

Витяев Е.Е.¹,
Неупокоев Н.В.

Математическая модель восприятия и образа

Существующие формализации образов и зрительного восприятия никак не связаны с психологией восприятия и с когнитивными процессами порождения образов. Тем самым нет связи между воспринимаемыми объектами и теми образами, которые они порождают. Для моделирования процессов порождения образов, их изменений и развития в процессе восприятия внешнего мира нужна адекватная формализация образа и восприятия, которая бы основывалась на психологии восприятия. В данной работе предлагается такая формализация, где образ и восприятие рассматриваются, в соответствии с существующими представлениями, как непрерывный процесс предвосхищения (предсказания) образом поступающих стимулов и проверка предсказаний на соответствие реальным стимулам. Показывается, что формализацией этого процесса являются неподвижные точки предсказаний. В работе приведена математическая модель неподвижных точек, алгоритм их обнаружения и эксперименты, демонстрирующие возможности данной модели. Проведено сравнение с сетями Хопфилда, которые являются наиболее похожими на неподвижные точки нейронными сетями.

Ключевые слова: Образ, восприятие, предсказание, интеллектуальный анализ данных, Data Mining.

Введение

Теория Функциональных Систем П.К.Анохина начинается с принципа опережающего отражения действительности. Мозг непрерывно во времени предвосхищает события окружающей среды и одновременно контролирует акцептором результатов действия правильность сделанных предсказаний.

В восприятии предвосхищение (антиципация) непрерывно во времени сравнивает «образ» («образ мира») с наличной стимуляцией и является процессом активного движения от «образа» к внешнему миру – непрерывным во времени процессом проверки предсказаний «образа» на соответствие стимулам внешнего мира. Только если все многочисленные предсказания будут совпадать с реальными стимулами непрерывно во времени, только тогда есть восприятие [1]. «Все это позволяет нарисовать следующую картину хода познавательной деятельности на уровне восприятия. Индивид всегда имеет некоторый образ или модель окружения, которая непрерывна во времени и пространстве и носит прогностический характер, т.е. в ней экстраполируются и воспроизводятся на языке чувственных модальностей ожидаемые результаты воздействия источника стимула на наши органы чувств» [1].

У. Найсером восприятие описано схемами и перцептивными циклами. «Если образы суть предвосхищения, они должны облегчать последующее восприятие. Перцептивная готовность это не скромный побочный продукт визуализации, это её суть. Иметь перцептивную установку в отношении чего-либо это значит иметь образ». «По моему мнению, важнейшими для зрения когнитивными структурами являются предвосхищающие схемы, подготавливающие индивида к принятию информации строго определенного, а не любого вида и, таким образом, управляющие зрительной активностью. ... В каждый момент воспринимающим конструируются предвосхищения некоторой информации, делающие возможным для него принятие её. ... Эта исследовательская активность направляется все теми же предвосхищающими схемами, представляющими собой своего рода планы перцептивных действий. ... Термин «восприятие» относится ко всему циклу, а не к какой-то отдельной его части» [2].

Насколько известно авторам, в настоящее время нет формализации восприятия, как непрерывного во времени процесса предвосхищения стимулов воспринимаемого объекта и проверки этих предвосхищений на соответствие стимулам внешнего мира. В данной работе мы предлагаем такую формализацию в виде неподвижных точек предвосхищений.

¹ Эта работа поддержана Российским Гуманитарным научным фондом, проект № 12-01-12026, Российским Фондом Фундаментальных Исследований № 11-07-00560-а, интеграционными проектами СО РАН № 3, 87, 136 и программой президента Российской Федерации поддержки научных школ НШ-276.2012.1.

Наиболее похожими подходами, в которых также есть неподвижная точка или «резонанс», являются работы Гроссберга и Хопфилда. Сети Хопфилда и экспериментальное сравнение с ними приведено далее в разделе 6.

В работах Гроссберга впервые говорится о «резонансе» между имеющейся моделью и поступающими стимулами – идентификация и распознавание объекта получаются в результате взаимодействия («резонанса») ожиданий «сверху-вниз» и сенсорной информацией «снизу-вверх» [3]. Принципиальное отличие нашего подхода от подхода Гроссберга состоит в том, что в нашей модели «резонанс» в виде неподвижной точки рассматривается внутри модели, а не вне её.

2. Формальная модель образа как неподвижной точки предвосхищений

Введем некоторые определения. Далее под предвосхищением будем понимать предсказание, а под схемой – совокупность закономерностей, предсказывающих, что будет воспринято в следующий момент времени при выполнении определенных перцептивных действий. Перефразируем высказывание Найсера в терминах закономерностей: в каждый момент времени воспринимающим извлекается из памяти весь опыт по восприятию данного объекта в виде совокупности закономерностей $Mem = \{P_1 \& \dots \& P_k \& A \Rightarrow P_0\}$. Эти закономерности означают, что, если мы воспринимаем признаки $P_1 \& \dots \& P_k$, то после осуществления перцептивного действия A , переводящего взгляд на признак P_0 , мы воспримем значение признака P_0 .

Предположим, что объекты восприятия определяются набором значений признаков (стимулов) x_1, \dots, x_n . Каждый признак x_i принимает некоторое множество значений $I_i = \{x_1^i, \dots, x_{k_i}^i\}$, $i = 1, \dots, n$. Будем предполагать, что признаки на объектах могут принимать по несколько значений. Тогда воспринимаемый объект, a описывается совокупностью подмножеств значений воспринимаемых признаков $X(a) = \{X_{j_1}(a), \dots, X_{j_m}(a)\}$, $X_{j_1}(a) \subset I_{j_1}, \dots, X_{j_m}(a) \subset I_{j_m}$, $X_{j_s}(a) \neq \emptyset$, $s = 1, \dots, m$. Восприниматься могут не все n признаков, а только m из них. Для каждого стимула (некоторого значения признака) определим предикат $P_j^i(a) \Leftrightarrow (x_j^i \in X_i(a))$. Предикат может быть с отрицанием $\bar{P}_j^i(a)$ или без него $P_j^i(a)$. Отрицание предиката в посылке правила означает, что нет данного стимула в поступающей информации. Предсказание отрицания предиката означает торможение соответствующего стимула. Предикат, который может быть, как с отрицанием, так и без него, обозначим через $\hat{P}_j^i(a)$.

