf{X}(n) = \{X_d(n), d = 1, \dots, D\}$. Сверху-вниз в эти же нейроны поступают маркирующие активизирующие сигналы, идущие от концепций-моделей $\mathbf{M}_h(\mathbf{S}_h, n)$; мы нумеруем модели индексом $h = 1, \dots, H$. Каждая модель зависит от параметров этой модели \mathbf{S}_h . В нейронной структуре мозга параметры кодируются синапсами, силой их связи или весом; математически они задаются набором чисел, $\mathbf{S}_h = \{S_h^a, a = 1, \dots, A\}$. Модели являются *представлениями* входных сигналов в следующем смысле. Пусть, например, сигнал $\mathbf{X}(n)$ поступает от сенсорного нейрона, активированного объектом h , параметры модели которого \mathbf{S}_h . Эти параметры могут описывать положение, ориентацию или освещение объекта h . Модель $\mathbf{M}_h(\mathbf{S}_h, n)$ *предсказывает* значение входного сигнала $\mathbf{X}(n)$ в нейроне n . Например, при зрительном восприятии нейрон n в зрительной коре получает сигнал $\mathbf{X}(n)$ из сетчатки и маркирующий сигнал $\mathbf{M}_h(\mathbf{S}_h, n)$ из концепции-модели объекта h . Нейрон n возбуждается, если сильны оба сигнала, идущие снизу-вверх из нижнего уровня и сверху-вниз из модели \mathbf{M}_h . Ниже описано, как различные модели конкурируют за входные сигналы, в то время как их параметры адаптируются для лучшего соответствия входным сигналам.

Обычное повседневное восприятие использует много уровней от сетчатки до восприятия объекта. В теории НМП одни и те же принципы нейродинамики, действуют на каждом уровне. Восприятие мелких деталей, обычных объектов или понимание сложных абстрактных концепций использует один и тот же механизм, излагаемый ниже. И в восприятии, и в познании участвуют модели и обучение. В восприятии модели соответствуют объектам; в познании – соотношениям и ситуациям.

Обучение состоит в увеличении меры близости между множествами моделей и сигналов $L(\{\mathbf{X}\}, \{\mathbf{M}\})$. Мера близости зависит от параметров моделей и от ассоциаций между входными, идущими снизу-вверх сигналами и сигналами от концепций-моделей. Для конкретности мы описываем здесь восприятие объектов упрощённо, как если бы оно происходило в одном уровне.

Сконструируем меру близости $L(\{\mathbf{X}\}, \{\mathbf{M}\})$, используя подходы теории вероятностей. При построении выражения для меры близости, необходимо учесть два принципа (почти очевидных). Во-первых, содержание поля зрения неизвестно до восприятия, и, во-вторых, в нем могут присутствовать разнообразные объекты. Важная информация может содержаться в любом из входных сигналов, следовательно, мера близости должна учесть все входные сигналы $\mathbf{X}(n)$,

$$L(\{\mathbf{X}\}, \{\mathbf{M}\}) = \prod_{n \in N} l(\mathbf{X}(n), \{\mathbf{M}\}). \quad (1)$$

Содержащееся здесь произведение по всем входным сигналам удовлетворяет первому принципу: все входные сигналы должны быть учтены (если хотя бы один член произведения равен нулю, мера близости равна нулю, и инстинкт к знанию не удовлетворен). Во-вторых, до восприятия мышлению неизвестно, какой сигнал произошёл от какого объекта. Следовательно, частичная мера близости

сти $l(\mathbf{X}(n), \{\mathbf{M}\})$ строится так, чтобы каждый потенциальный объект-концепция-модель мог сопоставляться всем возможным входным сигналам. Для этого необходимо рассмотреть все возможные альтернативы по моделям (то есть, провести суммирование по моделям) для каждого входного сигнала. Слагаемые в этой сумме – это условные меры близости между сигналом $\mathbf{X}(n)$ и моделью \mathbf{M}_h , $l(\mathbf{X}(n)|h)$. Эта мера «условна»: она предполагает условие, что сигнал n пришёл от объекта h , и что этот объект присутствует [49], следовательно, собирая эти слагаемые в общую меру близости, L , их следует умножить на величину $r(h)$, которая представляет вероятностную меру того, что объект h действительно присутствует. В результате получаем [50]:

$$L(\{\mathbf{X}\}, \{\mathbf{M}\}) = \prod_{n \in N} \sum_{h \in H} r(h) l(\mathbf{X}(n) | h). \quad (2)$$

Структура этого уравнения следует стандартным принципам теории вероятности: сумма берётся по альтернативам h , а произведение – по различным источникам данных n . Это выражение не обязательно является вероятностью, но у него вероятностная структура. Если обучение прошло успешно, эта мера близости приближается к вероятностной и ведёт к приблизительно байесовским (то есть оптимальным) решениям. Название «условная мера близости» для $l(\mathbf{X}(n)|h)$ (или просто $l(n|h)$) следует вероятностной терминологии. Если обучение успешно, $l(n|h)$ становится условной плотностью вероятности, вероятностной мерой того, что сигнал в нейрон n пришёл от объекта h . Тогда L является полной мерой правдоподобия того, что наблюдаемые сигналы $\{\mathbf{X}(n)\}$ пришли от объектов, описываемых моделями $\{\mathbf{M}_h\}$. Коэффициенты $r(h)$, называемые априорными в теории вероятности, содержат информацию о предварительных ожиданиях; их истинные значения неизвестны и подлежат обучению, как и остальные параметры моделей \mathbf{S}_h . Поскольку L является мерой правдоподобия, нейронное понятие *обучение* соответствует статистическому понятию оценки в теории оценок. Максимизация правдоподобия обычно ведёт к эффективным или статистически наилучшим оценкам.

Заметим, что в теории вероятности, произведение вероятностей предполагает, что данные независимы. Выражение (2) содержит произведение по n , однако сигналы $\mathbf{X}(n)$ не предполагаются независимыми. Зависимость между сигналами существует благодаря моделям; каждая модель $\mathbf{M}_h(\mathbf{S}_h, n)$ предсказывает ожидаемые значения сигналов для многих нейронов n . Независимы в выражении (2) лишь отклонения сигналов от модельных предсказаний. То есть, отклонения считаются случайными; неслучайные аспекты сигналов должны быть предсказаны моделями.

Во время процессов обучения-адаптации концепции-модели постоянно изменяются. В настоящей статье мы рассматриваем ситуации, когда функциональные формы моделей $\mathbf{M}_h(\mathbf{S}_h, n)$, фиксированы, и обучение-адаптация касается лишь параметров моделей \mathbf{S}_h . Более сложная структурная адаптация рассмотрена в [51,52]. Время от времени система образует новые модели, иногда старые концепции сливаются или уничтожаются. Это требует изменения меры близости (2); причина состоит в том, что большее число моделей всегда приведет к

лучшей подгонке данных. Это хорошо известная проблема, она решается тем, что меру (2) следует уменьшить, используя «скептическую штрафную функцию» $p(N, M)$, которая растёт вместе с числом моделей M , и скорость её роста возрастает для малого числа данных N . Например, асимптотически несмещённая оценка по максимуму правдоподобия (говоря простым языком, статистически верная в среднем) требует, чтобы (2) было умножено на $p(N, M) = \exp(-N_{par}/2)$, где N_{par} – это полное число адаптивных параметров во всех моделях (эта штрафная функция известна как информационный критерий Акаика, см. [8] для дальнейших ссылок).

5.2. Динамическая логика

Процесс обучения-адаптации состоит в оценке параметров моделей \mathbf{S} и в ассоциации сигналов с концепциями-моделями; эта оценка производится максимизацией меры (2). Отметим, что (2) содержит все возможные комбинации сигналов и моделей: если раскрыть сумму в (2) и перемножить все члены, то получим H^N слагаемых. Это огромное число, равное числу комбинаций между всеми входными сигналами (N) и всеми моделями (H). В этом состоит источник КС прежних алгоритмов. Например, широко известный алгоритм «проверка многих гипотез» пытается максимизировать меру L по параметрам и по ассоциациям между сигналами и моделями в два приёма. Вначале каждое из H^N слагаемых (т. е. каждая из ассоциаций сигналы-модели) максимизируется по параметрам. Затем выбирается наибольшее слагаемое, то есть, наилучшая ассоциация для лучшего набора параметров. Ясно, что этот алгоритм сталкивается с проблемой КС, количество вычислений остается порядка H^N .

Теория моделирующих полей решает эту проблему, используя динамическую логику [8,53]. Принципиальный аспект динамической логики состоит в том, что неопределённость или расплывчатость меры близости совпадает с неопределённостью или неточностью моделей. Сначала значения параметров неизвестны и неточность моделей велика; соответственно велика расплывчатость меры близости. В процессе обучения модели становятся более точными и мера близости более чёткой, значение меры L увеличивается. Приведем математическое описание этого механизма. Сначала зададим произвольные значения неизвестных параметров $\{\mathbf{S}_h\}$. Затем вычислим ассоциативные переменные $f(h|n)$,

$$f(h|n) = r(h) l(\mathbf{X}(n)|h) / \sum_{h' \in H} r(h') l(\mathbf{X}(n)|h'). \quad (3)$$

Уравнение (3) подобно байесовской формуле для апостериорных вероятностей; если $l(n|h)$ в результате обучения станут условными функциями правдоподобия, то $f(h|n)$ станут байесовскими вероятностями того, что сигнал n произошёл от объекта h . Динамическая логика НМП задаётся следующими уравнениями

$$df(h|n)/dt = f(h|n) \sum_{h' \in H} \{[\delta_{hh'} - f(h'|n)] \cdot [\partial \ln l(n|h') / \partial \mathbf{M}_{h'}] \partial \mathbf{M}_{h'} / \partial \mathbf{S}_h\} \cdot d\mathbf{S}_h / dt, \quad (4)$$

$$d\mathbf{S}_h/dt = \sum_{n \in N} f(h|n) [\partial \ln l(n|h) / \partial \mathbf{M}_h] \partial \mathbf{M}_h / \partial \mathbf{S}_h, \quad (5)$$

здесь

$$\delta_{hh'} = 1, \text{ если } h=h', \text{ или } \delta_{hh'} = 0, \text{ если } h \neq h'. \quad (6)$$

Параметр t – время внутренней динамики системы НМП (число итераций). Для условных мер близости обычно используются гауссовские функции,

$$l(n|h) = G(\mathbf{X}(n) | \mathbf{M}_h(\mathbf{S}_h, n), \mathbf{C}_h). \quad (7)$$

Здесь G – гауссовская функция со средним \mathbf{M}_h и матрицей ковариаций \mathbf{C}_h . Отметим, что в статистике часто используется «гауссовское предположение»; предполагается гауссовская плотность вероятности сигналов. Здесь нет подобного предположения. В (7) сигналы не предполагаются гауссовскими, лишь отклонения сигналов \mathbf{X} от моделей \mathbf{M} считаются гауссовскими, так как они возникают в результате многих случайных причин (неслучайные аспекты сигналов учтены в моделях). Если, тем не менее, отклонения не гауссовские, то следует использовать подходящие функции. Если нет информации об их функциональной форме, обычно (7) является хорошим выбором, и не ограничивает общность: взвешенная сумма гауссовских функций в (2) может аппроксимировать любую положительную функцию, каковой является мера близости.

Матрицы ковариаций \mathbf{C}_h и другие неизвестные параметры в (7) оцениваются эвристически. Первоначальные значения \mathbf{C}_h должны быть велики, в соответствии с неопределенностью знания моделей \mathbf{M}_h . По мере улучшения значений параметров и моделей ковариации уменьшаются до истинной величины различий между моделями и сигналами (существующими из-за погрешности измерений и неточности моделей). По мере того, что ковариации уменьшаются, меры близости становятся более четкими, приближаются к дельта-функциям; ассоциативные переменные (3) приближаются к четким значениям $\{0, 1\}$ и решения динамической логики сходятся к четкой логике (или к вероятностям, в зависимости от задачи). Этот процесс, в котором одновременно улучшаются значения параметров, а мера близости сходится к четкой логической функции, является сутью динамической логики. Таким образом, механизм динамической логики соединяет расплывчатую и четкую логику.

Динамическая эволюция от большой нечеткости к малой – причина для названия «динамическая логика». Математически этот механизм помогает избегать локальных максимумов во время сходимости [8] и соответствует психологическим свойствам [54].

В [8] доказана следующая теорема.

Теорема. Уравнения (3)–(6) определяют сходящуюся динамическую систему (НМП) со стационарными состояниями, определяемыми $\max \{\mathbf{S}_h\} L$.

Отсюда следует, что стационарные состояния системы НМП являются состояниями максимальной меры близости, удовлетворяющие инстинкт к зна-

нию. Если условные меры близости заданы, как плотности вероятности (или функции правдоподобия), стационарные значения параметров $\{S_h\}$ являются асимптотически несмещёнными и эффективными оценками [55]. Вычислительная сложность НМП линейна по N .

Говоря простым языком, это означает, что динамическая логика – это сходящийся процесс. Отдельные аспекты сходимости НМП обсуждаются ниже. Если функции правдоподобия (плотности вероятностей) используются как меры близости, значения параметров оцениваются эффективно (то есть, во многих случаях значения параметров не могут быть улучшены никакой другой процедурой). Приведенная выше теорема также утверждает, что мера близости увеличивается на каждой итерации.

5.3. Функционирование динамической логики. Пример

В следующем примере мы ищем «улыбающиеся» и «нахмуренные» модели в шуме (рис. 1). На рис. 1А показаны модели без шума, на рис. 1В – зашумленные модели (сигнал, в котором надо распознать модели). Каждая модель – это 3-х параметрическая параболическая форма (3 параметра описывают положение и кривизну модели-объекта). Размер всего изображения не очень велик, всего 100x100 точек; число объектов равно трем, но оно заранее неизвестно и подлежит определению. Поэтому следует попытаться подогнать к данным по меньшей мере 4 модели, чтобы решить, что 3 модели подходят лучше. Чтобы подогнать $4 \times 3 = 12$ параметров к 100x100 точкам методом полного перебора, потребовалось бы порядка 10^{32} вычислений. Чтобы применить НМП и динамическую логику для этого примера, следует разработать параметрические модели для ожидаемых объектов, которые в этом случае не слишком сложны [56]: равномерная модель для шума, круглые гауссовские формы для сильно нечётких, плохо различимых объектов и параболические модели для «улыбок» и «нахмурок». Число компьютерных операций в этом примере было около 10^{10} , что значительно меньше, чем при полном переборе.

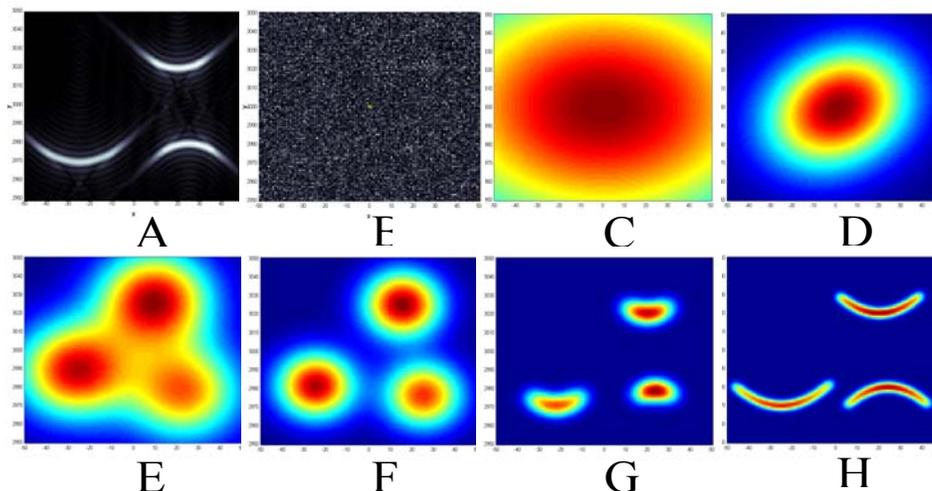


Рис.1 Процесс функционирования динамической логики: нахождение «улыбок» и «нахмурок» в шуме: (A) точная форма «улыбок» и «нахмурок» без шума; (B) реально измеренное изображение с шумом (сигнал слабее шума; отношение сигнала к шуму принимает значения между -2dB и -0.7dB ; то есть между $1/3$ и $1/5$); (C) начальная расплывчатая модель; от (D) до (H) показаны улучшающиеся модели на различных стадиях итерационного процесса (всего 22 итерации). После каждых 5 итераций алгоритм пытался увеличить или уменьшить число моделей. Между итерациями (D) и (E) алгоритм пытался подогнать больше одной модели к данным, решил, что ему требуется 3 гауссовские модели для «понимания» данных. Всего было использовано, как уже упоминалось, три типа моделей: одна равномерная модель, описывающая шум и переменное число круглых и параболических моделей. До приблизительно итерации (G) алгоритм «мыслил» в терминах простых круглых моделей, начиная с (G) и далее, алгоритм решил, что для описания данных требуются более сложные параболические модели. Итерации прекратились в (H), когда мера схожести (2) перестала увеличиваться. Этот пример описан подробнее в [56].

