ВЕРОЯТНОСТНОЕ ОБОБЩЕНИЕ ФОРМАЛЬНЫХ ПОНЯТИЙ И МОДЕЛЬ ВОСПРИЯТИЯ*

Витяев Е.Е., *д.ф.-м.н.*
Институт математики им. С. Л. Соболева СО РАН,
Новосибирский Государственный Университет
e-mail: vityaev@math.nsc.ru

1. Введение.

Формальные понятия возникающие в анализе формальных понятий FCA [1,2] (Formal Concept Analysis) могут быть определены через неподвижных точек импликаций. Мы обобщаем формальные понятия на вероятностный случай путем введения вероятностных импликаций и определения неподвижных точек для вероятностных импликаций. Обобщение осуществлено так, что при определенных условиях, оговоренных в теореме 1, вероятностные формальные понятия и формальные понятия совпадают. Проведен машинный эксперимент, показывающий, что вероятностные формальные понятия могут «восстанавливать» формальные понятия после наложения на данные шума [3,4]. Тем самым они становятся устойчивыми по отношению к искажению данных.

Вероятностные формальные понятия дают не только новый метод Data Mining автоматического формирования понятий в условиях шумов в данных, но и дают возможность получить формализацию процесса восприятия.

Далее приведено описание психологии восприятия и формальной модели восприятия, основанной на неподвижной точке предвосхищений, которая определяется, как и вероятностные формальные понятия, через неподвижные точки вероятностных импликаций. Машинный эксперимент, приведенный [5-6], подтверждает эффективность формальной модели восприятия на примере закодированных цифр.

Первоначально алгоритм обнаружения неподвижных точек вероятностных импликаций рассматривался нами как алгоритм «естественной» классификации и имел самостоятельное обоснование (см. работы по «естественной» классификации на сайте [7]). Алгоритм успешно

---

*Работа поддержана грантом РФФИ № 11-07-00560-а, интеграционными проектами СО РАН № 3, 87, 136 и программой президента Российской Федерации поддержки научных школ НШ-276.2012.1.

применялся для решения задач биоинформатики [8].

2. Вероятностное обобщение формальных понятий [3-4]

Определение 1. Формальным контекстом назовём тройку (G, M, I), где G и M – некоторые множества, а $I \subseteq G \times M$ – некоторое отношение между элементами G и M. Элементы G называются объектами контекста а элементы M – атрибутами контекста. На подмножествах $A \subseteq G$, $B \subseteq M$ контекста (G, M, I) определим операцию $'$:

$A' = \{m \in M \mid \forall g \in A (g, m) \in I\}$, $B' = \{g \in G \mid \forall m \in B (g, m) \in I\}$.

Если $g \in G$, то обозначение g' является сокращением для $\{g\}'$.

Определение 2. Понятием в контексте (G, M, I) называется пара (A, B), где $A \subseteq G, B \subseteq M, A' = B, B' = A$. Множество A называется объемом, а B – содержанием понятия (A, B).

Определение 3. Импликацией на некотором множестве M назовем упорядоченную пару подмножеств $A, B \subseteq M$, обозначаемую как $A \to B$. Скажем, что множество $T \subseteq M$ удовлетворяет импликации $A \to B$, если $A \not\subseteq T$ или $B \subseteq T$. Семейство подмножеств M удовлетворяет импликации, если каждое из множеств M удовлетворяет ей. Импликация $A \to B$ истинна на контексте $K = (G, M, I)$ (обозначение $K \vDash A \to B$), если $A, B \subseteq M$ и семейство множеств $\{g' \mid g \in G\}$ удовлетворяет $A \to B$. Множество всех импликаций истинных на контексте K обозначим через $\mathrm{Imp}(K)$. Скажем, что посылка импликации $A \to B$ ложна на K, если не существует $g \in G$ такого, что $A \subseteq g'$. Назовем импликацию $A \to B$ тавтологией, если $B \subseteq A$.