Закономерности определим как высказывания вида $\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$, в котором действие опущено, поскольку по значениям i_0, j_0 предиката $\hat{P}_{j_0}^{i_0}$ всегда ясно, на что надо обратить внимание и куда перевести взгляд (на признак i_0 и его значение j_0). Смысл действия в том, чтобы изменить воспринимаемую стимуляцию. В закономерности фактически записано, какое действие должно быть осуществлено – надо от восприятия признаков $\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k}$ перейти к восприятию признака $\hat{P}_{j_0}^{i_0}$. Закономерность представляет собой формализацию действия и одновременно предвосхищение результата действия. Все закономерности из Mem обнаруживаются в процессе самообучения семантическим вероятностным выводом, рассмотренным в разделе 4.

Будем говорить, что закономерность $\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$ *извлекается из памяти* при восприятии объекта a , если посылка $\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k}$ закономерности и, значит, все предикаты, входящие в нее, становятся истинными на объекте a , т.е. когда $x_{j_s}^{i_s} \in X_{i_s}(a)$, если предикат $P_{j_s}^{i_s}$ не имеет отрицания и $x_{j_s}^{i_s} \notin X_{i_s}(a)$, если предикат $\bar{P}_{j_s}^{i_s}$ имеет отрицание, $s = 1, \dots, k$. Обозначим через $LP(X(a)) \subseteq Mem$ множество закономерностей, извлекаемых из памяти при восприятии объекта a . Если закономерность $\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$ применима к воспринимаемому объекту a и её заключение $\hat{P}_{j_0}^{i_0}$ истинно в указанном смысле ($x_{j_0}^{i_0} \in X_{i_0}(a)$, если предикат $P_{j_0}^{i_0}$ не имеет отрицания и $x_{j_0}^{i_0} \notin X_{i_0}(a)$, если предикат $\bar{P}_{j_0}^{i_0}$ имеет отрицание), то будем говорить, что предвосхищение, осуществляемое закономерностью, *подтвердилось* на объекте a , в противном случае *опроверглось*.

Восприятие объекта a – это непрерывный цикл предсказаний одних свойств объекта по другим свойствам посредством всех извлечённых из памяти закономерностей и проверка того, что все эти

закономерности подтвердились. В этом случае перцептивный цикл завершен, и все предсказания стимулов $\hat{P}_{j_0}^{i_0}$ совпали с наличной стимуляцией. Если предсказывается отрицание стимула, то его не должно быть в образе – это процесс вытормаживания стимулов. Возможные противоречия в предсказаниях закономерностей решаются на основании специального критерия Krit максимальной согласованности предсказаний, приведённого в разделе 4.

Восприятие, как непрерывный цикл перцептивных действий и предсказаний вместе с проверкой предсказаний на совпадение с реальными стимулами, формально может быть описан неподвижной точкой предсказаний (своеобразным «резонансом» закономерных связей).

Определим оператор предсказания Pr, применённый к некоторому множеству стимулов $X = \{X_{i_1} \cup \dots \cup X_{i_m}\}$ (это могут быть признаки $X(a)$ объекта a , либо целая картина или сцена, описываемая множеством признаков X). Используя все «извлечённые из памяти» закономерности $(\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow P_{j_0}^{i_0}) \in LP(X)$ этот оператор предсказывает наличие стимулов $x_{j_0}^{i_0}$ в восприятии, если предикат $P_{j_0}^{i_0}$ не имеет отрицания и предсказывает отсутствие стимула $\bar{x}_{j_0}^{i_0}$, если предикат $\bar{P}_{j_0}^{i_0}$ имеет отрицание. Тогда оператор предсказания Pr может быть записан следующим образом:

$$Pr(X) = \Phi_{Krit}(X \cup \{x_{j_0}^{i_0} \mid (\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow P_{j_0}^{i_0}) \in LP(X)\} \cup \{\bar{x}_{j_0}^{i_0} \mid (\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow \bar{P}_{j_0}^{i_0}) \in LP(X)\}),$$

где функция Φ_{Krit} минимизирует возможные противоречия в предсказаниях, используя специальный критерий согласованности Krit закономерностей по предсказанию. Подробное определение Krit и Φ_{Krit} приведено в разделе 4.

В восприятии осуществляется не один цикл предвосхищений, как показывают эксперименты с тахистоскопом. Циклы предвосхищений должны пройти несколько раз, чтобы исчезли противоречия между предвосхищениями и реальными стимулами. Когда это достигнуто, то восприятие объекта a завершено и мы имеем неподвижную точку оператора Pr. Если восприятие началось с восприятия стимулов $X(a)$ объекта a , то после нескольких итераций предвосхищения оператором Pr получим неподвижную точку – образ объекта a . Обозначим n -кратное применение оператора Pr через Pr^n . Тогда восприятие стимулов объекта a в виде неподвижной точки оператора Pr будет определяться равенством $Pr^{n+1}(X(a)) = Pr^n(X(a))$, где n – этап стабилизации предвосхищений. Более детальное описание этого процесса дано в разделе 4.

Приведём машинные эксперименты, иллюстрирующие введенные понятия на примере закодированных цифр.

3. Эксперименты по автоматическому обнаружению неподвижных точек цифр.

Проиллюстрируем нахождение неподвижных точек восприятия с помощью цифр. Закодируем цифры горизонтальными, вертикальными и диагональными линиями, расположенными в квадратах от 1 до 24 как указано в табл. 1. Вместе с белым пространством они образуют 7 признаков клетки. На рис. 1 видно, что есть некоторые разрывы в цифрах, но на результат восприятия это никак не повлияет. Цифры с добавленной сеткой демонстрируют, что некоторые цифры кодируются однозначно, а другие, как например 4 или 1, имеют избыточную кодировку (вертикальная линия состоит из двух кодов).