Во время процесса адаптации начальные расплывчатые и неопределённые модели ассоциируются со структурами во входных сигналах, нечёткие модели становятся более определёнными и чёткими с каждой итерацией. Тип, форма и число моделей подбираются алгоритмом, в соответствии с уравнениями (3)-(7), так что внутреннее представление внутри системы соответствует входным сигналам: концепции-модели НМП моделируют структуры-объекты в сигналах. На рис. 1A показана точная форма «улыбок» и «нахмурок» без шума; на рис. 1B – реально измеренный сигнал с шумом (отношение сигнала к шуму принимает значения между -2dB и -0.7dB ; то есть между $1/3$ и $1/5$); рис. 1C – начальная расплывчатая модель, чем больше расплывчатость, тем больше неопределённость знаний об объектах; от рис. 1D до рис. 1H показаны улучшающиеся модели на различных стадиях итерационного процесса (всего 22 итерации). После каждых 5 итераций алгоритм пытался увеличить или уменьшить число моделей. Между итерациями D и E алгоритм решил, что ему требуется 3 гауссовские модели для «лучшей» подгонки. Всего было использовано, как уже

упоминалось, три типа моделей: одна равномерная модель, описывающая шум, и переменное число круглых и параболических моделей, для которых число моделей, положение и кривизна оцениваются из данных. До приблизительно итерации G алгоритм использовал простые круглые модели, начиная с G и далее, алгоритм решил, что для описания данных требуются более сложные параболические модели. Итерации прекратились в H, когда мера близости перестала увеличиваться.

5.4. Иерархическая организация НМП

Выше мы описали один уровень обработки сигналов в иерархической системе НМП. На каждом уровне иерархии имеются входные сигналы из нижних уровней, модели, меры близости L, эмоции, являющиеся изменениями величины L, и действия; действия внутри мышления включают адаптацию, то есть поведение, удовлетворяющее инстинкт к знанию – максимизацию меры близости в соответствии с уравнениями (3)-(6). Входные сигналы на каждом уровне X(n) представляют собой поля нейронных активностей (возбуждений). В результате обработки сигналов на каждом уровне формируются активированные модели или концепции h, воспринятые или распознанные во входных сигналах n; эти модели вместе с соответствующими инстинктивными сигналами и эмоциями могут активировать модели поведения и вести к действиям на данном уровне.

Активированные модели инициируют и другие действия. Они посылают входные сигналы вверх, на следующий уровень, где создаются формируются более общие, абстрактные модели. Выходные сигналы с данного уровня, служащие входными сигналами на следующем уровне, могут задавать степени активации моделей a_h , определяемые следующим образом

$$a_h = \sum_{n \in N} f(h|n). \quad (8)$$

Кроме этого, выходные сигналы могут включать параметры моделей. Иерархическая система НМП показана на рис. 2. В иерархии мышления, каждая концепция-модель приобретает «ментальные» смысл и цель на следующем уровне иерархии (в дополнение к другим целям). Например, рассмотрим концепцию модель «стул». «Поведенческая» цель этой концепции – в инициировании действия «сесть» (если наличествует такая телесная потребность), эта «телесная» цель находится на том же иерархическом уровне. Кроме того, у этой концепции есть «чисто ментальная» цель на более высоком уровне иерархии, цель участвовать в распознавании-формировании более общей концепции, скажем «концертный зал», модель которого включает ряды стульев.

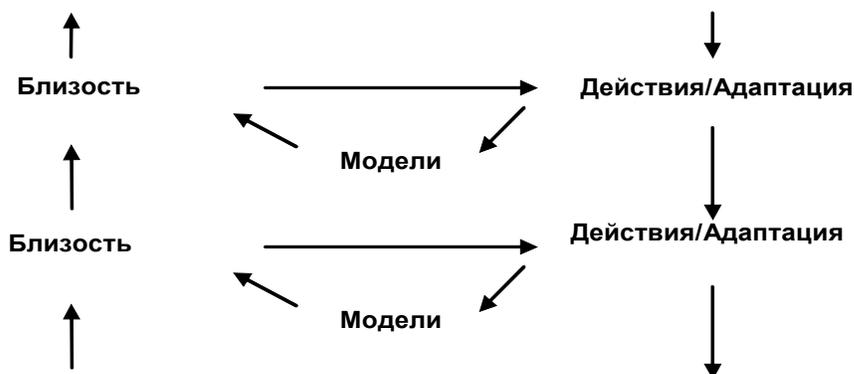


Рис.2. Иерархическая система НМП. На каждом уровне иерархии есть входные сигналы, модели, меры близости, эмоции и действия (в том числе адаптация, максимизирующая меры близости и удовлетворяющая инстинкт к знанию). Большие значения условных мер близости соответствуют концепциям, распознанным на данном уровне, и ведут к активации этих концепций. Сигналы активированных концепций – это выходные сигналы на данном уровне. Они становятся входными сигналами на следующем уровне, распространяя знание вверх по иерархии (на языке физической теории поля, это – пропагаторы концепций-знаний между уровнями иерархии).

Модели на верхних уровнях иерархии более общие и абстрактные, чем модели на нижних уровнях. Например, в зрительной системе, на нижнем уровне иерархии (говоря приблизительно) модели соответствуют ганглиозным нейронам сетчатки и выполняют такие же функции: они распознают простые элементы в поле зрения, линии, движущиеся точки. На более высоких уровнях модели соответствуют функциям, выполняемым в VI-м и в последующих уровнях зрительной коры, там происходит распознавание более сложных зрительных элементов, таких как границы контрастов, их направление, отделение эффектов освещения от геометрических форм, и т.д. Иерархическая структура зрительной коры изучена в деталях [9,57], эти модели могут быть использованы в НМП. В следующих, более высоких познавательных уровнях модели соответствуют объектам, ситуациям, взаимоотношениям между ситуациями, и т.д. [8]. Время от времени, как обсуждалось, система НМП формирует новую концепцию или уничтожает (забывает) старую. Более общий механизм НМП, используемый в сложных задачах, работает следующим образом. На каждом уровне система всегда хранит запас неопределённых, расплывчатых, неактивных концепций-моделей (с большими матрицами ковариаций C). Эти модели неактивны, поскольку их параметры не адаптированы к сигналам, и поэтому малы их меры близости к сигналам. Такое неактивное состояние самоподдерживается: ассоциативные переменные $f(h|n)$ для этих моделей малы, в соответствии с уравнением (5) их параметры не адаптируются, и они не участвуют в соревновании между моделями: их меры близости продолжают оставаться малыми для всех сигналов. Однако из-за большой степени расплывчатости (больших ковариаций) их меры близости не обращаются в точности в ноль. Поэтому неактивные модели могут потенциально «ухватить» каждый сигнал, который не попадает в более

специфичные, менее расплывчатые, активные модели. Когда накапливается достаточно сигналов, «выпавших» из активных моделей, и попавших в неактивную модель h , ее активационный сигнал a_h превышает определённый порог и модель активируется. Аналогично, когда слишком мало сигналов ассоциировано с определённой моделью, её активационный сигнал уменьшается ниже порогового значения и модель деактивируется. Пороги активации и деактивации обычно определяются на основе информации, существующей на более высоком иерархическом уровне (априорная информация, данные о системных ресурсах, количество активированных моделей различного типа, и т.д.). Активационные сигналы, создаваемые активными моделями на данном уровне $\{a_h\}$, формируют «нейронное поле», являющееся входными сигналами для следующего уровня, где формируются более абстрактные и общие концепции.

5.4. Интеграция языка и мышления

Создание интеллектуальных компьютерных систем требует математических методов интегрирования языка и мышления. В НМП это достигается интегрированием языковых и познавательных моделей [51,52]. В уравнениях (1)-(5) \mathbf{M}_h определяется как двойная модель,

$$\mathbf{M}_h = \{ \mathbf{M}_h^C, \mathbf{M}_h^L \}. \quad (9)$$

Здесь \mathbf{M}_h^C обозначает познавательную часть модели объекта или ситуации в мире (как модели в примере, представленном на рис. 1); \mathbf{M}_h^L – это языковая часть модели. Математические механизмы интеграции языка и познания требуют изменений в уравнениях (1)-(5), которые были сформулированы в [58,59]. Эта изменённая система НМП используется сегодня для создания будущих интегрированных систем роботов и «инфоботов», которые, возможно, заменят сегодняшние системы поиска информации на Интернете.

6. Заключение

Итак, в настоящей работе построена модель нейронных модельных полей, позволяющая преодолеть «проклятие размерности» в задачах распознавания и проинтерпретировать ряд особенностей мыслительных процессов.

Уравнения (3)-(6) описывают элементарный процесс восприятия или познания, который максимизирует знание. Знание измеряется мерой близости между концепциями-моделями и миром. В этом процессе большое число концепций-моделей соревнуются за входные сигналы, модели модифицируются, формируются новые; при этом формируются ассоциации между моделями и сигналами, в начале неопределённые, а в результате конкретные, указывающие какой сигнал соответствует какой концепции-объекту-модели. В процессе восприятия входные сигналы поступают от сенсорных органов и соответствуют моделям объектов и простых ситуаций во внешнем мире. В процессе познания

входные сигналы поступают от предыдущих уровней и соответствуют моделям соотношений между концепциями, распознанными на нижних уровнях.

Процессы эти описываются динамической логикой. Её важной математической чертой является соответствие между неопределённостью в моделях и нечёткостью-расплывчатостью ассоциаций $f(h|n)$. Пока параметры моделей не соответствуют сигналам, многие модели плохо ассоциированы со многими объектами одновременно. Постепенно одна модель выигрывает соревнование за некий набор сигналов, ассоциации становятся чёткими, и сигналы распознаются как концепции. Начальные нечёткие концепции становятся чёткими, приблизительно подчиняющимися логике.

При подготовке статьи к публикации В.Г. Редько обратил мое внимание на то, что роль эмоций и потребности накопления знаний (аналога инстинкта знаний) при формировании интеллектуального адаптивного поведения подчеркивается в работах А.А. Жданова [60].

Благодарности

Я с удовольствием выражаю благодарность тем, чьи мысли, идеи, поддержка, и финансирование способствовали написанию этой статьи: М. Арановский, Р. Брокетт, П. Вербос, Б. Виджерс, Д. Виньковецкая, Я. Виньковецкий, Г. Карпентер, Р. Деминг, Ю. Димитрин, В. Дмитриев, Л. Гарвин, И. Гинзбург, Р. Гудвин, С. Гренадер, С. Гроссберг, А. Кангелози, М. Карповский, Л. Заде, Г. Зайниев, С. Кашин, Р. Козма, М. Ксариос, Д. Левайн, Л. Левитин, Р. Линнехан, Т. Лугинбул, М. Мазо, А. Мейстель, С. Миддлтон, К. Мор, К. Мюц, А. Нахман, А. Овсич, Р. Пейн, И. Перловский, В. Петров, Н. Печерская, Ч. Плам, В. Редько, Ю. Румер, Д. Скатруд, Р. Стрейт, Е. Таборская, Т. Терновская, И. Терновский, Ф. Фонтанари, М. Франк-Каменецкий, У. Фрееман, К. Фукунага, А. Черняхов, Е. Шахнович, У. Шёндорф, Дж. Шогрен, Н. Шохирев.

Литература

1. For the discussions and further references see: Grossberg, S. (1988). *Neural Networks and Natural Intelligence*. MIT Press, Cambridge, MA. Albus, J.S., Meystel, A.M. (2001). *Engineering of Mind: An Introduction to the Science of Intelligent Systems*. Wiley, New York, NY. Meystel, A.M., Albus, J.S. (2001). *Intelligent Systems: Architecture, Design, and Control*. Wiley, New York, NY. Perlovsky, L.I. (2001). *Neural networks and intellect: using model based concepts*. Oxford University Press, NY.
2. *The American Heritage College Dictionary*, Houghton Mifflin, 3rd ed., 2000, Boston, MA. I would emphasize, these general dictionary formulations could only be taken as a starting point.
3. Newell, A. (1983). Intellectual issues in the history of artificial intelligence. In: *The Study of Information*, ed. F.Machlup & U.Mansfield, J. Wiley, New York, NY.
4. Minsky, M. (1988). *The Society of Mind*. MIT Press, Cambridge, MA.
5. Penrose, R. (1994). *Shadows of the Mind*. Oxford: Oxford University Press.

6. Josephson, B. (1997). An Integrated Theory of Nervous System Functioning embracing Nativism and Constructivism. In: *International Complex systems Conference*. Nashua, NH.: Sept 21-26.
7. Hameroff, S. (1994). *Toward a Scientific Basis for Consciousness*. Cambridge, MA.: MIT Press;
8. Perlovsky, L.I. (2001). *Neural Networks and Intellect: Using Model Based Concepts*. New York: Oxford University Press.
9. Grossberg, S. (1988). *Neural Networks and Natural Intelligence*. MIT Press, Cambridge, MA.
10. Grossberg, S., Levine, D.S. (1987). Neural dynamics of attentionally modulated Pavlovian conditioning: blocking, inter-stimulus interval, and secondary reinforcement. *Psychobiology*, **15**(3), pp.195-240;
11. Zadeh, L.A. (1997). Information granulation and its centrality in human and machine intelligence. In: *Proceedings of the Conf. on Intelligent Systems and Semiotics '97*. Gaithersburg, MD, pp. 26-30.
12. Meystel, A. (1995). *Semiotic Modeling and Situational Analysis*. AdRem, Bala Cynwyd, PA.
13. G. M. Edelman, G. Tononi, (1995), *A Universe of Consciousness: How Matter Becomes Imagination*, Basic Books, New York, NY.
14. Perlovsky, L.I. & McManus, M.M. (1991). Maximum Likelihood Neural Networks for Sensor Fusion and Adaptive Classification. *Neural Networks*, **4** (1), pp. 89-102.
15. Perlovsky, L.I. (1996). Mathematical Concepts of Intellect. In: *Proceedings, World Congress on Neural Networks*. San Diego, CA.: L. Erlbaum Assoc., pp. 1013-16.
16. Perlovsky, L.I. (2006). Toward physics of the mind: Concepts, emotions, consciousness, and symbols. *Physics of Life Reviews*, **3**(1), pp.22-55.
17. Aristotle, IV BCE, *Complete Works of Aristotle*, Ed.J.Barnes, Princeton, NJ, 1995.
18. Aristotle, IV BCE, *Rhetoric for Alexander*, Complete Works of Aristotle, Ed.J.Barnes, Princeton, NJ, 1995.
19. Aristotle, IV BCE, *Organon*, Complete Works of Aristotle, Ed.J.Barnes, Princeton, NJ, 1995, 18a28-19b4; 1011b24-1012a28.
20. Hilbert, David. (1928/1967). The foundations of mathematics. In J. van Heijenoort, Ed., *From Frege to Gödel*. Cambridge, MA: Harvard University Press, p.475.
21. Gödel, K. (1986). *Kurt Gödel Collected Works*, I. (Ed. S.Feferman et al). Oxford University Press.
22. Об альтернативном истолковании логики см. Gentzen, G. (1935,1936), "Untersuchungen uber das logische Schliessen". *Mathematische Zeitschrift*, 1935, B. 39, pp. 176-210, 405-443. "Die Widersprunghsfreiheit der reinen Zahlentheorie". *Math. Ann.* 1936, Vol. 112, N. 4, pp.493-565. О различных истолкованиях Гёделевской теории и развитии логики после Гёделя см. B. Marchal, 2005, Theoretical Computer Science and the Natural Sciences, *Physics of Life Reviews*, **2**(3), pp.1-38.
23. Perlovsky, L.I. (1998). Conundrum of Combinatorial Complexity. *IEEE Trans. PAMI*, **20**(6) p.666-70.
24. Bellman, R.E. (1961). *Adaptive Control Processes*. Princeton University Press, Princeton, NJ.
25. Minsky, M.L. (1975). A Framework for Representing Knowledge. In: *The Psychology of Computer Vision*, ed. P. H. Winston, McGraw-Hill Book, New York.
26. Winston, P.H. (1984). *Artificial Intelligence*. 2nd edition, Addison-Wesley. Reading, MA.
27. A simple example of an adaptive model is linear regression: the knowledge is encapsulated in using linear model, and in the choice of variables, the uncertainty and adaptivity is in the unknown parameters, fitted to the data. Whereas linear regression uses one model, model-based systems used a large number of models. For example, a