Замечание 1. Если $K = (G, M, I)$ – контекст, $A \to B$ – импликация на M, то $K \vDash A \to B \Leftrightarrow \forall m \in B (K \vDash A \to \{m\})$. В дальнейшем мы будем рассматривать импликации только вида $A \to \{m\}$ и использовать обозначение $A \to m$ для таких импликаций.

Если K – контекст, то для любой импликации $A \to m \in \mathrm{Imp}(K)$ найдется подмножество A_0, $\{A_0 \to m \in \mathrm{Imp}(K) \mid A_0 \subseteq A$, такое что для любого $A_1 \subseteq A$ из $A_1 \subset A_0$ следует $A_1 \to m \notin \mathrm{Imp}(K)\}$. Для контекста K обозначим через $\mathrm{MinImp}(K)$ множество всех импликаций вида $A_0 \to m$, истинных на K, в которых множество A_0 минимально в указанном смысле.

Любое множество импликаций L на множестве M порождает монотонный оператор $f_L : 2^M \to 2^M$, определяемый следующим образом: $f_L(X) = X \cup \{B \mid A \to B \in L, A \subseteq X\}$. Ясно, что для любого $X \subseteq M$ верно $f_L(X) = X \;\Leftrightarrow\; X \vDash L$.

Приведем несколько измененное предложение 20 из [1], которое определяет формальные понятия через импликации.

Предложение 1. Пусть $K = (G, M, I)$ – контекст, $T \subseteq \mathrm{Imp}(K)$ – множество тавтологий на M, а $F \subseteq \mathrm{Imp}(K)$ – множество импликаций, посылки которых ложны на K. Тогда для любого множества $B \subseteq M$ выполняется следующее:

1. $f_{\mathrm{MinImp}(K) \setminus T}(B) = B \Leftrightarrow B'' = B$;
2. если $B' \neq \varnothing$, то $f_{\mathrm{MinImp}(K) \setminus \{F \cup T\}}(B) = B \Leftrightarrow B'' = B$.

Из определения 2 следует, что подмножество $B \subseteq M$ является содержанием некоторого понятия в контексте K тогда и только тогда, когда $B'' = B$. Из пункта 1 предложения следует, что неподвижные точки оператора $f_{\mathrm{MinImp}(K) \setminus T} : 2^M \to 2^M$ совпадают с содержаниями понятий контекста K. Если среди импликаций из $\mathrm{MinImp}(K) \setminus T$, исключить импликации F, посылка которых ложна на K, то неподвижные точки оператора $f_{\mathrm{MinImp}(K) \setminus \{F \cup T\}} : 2^M \to 2^M$ совпадают с содержаниями понятий контекста K за исключением единственного понятия $(\varnothing, M)$ (в силу того, что для любого $B \subseteq M$ условие $B'' \neq M$, очевидно, влечет $B' \neq \varnothing$).

Выше было определено понятие истинности импликации на некотором отдельно взятом контексте. Опишем, как обобщить данное понятие путем введения вероятности на классе контекстов и специальных вероятностных импликаций, обнаруживаемых на классе контекстов семантическим вероятностным выводом [9].

Определение 4. Классом контекстов над множествами G и M назовем семейство $K = \{(G, M, I_j)\}_{j \in J \neq \varnothing}$, где для каждого $j \in J$ тройка (G, M, I_j) является контекстом. Будем использовать обозначение $K(G, M)$ для класса контекстов K над множествами G и M. Вероятностной моделью назовем пару $M = (K(G, M), \rho)$, где $G \neq$ и ρ – вероятностная мера на множестве K, удовлетворяющая условию:

$$\forall S_1, S_2 \subseteq G \times M, \forall (G,M,I) \in K \ (S_1 \not\subseteq I \text{ или } S_2 \subseteq I) \Longleftrightarrow$$
$$\rho(\{(G,M,I_j) \mid S_1 \cup S_2 \subseteq I_j\}) = \rho(\{(G,M,I_j) \mid S_1 \subseteq I_j\}).$$

Если $S \subseteq G \times M$, то вероятностью множества S на модели M назовем значение функции $\nu_M(S) = \rho(\{(G,M,I) \in K \mid S \subseteq I\})$. Для краткости будем называть пары $(K(G,M),\rho)$ *вероятностными моделями* или просто *моделями*.