Таблица 1. Кодировка полей цифр.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16
17	18	19	20
21	22	23	24

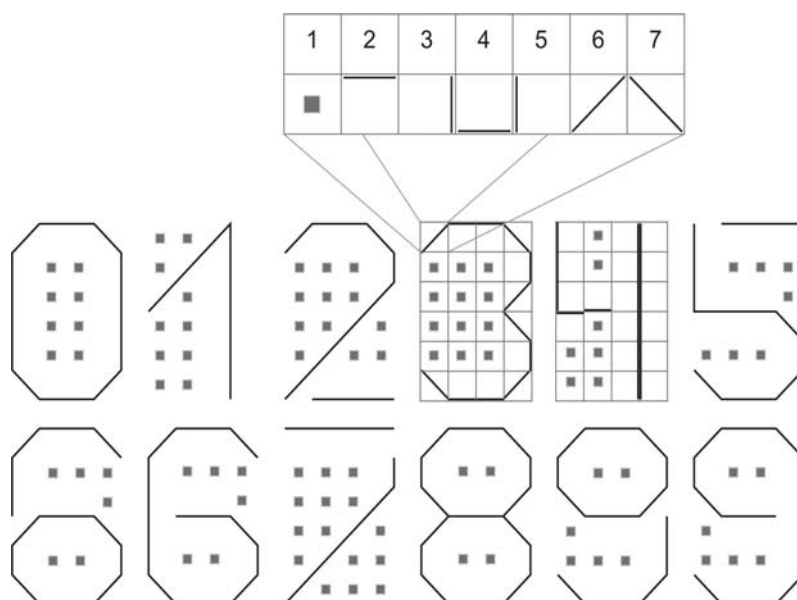


Рис. 2. Кодировка цифр.

Нами было проведено два эксперимента.

3.1. **Первый эксперимент** был проведен на обучающем множестве **A**, состоящем из 360 цифр (12 цифр рис. 1 дублировались в 30-ти экземплярах), которое представлялось как единое множество объектов без указания, где какая цифра. Эксперимент состоял в том, чтобы найти все неподвижные точки на этом множестве и убедиться, что они порождают образы, соответствующие нашим 12 цифрам.

На обучающем множестве **A** было обнаружено 55089 закономерностей. Алгоритм обнаружения закономерностей приведен в разделе 4. По этим закономерностям оператором предсказания P_r было построено 12 неподвижных точек, которые в точности соответствовали нашим цифрам. Алгоритм обнаружения неподвижных точек также приведен в разделе 4. Таким образом, на данном множестве образов алгоритм обнаружения неподвижных точек в точности решает задачу обнаружения заданных образов.

Рассмотрим на примере цифры 6 (рис. 2), что представляет собой неподвижная точка. Первая закономерность цифры 6, представленная в первом прямоугольнике после фигурной скобки, говорит, что, если в квадрате 13 (табл. 1) стоит признак 6 (обозначим это как 13-6), то в квадрате 3 должен стоять признак 2 (обозначим также как (3-2)). Предсказываемый признак обозначается точечной линией. Запишем эту закономерность как $(13-6 \Rightarrow 3-2)$. Нетрудно проверить по цифрам, что эта закономерность действительно имеет место на множестве **A**. Вторая закономерность

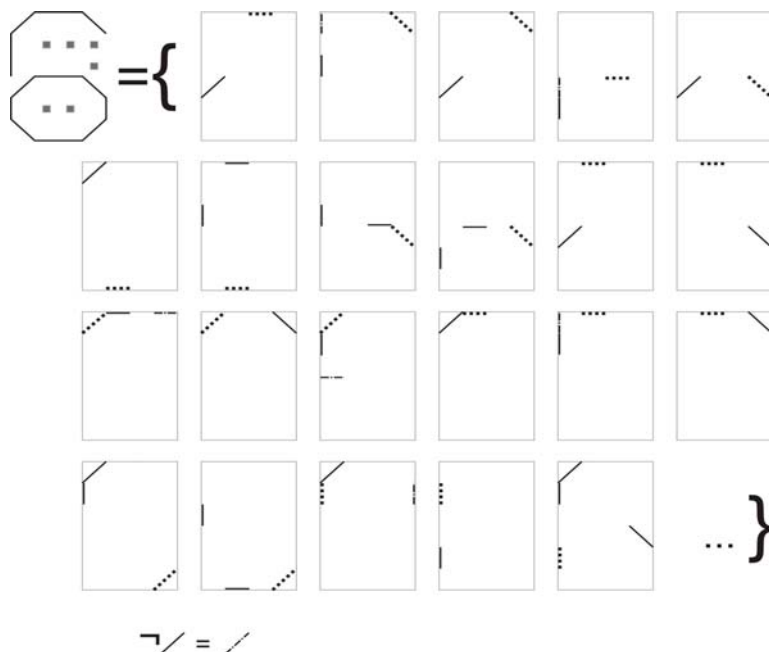


Рис. 1. Неподвижная точка цифры 6.

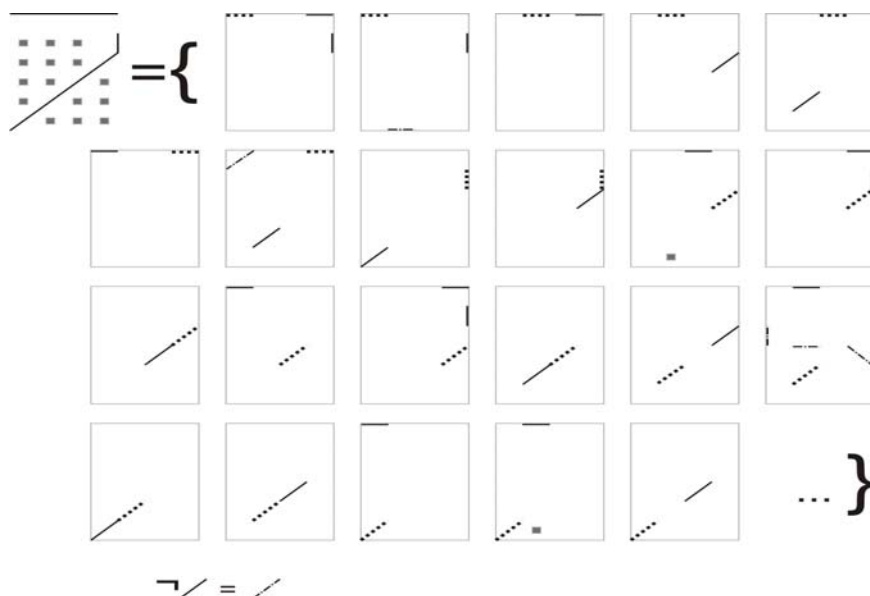


Рис. 4. Неподвижная точка цифры 7.