- scene is described using geometric models of many objects. Parameters may include: size, orientation angles, color, illumination conditions, etc. A simple, still nontrivial problem, causing problems in applications till today, is tracking multiple objects in presence of clutter [see ref 29].
28. Singer, R.A., Sea, R.G. and Housewright, R.B. (1974). Derivation and Evaluation of Improved Tracking Filters for Use in Dense Multitarget Environments, *IEEE Transactions on Information Theory*, IT-20, pp. 423-432.
 29. Perlovsky, L.I., Webb, V.H., Bradley, S.R. & Hansen, C.A. (1998). Improved ROTH Detection and Tracking Using MLANS. *AGU Radio Science*, **33**(4), pp.1034-44.
 30. Perlovsky, L.I. (1996). Gödel Theorem and Semiotics. In: *Proceedings of the Conference on Intelligent Systems and Semiotics '96*. Gaithersburg, MD, v.2, pp. 14-18.
 31. Kecman, V. (2001). *Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models (Complex Adaptive Systems)*. The MIT Press, Cambridge, MA.
 32. B. Marchal, 2005, Theoretical Computer Science & the Natural Sciences, *Physics of Life Reviews*, **2**(3), pp.1-38.
 33. D.S. Levine, S.J. Leven, Eds. (1992). *Motivation, Emotion, and Goal Direction in Neural Networks*, Erlbaum, Hillsdale, NJ.
 34. Descartes, Rene (1646). *Passions of the Soul*. Descartes deeply penetrated into the interaction between emotions and consciousness: "those who are the most excited by their passions are not those who know them best and their passions are... confused and obscure." He did not differentiate thoughts and emotions consistently: "of all the various kinds of thoughts... these passions," "the action and the passion are thus always one and the same thing." Descartes showed a bias toward unconscious and fused perception of emotions characteristic of a thinking psychological type.
 35. Kant, I. (1790). *Critique of Judgment*, tr. J.H. Bernard, Macmillan & Co., London, 1914.
 36. Sartre J. P., 1948, Emotions. See J. P. Sartre, *Existentialism and Human Emotions*, Citadel Press; Reissue edition, 1984.
 37. Jung, C.G., 1921, Psychological Types. In: *The Collected Works*, v.6, Bollingen Series XX, 1971, Princeton University Press, Princeton, NJ.
 38. Ortony, A, Clore G.L., Collins A., 1990, *The Cognitive Structure of Emotions*, Cambridge University Press
 39. LeDoux, J., 1998, *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*, Simon & Schuster, New York, NY.
 40. P. E. Griffiths, 1998, *What Emotions Really Are: The Problem of Psychological Categories*, University Of Chicago Press, Chicago, IL.
 41. Damasio, A.R. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain*. Grosset/Putnam, New York, NY.
 42. Perlovsky, L.I. (1998). *Cyberaesthetics: aesthetics, learning, and control*. STIS'98, Gaithersburg, M.D.; Perlovsky, L.I. (1999). Emotions, Learning, and Control. In: *Proc. International Symp. Intelligent Control, Intelligent Systems and Semiotics*, Cambridge MA, pp.131-137
 43. Perlovsky, L.I. (2006). Modeling Field Theory of Higher Cognitive Functions. In: *Artificial Cognition Systems*, Ed. A. Loula and R. Gudwin, Idea Group, Hershey, PA.
 44. Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford Univ. Press, New York, NY.
 45. Harlow, H.F., Mears, C. (1979). *The Human Model: Primate Perspectives*, Washington, DC: V. H. Winston and Sons.

46. Berlyne, D. E. (1960). *Conflict, Arousal, and Curiosity*, McGraw-Hill, New York, NY; Berlyne, D. E. (1973). *Pleasure, Reward, Preference: Their Nature, Determinants, and Role in Behavior*, Academic Press, New York, NY.
47. Festinger, L. (1957). *A Theory of Cognitive Dissonance*, Stanford, CA: Stanford University Press.
48. Carpenter, G.A., Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine, *Computer Vision, Graphics and Image Processing*, 37, 54-115.
49. Mathematically, the condition that the object h is present with 100% certainty, is expressed by normalization condition: $\int l(\mathbf{X} | h) d\mathbf{X} = 1$. We should also mention another normalization condition: $\int l(\mathbf{X}(n)) d\mathbf{X}(n) = 1$, which expresses the fact that each received signal n is present with 100% certainty. This statement requires a qualification: it does not mean that the system is equally certain about all signals it receives, as discussed in section 5.4, a lower level in the hierarchy is responsible for activating-sensing input signals for a higher level; the lower level “keeps track” of activation levels (certainties) of signals, which it sends up to the higher level.
50. This construction is not unique. Expression (2) is inspired by the notion of likelihood between models and signals. Another general type of similarity measure suitable for the knowledge instinct, inspired the notion of mutual information in the models about the signals, is discussed in [8]. Here we would like to mention a modification of (2) for a specific case. Sometimes a set of observations N is more convenient to describe mathematically as a continuous flow of signals, for example, a flow of visual stimuli in time and space; then, it is convenient instead of eq.(1) to consider its continuous version,

$$L = \exp \int_N \ln \left(\sum_{h \in H} r(h) l(\mathbf{X}(n) | h) \right),$$

where N is a continuum of bottom-up signals, such as in time-space.

51. Perlovsky, L.I. (2004). Integrating Language and Cognition. *IEEE Connections*, Feature Article, 2(2), pp. 8-12.
52. Perlovsky, L.I. (2006). Symbols: Integrated Cognition and Language. In: *Computational Semiotics*, Ed. A. Loula and R. Gudwin, Idea Publishing, Hershey, PA.
53. Perlovsky, L.I. (1996). Mathematical Concepts of Intellect. In: *Proc. World Congress on Neural Networks*, San Diego, CA; Lawrence Erlbaum Associates, NJ, pp.1013-16;
- Perlovsky, L.I.(1997). Physical Concepts of Intellect. *Proc. Russian Academy of Sciences*, 354(3), pp. 320-323.
54. A desirable property of convergence of model parameters to their correct values for objects they represent, or to the global maximum of the likelihood is not necessary in practical computer applications or in workings of the mind all the time. Perceptions are not always correct at a first sight, sometimes, we have to move, to change a point of view, before we unambiguously recognize an object. Same is true in computer applications. The practical aspects of efficient convergence are considered in [8, 29].
55. Cramer, H. (1946). *Mathematical Methods of Statistics*, Princeton University Press, Princeton NJ.
56. Linnehan, R., Mutz, Perlovsky, L.I., C., Weijers, B., Schindler, J., Brockett, R. (2003). Detection of Patterns below Clutter in Images. In: *Int. Conf. on Integration of Knowledge Intensive Multi-Agent Systems*, Cambridge, MA Oct.1-3, 2003.
57. Zeki, S. (1993). *A Vision of the Brain* Blackwell, Oxford, England.
58. Perlovsky, L.I. (2004). Integrating Language and Cognition. *IEEE Connections*, Feature Article, 2(2), pp. 8-12.

Л.И. Перловский

59. Perlovsky, L.I. (2006). Neural Networks, Fuzzy Models and Dynamic Logic. Chapter in R. Köhler and A. Mehler, eds., *Aspects of Automatic Text Analysis* (Festschrift in Honor of Burghard Rieger), Springer, Germany, pp.363-386.
60. Жданов А.А. Бионический метод автономного адаптивного управления // *От моделей поведения к искусственному интеллекту* (под ред. Редько В.Г.). М.: УРСС. Серия "Науки об искусственном". 2006. С. 343-385.

Ионотропные и метаботропные реакции нейрона как инструменты нейронной памяти

А. Н. Радченко

Санкт-Петербургский институт информатики и автоматизации РАН

radch@spiras.nw.ru

Аннотация. Рецепторы, взаимодействуя с мембраной, медиаторами, ионными каналами и между собой, образуют на плазматической мембране нейрона мозаики ионотропных и метаботропных рецептивных кластеров (ИРК и МРК). Эти кластеры имеют три состояния — два активных с дискретными конформационными переходами и одно десенсибилизированное, оно представлено континуумом аналоговых состояний. Конформационные переходы ИРК и МРК трансмембранно передают сигналы, но разными способами: первые генерируют постсинаптический потенциал (ПСП), вторые — активируют вторичный посредник. В активном состоянии МРК функционируют как бистабильные (W) или моностабильные (R) молекулярные триггеры — в зависимости от мембранного потенциала. Это позволяет МРК записывать или воспроизводить энграммы. Исследовано формирование ИРК и МРК и их реакции на электрические и химические стимулы. Совместное их действие переводит МРК в третье (десенсибилизированное) состояние, где адгезия обездвиживает рецепторы. В ИРК такое уменьшение конформационной подвижности иммобилизует воротные заряды и инактивирует проводимость каналов, а в МРК — формирует элемент энграммы. Конформационные свойства МРК объясняют механизмы кратковременной и долговременной памяти и расслоение энграмм по возрасту. Вместе с синаптическим окружением мозаика МРК образует синаптически адресуемую запоминающую среду с опознающей выборкой информации, где запись (W) и воспроизведение (R) чередуются соответственно с динамикой корковых потенциалов.

1. Введение

Действие синапсов не ограничено субсинаптическими рецепторами и каналами. Медиаторы распространяются за пределы синапсов (spillover), интерферируя в нейроно-глиальных щелях и действуя на перисинаптические рецепторы. Их конформационные изменения активируют метаболические реакции нейрона, переходящие в эндогенные спайки. В отличие от субсинаптически вызванных (ионотропных) реакций, эндогенные спайки называют также фоновыми, паракринными или спонтанными. Вопреки названию эта активность детерминирована, зависит от входных сигналов и определяет выходные реакции нейрона [1]. Спонтанная активность исчезает под наркозом и при охлаждении, но всегда сопровождает состояние бодрствования. Ионотропная активность сохраняется в условиях наркоза и охлаждения, когда информационные функции мозга выключены. Это интересно, т. к. модели нейронных сетей и их приложения к искусственному интеллекту основаны только на субсинаптических реакциях нейронов. Цель данной статьи — разработать основы теории реальных нейронных сетей с учетом не только ионотропной, но и метаботропной активности, а также их взаимодействия.

Далее рассмотрены три аспекта проблемы.

Первый — анализирует чувствительность метаботропных и ионотропных рецептивных кластеров (МРК и ИРК) к медиаторам и электрическому полю в

связи с локализацией ИРК и МРК на мембране. Благодаря гистерезисно-конформационным свойствам МРК действуют как триггеры, запуская генерацию эндогенного спайка при взаимодействии с медиаторами и электрическим полем — постсинаптическими потенциалами (ПСП) ближних синапсов. Мозаика из 10^3 – 10^5 таких молекулярных триггеров подобно мелким островкам покрывает плазматическую мембрану нейрона. Каждый МРК является детектором локальной внеклеточной концентрации медиаторов. Множество МРК вместе с адресующими их окрестными синапсами формирует запоминающую среду, сочетающую в себе свойства оперативной и постоянной памяти. Адресацией называется локализация на данном МРК совокупности химических стимулов, достаточная для изменения его дискретного конформационного состояния.

Второй аспект посвящен консолидации памяти — исследованию гистерезисно-конформационных свойств МРК, переключаемых химическими и электрическими стимулами. В зависимости от мембранного потенциала МРК действует как бистабильный или моностабильный триггер, обеспечивая функции записи (W) и воспроизведения (R) энграмм соответственно. Показано, как сочетание химического и электрического стимулов присоединяет к конформационному процессу адгезионно-полимеризационные силы, переводя МРК в континуум монотонных состояний. Они определяют степень консолидации этого элемента энграммы и соответственно кратко- или долговременную память. Гиперполяризация мембраны стирает эту информацию, возвращая кластеру гистерезисные свойства. В экспериментах с воротными токами такую операцию называют кондиционированием.

Третий аспект касается адресации МРК окружающими его возбуждающими и тормозными синапсами. Их пространственная конфигурация, пре- и постсинаптические задержки, а также порогово-гистерезисные свойства МРК определяют набор паттернов, которые способны обеспечить когерентный приход порций возбуждающих синаптических медиаторов к МРК и исключить когерентные тормозные стимулы. В адресном пространстве такой набор паттернов образует область, называемую далее адресным доменом. Попадание в эту область паттерна внешних сигналов определяет выборку (адресацию) данного элемента памяти, а размер этой области (число паттернов) определяет ее опознающие свойства.

Переходя к подробному изложению, отметим, что рассмотренные свойства МРК — этого молекулярного триггера — состоят в том, что два из его состояний характеризуются различными энергетическими уровнями, а их разность в виде фонона (конформона), освобождаясь в одном из конформационных переходов, способна инициировать цепь взрывообразно расширяющихся внутриклеточных метаболических реакций. Такой квантово-энергетический аспект работы МРК роднит наш подход с новым физическим направлением исследований в информатике, называемым квантовые компьютеры [2]. Молекулярные триггеры — это агрегаты рецептивных молекул, сшитые димерными связями. Принцип димеризации рецепторов переключается с димерной парадигмой иммунологической памяти [3]. Димерные молекулярные агрегаты другого типа (ДНК, РНК), как известно, являются основой также и генетической памяти. Новый подход к анализу молекулярных, биофизических и информационных процессов в реальных нейронных сетях открывает в этой области знаний пер-

спективу, подобную той, которая была реализована в генетике после открытия молекулярных основ генетического кода.

2. Локализация нейрорецепторов относительно синапсов и разделение ионотропных и метаботропных функций

2.1. Конформационная чувствительность МРК

Обоняние обнаруживает несколько молекул, зрение — несколько квантов света. Такую чувствительность обеспечивают метаботропные рецепторы и внутриклеточные механизмы химического усиления [4]: "конформация одной рецептивной молекулы активирует много молекул трансдуцина⁴, каждая из них инактивирует сотни молекул цГМФ". За счет такого взрывообразного высвобождения метаболической энергии первичный сигнал возрастает в 10^4 – 10^7 раз. Этого достаточно для передачи сигнала между молекулярно-мембранным и мембранно-клеточным уровнями. Метаболический процесс завершается эндогенной деполяризацией (ЭД) начального сегмента аксона и генерацией "спонтанного" спайка. Этот механизм действует как в сенсорных, так и центральных нейронах. Но рецепторы сенсорной клетки настроены на специфические внешние влияния, а рецепторы нейрона — на молекулы лигандов или их комбинаций. Усилительный механизм ЭД позволяет нейрону, как и сенсорной клетке, легко преодолевать грубость механизма прямой деполяризации (ПД), обусловленную, как известно, низким емкостным импедансом клетки.

Действие ЭД и ПД отличается локализацией пусковых стимулов на мембране нейрона. В случае ЭД спайк инициируется одним из многих тысяч МРК в результате интерференции тормозных и возбуждающих медиаторов, секретлируемых двумя–тремя десятками разномодалых синапсов из ближней окрестности МРК. Пусковой стимул в этом химическом коктейле, действующем на МРК, локален, далее требуется внутриклеточное усиление. В случае ПД наблюдается обратное отношение — для деполяризации начального сегмента аксона необходима суммация многих локальных возбуждающих постсинаптических потенциалов (ВПСП). Обычно они рассредоточены по мембране. Метаболическая поддержка здесь не требуется. Надпороговый суммарный ВПСП, стремящийся направить передачу сигнала мимо хемочувствительного усилительного механизма, подобен его "короткому замыканию". Информация, представленная на мембране нейрона конформационными состояниями МРК, при этом не используется.

Для эндогенного возбуждения нейрона нужна потенциация эндогенного усилительного механизма. Ее вызывает, например, краткая гиперполяризация: по ее окончании нейрон генерирует серию эндогенных спайков (off-реакция). Потенциацию другого рода — долговременную — можно вызвать, интенсифицируя метаболические процессы ритмическим раздражением клетки. С помощью усилительного механизма реакция на конформационный сигнал передается на расстояние порядка 300 - 400 мкм от удаленных дендритов к соме и аксо-

⁴ В данном примере это родопсин.

ну без возбуждения мембраны дендритов [5]. Периферические синапсы, не вызывая достаточных суммарных ВПСП, способны инициировать лишь конформационные эффекты и последующее эндогенное усиление, перерастающее в ПД и генерацию «спонтанного» спайка. Этот неудачный термин порожден трудностями регистрации локальных конформационных изменений и традицией судить о реакции по ее электрическим проявлениям. Кроме того, эндогенное усиление часто сопровождается заметной деполяризацией клетки, которая появляется в форме пьедестала под спайком, маскируя инициирующий процесс. Поэтому процессы МРК–усиление–ЭД–ПД–спайк и ИРК–ПД–спайк трудно различить, особенно средствами электрической регистрации. Исключение показано на рис. 1, где метаботропные спайки не маскируются деполяризацией и возникают на фоне гиперполяризации [6].



Рис. 1. Спайк, возникающий на фоне развития ТПСП [6]. Калибровка 250 мкв и 20 мсек. Осциллограмма показывает относительную независимость пусковых механизмов "спонтанной" активности от поляризации мембраны.