Пусть $M = (K(G,M),\rho)$ – вероятностная модель и $A \to m$ – некоторая импликация на M. Подстановочным случаем импликации $A \to m$ на M назовем пару $\langle g, A \to m\rangle$, $g \in G$. Вероятность пары $\langle g, A \to m\rangle$ на модели M определим как:

$$\mu_M(\langle g, A \to m\rangle) = \begin{cases} \dfrac{\nu_M(S \cup \{<g,m>\})}{\nu_M(S)}, & \nu_M(S) \neq 0, \\ & S = \{<g,a> \mid a \in A\} \\ \text{не определено} & \text{в противном случае} \end{cases}$$

Вероятность импликации $A \to m$ на M определим как:

$$\eta_M(A \to m) = \begin{cases} \text{не определено,} \\ \text{если} \forall g \in G \mu_M(\langle g, A \to m\rangle) \text{ не определено} \\ \inf_{g \in G} \mu_M(\langle g, A \to m\rangle) \text{ в противном случае} \end{cases}$$

Определение 5. Пусть $M = (K(G,M),\rho)$ – модель, $imp(M)$ – множество импликаций вероятность которых на M определена. Вероятностными закономерностями на M назовем импликации $A \to m \in imp(M)$, для которых, если $A_0 \to m \in imp(M)$ и $A_0 \subset A$, то $\eta_M(A_0 \to m) < \eta_M(A \to m)$, $\eta_M(A \to m) \neq 0$.

Импликацию $A \to m \in imp(M)$ назовем максимально специфичной вероятностной закономерностью на M, если она есть вероятностная закономерность на M, $A \neq \{m\}$, и не существует другой вероятностной закономерности $A_0 \to m$ на M, что $A \subset A_0$ и $A_0 \to m$ не тавтология.

Определение 6. Пусть $M = (K(G,M),\rho)$ – вероятностная модель, $S(M)$ – множество всех максимально специфичных вероятностных закономерностей на M. Импликацию $A \to m \in S(M)$ назовем сильнейшей вероятностной закономерностью на M, если значение ее вероятности на M максимально среди всех импликаций $B \to m \in S(M)$. Обозначим через $D(M)$ множество всех сильнейших вероятностных закономерностей на модели $\mathcal{M}$.

Определение 7. Пусть $\mathcal{M} = (K(G,M),\rho)$ – вероятностная модель. Вероятностным понятием контекста $(G,M,I) \in K$ в модели $\mathcal{M}$ назовем пару множеств (A,B), удовлетворяющих условиям:

1. $A \subseteq G$, $B \subseteq M$,
2. $f_{D(\mathcal{M})}(B) = B$,
3. $\exists E \subseteq B(\overline{f}_{D(\mathcal{M})}(E) = B$ и $E \neq \varnothing \neq E')$,
4. $A = \bigcup\{E' \mid \varnothing \neq E \subseteq B, \overline{f}_{D(\mathcal{M})}(E) = B\}$,

где $'$ – операция в рамках контекста (G,M,I). Множество A назовем объемом, а B – содержанием вероятностного понятия (A,B).

Теорема 1. Рассмотрим контекст $K = (\varnothing \neq G,M,I)$, и вероятностную модель $\mathcal{M} = (\{K\},\rho)$. Тогда для любых непустых подмножеств $A \subseteq G$ и $B \subseteq M$ пара (A,B) является понятием в контексте K тогда и только тогда, когда она является вероятностным понятием контекста K в модели $\mathcal{M}$.

3. Психология образа

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 [10]. «Все это позволяет нарисовать следующую картину хода познава-тельной 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» [10].