говорит, что из признака (9-5) и отрицания значения 5 первого признака $\neg(1-5)$ (первый признак не должен быть равен 5) следует признак (4-7). Отрицание $\neg|$ значения $|$ с номером 5 некоторого признака или отрицание $\neg/$ значения $/$ с номером 6 некоторого другого признака обозначается пунктирной линией (точка-пунктир), как показано в нижней части рисунков 2 и 3. В результате получаем закономерность $(9-5 \& \neg(1-5) \Rightarrow 4-7)$. Последующие 3 закономерности в первой строке цифры 6 на рис. 2 будут соответственно закономерности $(13-6 \Rightarrow 4-7)$, $(17-5 \& \neg(13-5) \Rightarrow 4-7)$, $(13-6 \Rightarrow 16-7)$.

На рис. 2 видно, что признаки цифры 6 взаимно предсказывают друг друга, образуя неподвижную точку. На рис. 3 приведён пример неподвижной точки цифры 7.

3.2. Второй эксперимент состоял в том, чтобы обнаружить эти же образы в условиях недостатка информации. В качестве данных было взято множество **В**, включающее те же 12 цифр в 30-ти экземплярах каждая, но с пробелами в информации – в каждой цифре отсутствовал один случайно выбранный признак, как показано на рис. 4. К полученным 360 цифрам была добавлена отрицательная выборка, состоящая из 1050 объектов со случайными признаками. На этих данных было обнаружено 73458 закономерностей. По этим закономерностям оператором предсказания P_t также были обнаружены неподвижные точки, которых оказалось 14. Из них 12 цифр это цифры рис. 1 и для цифр 6 и 9 ещё по одной неподвижной точке, содержащих пробелы в 13-м признаке для цифры 6 и в 12-м признаке для цифры 9.

На рис. 5 показана неподвижная точка цифры 6, полученная по этим данным. На рис. 5 видно,

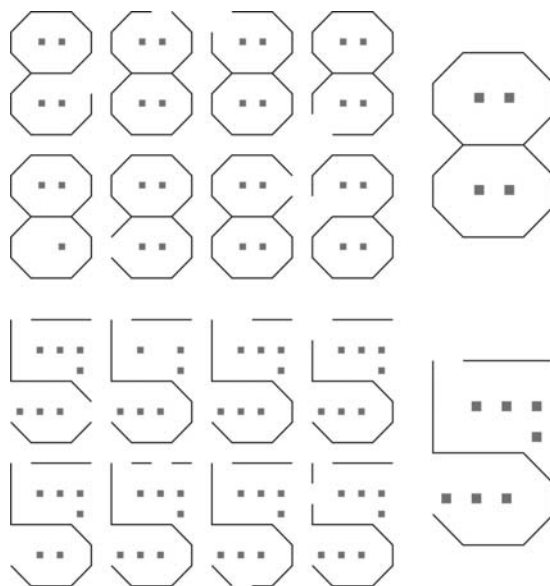


Рис. 3. Цифры с отсутствующими признаками.

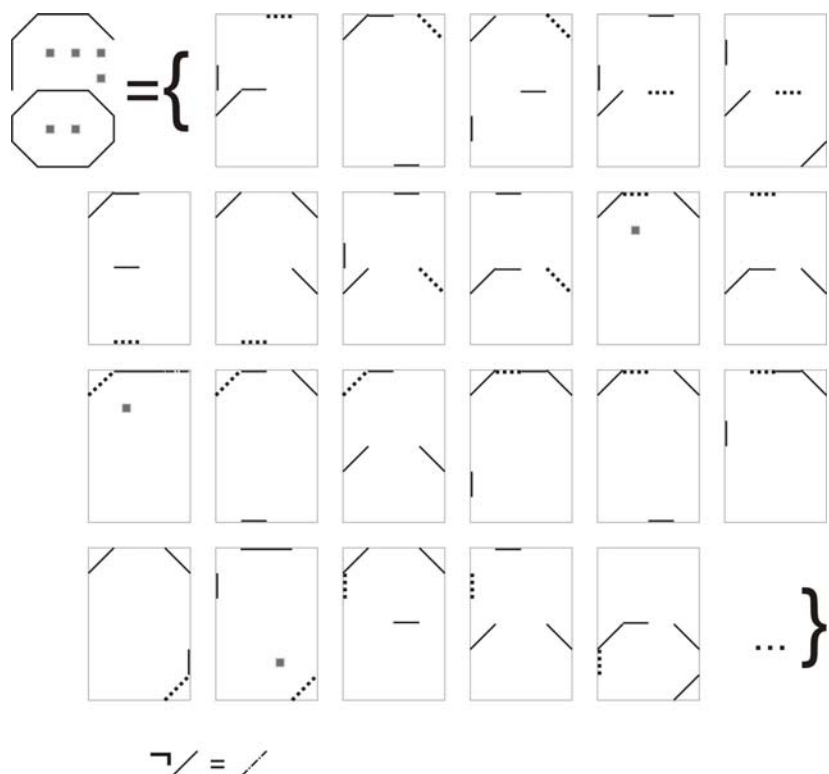


Рис. 5. Неподвижная точка цифры 6 по данным с неполной информацией.

что закономерности здесь более сложные, чем на рис. 2 так как данные более сложные и для более надёжного взаимного предсказания признаков цифр требуются закономерности в более точном условии, содержащем больше предикатов в посылке правила.

Отнесение некоторой цифры с недостающим признаком к некоторому образу (распознавание цифры) осуществляется тем же оператором предсказания P_r , путем построения неподвижной точки, отправляясь от признаков X этой цифры. Например, для цифры 5 первыми двумя закономерностями верхней строки рис. 6 предсказывается недостающий признак – хвостик в верхнем правом углу, которого нет в исходной цифре.