Часто нейроны генерируют спайки, синаптическая активация которых не установлена. Поэтому к эндогенным спайкам естественно причислить пейсмекерную активность. Усилительные свойства пейсмекерного механизма отмечены экспериментально [7]. Отметим и другое сходство: пейсмекерные спайки инициируются кальцием, освобождаемым из внутриклеточного депо, а в случае метаботропных спайков тот же кальций может инициировать усиление как вторичный посредник. На рис. 2 приведен пример "спонтанной" активности в клетках культуры гиппокампа на стадии, когда они не имеют синаптических контактов, но имеют рецепторы и каналы. Здесь несинаптический процесс инициации спайка похож на пейсмекерный, но активность клеток псевдослучайна [8]. Авторы установили, что спайк сопровождается открытием одиночного ионного канала. По-видимому, он встроен в МРК, но заведомо не достаточен для ПД. Однако спайк возникает: конформационный переход МРК запускает эндогенный усилительный механизм, который деполяризует клетку. Деполяризация, как видно из рис. 2 может быть достаточной или недостаточной для открытия канала. Но если он открылся, то деполяризация усугубляется, облегчая генерацию спайка.

Интересно, что авторы этого исследования видят детерминистский смысл такой "спонтанной" импульсации в выполнении клеткой своей генетической программы развития — установления синаптических (или десмосомных) связей с другими нейронами. Благодаря эндогенным механизмам клетка может осуществлять сигнализацию не только внутрь клетки (outside-in signalling), но и изнутри наружу (inside-out signalling), как в данном примере.

Другим примером могут служить клетки Пуркинье (КП) мозжечка, где "спонтанная" активность представлена *обыкновенными спайками*, а вызванная — прямой деполяризацией — *сложными спайками*. Первая — производится параллельными волокнами клеток-зерен (без заметных ПСП), вторая — лианнными волокнами. Простые и сложные спайки резко отличаются друг от друга, поэтому в этом редком случае трудностей классификации инициирующего процесса (ЭД или ПД) не возникает. Впрочем, подобные данные многочисленны: еще в шестидесятых годах XX-го века Ч. Экклз отметил, что "*существует много примеров химической передачи возбуждения, при которой постсинаптические потенциалы не генерируются*" [9].

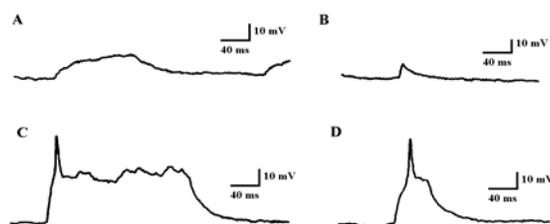


Рис. 2. Спонтанные потенциалы в клетках культуры гиппокампа [8]. А–В — платообразные потенциалы двух клеток. С–D — плато-потенциалы тех же клеток, индуцирующие спайки. Записи А и С принадлежат одной клетке, также как и В и D.

Встроенные в МРК каналы продуцируют пьедестал под метаботропными спайками. Возможно, из-за этого внесинаптические НМДА рецепторы считают ионотропными. Различить вызванную и "спонтанную" активность трудно, но возможно. Например, малая искусственная поляризация действующего нейрона не влияет на латентность эндогенного спайка, но смещает временную позицию вызванных спайков. М.Н. Ливанов [10] отмечает, что вызванные спайки не всегда возникают даже при надпороговой деполяризации. Однако при значительной искусственной деполяризации вблизи неизменного эндогенного спайка может появиться обрамляющая его пачка импульсов. Деполяризация, несомненно, облегчает генерацию эндогенных спайков. Но только они несут информацию о конформационных переходах МРК. Безупречным критерием спонтанной активности считают охлаждение или наркоз: их действие выключает метаботропную, но сохраняет вызванную активность. Информативность пространственно-временного кода определяется, в основном, относительными позициями спайков смежных нейронов, т.е. их латентными периодами, которые не зависят от дополнительной поляризации или влияющей на нее интенсивности раздражителя. Огромное количество информации, поставляемое в мозг, можно передать только с помощью фазово-импульсного кодирования.

Далее показано, что МРК действуют также как элементы памяти. Они реагируют на внешние стимулы и передают в клетку сигналы, которые активируют вторичный посредник. Он запускает серию эндогенных реакций, которыми клетка отвечает на химические или/и электрические стимулы. Имеются содержательные обзоры таких реакций [11, 12]. Главное свойство метаботропных рецепторов — их *высочайшая чувствительность и усиление конформационного*

сигнала в сотни тысяч раз. Упомянутый выше трансдуцин (родопсин) аналогичен многим нейрорецепторам, сокращенно GPCR — guanine protein-coupled receptors, а циклический нуклеотид цГМФ действует как *вторичный посредник*, запуская генерацию эндогенного спайка.

2.2. Структура нейрорецептивных молекул и их агрегация

Полипептидная цепь GPCR-рецептора пересекает мембрану, соединяя внешнюю и внутреннюю среды клетки. N-терминаль цепи оказывается вне клетки, а С-терминаль — внутри ее. Длинные концевые цепи этого интегрального белка в основном гидрофильны, а срединная часть содержит гидрофобные фрагменты. Они встраиваются в мембрану, где закручиваются в спирали, образуя один или несколько трансмембранных доменов. Семь таких спиралей ТМ1–ТМ7, характерных для GPCR рецепторов, показаны на рис. 3 А.

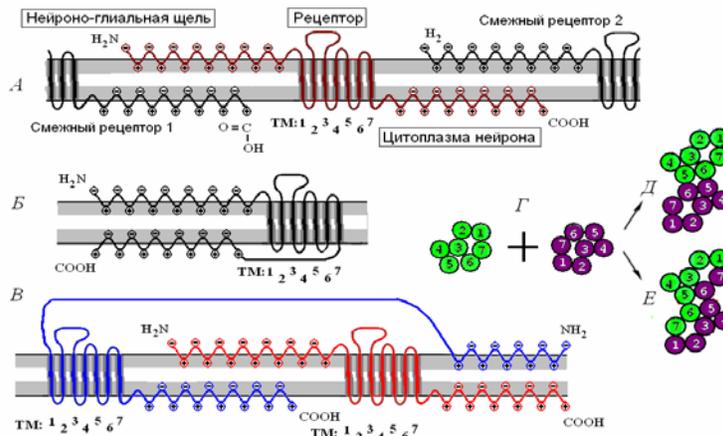


Рис. 3. Латеральное и трансмембранное взаимодействие рецепторов, формирующее рецептивные кластеры. Электронно-плотные 2.5 нм слои мембраны образованы головками фосфолипидов и растянутыми полипептидными цепями. Наибольший размер светлого слоя равен расстоянию между головками фосфолипидов (6.5–7.5 нм), наименьший — определяется переменной глубиной погружения заряженных групп (+) и (–) в липидный слой с обеих его сторон. А — Структура одиночного рецептора в мембране и его взаимодействие с окружающими средами и другими рецепторами. Б — Мономерная рецептивная молекула. В — Рецептивный димер. Г — Варианты объединения двух мономеров в контактный (Д) и комбинационный (Е) димеры [13].

Внешняя часть молекулы, взаимодействующая с медиаторами, имеет *участки-сенсоры*, которые определяют специфику рецепции. Биохимическими и генетическими методами выявлены молекулярные конструкции цепей многих рецепторов, определены места связывания с лигандами, выяснены конформационные механизмы рецепции на внеклеточной стороне мембраны. Для разных рецепторов также идентифицированы *участки-эффекторы* на стороне цитоплазмы. Их активные группы инициируют выходные реакции, в том числе, ка-

тализ вторичного посредника и характер сигнализации. От них зависит выбор метаболической реакции, например, потенциации усилительного механизма клетки или запуск спайка.

В приведенном примере межспиральные участки рецептора соединяются тремя внешними и тремя внутриклеточными петлями. Они и длинные концевые части молекулы стелятся по обеим сторонам мембраны, формируя вторичные структуры на поверхности раздела фаз. Амфифильные и гидрофобные аминокислотные остатки этих цепей частично погружены в билипидный матрикс (БЛМ) с обеих его сторон, а гидрофильные — препятствуют полному погружению. Поэтому цепи расщепляются на подслои, растягивая полипептидные связи: гидрофильные участки удерживаются в водных фазах снаружи и внутри клетки, а гидрофобные — погружаются в БЛМ с обеих его сторон навстречу друг другу. Рис. 3 А схематизирует такое расщепление и показывает, что погруженные подслои сближаются в БЛМ и могут трансмембранно взаимодействовать с такими же цепями смежных рецепторов. Пример гидрофобных профилей двух рецептивных молекул показан на рис. 4.

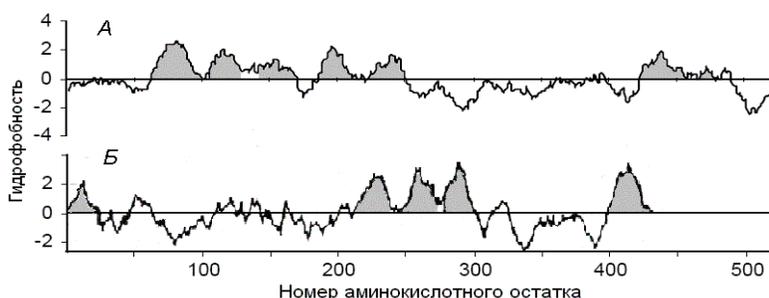


Рис. 4. Гидрофобные профили двух рецепторов. Затемненные участки соответствуют трансмембранным доменам. А — родопсин (семь доменов), Б — рецептор ацетилхолина (пять доменов) [14].

Подвижность боковых групп белковой цепи не зависит от ее вторичной структуры [15]. Сближаясь в центре БЛМ, погруженные подслои взаимодействуют с противоположащими подслоями смежных рецепторов. Глубина погружения зависит от гидрофобности аминокислотных остатков, их чередования в пептидной цепи, трансмембранного взаимодействия с оппозитными остатками, ионизации противоположащих пар и мембранного потенциала. Наиболее глубоко погружаются звенья, не перемежающиеся гидрофильными остатками. Например, большая часть молекулы бактериородопсина «тонет» в БЛМ [16]. Погружение можно углублять, нейтрализуя часть внеклеточных зарядов протонами (понижением pH). Кальциевые и другие бивалентные ионы, не только нейтрализуют эти заряды, но перезаряжают участки положительно. Нейтрализованные участки рецептивной молекулы теряют гидрофильность и могут погрузиться дополнительно (аналог гиперполяризации мембраны). Кальций-связывающие участки рецептора вне и внутри клетки влияют на его конформационную активность. Встречное погружение вне- и внутриклеточных цепей в мембрану может

удвоить их сближение, вызвать ионизацию и усилить трансмембранное взаимодействие. Такое же трансмембранное притяжение смежных рецепторов объединяет их в кластеры. Оно же влияет на их конформационные свойства.

Взаимодействуя, участки молекулы образуют вторичные и третичные структуры. В одиночном рецепторе (пример на рис. 3 Б) наибольшее трансмембранное взаимодействие можно ожидать от длинных концевых цепей доменов ТМ1 и ТМ7. Они замыкают часть молекулы в кольцо, где оказываются остальные домены. Их короткие междоменные цепи связываются слабее, образуя структуры, близкие к гексагональным (рис. 3 Г). Пары мономеров ТМ1–ТМ7 способны к димеризации (пример на рис. 3 Д). Смежные рецепторы с разобщенными концевыми группами доменов ТМ1 и ТМ7 могут взаимодействовать сильнее, образуя кольца, например, из 14-ти доменов (рис. 3 В). Эти домены также собираются в упаковку, близкую к гексагональной (рис. 3 Е). Чем выше концентрация рецепторов в мембране, тем более вероятны большие рецептивные олигомерные структуры. Они содержат сотни рецепторов и имеют размеры порядка 0.5–2 мкм [17]. В гексагональной структуре каждый рецептор может контактировать с шестью смежными рецепторами, но часть этих связей может быть занята каналами. В результате рецепторы объединяются в кластеры (side-to-side receptor association), а удерживающие их трансмембранные связи формируют оппозитные решетки «воротных» зарядов (ВЗ).

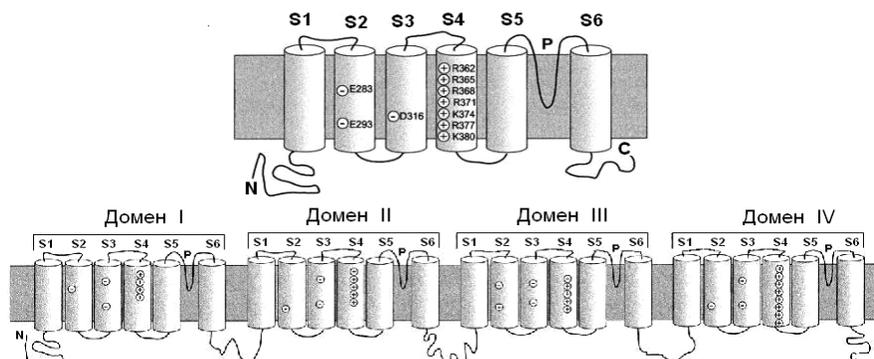


Рис. 5. Формирование шекеровского ИК объединением четырех белковых молекул [19]. Их спиральные участки изображены в форме цилиндров.

Главной проблемой клеточной сигнализации является *причинно-следственная связь между конформациями сенсорных и исполнительных цепей* рецептора. Этот механизм исключительно важен не только для понимания функций рецепторов, но и каналов [18]. Причину трансмембранной передачи сигналов обычно усматривают во взаимодействии и конформациях промежуточных звеньев рецептивной молекулы — спиральных протомеров, для чего анализируют их взаимное положение и взаимодействие их зарядов. Наш подход отличается тем, что переориентация спиральных протомеров в конформационном переходе считается не причиной, а следствием прямого трансмембранного электростатического взаимодействия концевых цепей и отчасти межспиральных петель рецептора, точнее — их погруженных в БЛМ участков, несущих ВЗ. Это

взаимодействие охватывает смежные рецепторы, служит причиной их димеризации и олигомеризации, а также делает конформационные изменения объединенных молекул кооперативными. Есть данные [13], что недимеризованные рецепторы не способны управлять ионными каналами. Понятно, что кольцо не связанных друг с другом рецепторов не сможет передать охватываемому каналу своих конформационных изменений. Поэтому для управления каналом предпочтительны многократные связи типа *E* (рис. 3), — они удерживают кольцо от разрыва. Пример трансляции конформационных изменений от рецепторов к каналу будет показан далее на рис. 10.

Молекулы ионных каналов (ИК) имеют сходные с рецепторами спиральные участки внутри мембраны и могут взаимодействовать с рецепторами и между собой по той же схеме. Рис. 5 показывает структуру одного из калиевых каналов, объединяющего четыре белковые молекулы. Аналогично ИК объединяются с рецепторами, образуя гетероолигомеры — канално-рецептивные структуры. В них конформации рецепторов могут управлять каналами, хотя некоторые каналы, не окруженные рецепторами, способны самостоятельно изменять свои конформации. В примере, представленном на рис. 5 [19], конформации канала предполагались результатом действия электрического сенсора, перемещающего спиральные части. Мы не считаем, что электростатические силы между спиралями достаточны для конформационных изменений канала, но данный пример показывает сходство строения рецепторов и каналов и говорит о механизмах их возможной агрегации в общие структуры. Такие структуры не всегда могут быть представлены в виде гексагональной упаковки протомеров, но гексагональная упаковка приближенно описывает реальную ситуацию.

2.3. Позиционирование рецепторов под синапсами и в синаптических ловушках

Есть данные, что одни и те же рецепторы могут действовать как метаботропные или ионотропные в зависимости от локализации. Она, в свою очередь, зависит от латеральной подвижности рецепторов в мембране и спilloвера — дивергенции медиаторов за пределы синапсов [20, 21], исследованной во многих работах [22, 23, 24]. Рассматривая распространение медиаторов от синапса к МРК (рис. 6), следует учитывать действие транспортеров, эвакуирующих медиаторы из синаптической щели, и еще два источника — секрецию медиаторов из варикозных образований аксонов и несинаптических терминалей непосредственно в нейроно-глиальные щели. Дивергирующие медиаторы и их дериваты оттесняют рецепторы и каналы от синапсов, направляя дрейф этих интегральных белков к центру синаптической ловушки. Возможны также сильные, но кратковременные влияния на этот процесс электрофоретических сил: ионизированные молекулы медиаторов могут оказаться между близко расположенными (в среднем на расстоянии 4 мкм) ВПСП и ТПСП. Напряженность поля между ними может достигать десятков киловольт на метр.

Структурообразующими элементами в разделении мозаик ИРК и МРК являются постсинаптические уплотнения (ПСУ). Входящие в него белки, например, SAP97 (от synapse-associated protein), PSD95 (от postsynaptic density) и, возможно другие, связывают С-терминалы рецепторов и каналов с ПСУ, удерживают

живая их в синаптической щели [25]. Эта фиксация обеспечивает высочайшую концентрацию рецепторов и каналов под синапсом и обеспечивает их мощное ионотропное действие. ПСУ препятствует метаботропным рецепторам контактировать с цитоплазмой и, следовательно, активировать вторичный посредник.