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. Мы предлагаем такую формализацию, основанную на неподвижной точке предвосхищений, которая определяется, также как и вероятностные формальные понятия, через неподвижные точки вероятностных импликаций.

Наиболее похожими подходами, в которых также есть неподвижная точка или «резонанс», являются работы Гроссберга и Хопфилда.

Гроссберг говорит 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» [11]. 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.

4. Формальная модель восприятия как неподвижной точки предвосхищений [5-6]

Введем некоторые определения. Далее под предвосхищением будем понимать предсказание, а под моделью – совокупность закономерностей, предсказывающих, что будет воспринято в следующий момент времени при выполнении определенных персептивных действий. В терминах закономерностей восприятие в каждый момент времени извлекает из памяти весь опыт по восприятию данного объекта в виде совокупности закономерностей $\mathrm{Mem} = \{P_1 \& \ldots \& P_k \& A \Rightarrow P_0\}$. Эти закономерности означают, что, если мы воспринимаем признаки $P_1 \& \ldots \& P_k$, то после осуществления персептивного действия A, переводящего взгляд на признак P_0, мы воспримем значение признака P_0.

Предположим, что объекты восприятия определяются набором значений признаков (стимулов) $x_1,\ldots,x_n$. Каждый признак x_i принимает некоторое множество значений $I_i = \{x_1^i,\ldots,x_{k_i}^i\}, i = 1,\ldots,n$. Будем предполагать, что признаки на объектах могут принимать по нескольку значений. Тогда воспринимаемый объект *a* описывается совокупностью подмножеств значений воспринимаемых признаков $X(a) = \{X_{j_1}(a),\ldots,X_{j_m}(a)\}$, $X_{j_1}(a) \subset I_{j_1},\ldots,X_{j_m}(a) \subset I_{j_m}\}$, $X_{j_s}(a) \neq \varnothing$, s = 1,...,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 $\overline{P}_j^i(a)$ или без него $P_j^i(a)$. Заметим, что при определении вероятностных формальных понятий не использовалось отрицание, поэтому приводимая далее формализация является более общей. 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 $\hat{P}_{j_1}^{i_1} \& \ldots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$, 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. Закономерность формализует одновременно действие и предвосхищение результата действия. Все закономерности из Mem обнаруживаются в процессе самообучения семантическим вероятностным выводом, описанным далее.

Будем говорить, что закономерность $\hat{P}_{j_1}^{i_1} \& \ldots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$ *извлекается из памяти* при восприятии объекта *a*, если посылка $\hat{P}_{j_1}^{i_1} \& \ldots \& \hat{P}_{j_k}^{i_k}$ 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 $\overline{P}_{j_s}^{i_s}$ имеет отрицание, s = 1,…k. Обозначим через LP(X(a)) $\subseteq$ Mem множество закономерностей, извлекаемых из памяти при восприятии объекта *a*. Если закономерность $\hat{P}_{j_1}^{i_1} \& \ldots \& \hat{P}_{j_k}^{i_k} \Rightarrow \hat{P}_{j_0}^{i_0}$ 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 $\overline{P}_{j_0}^{i_0}$ 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, определенного далее.

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

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

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

где функция Φ_{Krit} минимизирует возможные противоречия в предсказаниях, используя специальный критерий согласованности Krit закономерностей по предсказанию. Определение критерия Krit и функции Φ_{Krit} приведено далее.

В восприятии осуществляется не один цикл предвосхищений. 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 предвосхищений оператором Pr получим неподвижную точку – восприятие образа объекта *a*. В этой неподвижной точке достигается единство двух сущностей – совокупности признаков и закономерностей, которым они удовлетворяют.

Обозначим n-кратное применение оператора Pr через Pr^n. Тогда восприятие стимулов объекта *a* в виде неподвижной точки оператора Pr будет определяться равенством $\mathrm{Pr}^{n+1}(X(a)) = \mathrm{Pr}^n X(a)$, где n – этап стабилизации предвосхищений.