На второй и третьей строке рис. 6 представлены закономерности, предсказывающие признак в квадрате 1. Во второй строке представлены закономерности, которые предсказывают значение признака 5 (вертикальная линия), а в третьей строке приведена единственная обнаруженная закономерность, предсказывающая значение признака 6 (косая линия) для первого квадрата. В процессе распознавания и восстановлении недостающей информации в цифре, могут возникать противоречия, как в данном случае, предсказывающие различные значения признака в квадрате с

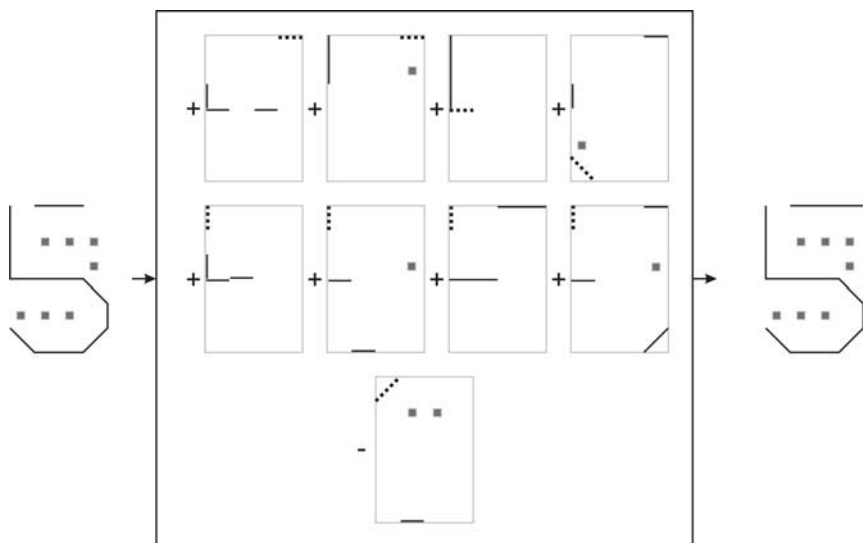


Рис. 6. Восстановление образа посредством неподвижной точки восприятия.

недостающей информацией. Эти противоречия разрешаются критерием максимальной согласованности предсказаний и оператором $\Phi_{\text{крит}}$ модификации стимулов (см. определения в разделе 4). Критерий согласованности закономерностей в этом случае отдает предпочтение более сильному предсказанию (по четырем закономерностям второй строки) вертикальной линии, игнорируя (вытормаживая) предсказание косой линии. При этом критерий $\Phi_{\text{крит}}$ учитывает и другие закономерности, которые взаимно предсказывают признаки цифры 5, как например, третья и четвертая закономерности первой строки рис. 6. Тогда цифра 5 восстанавливается полностью и правильно. В нашем эксперименте все 360 цифр с пробелами в информации были распознаны правильно.

4. Математическое определение неподвижной точки

Для полного определения оператора предсказания Pr , приведенного в конце 2-го раздела, необходимо определить множество закономерностей Mem , критерий максимальной согласованности предсказаний Krit и функцию $\Phi_{\text{крит}}$.

Обозначим через $\Pi = \{\hat{P}_j^i, i = 1, \dots, n; j = 1, \dots, k_j\}$ множество всех предикатов, фиксирующих поступающие стимулы (напомним, что под обозначением \hat{P}_j^i мы понимаем либо сам предикат P_j^i , либо его отрицание \bar{P}_j^i).

Все закономерности из Mem получаются *семантическим вероятностным выводом* (СВВ) [4-6]. Описательно СВВ рассмотрен в [6], где показано, что он формализует правило Хебба образования условных связей на уровне нейрона. Семантический вероятностный вывод устроен так, что он автоматически включает в закономерность все стимулы, которые могут усилить предсказание (увеличить условную вероятность) интересующего нас стимула (например, того, на который будет переведён взгляд).

Формально СВВ определяется как последовательность правил $R_1 \sqsubset R_2 \sqsubset \dots \sqsubset R_m$, которые от номера к номеру все сильнее предсказывают интересующий нас стимул, заданный предикатом $P_0 \in \Pi$. Эта должна удовлетворять следующим условиям:

1. $R_i = (P_1^i \& \dots \& P_{k_i}^i \Rightarrow P_0), P_j^i \in \Pi, i = 1, \dots, m$;
2. R_i – подправило правила R_{i+1} , т.е. $\{P_1^i, \dots, P_{k_i}^i\} \subset \{P_1^{i+1}, \dots, P_{k_{i+1}}^{i+1}\}$;
3. $\text{Prob}(R_i) < \text{Prob}(R_{i+1}), i = 1, 2, \dots, m-1$, где $\text{Prob}(R_i) = \text{Prob}(P_0 / P_1^i \& \dots \& P_{k_i}^i)$ – условная вероятность правила;
4. R_i – *вероятностные законы*, т.е. для любого подправила $R' = (P_1 \& \dots \& P_k \Rightarrow P_0)$ правила R_i , $\{P_1, \dots, P_k\} \subset \{P_1^i, \dots, P_{k_i}^i\}$ выполнено неравенство $\text{Prob}(R') < \text{Prob}(R_i)$;
5. R_m – *максимально специфический закон*, для которого цепочка правил $R_1 \sqsubset R_2 \sqsubset \dots \sqsubset R_m$ не может быть продолжена, т.е. для R_m не существует правила R_{m+1} удовлетворяющего условиям 1-4.

Предикат $P_0 \in \Pi$ может предсказываться различными семантическими вероятностными выводами, поэтому полное множество правил предсказывающих предикат $P_0 \in \Pi$ образует решетку $\text{Lat}(P_0)$ семантических вероятностных выводов. Полное множество закономерностей $\text{Mem}(P_0)$, которые участвуют в предсказании предиката $P_0 \in \Pi$ состоит из всех вероятностных законов, входящих в $\text{Lat}(P_0)$. Вся память Mem есть объединение всех закономерностей $\text{Mem}(P_0)$ для всех предсказываемых предикатов $P_0 \in \Pi$.