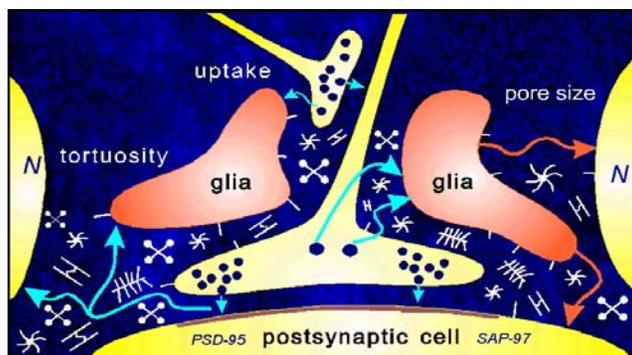


Рис. 6. Спилонер: распространение медиаторов за пределы синаптической щели [24].

В синаптической щели удерживаются, например, NMDA-рецепторы, а AMPA-рецепторы постепенно вытесняются из нее [26–28]. Вероятно, что вместе с рецепторами секреция медиаторов вытесняет из синаптической щели и некоторое количество незакоренных каналов. Вместе они способны дрейфовать от окрестных синапсов к квазистационарным точкам перисинаптической мембраны, где статистическая сумма гидравлических градиентов за многодневный период равна нулю. Эти локусы, называемые эффективными центрами *синаптических ловушек* [29], благоприятны для кластеризации незакоренных белков и образования МРК. Это можно проиллюстрировать на примере глутаматных рецепторов: *вокруг синапсов и МРК формируются зоны, свободные от рецепторов* [30, 31].

К. Прибрам [4], цитируя Ч. Шеррингтона, многократно призывал искать объяснение сложности разумного поведения сложностью микроструктур синаптических соединений. Мы упростили задачу, разделив синаптические структуры на множество синаптических ловушек: они полно учитывают не только морфологию в форме случайно зафиксированных тормозно-возбудительных синаптических конфигураций, но также все пре- и постсинаптические задержки. Множество разнородных синапсов и мозаику МРК удобно рассматривать как систему источников и стоков. В МРК могут включаться и каналы. Они уменьшают число связей между рецепторами, но продуцируют пьедестал, облегчающий генерацию метаботропных спайков и маскирующий их метаботропное происхождение. Эти электрические реакции, разумеется, слабее, чем в ИРК. Тем не менее, встроенные в МРК каналы дают повод считать внесинаптические НМДА рецепторы ионотропными. Похоже, что классификация рецепторов по признаку электрогенеза недостаточна. Целесообразно дополнить ее способностью рецептора активировать внутриклеточный усилитель.

Мощные синапсы сильнее отгесняют МРК, приближая его к более слабым синапсам. Это уравнивает влияние разнородных синапсов в «ловушке» на реактивность МРК и стабилизирует его адресные свойства в синаптической ловушке. *Если эффективность синапсов меняется, то МРК медленно смещается в синаптической ловушке так, что эти изменения компенсируются.* Конформации МРК в ловушке индуцируются не только экстрасинаптическими медиаторами данного нейрона, но и медиаторами синапсов смежных клеток. В механизме позиционирования МРК участвуют также медиаторы, которые секретируются в нейроно-глиальные щели несинаптическими терминалями и варикозами аксонов [32]. Хотя в адресации МРК участвуют все окрестные источники медиаторов, нам удобно называть такую структуру *синаптической* ловушкой. Первопричина функционального отличия МРК от ИРК обусловлена связыванием рецепторов и каналов с ПСУ. На перисинаптической мембране ПСУ не препятствует запуску усилительного механизма клетки и обуславливает высочайшую чувствительность клетки к конформациям МРК.

Резкое уменьшение концентрации медиаторов из-за обратного захвата и потерь при транспортировке, а также из-за присутствия тормозных медиаторов определяет *условие активации МРК: необходим когерентный приход возбуждающих медиаторов к МРК от многих разноудаленных окрестных синапсов и несинаптических источников.* Тормозные источники медиаторов в окрестности МРК ограничивают его реактивность, уменьшают число распознаваемых паттернов, т. е. увеличивают селективность адресации МРК кодовой структурой нервных импульсов. Для каждой синаптической ловушки, окружающей МРК, существует набор паттернов, который обеспечивает упомянутую когерентность. Любой из этих паттернов может вызвать конформационный переход МРК, совокупность активных паттернов называется *адресным доменом* данного МРК (подробнее см. рис. 16). Оптимальные размеры адресного домена и характеристики когерентной адресации МРК были исследованы ранее [30–33].

Итак, синаптическая ловушка представляет собой конфигурацию возбуждающих и тормозных синапсов, в функциональном центре которой встроено МРК — детектор локальной пороговой внеклеточной концентрации одного из антагонистических медиаторов. Каждый синапс входит в несколько смежных ловушек, но МРК принадлежит только одной ловушке. Далее показано, что **МРК в синаптической ловушке функционирует как элемент памяти.**

2.4. Структурно-функциональные особенности ИРК и МРК

Итак, рецепторы обретают метаботропные свойства, перемещаясь из субсинаптической в перисинаптическую мембрану — в направлении дивергенции медиаторов и их дериватов. Оставшиеся в синаптической щели каналы и связанные с ними рецепторы образуют ИРК, они обеспечивают мощные ионотропные реакции нейрона. Вытесняемые из синаптической щели рецепторы и не удержавшиеся в ней каналы кластеризуются в квазистационарных областях внесинаптической мембраны. В результате на суб- и внесинаптических мембранах образуются мозаичные рецептивно-канальные структуры ИРК и МРК. Они имеют разную локализацию и насыщенность каналами. На рис. 7 эта ситуация

схематизирована на примере гексагональных структур (см. также рис. 3 Г, Д и Е). Схема двухканального управления нейроном — субсинаптического и внесинаптического — показана в правой части рисунка, где медиаторы активируют ИРК и МРК, соответственно, в синаптических и нейроно-глиальных щелях.

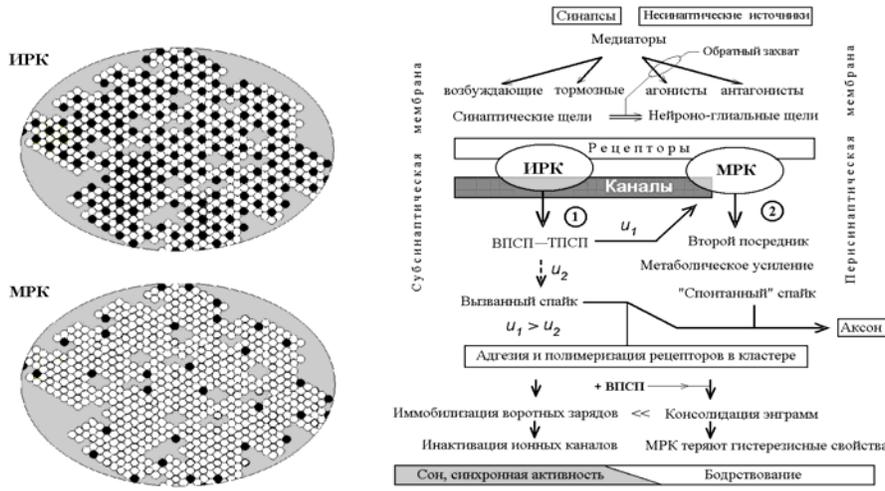


Рис. 7. Агрегация рецепторов и каналов (светлые и темные точки) в ИРК и МРК и схема формирования «вызванных» и «спонтанных» спайков электрическими (1) и химическими (2) стимулами. Различие рецептивно-канальных структур и функций ИРК и МРК обусловлено их различной локализацией на мембране относительно ПСУ.

Для ионотропного управления наиболее благоприятна структура, в которой соотношение рецепторов и каналов близко к 3:1. В случае МРК, где каналы не нужны, это отношение круто растет, и упаковка рецепторов уплотняется. Из рис. 7 следует, что в ИРК смежность рецепторов и число связей рецептор-рецептор приближается к четырем, а в МРК — к шести, что отмечено значком << в нижней части схемы рис. 7. Таким образом, главное отличие ионотропных и метаболитных рецептивных кластеров, состоящее в их различной локализации на мембране и резко различной чувствительности, дополняется третьим свойством — усилением межрецепторного взаимодействия в МРК. Оно влияет на конформационную подвижность рецепторов в кластере. В ИРК такое обездвижение рецепторов проявляется как иммобилизация воротных зарядов (ВЗ), десенситизация рецепторов и инактивация проводимости ИК [37]. Далее будет показано, что иммобилизация ВЗ является инструментом консолидации энграмм, а усиление межрецепторного взаимодействия по сравнению с ИРК делает консолидацию МРК прочнее.

Есть основания считать, что конформационная подвижность рецепторов в одном из конформационных состояний снижается из-за адгезии [38] и полимеризации [39]. Адгезионные и полимеризационные процессы объединяют рецепторы в «пэтки», «шапки» [40] и другие структуры [17], эволюционно связанные с эндоцитозом. Четырехкратную смежность рецепторов в ИРК естественно со-

относит к интенсивности иммобилизации ВЗ и инактивации ИК, а шестикратную смежность рецепторов в МРК – с консолидацией энграмм. Далее будет показано, что *управление конформационной подвижностью рецепторов в МРК — основа нейронной памяти.*

Схема рис. 7 сравнивает эффективность электрического и химического управления. Одиночный ВПСР субсинаптического канала всегда недостаточен для инициации спайка, в то время как конформация одиночного МРК – достаточна. Более того, сумма множественных ВПСР на дендритах аннигилирует при суммации с ТПСР одновременно возбуждаемых интернейронов. Это практически исключает электрогенное возбуждение спайка со стороны дендритов. Но оно возможно от сомы, где баланс ВПСР–ТПСР обычно отсутствует. ПСР сомы распространяются в обоих направлениях: антидромно (компонент u_1 на рис. 7) — на множество МРК, и ортодромно — к аксону (компонент u_2), где ВПСР может конформировать МРК и с помощью вторичного посредника иницировать спайк. Электрически и/или химически индуцированные конформационные переходы МРК управляют метаболизмом клетки и ее *эндогенной активностью* с помощью вторичных посредников. Таким образом, и химические и электрические стимулы могут иницировать «спонтанные» спайки, хотя в случае заметных ПСР этот термин можно оспаривать. Итак, эффективность МРК оказывается выше, чем ИРК.

Из рис. 7 видно, что эволюционно поздний и малочувствительный *ионотропный* механизм пытается *равноправно* взаимодействовать со сверхчувствительным *метаботропным* механизмом благодаря: 1) высочайшей концентрации в синаптической щели медиаторов, рецепторов и каналов, 2) резкому снижению этой концентрации за пределами синаптической щели путем обратного захвата медиаторов и 3) ослаблению конформационной чувствительности метаботропных рецепторов из-за их объединения в кластеры. *Малая концентрация медиаторов за пределами синапсов возмещается высочайшей чувствительностью метаботропных рецепторов.* При этом речь идет не о самих рецепторах, где отношения могут быть и обратными, а о том, что в одном случае в процесс вовлекается усилительный механизм клетки, а в другом — нет. От прямой деполаризации клетки и вызванных спайков *эндогенное возбуждение («спонтанные» спайки) отличается тем, что химический пусковой стимул и вызванный им конформационный переход локализованы на конкретном МРК.* Его адресацию обеспечивает возбужденно-тормозная конфигурация окружающих синапсов. Малую концентрацию медиаторов за пределами синаптической щели на МРК обеспечивает обратный захват молекул медиатора. Подразумевается их «штучная» интеграция на МРК, от одного–двух десятков окрестных синапсов и несинаптических источников. Согласование разных уровней чувствительности химического и электрического механизмов МРК явилось, очевидно, не простой задачей для эволюции. В малой окрестности МРК колебания концентраций медиаторов контрастны, информационно зависимы и информационно значимы. Контрастность и селективность адресации МРК обеспечивается тормозным медиатором. Его участие в синаптическом процессе способствует обострению настройки — минимизирует число одновременно адресуемых МРК.

На рис. 8 приведена схема активации ионотропной и метаботропной активности нейрона, которую мы будем использовать, связывая конформацион-

ные изменения ИРК и МРК с перемещением воротных зарядов. Внешний сигнал передается внутрь клетки с помощью электрического (А) или химического стимулов (Б). В первом случае деполяризация мембраны (ВПСП) ослабляет внутримембранную электростатическую «пружину», сжимающую внутримембранные подслои рецепторов. Как следствие, в МРК происходит скачкообразный деполяризационный конформационный переход (ДКП «a → b»). Внутримембранные слои МРК разобзаются и его приповерхностные слои выталкиваются из БЛМ в обе стороны. При этом энергия «пружины» освобождается.

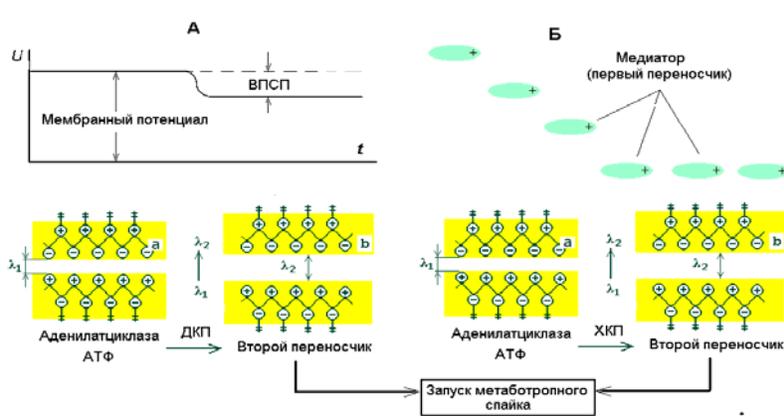


Рис. 8. Передача сигналов в клетку хемо- и электровзависимыми конформациями МРК. Электрическая (А) и химическая (Б) стимуляция конформаций МРК. Оба стимула увеличивают расстояния между ВЗ ($\lambda_1 \rightarrow \lambda_2$), вызывая соответственно ДКП или ХКП. Оба трансмембранных перехода выталкивают цепи в цитоплазму, активируя вторичный посредник и запуск эндогенного слайка.

Аналогичные процессы происходят при химической активации рецептивного кластера. Молекулы медиатора «салятся» на внеклеточную часть кластера, конформируя ее и частично вытягивая из БЛМ. Такое противодействие электростатическому трансмембранному притяжению зарядов индуцирует химический конформационный переход (ХКП) и «выстреливание» сопряженных частей кластера вне и внутрь клетки. В результате внешнее действие первичного медиатора транслируется внутрь клетки. Оба фактора способны катализировать синтез *вторичного посредника*, например, цАМФ или родственного ему цГМФ и активацию внутриклеточного усилительного механизма. Для этого процесса известен коэффициент усиления: в результате реакции средняя концентрация цАМФ оказывается в 10000 раз больше концентрации первичного медиатора. Похоже, что такое усиление с учетом дальнейшего расширения процесса достаточно для преобразования конформационного перехода в спайк. В других случаях эффект может быть ограничен синтезом цАМФ. Он может быть важен для потенциации клетки и запуска других метаболических механизмов клетки.

Из-за чувствительности МРК к стимулам двух модальностей затрудняется классификация рецепторов. Например, несинаптические НМДА рецепторы

считаются ионотропными, хотя они способны активировать вторичный посредник. Ионотропные реакции этих рецепторов можно объяснить тем, что в МРК легко встраиваются ИК. Если это возбуждающие ИК, то результирующая деполяризация усиливает действие второго посредника.

Конформационные изменения инициируются не только химическими и электрическими, но также механическими стимулами. Например, GPCR рецепторы эндотелиальных клеток кровеносных сосудов изменяют свои конформации в зависимости от артериального давления [41].

На каждой из сторон БЛМ полипептидный каркас жестко связывает расщепленные компоненты рецепторов, заставляя немембранные части повторять вариации липидной зоны рецептивного кластера. Эти части выталкиваются из БЛМ или частично погружаются в него соответственно при деполяризации (ДКП или ХКП) или гиперполяризации (ГКП) мембраны. Внеклеточные цепи реагируют на медиаторы, внутриклеточные – управляют синтезом вторичного посредника, а соединяющие их трансмембранные части (ионизированные слои, взаимодействующие внутри БЛМ) передают внешний сигнал внутрь клетки.

Выбор реакции клетки после синтеза цАМФ может направляться химическим средством внутриклеточных частей МРК, которые в результате ДКП или ХКП выталкиваются внутрь клетки из БЛМ. Сигнальное значение для этого процесса могут иметь сопутствующие химические группы внутримембранной части МРК. Они селективируют одну из многих возможных внутриклеточных реакций, а АТФ снабжает эти реакции энергией. Похоже, это различие определяет долговременные формы потенциации и депрессии. Известно, что связь большинства ферментов с мембранами и определенный порядок их расположения определяет не только их специфическое действие, но и последовательность реакций метаболического цикла. Этот порядок включает конформацию и ориентацию активных центров белков относительно БЛМ.

