Витяев Е.Е.
Мартынович В.В.

Вероятностные формальные понятия на контекстах с отрицаниями и их непротиворечивость

Аннотация: *Дано вероятностное обобщение формальных понятий, которое устойчиво к шумам в данных и дает возможность учитывать, как наличие признака, так и его отсутствие (контекст с отрицанием). Данное обобщение получено исходя из определения формальных понятий через неподвижные точки импликаций. В этом определении импликации заменены на вероятностные законы, включающие отрицание признака. Доказано, что неподвижная точка, основанная на вероятностных импликациях непротиворечива и тем самым определяет вероятностное формальное понятие. Получено соответствие между вероятностными формальными понятиями и классическими формальными понятиями. Ранее экспериментально было подтверждено, что, если взять формальные понятия, обнаруженные на некоторых данных и затем в эти данные ввести случайный шум, то на зашумленных данных можно снова получить исходные формальные понятия. Приводятся результаты эксперимента, демонстрирующие формирование вероятностных формальных понятий.*

Ключевые слова: *Data Mining, формальные понятия, ассоциативные правила, семантический вероятностный вывод*

Abstract: *The probabilistic generalization of formal concept is presented: construction is resistant to noise in the data and give one an opportunity to consider contexts with negation (object-attribute relation allows both attribute presence and it's absence). This generalization is obtained from the notion of formal concepts through its definition as fixed points of implications. Then implications, possibly with negations, are replaced by probabilistic laws. We prove that the fixed points based on the probabilistic implications are consistent and thus determine the probabilistic formal concepts. Correspondence between probabilistic formal concepts and classic formal concepts is also presented. Previously the following was confirmed experimentally: if the formal concepts are found on some data and are swayed with some random noise, then the noisy data context allows the complete reconstruction of initial formal context's concepts. In the end, the demonstration for the probabilistic formal concepts formation is given.*

Keywords: *Data Mining, formal concepts, associative rules, semantic probabilistic inference*

1 Введение

В анализе формальных понятий (далее FCA) [Ganter B., 2003; Ganter B., Wille R., 1999], понятия выступают в роли классификационной единицы. Однако, имеется ряд вопросов, оставленных в классической теории без должного внимания:

1. Предсказательная силы таких понятий: возможность классификации новых объектов, не входящих в контекст;
2. Устойчивость понятий относительно возможных ошибок в исходных данных;
3. Устранение случайных атрибутов из содержаний различных понятий.

В настоящее время не существует индуктивной парадигмы для FCA, способной ответить на данные вопросы. Цель данной работы – разработать индуктивное обобщение FCA, отвечающее на эти вопросы. Первый шаг в разработке индуктивной парадигмы был сделан в [Demin A., Ponomaryov D., Vityaev E., 2011; Demin A., Ponomaryov D., Vityaev E., 2012], где было сформулировано вероятностное обобщение FCA. Следующий шаг по направлению создания нового метода кластеризации, в котором классы не пересекаются, должен состоять в разработке вероятностного обобщения FCA содержащего отрицания. Для этого мы использовали определение формальных понятий в виде неподвижных точек импликаций с отрицанием. Аналогично работе [Demin A., Ponomaryov D., Vityaev E., 2011; Demin A., Ponomaryov D., Vityaev E., 2012], далее мы определили вероятностные формальные понятия через неподвижные точки вероятностных импликаций с отрицанием. Однако в этом случае потребовалось доказательство непротиворечивости полученных неподвижных точек. В результате было получено индуктивное вероятностное обобщение формальных понятий, которое:

- имеет предсказательную силу – новый объект может быть отнесен к одному из имеющихся классов;
- минимизирует описание классов путем элиминации случайных атрибутов;
- устойчиво относительно шумов в данных.

Примеры таких вероятностных формальных понятий приведены в [Demin A., Ponomaryov D., Vityaev E., 2012; Vityaev E., Lapardin K., Khomicheva I., Proskura A., 2014; Neupokoev N., Vityaev E., 2014].

2 Анализ формальных понятий

Здесь мы дадим краткий экскурс в анализ формальных понятий. Для более глубокого изучения можно обратиться, например, к классическим трудам [Ganter B., 2003; Ganter B., Wille R., 1999].

FCA изучает набор объектов G, каждый из которых может обладать некоторыми свойствами из фиксированного набора M. Для описания «объект g обладает свойством m» служит отношение $I \subseteq G \times M$.