Алгоритм обнаружения закономерностей из Mem в точности следует данному определению и состоит из следующих шагов:

1. На вход алгоритма подается обучающее множество воспринятых объектов $A = \{a_1, \dots, a_m\}$, для которых известны все стимулы в виде значений всех предикатов из Π ;
2. для обнаружения всех вероятностных законов, входящих в $\text{Lat}(P_0)$, фиксируем некоторый предикат $P_0 \in \Pi$;
3. для предиката $P_0 \in \Pi$ осуществляем базовый перебор глубины d правил множества $\text{Base}_d(P_0) = B_1(P_0) \cup \dots \cup B_d(P_0)$, где $B_1(P_0) = \{(P \Rightarrow P_0), P \in \Pi\}, \dots, B_d(P_0) = \{(P_1 \& \dots \& P_d \Rightarrow P_0), P_1, \dots, P_d \in \Pi\}$;

4. перебираем все правила из каждого множества $B_i(P_0), i=1, \dots, d$ и удаляем те из них, которые не являются вероятностными законами. Для проверки, является ли некоторое правило вероятностным законом, применяем специальный оператор LP, проверяющий выполнение пункта 4 определения СВВ. При этом вероятностные неравенства $\text{Prob}(R') < \text{Prob}(R_i)$ из пункта 4 проверяем по имеющимся данным точным критерием независимости Фишера [7]. В результате получим множества $BP_i(P_0) = LP(B_i(P_0)), i=1, \dots, d$ правил содержащих только вероятностные законы. В результате базового перебора глубины d получим множество $\text{Mem}(P_0) = BP_1(P_0) \cup \dots \cup BP_d(P_0)$;

5. Далее осуществляем направленный перебор правил, который состоит в уточнении правил из $BP_d(P_0)$. Для этого правила наращиваются оператором уточнения T следующим образом:

$$T(BP_d(P_0)) = \{(P'_1 \& \dots \& P'_{d+1} \Rightarrow P_0) \mid (P_1 \& \dots \& P_d \Rightarrow P_0) \in BP_d(P_0), \\ \{P_1, \dots, P_d\} \subset \{P'_1, \dots, P'_{d+1}\}, P_1, \dots, P_{d+1} \in \Pi\}.$$

6. К уточнённым правилам снова применяем оператор проверки на вероятностные законы LP и получим новое множество вероятностных законов $P_{d+1}(P_0) = LP(T(BP_d(P_0)))$, выходящих за базовый перебор.

7. Дальнейшее наращивание множества вероятностных законов $\text{Mem}(P_0)$ происходит рекурсивно применением операторов T и LP, $P_{d+i}(P_0) = LP(T(P_{d+i-1}(P_0)))$, $i > 1$. Наращивание заканчивается на этапе t , когда уже не хватает данных для статистической проверки критерия Фишера. Тогда $P_{d+t}(P_0) = LP(T(BP_{d+t-1}(P_0))) = P_{d+t-1}(P_0)$;

8. Все множество вероятностных законов $\text{Mem}(P_0)$ получается объединением всех вероятностных законов, полученных как в результате базового перебора, так и направленного:

$$\text{Mem}(P_0) = BP_1(P_0) \cup \dots \cup BP_d(P_0) \cup P_{d+1}(P_0) \cup \dots \cup P_{d+t-1}(P_0);$$

9. Для определения всего множества Mem, переходим к пункту 2 и выбираем новый, не рассмотренный ранее предикат $P_0 \in \Pi$.

Определим функцию Φ_{Krit} минимизации возможных противоречий в предсказаниях и критерий согласованности предсказаний Krit.

Оператор предсказания Pr предсказывает два множества стимулов, которые должны присутствовать $x_{j_0}^{i_0}$ и которые должны отсутствовать $\bar{x}_{j_0}^{i_0}$:

- $\text{Pr}^+(X) = \{x_{j_0}^{i_0} \mid (\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow P_{j_0}^{i_0}) \in LP(X)\}$;
- $\text{Pr}^-(X) = \{\bar{x}_{j_0}^{i_0} \mid (\hat{P}_{j_1}^{i_1} \& \dots \& \hat{P}_{j_k}^{i_k} \Rightarrow \bar{P}_{j_0}^{i_0}) \in LP(X)\}$.

Тогда оператор предсказания Pr примет вид:

$$\text{Pr}(X) = \Phi_{\text{Krit}}(X \cup \text{Pr}^+(X) \cup \text{Pr}^-(X))$$

Определим функцию Φ_{Krit} на множестве стимулов $Y = X \cup \text{Pr}^+(X) \cup \text{Pr}^-(X)$. Она либо добавляет один элемент $x_{j_0}^{i_0}$ во множество X, либо удаляет один элемент $\bar{x}_{j_0}^{i_0}$ из множества X. При этом, она учитывает предсказания не отдельных правил из $LP(X)$, а их взаимную согласованность по предсказаниям (их закономерный «резонанс»).

Для этого используется специальный критерий Krit взаимной согласованности закономерностей по предсказанию, который определяется следующим образом. Пусть $S(X) \subset LP(X)$ – множество закономерностей, подтверждающихся на интересующем нас наборе стимулов X, а $F(X) \subset LP(X)$ – множество закономерностей, опровергающихся на наборе X. Тогда критерий Krit есть сумма весов подтверждающихся закономерностей минус сумма весов опровергающихся закономерностей:

$$\text{Krit}(X) = \sum_{R \in S(X)} \mu(R) - \sum_{R \in F(X)} \mu(R), \text{ где } \mu(R) = -\log(1 - \text{Prob}(R)).$$

Функция $-\log(1 - \text{Prob}(R))$ учитывает не саму вероятность, а её близость к 1. Логарифм берется потому, что рассматривается логарифм критерия и величины суммируются, а не перемножаются.

Функция Φ_{Krit} , при добавлении/удалении какого-то элемента $x_{j_0}^{i_0} / \bar{x}_{j_0}^{i_0}$ во множество X, должна строго увеличивать взаимную согласованность всех применимых к $X \cup x_{j_0}^{i_0}$ или к $X \setminus \bar{x}_{j_0}^{i_0}$ закономерностей и должно выполняться либо неравенство $\text{Krit}(X) < \text{Krit}(X \cup x_{j_0}^{i_0})$, либо неравенство

$\text{Krit}(X) < \text{Krit}(X \setminus \bar{x}_{j_0}^{i_0})$. В противном случае множество X остается без изменений. В обоих случаях нас интересует такое добавление/удаление элемента, которое максимально увеличивает критерий. Эти величины равны соответственно:

$$\delta^+(X) = \max_{x_{j_0}^{i_0} \in \text{Pr}^+(X)} \{\text{Krit}(X \cup x_{j_0}^{i_0}) - \text{Krit}(X)\}, \quad \delta^-(X) = \max_{\bar{x}_{j_0}^{i_0} \in \text{Pr}^-(X)} \{\text{Krit}(X \setminus \bar{x}_{j_0}^{i_0}) - \text{Krit}(X)\}.$$