Известны три конформационные состояния нейрорецепторов. После удаления агонистов эти состояния могут сохраниться [42]. По нашим данным они определяют закрытие, открытие и инактивацию ионных каналов [37]. Но управлять ионными каналами, согласно данным [13], способны только димеризованные рецепторы. Поэтому наиболее интересны трансмембранные конформационные переходы внутри БЛМ. Обычно их анализ ограничивают межспиральным взаимодействием внутри БЛМ. ***Наша модель расширяет обычный подход, допуская, что перемещение внутримембранных спиралей — не причина, а следствие конформационной подвижности немембранных цепей рецептора и его ионизированных групп (ВЗ). Это единственное допущение*** позволит нам исследовать конформации рецептивного кластера как запоминающего элемента и проследить функции обширного множества этих кластеров на разных уровнях механизма памяти.

Подчеркнем, что химическое управление влияет на конформационные переходы прямо, в то время как электрические стимулы возникают опосредованно – через механизм активации ионных каналов и результирующие ПСП. Это двойное – прямое и косвенное – управление МРК реализуется хемо- и ионотропным действием медиаторов. Ионотропное действие косвенно и многоступенчато ослабляет или натягивает внутреннюю "электростатическую пружину" МРК: Ее переменное натяжение обуславливают вариации ПСП, продуцируемые открытием или закрытием ионных каналов.

Если ионотропное управление электростатическими силами внутри рецептора конформирует кластер сжимающими усилиями в зависимости от величины ПСП, то метаботропные стимулы, действующие снаружи и/или изнутри клетки, управляют растяжением половин кластера в разные стороны от БЛМ. Тем самым медиаторы противопоставляют действию сжимающей электростатической "пружины" силы другой природы и другого направления. Молекулярное моделирование такого "рецептора растяжения" [43], показывает, как в масштабе межатомных связей реализуется конформационное управление открытием и закрытием конкретного ионного канала. Но противоборство сжимающих и растягивающих сил внутри кластера обеспечивает лишь инициацию конформационных переходов. Важная для механизмов памяти фиксация одного их конформационных состояний достигается благодаря агрегации рецепторов и включению в этот механизм сил адгезии и полимеризации. Существенную ясность в конформационные процессы вносит математическое моделирование.

3. Конформационные состояния и переходы рецептивных кластеров

Взаимодействие нейрорецепторов с мембраной, электрическим полем, окрестными синапсами и внеклеточными средами вызывает реверсивные или бистабильные конформационные переходы. Они подобны переключениям соответственно моно- или бистабильного молекулярного триггера, управляемого химическими и электрическими стимулами. Этими триггерными свойствами можно управлять, изменяя мембранный потенциал: его снижение переводит рецептивные кластеры из моностабильного конформационного режима в бистабильный режим и наоборот. Кроме того, возможен перевод рецептивного кластера в третье состояние. Показано, как сочетание химического и электрического стимулов включает в конформационный процесс адгезионно-полимеризационные силы, обуславливая переход кластера в состояние, задерживаясь в котором кластер теряет триггерные (гистерезисные) свойства, что делает его элементом энграммы. Гиперполяризация мембраны возвращает кластеру гистерезисные свойства, стирая эту информацию. Найдена зависимость функций записи-воспроизведения энграмм от мембранного потенциала и границ конформационных переходов нейрорецепторов.

3.1. Конформационная модель

На границах сред с различными диэлектрическими постоянными, в нашем случае между мембраной и растворами, возникают доннановские потенциалы. Их полярность определяется нернстовским потенциалом и терминальными группами рецептора. Во внешнем и внутриклеточном растворах терминальные группы рецептора представлены радикалами $-\text{NH}_3^+$ и $-\text{COO}^-$. Доннановские потенциалы *компенсируются* противоионами вне и внутри клетки. Но в центральной электронно-прозрачной зоне мембраны, где противоионы отсутствуют, между доннановскими потенциалами возникает мощное электростатическое поле (порядка 10^8 В/м [38]). В нем погруженные оппозитные аминокислотные группы ионизируются, образуя трансмембранные диполи. Было показано

[37], что они являются воротными зарядами (ВЗ). *Некомпенсированные* доннановские потенциалы внутри БЛМ и мембранный потенциал u определяют электростатический профиль мембраны (рис. 9). Электростатическое притяжение решеток ВЗ дополнительно погружает расщепленные цепи рецептора в мембрану, смещая гидрофильно-гидрофобные границы внутрь бислоя.

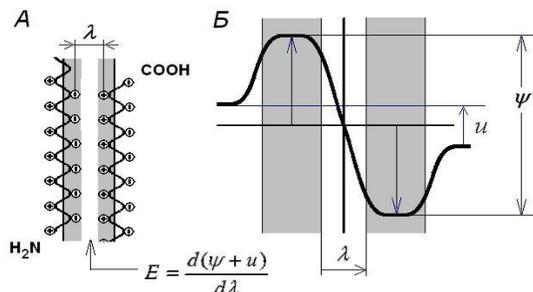


Рис. 9. Трансмембранное взаимодействие смежных рецепторов. *А* — схема взаимодействия решеток зарядов. *Б* — распределение сумм двух доннановских и нернстовского потенциалов внутри мембраны [45].

Погружение цепей рецептора в БЛМ перемещает межфазные границы. Похоже, что они проходят не по границам БЛМ, а внутри его — по зарядам противоположащих подслоев. Электростатическое притяжение решеток ВЗ погружает расщепленные петли рецепторов в БЛМ, чем смещает гидрофильно-гидрофобные границы внутрь бислоя. Следовательно, межфазные границы проходят не по границам бислоя, а внутри БЛМ — по зарядам противоположащих подслоев. Поликар [44] отмечал, что средняя электронно-прозрачная «липидная» зона клеточной мембраны имеет толщину 2.5–5.0 нм, варьирующуюся в зависимости от условий. По данным [45] расстояние между осмиофильными (т.е., вероятно, и гидрофильными) слоями меняется от 1.4 до 2.2 нм при изменении рН среды от 7 до 9. В обоих случаях данные искажены осреднением и далее используются как ориентировочные.

Рис. 3, 8 и 9 показывают, что **рецептивный кластер подобен заряженному конденсатору с пластинами на упругой подвеске**. Сотни тысяч таких МРК в виде малоподвижных островков оккупируют межсинаптическую мембрану нейрона. Полипептидный каркас связывает расщепленные компоненты рецепторов, заставляя их внеклеточные части зеркально повторять перемещения зарядов в центральной липидной зоне. Деполяризация выталкивает концевые части рецепторов из БЛМ вовне и внутрь клетки одновременно. При гиперполяризации они погружаются в БЛМ встречно. Эти перемещения сопровождаются деполяризационными и гиперполяризационными конформационными переходами (ДКП и ГКП), они определяют закрытие, открытие и инактивацию ионных каналов.

Важные детали ДКП и ГКП, в том числе существование трех конформационных состояний рецептивного кластера, выявляют его вольт-

конформационные характеристики (ВКХ). Они выводятся из баланса электростатических P_e , химических P_c и упругих сил P_m , действующих на кластер:

$$P_e + P_c + P_m = 0 \quad (1)$$

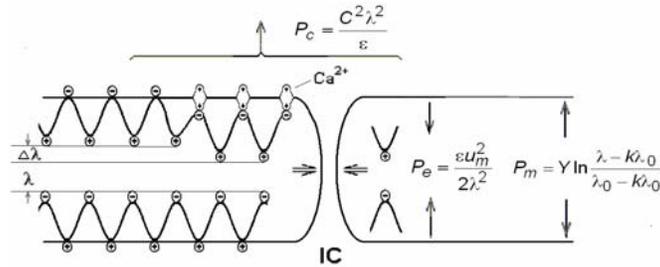


Рис. 10. Баланс трансмембранных сил на рецептивном кластере. Полые стрелки – внутрикластерное давление. Оно передается ионному каналу (IC), если рецепторы, окружающие канал, объединены в кластер [13]).

Рис. 10 поясняет направление сил и создаваемое ими внутримембранное давление. Эти силы зависят от переменных E , ρ , λ , k и Y — напряженности поля, концентрации зарядов в среде, расстояния между решетками зарядов, конформационного ограничения и модуля Юнга боковых групп полипептидных цепей. Если рецепторы в кластере, в котором находится канал, связаны между собой по схеме рис. 3 Е, то создаваемое ими внутримембранное давление будет передаваться ионному каналу (IC).

$$P_e + P_c = \frac{\epsilon}{2} \left(\frac{U_m}{\lambda} + \frac{\rho}{\epsilon} \cdot \lambda \right)^2, \quad P_m = Y \int_{\lambda_0 - k \lambda_0}^{\lambda - k \lambda_0} \frac{d\lambda}{\lambda} = Y \ln \frac{\lambda - k \lambda_0}{\lambda_0 - k \lambda_0} \quad (2)$$

Здесь в скобках — компоненты напряженности поля U_m / λ и $\rho \lambda / \epsilon$, вызванные соответственно мембранным и доннановскими потенциалами, ϵ — диэлектрическая постоянная, λ и λ_0 текущее и начальное расстояния между зарядами.

Химический компонент P_c выводится из уравнения Пуассона.

$$\frac{\partial^2 \psi}{\partial \lambda^2} = -\frac{\rho}{\epsilon} \quad (3)$$

Его интегрирование по λ дает доннановские компоненты напряженности поля E_c :

$$E_c = -\frac{\partial \psi}{\partial \lambda} = \frac{\rho}{\epsilon} \lambda \quad (4)$$

Противодействие перемещению зарядов P_m выведено из закона Гука, где k — ограничение подвижности зарядов ($0 < k \leq 1$); $k\lambda_0$ — расстояние предельного сближения зарядов (уменьшается при взаимодействии с ионами H^+ и Ca^{2+}); Y — модуль Юнга. Подставляя эти данные в баланс сил, и вводя обозначение $y = \lambda / \lambda_0$, получим уравнение:

$$\left(\frac{u_m}{\lambda_0 y} + \frac{\rho}{\varepsilon} \cdot \lambda_0 y\right)^2 = 2 \frac{Y}{\varepsilon} \ln \frac{1-k}{y-k} \quad (5)$$

Откуда

$$u_m = B y \sqrt{\ln \frac{1-k}{y-k}} - A y^2 \quad (6)$$

где $A = \lambda_0 \rho / \varepsilon$ и $B = \lambda_0 \sqrt{2Y / \varepsilon}$ — константы. Для $\lambda_0 = 1.4 \cdot 10^{-9}$ м, $\varepsilon = 10.8.85 \cdot 10^{-12}$ Ф/м [38], и $Y = 0.67 \cdot 10^6$ Н/м² [39], получим $B = 0.12$. Временно исключив химические влияния ($A = 0$), можно построить вольт-конформационные характеристики (ВКХ), $y = f(B, u_m, k)$, представленные на рис. 11.

Пример ВКХ при $k = 0.23$ демонстрирует, как увеличение u_m сближает заряженные подслои МРК, причем в позиции 1 происходит скачок — *гиперполяризационный конформационный переход* (ГКП) в позицию 2. Дальнейшее увеличение МП асимптотически уменьшает расстояние между зарядами, приближая его к конформационному ограничению $k\lambda_0$. Возврат кластера в разобленную конформацию — *деполяризационный конформационный переход* (ДКП) — происходит при меньшем мембранном потенциале, чем ГКП. При $k < 0.308$ ВКХ содержит гистерезисный участок 1-2-3-4-1 (заштрихован). Его ширина зависит от ограничения на конформационную подвижность зарядов k . В разобленной конформации, когда подслои МРК выталкиваются из мембраны, баланс сил дополняется *поверхностными силами адгезии и полимеризации* [38, 39]. Они консолидируют кластер, фиксируя возросшее значение k . Анализ формулы (6) показывает [37], что гистерезисные свойства ВКХ исчезают при

$$k = k_s = 2 / (e^{3/2} + 2) \cong 0.308 \quad (7)$$

Нарастающие конформационные ограничения $k = 0.15, 0.23, 0.308, 0.4, 0.6, 0.8$ определяют консолидационную эволюцию МРК. В пределах гистерезисного участка МРК действует как **молекулярный триггер**. Его состояния «а» и «б» есть следствия деполяризационного и гиперполяризационного конформационных переходов (ДКП и ГКП) кластера. В зависимости от мембранного потенциала (МП) и ограничений на подвижность ВЗ, свойства рецептивного кластера *четырежды* качественно меняются: в области W ($V_{иДКП} < V_{и} < V_{иГКП}$) он действует как **бистабильный триггер**, в области R ($0.308 > V_{и} > V_{иГКП}$) — как

одновибратор, в случае $k = 0.308$ кластер становится **триггером Шмитта**, а далее переходит к **континууму монотонных ВКХ** ($k > 0.308$), которые соответствуют консолидации разной прочности.

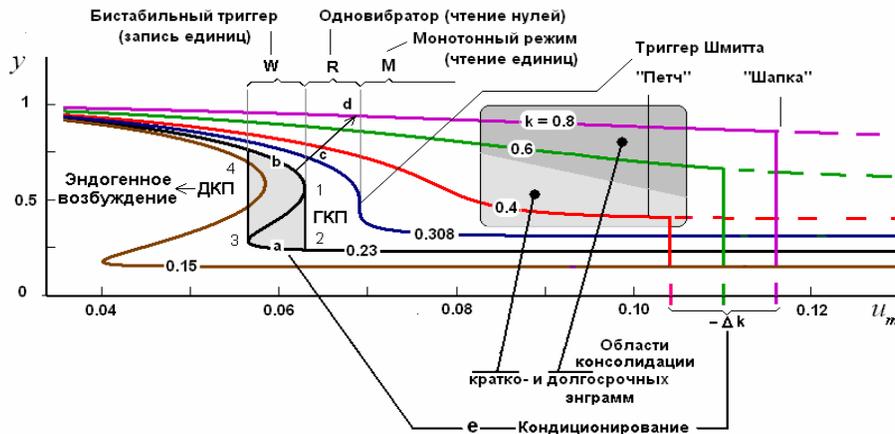


Рис. 11. Характеристики молекулярного триггера (MPK) в зависимости от мембранного потенциала (u_m) и их эволюция в зависимости от уменьшающейся конформационной подвижности воротных зарядов ($1 - k$).

Действие химических сил на ВКХ представлено на рис. 12. Они заданы переменной концентрацией медиатора $A_0 < A_1 < \dots < A_3$ [$0 < 0.02 < 0.04 < 0.06$ — безразмерные величины] в центре синаптической ловушки, ограничением на конформационную подвижность ВЗ, $k = 0.23$, и упругими свойствами МРК полипептидных цепей, $B = 0.12$. Повышение концентрации медиатора деформирует ВКХ. При пересечении кривой (A_2) с мембранным потенциалом u_0 возникает **химический конформационный переход** (ХКП). Его можно вызывать для любого значения абсциссы в пределах от ДКП до потенциала, который определяется точкой перегиба ВКХ для $k = 0.308$. ХКП сопровождается расширением гистерезисной петли (на выделенных петлях расширение от 5.4 мВ до 20 мВ), что важно для удержания МРК в разобленном состоянии, **b**, для консолидации бита памяти. Моменты адресации элементов мозаики МРК сопровождаются расширением гистерезисной петли адресованного кластера.

Освобожденный квант энергии (фонон, конформон), активирующий эндогенный усилительный механизм, пропорционален площади гистерезисных петель. Гистерезис заметно больше при ХКП, чем при ДКП. **Прирост энергии обеспечивается химическим средством медиатора и рецепторов МРК.** Оба механизма — ДКП и ХКП — инициируют удлинение электростатической «пружинки». Но ДКП является результатом ее ослабления, а ХКП — результатом ее дополнительного растяжения. *Химический запуск эндогенного спайка оказывается более мощным, чем в случае деполяризации.* Кривые на рис. 12 носят качественный характер, т.к. формулы (3)–(6) не учитывают механизм действия медиатора по принципу «ключ–замок».

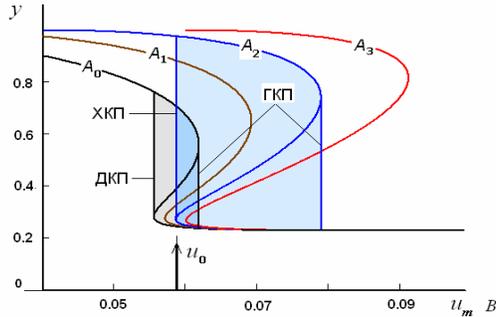


Рис. 12. Химическая инициация конформационного перехода МРК при повышении концентрации медиатора $A_0 < A_1 < A_2 < A_3$.