Определение 1. *Формальным контекстом называется упорядоченная тройка* (G, M, I)*, где* G *и* M *- множества произвольной природы, а* $I \subseteq G \times M$.

Далее формальный контекст мы будем называть просто контекстом. На таком контексте определим операцию $'$ (взятие общего) следующим образом:

Определение 2. $A \subseteq G$, $B \subseteq M$. *Тогда:*

1. $A' = \{m \in M | \forall_{g \in A} (g, m) \in I\}$
2. $B' = \{g \in G | \forall_{m \in B} (g, m) \in I\}$
3. $g' = \{g\}' = \{m \in M | (g, m) \in I\}$

Ключевым объектом FCA является «формальное понятие».

Определение 3. *Упорядоченная пара* (A, B) *называется понятием, если* B *- общие признаки для объектов из* A*, а* A *- в точности все объекты обладающие признаками из* B*. Иначе говоря,* $A' = B$*, а* $B' = A$.

Естественным образом возникает порядок на понятиях:

Определение 4. *Если* (A_1, B_1) *и* (A_2, B_2) *- понятия заданного контекста, то* $(A_1, B_1) \leq (A_2, B_2) \Leftrightarrow A_1 \subseteq A_2$.

Здесь и далее мы немного окунемся в теорию FCA [Ganter B., 2003; Ganter B., Wille R., 1999; Ganter B., Obiedkov S., 2004], но лишь ровно настолько, насколько это необходимо для разработки предлагаемой в статье контрукции.

Следствие 1. *Пусть* $A_1 \subseteq A_2 \subseteq G, B_1 \subseteq B_2 \subseteq M$. *Тогда*

1. $A_2' \subseteq A_1', B_2' \subseteq B_1'$
2. $A \subseteq A'', B \subseteq B''$
3. (A, B) *- понятие* $\Rightarrow B'' = B$

Теорема 1. *Множество понятий с определенным на нем отношении порядка* $\leq$ *является полной решеткой, где операторы* inf *и* sup *заданы как [Ganter B., 2003]:*

$$\bigwedge_{j \in J}(A_j, B_j) = \left(\bigcap_{j \in J} A_j, (\bigcup_{j \in J} B_j)''\right)$$
$$\bigvee_{j \in J}(A_j, B_j) = \left((\bigcup_{j \in J} A_j)'', \bigcap_{j \in J} B_j\right)$$

В реальности атрибуты не могут быть разбиты абсолютно произвольным образом. Как правило, они образуют многочисленные взаимосвязи для описания которых служит понятие «импликации». A, B в опредлениях ниже понимаются как наборы атрибутов $\subseteq M$.

Определение 5. *Импликацией мы называем упорядоченную пару* (A, B) *и пишем* $A \to B$.

Определение 6. $A \to B$ *истина на* $K = (G, M, I)$, *если* $\forall g \in G(A \nsubseteq g'$ *или* $B \subseteq g')$. *Множество всех истиных импликаций мы будем обозначать как* $Imp(K)$.

Определение 7. *Импликацию* $A \to B$ *назовем нетривиальной на* K, *если* $B \nsubseteq A$ *и* $A' \neq \varnothing$. *Множество всех нетривиальных истиных импликаций на* K *мы будем обозначать* $ntImp(K)$.

Определение 8. *Для любого набора импликаций* L *можно построить оператор непосреднного вывода* f_L, *добавляющий заключения всех импликаций к множеству-операнду:*

$$f_L(X) = X \cup \{B \mid A \subseteq X,\ A \to B \in Imp(K)\}$$

Последовательно применяя оператор непосредственного вывода к какому-либо множеству X, мы постепенно приближаемся к его замыканию [Ganter B., 2003; Demin A., Ponomaryov D., Vityaev E., 2012].

Определение 9. *Оператор* cl_L, *замыкающий множество X относительно непосредственного вывода, есть* $cl_L(X) = f_L^{\infty}(X)$.

Теорема 2. *Для любого множества* $M \subseteq M$ *выполнено следующее [Demin A., Ponomaryov D., Vityaev E., 2012]:*

1. $f_{Imp(K)}(B) = B \Leftrightarrow B'' = B$;
2. *Если* $B' \neq \varnothing$, *то* $f_{ntImp(K)}(B) = B \Leftrightarrow B'' = B$.

3 Многозначные контексты

Как нетрудно заметить, рассуждения предыдущего раздела допускают лишь теоретико-множественные высказывания об атрибутах рассматриваемого объекта. Ничего не говорится о степени обладания тем или иным атрибутам, а единственные доступные конструкты - подмножества атрибутов объекта.