Функция Φ_{Krit} добавляет/удаляет элемент из множества X , который максимизирует соответствующее значение. Эти элементы определяются следующим образом:

$$x_{j_0}^{i_0}(X) = \arg \max_{x_{j_0}^{i_0} \in \text{Pr}^+(X)} (\text{Krit}(X \cup x_{j_0}^{i_0})), \quad \bar{x}_{j_0}^{i_0}(X) = \arg \max_{\bar{x}_{j_0}^{i_0} \in \text{Pr}^-(X)} (\text{Krit}(X \setminus \bar{x}_{j_0}^{i_0}))$$

При каждом применении оператора предсказания Pr функция Φ_{Krit} не одновременно добавляет/удаляет элемент из множества X , а выбирает тот из них, который максимально увеличивает критерий, т.е. добавляет элемент $x_{j_0}^{i_0}(X)$, если $\delta^+(X) > \delta^-(X)$, $\delta^+(X) > 0$ и удаляет элемент $\bar{x}_{j_0}^{i_0}(X)$, если $\delta^-(X) > \delta^+(X)$, $\delta^-(X) > 0$.

Итак, функция модификации Φ_{Krit} определяется следующим образом:

$$\Phi_{\text{Krit}}(X) = \begin{cases} X \cup x_{j_0}^{i_0}(X), & \text{если } \delta^+(X) > \delta^-(X), \delta^+(X) > 0, \\ \quad x_{j_0}^{i_0}(X) = \arg \max_{x_{j_0}^{i_0} \in \text{Pr}^+(X)} (\text{Krit}(X \cup x_{j_0}^{i_0})) \\ X \setminus \bar{x}_{j_0}^{i_0}(X), & \text{если } \delta^-(X) \geq \delta^+(X), \delta^-(X) > 0, \\ \quad \bar{x}_{j_0}^{i_0}(X) = \arg \max_{\bar{x}_{j_0}^{i_0} \in \text{Pr}^-(X)} (\text{Krit}(X \setminus \bar{x}_{j_0}^{i_0})) \\ X, & \text{если } \delta^+(X) \leq 0 \text{ и } \delta^-(X) \leq 0. \end{cases}$$

Неподвижная точка $\text{Pr}^{n+1}(X) = \text{Pr}^n(X)$ получается в третьем случае, когда добавление/удаление элемента не увеличивает критерий.

Алгоритм обнаружения неподвижных точек точно следует данным определениям:

1. на вход алгоритма подается множество воспринятых в данный момент стимулов X и множество закономерностей $\text{LP}(X)$;
2. обнаруживаются множества $\text{Pr}^+(X)$ и $\text{Pr}^-(X)$;
3. применяется оператор предсказания $\text{Pr}(X) = \Phi_{\text{Krit}}(X \cup \text{Pr}^+(X) \cup \text{Pr}^-(X))$ к множеству X , в результате чего оно изменяется на один элемент;
4. применением оператора предсказания до тех пор пока не получим неподвижную точку $\text{Pr}^{n+1}(X) = \text{Pr}^n(X)$.
5. переходим к пункту 1 и выбираем новое множество X . Выбор различных совокупностей множеств X зависит от решаемой задачи. В нашем примере с цифрами мы брали последовательно множества $X(a) \in \mathbf{A}$ для экспериментов со множеством \mathbf{A} и $X(a) \in \mathbf{B}$ для экспериментов со множеством \mathbf{B} .

5. Сравнение с сетями Хопфилда

Нами выбрана нейронная рекуррентная сеть Хопфилда, в силу того, что она обладает цикличностью, т.е. неоднократной обработкой входного сигнала, в результате которой сигнал сводится к неподвижной точке или точке равновесия [8]. Сеть в режиме обучения перебирает все тренировочные объекты, осуществляя вычисление матрицы весов

$$w_{ij}^1 = \frac{1}{n} x_i^1 x_j^1 (i \neq j),$$

где n – длина входного сигнала, равная числу нейронов в сети.

Для каждого следующего объекта вычисляется матрица и складывается с полученной ранее: $W = W^1 + W^2 + \dots + W^l$, где l – число объектов в тренировочном множестве. Таким образом, матрица весов формируется по формуле:

$$w_{ij} = (1 - \delta_{ij}) \frac{1}{n} \sum_{k=1}^l x_i^k x_j^k$$

Все вычисления по распознаванию сводятся к простому умножению входного вектора на матрицу весов $h(0) = y, h(1) = \text{sign}(Wy), h(2) = \text{sign}(Wh(1))$, пока не будет найдена неподвижная точка

$Output = h(t+1) = h(t)$. Существование неподвижной точки гарантируется уменьшением энергии сети $E(t) = -\frac{1}{2} \sum w_{ij} h_i(t) h_j(t)$. Для каждой итерации d верно $E(d+1) - E(d) \leq 0$, откуда следует, что существует t такое, что $h(t+1) = h(t) = \text{sign}(Wh(t))$.

Для сетей Хопфилда возможна катастрофа памяти, которая происходит в результате переполнения сети, когда обучение сети состоит из большого числа классов, и поэтому на некотором этапе после обучения матрица связей “разрушается”, т.е. теряется суперпозиция предыдущих шагов обучения, которую хранила матрица W . В [9] указывается, что число различных объектов, которые может запомнить сеть без катастрофы памяти равно $m(n) = \frac{n}{2 \log n}$.

В работе [10] исследован способ обхода катастрофы памяти. Этим метод память продолжает функционировать и за пределами границы, но имеет свойство “забывать” паттерны с низким порогом, и для восстановления забытых паттернов их нужно предъявлять заново, чтобы установить новый порог. Такой способ поведения сети неэффективен.

Покажем, что и в нашем примере происходит катастрофа памяти. Проведем соответствующий эксперимент. Для этого проинтерпретируем 7-значный вектор признаков, кодирующий наши буквы, как биполярный (см. рис. 7). Для этого каждое значение признака представим набором из семи единиц и минус единиц, где (-1) означает отсутствие значения, (1) – наличие значения (рис. 7.а). Тогда, например, вместо признака 6 получим набор (-1 -1 -1 -1 -1 1 -1), вместо 2 получим набор (-1 1 -1 -1 -1 -1 -1). В итоге вместо, например, вектора (6 2 2 7 1 1 3 1 1 6 1 1 6 1 1 6 4 4 4), описывающую цифру в порядке номеров клеток в соответствии с таблицей 1, получим вектор длины 168 вида

(-1 -1 -1 -1 -1 1 -1 | -1 1 -1 -1 -1 -1 -1 | -1 1 -1 -1 -1 -1 -1 | -1 -1 -1 -1 -1 -1 1...)