Представленное на рис. 11 разделение значений мембранного потенциала на области W, R или M соответствует квазистационарным значениям МП. Они определяют конформационные режимы МРК для данного нейрона. В области W оба конформационных состояния **a** и **b** стабильны (бистабильность). Переходы между ними ДКП (ХКП) и ГКП происходят под влиянием ПСП или медиаторов. То же происходит в области R, но здесь стабильно только состояние **a**, а состояние **b** удерживается лишь на время действия ПСП. По его окончании возвращается состояние **a** (моностабильность). В области M возможны лишь аналоговые упругие деформации МРК без конформационных переходов. Конформационные режимы изменяются (переключаются) относительно медленно, а конформационные состояния — быстро. Для консолидации МРК ("1" энграмма) благоприятен режим бистабильных конформаций — область W, где химическая адресация МРК может сочетаться с деполяризационным смещением мембранного потенциала (ВПСП).

Перечислим свойства рецептивных кластеров.

1. Дискретные конформационные состояния МРК соответствуют двум гистерезисным состояниям трансмембранного триггера (ВЗ сближены — «**a**») и (ВЗ разобщены — «**b**»). В состоянии «**a**» активные центры рецепторов изолированы от сред по обе стороны БЛМ, а в состоянии «**b**» взаимодействуют с этими средами.

2. Конформационные состояния «**a**» и «**b**» соответствуют разным энергетическим уровням. Энергия, аккумулируемая в межуровневом переходе ГКП и освобождаемая в ДКП или ХКП (площадь гистерезисной петли), зависит от параметра k .

3. Эти энергетические (и пусковые) свойства теряются в процессе иммобилизации ВЗ, когда $k \rightarrow 1$, и гистерезисная петля исчезает (континуум монотонных ВКХ от «**c**» до «**d**»).

4. Химически инициированный конформационный переход (ХКП) освобождает больший (по сравнению с ДПК) квант энергии. Поэтому химические механизмы адресации МРК являются основными для эндогенного возбуждения нейрона.

5. Освобождение энергии, запасенной МРК, можно инициировать тремя стимулами — взаимодействием кластера метаботропных рецепторов с агони-

стами (рис. 12), удалением ионов кальция, например, действием кальмодулина ($-\Delta\lambda$ на рис. 10) и деполяризацией (рис. 11).

6. Бистабильные конформационные режимы W благоприятны для записи информации, а моностабильные R — для ее считывания. Переключение этих режимов зависит от смещения мембранных потенциалов относительно ГКП (см. рис. 11).

Итак, переключение рецептивного кластера из состояния « a » в « b » и переход в « c » и « d », сопровождается обездвижением зарядов (на рис. 11 параметр k изменяется от 0.23 до 0.8, $\Delta k = 0.77$). Переход МРК от гистерезисного к монотонному состоянию задается ВПСП. Его длительность должна обеспечить инициацию адгезионного процесса. В противном случае и, тем более, при ТПСП гистерезисное состояние сохраняется. В зависимости от полярности ПСП состояние МРК будет представлять элемент энграммы «1» и «0». Область консолидации — возрастающие значения $k = 0.23, 0.308, 0.4, 0.6, 0.8$ — выделены на рис. 11 затенением.

3.3. Эволюция конформационных состояний МРК

Дихотомическое управление конформационными процессами МРК, представленное как формирование элементов «0» или «1» энграммы, схематизировано на рис. 13. Запись нуля соответствует нижней части рисунка, где обратимые конформационные переходы между двумя дискретными состояниями МРК ($a \leftrightarrow b$) происходят в режиме R , т.е. при общем гиперполяризационном смещении мембранного потенциала (МП). Здесь МРК действует как одновибратор: стимул вызывает деполяризационный конформационный переход (ДКП), а его снятие — гиперполяризационный конформационный переход (ГКП) — возврат в состояние a . Амплитуда стимула, очевидно, должна превышать ширину гистерезисной петли на вольт-конформационной характеристике (ВКХ), а конформационное состояние b должно быть кратковременным. Совпадение адресных сигналов с ТПСП (или отсутствием ВПСП) восстановит начальную конформацию по схеме $a \rightarrow b \rightarrow a$. Такой МРК сохранит гистерезисные свойства, что означает нуль энграммы — «0».

В режиме W , т.е. при понижении МП и бистабильных конформациях a — b , возможен необратимый переход МРК из b в континуум монотонных состояний c — d по схеме $a \rightarrow b \rightarrow c \rightarrow d$ (верхняя часть рис. 13). Это произойдет, если состояние b сочетается с ВПСП в течение времени, достаточного для потери конформационной подвижности МРК. Условие совмещения химического и электрического стимулов обозначено как $(X+U)$. Напомним, что совмещение адресации с ВПСП усугубляет выталкивание цепей МРК из БЛМ и дольше удерживает их в зоне, где в баланс сил включаются адгезия и полимеризация [35, 36]. Они уменьшают конформационную подвижность МРК, и его вольт-конформационная характеристика (ВКХ) становится монотонной, переходя границу конформационного гистерезиса ($k > 0.308$, см. [37]). Это позволяет записать единичный элемент энграммы — «1».

Независимо от конформационных режимов R или W и эволюции конформационных состояний МРК, а именно: $a \leftrightarrow b$ (запись нуля) или $a \rightarrow b \rightarrow c \rightarrow d$ (запись единицы), в обоих случаях переход $a \rightarrow b$ освобождает квант энергии. Он активирует вторичный переносчик, вызывая при записи эндогенное возбуждение (ЭВ) и «спонтанный» спайк. При воспроизведении спайки будут возникать

только на гистерезисных МРК, т.е. нули энграмм будут воспроизводиться как единицы, а единицы – как нули. Таким образом, энграмма воспроизводится инверсно по отношению к записи. Повторная инверсия, например, с помощью тормозных синапсов восстанавливает запись.

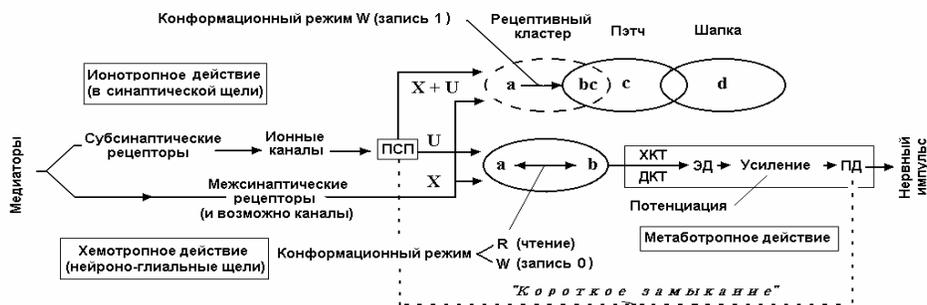


Рис. 13. Эволюция конформационных состояний МРК: X и U — химические и электрические стимулы, X+U — их совмещение. Режимы записи W (верхний канал) или чтения R (нижний канал) задаются смещением МП нейрона. Он запускается химически или электрически индуцируемыми конформационными переходами ХКП и/или ДКП. Пунктир показывает прямую деполяризацию (ПД) нейрона в обход эндогенного усилителя (вызванный спайк).

3.2. Формирование элемента энграммы и стадии экзоцитоза. Прочность консолидации кратковременной и долговременной памяти

После конвергенции рецепторов к центру синаптической ловушки сформированный МРК выполняет функции элемента памяти. Его конформационные переходы МРК вида $a \leftrightarrow b$ скачкообразно изменяют расстояния между ВЗ ($\lambda_1 \leftrightarrow \lambda_2$). При гиперполяризации притяжение ВЗ возрастает, инициируется ГКП, цепи погружаются в БЛМ (λ_1), и их контакт с обеими внеклеточными средами уменьшается. Связывание МРК с агонистом или деполяризация мембраны, т. е. ХКП или ДКП, выталкивает цепи рецепторов из БЛМ (λ_2) наружу и внутрь клетки, заставляя МРК контактировать с обеими средами. Это инициирует два процесса: 1) в цитоплазме активируется вторичный посредник и генерируется эндогенный спайк; 2) на поверхностях БЛМ активируются процессы адгезии и полимеризации. Если они длительны, то трансмембранная подвижность рецепторов (ВЗ) уменьшается или блокируется (процесс $a \rightarrow b \rightarrow c$ на рис. 11). Далее полимеризация увеличивает жесткость "подвески" ВЗ ($c \rightarrow d$), и кластер теряет гистерезисные свойства. Его вольт-конформационная характеристика (ВКХ) становится монотонной (область М), что соответствует записи "1" энграммы.

Секреция медиаторов и реакции рецептивных кластеров соотносятся как экзо- и эндоцитоз. Механизм памяти следует начальным фазам эндоцитоза: после конформационного перехода $\lambda_1 \rightarrow \lambda_2$ МРК преобразуется в "пэтч" и "шапку", где адгезия и затем полимеризация с помощью, например, клатрина постепенно увеличивают жесткость подвески ВЗ. Далее "шапка" преобразуется в "ямку",

она инвагинирует в клетку, образуя везикулу, происходит ее отрыв и т. д. в соответствии с рис. 14, где эти стадии сопоставлены с эволюцией ВКХ. Механизм памяти эксплуатирует только часть трофических функций клетки. Начальной стадией консолидации, похоже, является "пэтч", а достаточной — "шапка". При этом дискретные конформационные свойства МРК заменяются монотонными (кратковременная память), МРК теряет способность запасать (ГКП) и отдавать (ДКП или ХКП) энергию. Если гистерезисные свойства МРК не были восстановлены электрической активностью мозга, индуцирующей ГКП, то дальнейшая полимеризация рецепторов постепенно упрочнит МРК, и память будет долговременной.

Продолжение трофического цикла не изменяет результат: бит информации, соответствующий "1", сохраняется в памяти также после вакуолизации и распада вакуоли, — важно, что центр данной синаптической ловушки перестает быть источником "спонтанной" активности. Для постоянного хранения "1" желательнее, чтобы вакуоль не отрывалась от мембраны или рецепторы, регенерированные на последующих стадиях эндоцитоза, не возвращались бы полностью в данную синаптическую ловушку. В этом случае забывание можно представить как появление и кластеризацию на месте бывшей "шапки" новых рецепторов. Тогда память окажется кратко- или среднесрочной.

Такой возврат маловероятен, т. к. регенерированные рецепторы, скорее всего, равномерно распределятся между многими тысячами МРК данного нейрона. Более того, есть данные, что такие рецепторы встраиваются в уже существующие кластеры, а не в освободившееся место бывшей "шапки". Такие изменения в синаптической ловушке способны удерживаться всю жизнь. Однако для чтения энграмм в них должны сохраниться не только единицы, но и нули — гистерезисные МРК, т. е. нейроны с МРК с ВАХ в областях R и W.

При отсутствии такого смещения (режим R) или смещении в обратном направлении процесс не достигает стадии "пэтча", и консолидация не успевает начаться. Гистерезисные свойства МРК в этом случае сохраняются, что соответствует записи "0" ($a \rightarrow b \rightarrow a$). Для записи "0" благоприятен моностабильный режим — область R, где конформации МРК реверсивны: при повышенном МП нейрон "отказывается от пищи", прерывая эндоцитоз. Стабильность нулей энграммы (интактные гистерезисные свойства МРК) и стабильность ее единиц (монотонные состояния МРК) одинаково важны для памяти. Поэтому необходимо учитывать все дестабилизирующие факторы.

Ионотропные и метаболотропные реакции нейрона

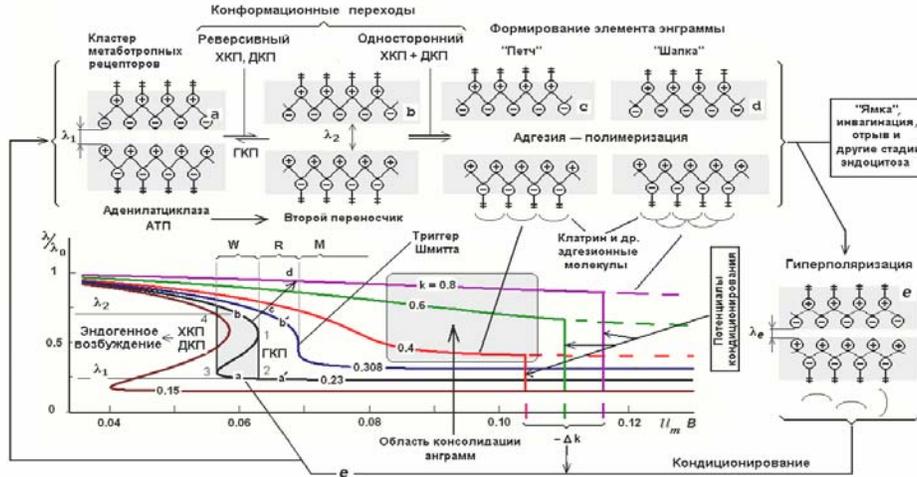


Рис. 14. Эволюция конформационных состояний МРК от записи (W) к чтению (R) и хранению (M) энграмм [4, 8]. Химические или электрические стимулы в области R вызывают реверсивные конформации $a \leftrightarrow b$ (ХКП \rightarrow ГКП или ДКП \leftrightarrow ГКП). Сочетание ХКП и ВПСП в области W включает в процесс адгезию и полимеризацию ($a \rightarrow b \rightarrow c \rightarrow d$), которые модифицируют МРК в "пятч" и затем "шапку" (показаны скобками). Возможна дихотомия: 1) инвагинация, образование "ямки" и последующие стадии эндоцитоза, 2) возврат ($d \rightarrow e \rightarrow a$) – распад "пятча" или разрушение "шапки" гиперполяризацией.



Рис. 15. Расслоение энграмм по возрасту и потенциалам кондиционирования, $u_0 < u_1 < \dots < u_i$. Параметры возрастающих конформационных ограничений k и жесткости полипептидных цепей Y соотнесены с адгезией и полимеризацией рецепторов в МРК.

Рассмотренные стадии эндоцитоза — кластеризация, образование "пятчей", "шапок", "ямки" и т. д. — присущи всем клеткам. Нервные клетки отличаются тем, что они подвергаются электрическим влияниям, например во время синхронной ритмической активности мозга. Ее гиперполяризационные полу-волны могут деполимеризовать рецепторы и вернуть МРК в начальное гистерезисное состояние, $d \rightarrow e \rightarrow a$. Параметры такого "кондиционирования" мембран изучались при исследовании воротных зарядов [19]. От длительности процес-

сов адгезии полимеризации на интервале $c-d$ зависит потенциал кондиционирования, а, следовательно, возраст и прочность энграмм. Четкое разграничение краткосрочной и долговременной памяти вряд ли возможно, но свежие энграммы можно условно соотнести с ограничением конформационной подвижности ($k \rightarrow 1$), а прочно консолидированные — с возросшей жесткостью полипептидных цепей (увеличение модуля Юнга Y в процессе полимеризации рецепторов). Эти отношения схематизированы на рис. 15.

Независимо от конформационных режимов R или W и эволюции конформационных состояний МРК, а именно: $a \leftrightarrow b$ (незавершенность консолидации – запись "0") или $a \rightarrow b \rightarrow c \rightarrow d$ (запись "1"), в обоих случаях переходы $a \rightarrow b$ освобождают кванты энергии, пропорциональные площади гистерезисной петли. Они содействуют активации вторичного посредника, инициируя при записи эндогенное возбуждение и "спонтанные" спайки. Однако при воспроизведении спайки будут возникать только при адресации гистерезисных МРК, т.е. нули энграмм будут воспроизводиться как единицы, а единицы – как нули. *Инверсия энграммы при воспроизведении позволяет игнорировать заключительные стадии эндочитоза: сохранится или исчезнет в синаптической ловушке "шапка" или "ямка" – совершенно безразлично для записи "1"*. В любом случае она будет воспроизводиться инверсно как "0", т. е. правильно. Однако остальные МРК с гистерезисными ВКХ (т. е. "0" и интактные кластеры), должны быть сохранены — они обеспечивают правильное инверсное воспроизведение "0" энграммы как "1". Первоначальная форма записи легко восстанавливается, например, повторной инверсией при передаче информации через тормозные синапсы.

Консолидация МРК, сопровождаемая увеличением k (адгезия) и Y (полимеризация) в соответствии с рис. 15, стратифицирует энграммы по потенциалу кондиционирования. Старые энграммы будут устойчивее к разрушающим их гиперполяризационным воздействиям. Можно ожидать, что длительность консолидационных процессов — вся жизнь [46]. Но у нейронов с возросшим количеством "пэтчей", "шапок" и "ямок" токопроводящая поверхность уменьшается, и поддержание обмена достигается повышением МП. Это приближает застарелые энграммы к границе кондиционирования, где амнестические процессы, индуцируемые гиперполяризационными полуволнами ритмической активности мозга, "заставляют бледнеть изображения в кладовой памяти" [14].