Определим теперь множество закономерностей Mem, критерий максимальности согласованности предсказаний Krit и функцию Φ_{Krit}. Обозначим через $\Pi = \{\hat{P}^i_j, i = 1,...n; j = 1,...k_i\}$ множество всех предикатов, фиксирующих поступающие стимулы.

Все закономерности из Mem получаются *семантическим вероятностным выводом* (СВВ) [9,12-13]. Описательно СВВ рассмотрен в [13], 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 $R_1 \sqsubset R_2 \sqsubset \ldots \sqsubset R_m$, которые от номера к номеру все сильнее предсказывают интересующий нас стимул, заданный предикатом $P_0 \in \Pi$. Эта последовательность должна удовлетворять следующим условиям:

1. $R_i = (P^i_1 \& ... \& P^i_{k_i} \Rightarrow P_0), P^i_j \in \Pi, i = 1,\ldots,m;$
2. R_i – подправило правила R_{i+1}, т.е. $\{P^i_1,\ldots,P^i_{k_i}\} \subset \{P^{i+1}_1,\ldots,P^{i+1}_{k_{i+1}}\}$;
3. $\mathrm{Prob}(R_i) < \mathrm{Prob}(R_{i+1}), i = 1,2,\ldots,n-1$, где
   $\mathrm{Prob}(R_i) = \mathrm{Prob}(P^i_0 / P^i_1 \& \ldots \& P^i_{k_i})$ – условная вероятность;
4. R_i – *вероятностные законы*, т.е. для любого подправила
   $R' = (P_1 \& \ldots \& P_k \Rightarrow P_0)$ правила R_i, $\{P_1,\ldots,P_k\} \subset \{P^i_1,\ldots,P^i_{k_i}\}$ выполнено неравенство $\mathrm{Prob}(R') < \mathrm{Prob}(R_i)$;
5. R_m – *максимально специфический закон*, для которого цепочка правил $R_1 \sqsubset R_2 \sqsubset \ldots \sqsubset R_m$ не может быть продолжена – 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 $Lat(P_0)$ семантических вероятностных выводов. Полное множество закономерностей $Mem(P_0)$, которые участвуют в предсказании предиката $P_0 \in \Pi$, состоит из всех вероятностных законов, входящих в $Lat(P_0)$. Вся память Mem есть объединение всех закономерностей $Mem(P_0)$ для всех предсказываемых предикатов $P_0 \in \Pi$.

Определим функцию Φ_{Krit} минимизации возможных противоречий в предсказаниях и критерий согласованности предсказаний Krit. Оператор предсказания Pr предсказывает два множества стимулов, в которые должны присутствовать/отсутствовать $x_{j_0}^{i_0} / \overline{x}_{j_0}^{i_0}$:

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

$$Pr^{-}(X) = \{\overline{x}_{j_0}^{i_0} \mid (\hat{P}_{j_1}^{i_1} \& \ldots \& \hat{P}_{j_k}^{i_k} \Rightarrow \overline{P}_{j_0}^{i_0}) \in LP(X)\}.$$

Тогда оператор предсказания Pr примет вид:

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

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

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:

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

Функция $-\log(1 - Prob(R))$ 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

$x_{j_0}^{i_0} / \overline{x}_{j_0}^{i_0}$ во множестве X, должна строго увеличивать взаимную согласованность всех применимых к $X \cup x_{j_0}^{i_0}$ или к $X \setminus \overline{x}_{j_0}^{i_0}$ закономерностей и должно выполняться либо неравенство $\mathrm{Krit}(X) < \mathrm{Krit}(X \cup x_{j_0}^{i_0})$, либо неравенство $\mathrm{Krit}(X) < \mathrm{Krit}(X \setminus \overline{x}_{j_0}^{i_0})$. В противном случае множество X остается без изменений. В обоих случаях нас интересует такое добавление/удаление элемента, которое максимально увеличивает критерий. Эти изменения критерия равны соответственно:

$$\delta^{+}(X) = \max_{x_{j_0}^{i_0} \in \Pr^{+}(X)} \{\mathrm{Krit}(X \cup x_{j_0}^{i_0}) - \mathrm{Krit}(X)\},$$

$$\delta^{-}(X) = \max_{\overline{x}_{j_0}^{i_0} \in \Pr^{-}(X)} \{\mathrm{Krit}(X \setminus \overline{x}_{j_0}^{i_0}) - \mathrm{Krit}(X)\}.$$