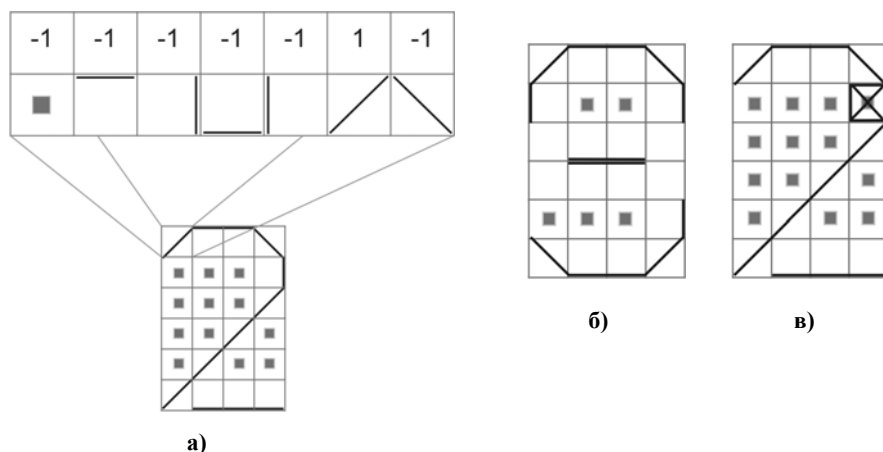


Рис. 7. а) Замена каждого символа на соответствующий биполярный блок; б) распознавание сети Хопфилда со сведением всех объектов к одному классу и в) созданием классов со специфическим признаком

Преобразуем наши обучающие данные **A** и **B** в соответствии с данной кодировкой. Мы провели два эксперимента при $n=168$ на полученных обучающих данных:

1. Классический вариант сети сводит все объекты, как обучающего множества **A**, так и обучающего множества **B** к одному классу, не совпадающему ни с одним объектом этих множеств (рис. 7.б). Пустая клетка означает, что все значения признаков равны -1;

2. Сеть из пакета MATLAB [11] справляется с обучающими данными **A** и даёт для них 12 классов, как и алгоритм неподвижных точек, но не справляется с данными **B**. Для данных **B** все пропуски информации превращаются алгоритмом в специфические признаки, когда все признаки равны 1, т.е. в клетке с пропуском одновременно присутствуют все семь признаков (рис. 7.в). Так как пропуски есть в каждой цифре, то во втором эксперименте сеть пакета MATLAB даёт столько классов, сколько цифр во множестве **B**.

Результаты сведены нами в таблицу 2.

Таблица 2. Результаты сравнения с сетью Хопфилда

Тип	Обучение	Результат
-----	----------	-----------

Классическая сеть Хопфилда	A	1 класс
	B	1 класс
newhor из MATLAB	A	12 классов
	B	Для каждого объекта порождается свой класс
Алгоритм неподвижных точек	A	12 классов
	B	14 классов

6. Выводы

Рассмотренный в работе пример с цифрами, хотя и является модельным, но тем не менее является достаточно сложным, что позволяет говорить о возможности решения нетривиальных прикладных задач алгоритмом неподвижных точек. Первоначально данный алгоритм рассматривался нами как алгоритм «естественной» классификации и имел самостоятельное обоснование (см. работы по «естественной» классификации на сайте [12]). В таком виде алгоритм успешно применялся для решения задач биоинформатики [13].

ЛИТЕРАТУРА

- Смирнов С.Д.** Психология образа. МГУ, М., 1985, с.231
- Найсер У.** Познание и реальность. “Прогресс”, М. 1981, с. 229.
- Carpenter, G.A. & Grossberg, S.**, Adaptive Resonance Theory, In Michael A. Arbib (Ed.), The Handbook of Brain Theory and Neural Networks, Second Edition, Cambridge, MA: MIT Press, 2003, pp. 87-90.
- Витяев Е.Е.** Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов // Новосибирский гос. ун-т. Новосибирск, 2006. с.293.
- Evgenii Vityaev.** The logic of prediction. In: Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference (August 16-19, 2005, Novosibirsk, Russia), edited by S.S. Goncharov, R. Downey, H. Ono, World Scientific, Singapore, 2006, pp.263-276
- Витяев Е.Е., Перловский Л.И., Ковалерчук Б.Я., Сперанский С.О.** Вероятностная динамическая логика мышления. Нейроинформатика, 2011, том 5, № 1, стр. 1-20
- Кендал М., Стьюарт А.** Статистические выводы и связи. М.: Наука, 1973. С. 899.
- John Hertz, Anders Krogh, Richard G. Palmer** Introduction To The Theory Of Neural Computation. - Addison-Wesley Pub. Co., 1991 P.327.
- McEliece, Robert J. and Posner, Edward C. and Rodemich, Eugene R. and Venkatesh, Santosh S.** The capacity of the Hopfield associative memory. IEEE Transactions on Information Theory, 33 (4). 1987 pp. 461-482.
- Карандашев Я.М., Крыжановский Б.В., Литинский Л.Б.** Взвешенные паттерны и устранение «катастрофы памяти» в модели Хопфилда. //XIV Всероссийская научно-техническая конференция “Нейроинформатика-2012” : Лекции по нейроинформатике. - М.: НИЯУ МИФИ, 2012. с.152-189.
- Li J., A.N. Michel, W. Porod** Analysis and synthesis of a class of neural networks: linear systems operating on a closed hypercube. IEEE Transactions on Circuits and Systems, Vol. 36, No. 11, November 1989, pp. 1405-1422.
- Scientific Discovery:** <http://www.math.nsc.ru/LBRT/logic/vityaev>
- Vityaev E.E., Lapardin K.A., Khomicheva I.V., Proskura A.L.** Transcription factor binding site recognition by regularity matrices based on the natural classification method. *Intelligent Data Analysis*. v.12(5), IOS Press, 2008, 495-512.