3.3. Адресный механизм: паттерны когерентной адресации и опознающая выборка

Пространственные параметры каждой синаптической ловушки и ее временные свойства (пре- и постсинаптические задержки) определяют индивидуальную настройку МРК на "свои" афферентные паттерны [35,36,48]. Паттерны, конформирующие МРК, можно выписать как табличную форму булевой функции. На рис. 16 расстояния от МРК к окрестным синапсам пронумерованы величиной задержки сигналов. Зная задержки, легко указать паттерн когерентного прихода всех возбуждающих сигналов к МРК и отсутствие хотя бы одного тормозного. Один из эффективных паттернов выписан под рисунком. На расстояниях (задержках) 14, 16, 17 и 19 (отмечены буквой Z) не оказалось ни возбуждающих, ни тормозных синапсов. Эти позиции кода можно заполнять произвольно. В результате в булеву форму войдет 16 паттернов.

В общем случае, когда Z -позиций много и конформационный порог МРК невысок, допустимо пропадание части импульсов паттерна или появление лишних импульсов (единиц) на месте нулей. Такие паттерны — точки булевого (адресного) пространства — присоединяются к перечисленным паттернам, образуя выпуклое подмножество — *адресный домен*. Число переменных этой булевой функции равно числу нервных волокон, влияющих на конформацию МРК. Однако в общем случае эту булеву функцию нужно рассматривать в расширенном пространстве, включив в него все несущественные переменные. Это поможет отразить конформации всех МРК данного нейрона и преобразование, заданное на многомерном входном потоке нервных импульсов.

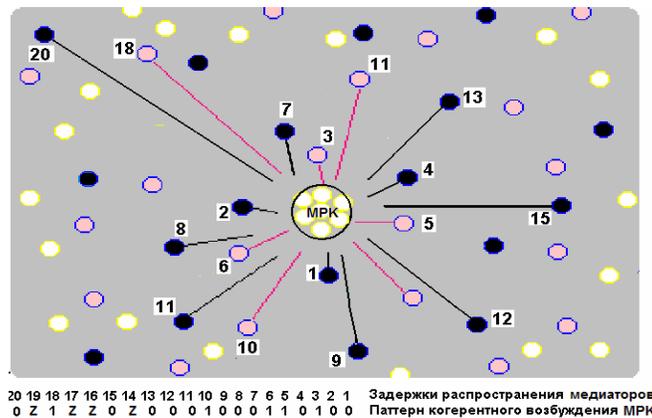


Рис. 16. МРК в ловушке возбуждающих (серые кружки) и тормозных (зачернены) синапсов Их номера отображают задержки прихода сигналов к МРК. Свободные рецепторы (белые кружки) внутри синаптической ловушки отсутствуют. Внизу выписан паттерн, обеспечивающий когерентный приход возбуждений к МРК. Z -позиции паттерна могут быть любыми — 0 или 1.

В общем случае конфигурация паттерна, адресующего данный МРК, зависит от активности многих окрестных волокон, включая синапсы смежных нейронов и ближние несинаптические терминалы. В результате адресный вектор заменяется матрицей с числом строк, равным числу конвергирующих волокон. В каждой строке адресной матрицы будут значимые позиции с элементами 0 или 1. Их число равно количеству тормозных и возбуждающих синапсов (и несинаптических источников), активируемых данным волокном прямо или через вставочные нейроны. Позиции значимых 0 и 1 в каждой строке матрицы зависят от суммарных афферентных задержек, включая время распространения медиаторов от возбуждающих и тормозных источников в нейроно-глиальных щелях. Для каждого МРК существует набор адресующих его многомерных паттернов нервных импульсов – адресный домен. Условие когерентности и задающие его уникальные пространственно-временные особенности окрестных синапсов определяют систему адресации нейронной памяти.

Анализируя адресацию каждого МРК нейрона и нейронного ансамбля, получим их преобразующие функции, как дизъюнктивную булеву форму. Ее

единицы соответствуют точкам пространства, покрытым адресными доменами, а нули – свободным точкам. Скользящая последовательность афферентных паттернов порождает траекторию адресных точек в булевом (адресном) пространстве. Попадание этих точек в покрытые и непокрытые доменами области этой функции заставляет нейронный ансамбль воспроизводить энграмму, генерируя "1" и "0" соответственно.

Заключение

Движение медиаторов в нейроно-глиальных щелях и заякоренность каналов и рецепторов на постсинаптическое уплотнение определяет суб- и экстра-синаптические локусы кластеризации нейрорецепторов и их функции — ионотропные или метаботропные. Рецепторы и каналы, удерживаемые постсинаптическим уплотнением, оказываются ионотропными. Рецепторы, вытесненные из синаптической щели, кластеризуются у квазистационарных точек синаптических ловушек, образуя на нейроне мозаику из 10^4 – 10^5 метаботропных рецептивных кластеров (МРК). Каждый МРК подобен молекулярному триггеру, один из его конформационных переходов вызывает эндогенное возбуждение нейрона и «спонтанный» спайк.

Мозаика МРК характеризует запас метаботропной активности нейрона и потенциальную емкость его памяти. Тормозно-возбудительная конфигурация синапсов в окрестности каждого МРК вместе с пре- и постсинаптическими задержками определяет конформирующий его импульсный паттерн — адрес этого МРК.

Конформации МРК зависят от концентрации медиатора, мембранного потенциала и конформационной подвижности рецепторов в кластере. Его вольт-конформационная характеристика (ВКХ) может быть гистерезисной (моностабильная или бистабильная конформации) или монотонной. В первом случае обе конформации дискретны и обратимы; третья конформация монотонна и необратима, но ее можно разрушить гиперполяризацией («кондиционирование»), восстановив гистерезис.

Медиаторы, интерферирующие за пределами синапсов, управляют локальными конформациями МРК, а поляризация мембраны определяет, будут ли эти трансмембранные триггеры бистабильными или моностабильными. Химическая селекция МРК *в бистабильном режиме* в сочетании с ПСП определяет формирование элемента энграммы. Сочетание адресации с ВПСП переводит МРК в третье (негистерезисное) конформационное состояние. При отсутствии ВПСП или действии ТПСП МРК возвращается в исходное гистерезисное состояние. В первом случае формируется «1» энграммы, во втором – «0».

Дополнительным условием формирования «1» является длительность удержания кластера в «разобщенной» (деполяризационной) конформации. Задержка должна быть достаточной, чтобы адгезия и полимеризация рецепторов [35, 36] обездвигили их димерные связи (иммобилизация воротных зарядов, возможно, с участием клатрина [37]). Задержкой управляют вариации ПСП (ВПСП – удлиняет, ТПСП – укорачивает).

В моностабильном режиме такие задержки отсутствуют, что позволяет читать энграммы без их разрушения. Этот режим минимизирует время разоб-

щенного состояния кластера, предотвращая полимеризационные изменения. Однако в этом режиме «1» и «0» энграмм при считывании будут представлены инверсно как «0» (гистерезис отсутствует и спайк не возникает) и «1» (возникающий ХКП инициирует эндогенный спайк). Для восстановления записанной последовательности в первоначальном виде следует прибегнуть к повторной инверсии, например, с помощью тормозных синапсов, как это сделано в клетках Пуркиньи и Гольджи мозжечка.

Подводя итог, отметим, что наиболее быстрыми в этих процессах должны быть конформационные механизмы (i). Их скорость не зависит от временной интеграционной базы адресного механизма, которая определяется скоростью распространения медиаторов и временем их удержания в нейроно-глиальных щелях.

Более длительно протекают процессы консолидации. Время удержания разобщенного конформационного состояния «b» должно быть достаточным для начала адгезионно-полимеризационных процессов (ii). Это критическая точка консолидации энграмм. Последующий период (iii), после которого процесс станет необратимым, отделяет кратковременную память от долговременной. Однако время последующего упрочнения энграммы не ограничено (iiii).

Наиболее инерционными могут быть переключения режимов записи-воспроизведения энграмм. Эти задержки не должны существенно удлинять поиск и воспроизведение новых энграмм.

Литература

1. Nabliz J. J., Jonston D. Cell. Mol. Endogenous nature of spontaneous bursting in hippocampal pyramidal neurons // Neurobiol., 1981. Vol. 1, P. 325–334.
2. Валиев К.А., Кокин А. А. Квантовые компьютеры: надежды и реальность. Москва-Ижевск. 2001. С. 352.
3. Ашмарин И. П. Загадки и откровения биохимии памяти. Издательство Ленинградского университета. 1975. 160 с.
4. Прибрам К. Языки мозга. М.: Прогресс. 1975. 467 с.
5. Mednikova Yu. S., Karnup S. V., Zhadin M. N. The two models of dendritic propagation. Proceedings of the 1-st International symposium on electrical activity of the brain // Mathematical models and analytical methods. Puschino, 1997. P. 79.
6. Шуранова Ж.П., Гвоздикова З.М. Реакция нейронов сенсомоторной области коры на ее прямое электрическое раздражение. В сб.: "Исследование организации нейронной деятельности коре больших полушарий головного мозга "Наука", М., 1971. С. 158-180.
7. Соколов Е.Н. Пейсмекерный потенциал в нейронной организации В сб. Системный анализ интегративной деятельности нейрона. М: Наука, 1974, С. 41-52
8. Johansson S., Arhem P. Single channel currents trigger axon potentials in small cultured hippocampal neurons (1994). Proceedings National Acad. of Sci. USA. V. 91, No 5, pp. 1761-1765.
9. Экклз Дж. Тормозные пути центральной нервной системы. М.: Мир, 1971. 168 с.
10. Ливанов М.Н. Пространственная организация процессов головного мозга. Наука. М: 1972, 132 с
11. Kullmann D. M., Ruiz A., Rusakov D. M., Scott R., Semyanov A., Walker M. C. Presynaptic, extrasynaptic and axonal GABAA receptors in the CNS: where and why? // Progress in Biophysics and Molecular Biology, 2005, pp. 8733–8746.

12. Семьянов А. Диффузная внесинаптическая нейротрансмиссия посредством глутамата и ГАМК <<http://www.biomednet.ru/content/view/57/57>>
13. Maggio R., Barbier P., Fornai F., Corsini G. U. Functional role of the third cytoplasmic loop in muscarinic receptor dimerization // *J. Biol. Chem.*, 1996. Vol. 271, pp. 31055–31060.
14. Soyer O., Dimmic M. W., Neubig R. R., Goldstein R. A. Using evolutionary methods to study G-protein coupled receptors // *University of Michigan. Ann. Arbor.*, 2005: MI 48109–1055.
15. Хаггис Дж., Михи Д., Мюир А., Робертс К., Уокер П. Введение в молекулярную биологию. Мир, М. 1967.
16. Бергельсон Л.Д. Биологические мембраны. М.: Наука, 1975.
17. Lim R., Alvarez F. J., Walmsley B. Quantal size is correlated with receptor cluster area at glycinergic synapses in the rat brainstem // *Journal of Physiology*. 1999. Vol. 516.2, pp. 505–512.
18. Zoli M., Agnati L. F., Hedlund P., Li X. M., Ferre S. and Fuxe K. Receptor-receptor interactions as an integrative mechanism in nerve cells. // *Mol. Neurobiol.* 1993. V. 7. P. 293–334.
19. Bezanilla F., Perozo E., Papazian D. M. Stefani E. Molecular basis of gating charge immobilization in Shaker potassium channels // *Science*. 1991. Vol. 254, no 5032. pp. 679–683.
20. Lujan R., Roberts J. D., Shigemoto R., Ohishi H., Somogyi P. Differential plasma membrane distribution of metabotropic receptors mGluR1 alpha, mGluR2 and mGluR5, relative to neurotransmitter release sites // *J. Chem. Neuroanat.* 1997. Vol. 13, pp. 219–241.
21. Blanpied T. A., Scott D. B., Ehlers M. D. Dynamics and regulation of clathrin coats at specialized endocytic zones of dendrites and spines // *Neuron*. 2002. Vol. 36, pp. 435–449.
22. Kullmann D. M., Ruiz A., Rusakov D. M., Scott R., Semyanov A., Walker M. C. Presynaptic, extrasynaptic and axonal GABA_A receptors in the CNS: where and why? // *Progress in Biophysics and Molecular Biology*, 2005, pp. 8733–8746.
23. Agnati L. F., Zoli M., Stromberg I., Fuxe K. Intercellular communication in the brain: Wiring versus volume transmission // *Neurosci.* 1995. V. 69, pp. 711–726.
24. Sykova E., Mazel T., Vargova L., Vorisek I., Prokopova-Kubinova S. Extracellular space diffusion and pathological states // *Progress in Brain Research*. 2000. Vol. 125, pp. 155–178.
25. Roche K. W., Standley S., McCallum J., Ly C. D., Ehlers M. D., Wenthold R. J. Molecular determinants of NMDA receptor internalization [Электронный документ] // *Nature Neuroscience*. 2001, pp. 794–802, <doi:10.1038/90498>.
26. Klauck T. M., Scott J. D. The postsynaptic density: a subcellular anchor for signal transduction enzymes // *Cell Signal* 1995. Vol. 7, pp. 747–757.
27. Kennedy, M.B. The postsynaptic density. *Curr. Opin. Neurobiol.* 1993. Vol. 3, pp. 732–737.
28. Ehlers M. D., Mammen A. L., Lau L. F., Huganir R. L. Synaptic targeting of glutamate receptors. *Curr. Opin. Cell Biol.* 1996. Vol. 8, pp. 484–489.
29. Радченко А.Н. Двухканальное управление нейронной памятью // Труды международной конференции «Устойчивость и процессы управления (SCP'2005)». Том 2. С. 1173–1182.
30. Hanson J. E., Smith Y. Cell-adhesion molecules in memory formation // *Journal of Neuroscience*. 1999. Vol. 19(15), pp. 6488–6496.
31. Kruk P. J., Korn H., Faber D. S. The effects of geometrical parameters on synaptic transmission: a Monte Carlo simulation study // *Biophysical Journal*. 1997. Vol. 73, pp. 2874–2890.

32. Nieuwenhuis R. Comparative aspects of volume transmission, with sidelight on other forms of intercellular communication // *Progress in Brain Research*. 2000. Vol. 125, pp. 49–126.
33. Радченко А. Н. Оптимизация распределенной ассоциативной памяти со случайной организацией // *Биофизика*. 1993. Т. 38, вып. 2. С. 294–299.
34. Радченко А. Н. Межуровневые отношения в нейронной памяти: внесинаптическая рецепция медиаторов, потенциация, спонтанная активность // *Успехи физиологических наук*. 2002. Том 33, № 1. С. 58–76.
35. Радченко А. Н. Ассоциативная память. Нейронные сети. Оптимизация нейропроцессоров. СПб.: Наука, 1998, 261 с.
36. Радченко А. Н. Информационные механизмы нейронной памяти и модели амнезий. СПб.: Анатолия, 2002. 297 с.
37. Радченко А. Н. Влияние подвижности воротных зарядов на вольтаперные характеристики возбудимых мембран // *Биофизика*. 1996. Т. 41, № 3. С. 625–635.
38. Kloboucek A., Behrisch A., Faix J. and Sackmann E. Adhesion-Induced Receptor Segregation and Adhesion Plaque Formation: A Model Membrane Study // *Biophys J*. 1999. Vol. 77, pp. 2311–2328.
39. Schmidt R. Cell-adhesion molecules in memory formation // *Behavior Brain Research*. 1995. Vol. 66, pp. 65–72.
40. Mueller V. J., Wienisch M., Nehring R. B., Klingauf J. Monitoring Clathrin-Mediated Endocytosis during Synaptic Activity // *The Journal of Neuroscience*. 2004. V. 24(8), pp. 2004–2012.
41. Chachisvilis M., Zhang Y-L., Frangos J. A. G protein-coupled receptors sense fluid shear stress in endothelial cells. *Proc. Natl Acad. Sci. USA*. 2006. V. 103, pp. 15463–15468.
42. Avissar S., Amitai G. and Sokolovsky M. // Oligomeric structure of muscarinic receptors is shown by photoaffinity labeling: Subunit assembly may explain high- and low-affinity agonist states," *Proc. Natl. Acad. Sci. USA*. 1983. V. 80, pp. 156–159
43. Sukharev S., Durell S. R. and Guy H. R. Structural Models of the MscL Gating Mechanism. *Biophysical J*. 2001. V. 81, pp. 917–936
44. Поликар А. Молекулярная цитология мембранных систем животной клетки. М.: Мир, 1982. 158 с.
45. Coster H. G. L. Electromechanical stresses and the effect of pH on membrane structure // *BBA*. 1975. Vol. 382, pp. 142–146.
46. Вейн А.М., Каменецкая Б. И. Память человека. Наука. М: 1973. 209 с.
47. Радченко А.Н. Настройка соматодендритных мембран на афферентные паттерны. Проблемы нейрокибернетики Ростов-на-Дону: 1999. — С. 149–152.
48. Radchenko A.N. Transmembrane receptive dimers as molecular triggers having chemical and electrical inputs. *Nuclear Instruments & Methods in Physics Research. Section A*. V. 502 (2003), pp. 745–748.

Статья поступила 15 мая 2006 г.
После доработки 23 сентября 2006 г.