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

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

При каждом применении оператора предсказания Pr функция Φ_{Krit} не одновременно добавляет/удаляет элемент из множества X, а выбирает тот их 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 $\overline{x}_{j_0}^{i_0}(X)$, если $\delta^{-}(X) > \delta^{+}(X)$, $\delta^{-}(X) > 0$.

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

$$\Phi_{\mathrm{Krit}}(X) = \left\{ \begin{array}{r} X \cup x_{j_0}^{i_0}(X), \text{ если } \delta^{+}(X) > \delta^{-}(X), \delta^{+}(X) > 0, \\ x_{j_0}^{i_0}(X) = \arg\max\limits_{x_{j_0}^{i_0} \in \Pr^{+}(X)} (\mathrm{Krit}(X \cup x_{j_0}^{i_0})) \\ X \setminus \overline{x}_{j_0}^{i_0}(X), \text{ если } \delta^{-}(X) \geq \delta^{+}(X), \delta^{-}(X) > 0, \\ \overline{x}_{j_0}^{i_0}(X) = \arg\max\limits_{\overline{x}_{j_0}^{i_0} \in \Pr^{-}(X)} (\mathrm{Krit}(X \setminus \overline{x}_{j_0}^{i_0})) \\ X, \text{ если } \delta^{+}(X) \leq 0 \text{ и } \delta^{-}(X) \leq 0. \end{array} \right\}$$

Неподвижная точка $\Pr^{n+1}(X) = \Pr^{n}(X)$ получается в третьем случае, когда добавление/удаление элемента не увеличивает критерий. Алгоритм обнаружения неподвижных точек и машинный эксперимент, подтверждающий эффективность данной модели восприятия, приведены в [5-6].

Список литературы

1. Ganter B., Wille R. Formal Concept Analysis: Mathematical Foundations. — Springer Verlag, 1999.
2. Formal Concept Analysis: Foundations and Applications. / Edited by B. Ganter, G. Stumme, R.Wille — Springer Verlag, 2005.
3. Витяев Е.Е., Демин А.В., Пономарёв Д.К. Вероятностное обобщение формальных понятий // Программирование, Т.38, №5, 2012, с.219-230
4. Alexander Demin, Denis Ponomarev, Evgenii Vityaev. Probabilistic Concepts in Formal Contexts // Preliminary Proceedings of the Ershov Informatics Conference PSI Series, 8-th Edition (June 27 – July 1, 2011, Novosibirsk), Novosibirsk, 2011, pp 29-38
5. Витяев Е.Е., Неупокоев Н.В. Формальная модель восприятия и образа как неподвижной точки предвосхищений. Нейроинформатика, 2012, том 6, № 1, стр. 28-41
6. Витяев Е.Е., Неупокоев Н.В. Математическая модель восприятия и образа // Информационные технологии в гуманитарных исследованиях, Вып.17, ИАЭТ СО РАН, Новосибирск, 2012, 63-72.
7. Scientific Discovery: http://www.math.nsc.ru/LBRT/logic/vityaev
8. 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.
9. 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
10. Смирнов С.Д. Психология образа. МГУ, М., 1985, с.231
11. 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.
12. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов // Новосибирский гос. ун-т. Новосибирск, 2006. с.293.
13. Витяев Е.Е., Перловский Л.И., Ковалерчук Б.Я., Сперанский С.О. Вероятностная динамическая логика мышления. Нейроинформатика, 2011, том 5, № 1, стр. 1-20