Такой подход, ввиду выразительной скудности, в большинстве практических задач ведет к плохо соотносящимся с действительностью интерпритациям, а иногда и результатам [Demin A., Ponomaryov D., Vityaev E., 2012; Missaoui R., Kwuida L., 2011]. Особенно это относится к комбинациям атрибутов и к рассуждениям в терминах импликаций [Missaoui R., Kwuida L., 2011].

Существует несколько различных подходов, связанных с расширением понятия отношения I. Классические примеры таких построений можно найти в работах [Ganter B., Obiedkov S., 2004; Missaoui R., Kwuida L., 2011]. В данной главе мы обогатим контексты, снабдив каждую пару (g, m) степенью принадлежности аттрибута объекту.

Итак, пусть каждый атрибут m имеет свое множество допустимых значений V_m. Для описания принадлежности аттрибута объекту теперь необходимо задать значение v этого атрибута m на выбранном объекте g, $(g, m, v) \in I$, где $v \in V_m$.

Определение 10. *Пусть G - множество объектов и M - множество атрибутов и для каждого атрибута заданы множества принимаемых атрибутов значений V_m. Многозначным контекстом K называется тройка*

$$(G, M, \{I_m : G \to V_m \mid m \in M\})$$

По сути $I_m(g)$ задают значения атрибута m для объекта g и являются означиваниями качеств из набора M для каждого из рассматриваемых объектов.

Нетрудно усмотреть взаимосвязь между многозначными контекстами и обычными контекстами, выраженную в следующей конструкции. Каждый атрибут вместе со своим конкретным значением может быть рассмотрен как новый независимый атрибут. Т.е. для каждого атрибута m рассмотрим множество пар (m, v) где $m \in M$, а $v \in V_m$. Свойство объекта принимать качество v на атрибуте m можно теперь описать как $(g, (m, v)) \in I$.

Определение 11. *Клоном $K = (G, M, \{I_m\})$ мы называем $K^* = (G, M^*, \cup I^*_m)$, где $M^* = \{(m, v) \mid m \in M, v \in V_m\}$, а $I^*_m = \{(g, (m, v)) \mid I_m(g) = v\}$. Для краткости, мы будем говорить что $g \in G$ обладает атрибутом m_v, если $I_m(g) = v$ или, что тоже самое, $(m, v) \in g'$ в рамках K^*.*

На клоне K^* естественным образом возникают все конструкции, а также утверждения и теоремы, данные классическим анализом формальных понятий. Приняв их верными, не умаляя общности, мы можем говорить например о формальных понятиях, определнных на многозначных контекстах. Достаточно в соответствующих текстах допустить вместо m вхождения m_v. Далее говоря о классических контрукциях на K мы имеем в виду именно те же самые конструкции относительно его клона K^*. Например,

Определение 12. *Формальным понятием на многозначном контексте K является пара (A, B), где $A \subseteq G$, $B \subseteq M^*$, такая что (A, B) - формальное понятие на K^*.*

4 Формулы на бинарных контекстах

Естественным является желание рассмотреть предложенную структуру в простейшем случае. Каждый атрибут снабдим интерпритацией предикатной природы, отождествив значение 1 с наличием соответствующего атрибута и 0 с его отсутствием.

Определение 13. *Бинарный контекст есть многозначный контекст, для которого* $\forall m(I_m : G \to \{0,1\})$. *Атрибуты клона* K^* *мы будем означать как* m *и* $\overline{m}$ *для* $(m,1)$ *и* $(m,0)$ *соответственно.*

Наша ближайшая задача - построить на произвольном бинарном контексте формальную систему, аналогичную логике первого порядка. Это даст нам большую гибкость в описании, а также проиллюстрирует непосредственную применимость аппарата логики первого порядка к исследованию бинарных контекстов.

Определение 14. *Для бинарного контекста* $K = (G, M)$ *определим сигнатуру* σ*:*

1. *Множество предикатов R - в точности множество всех* I_m*, утверждающих наличие соответствующего атрибута или отрицающих таковое;*

2. *Пустое множество функциональных символов F* $= \varnothing$*;*

Все понятия атома, терма, литеры, формулы и т.д. определяются классическим образом. Формула, возникающая в рамках описанной сигнатуры оперирует логическими связками $\&, \vee, \to, \neg$ и предикатами, описывающими наличие или отсутствие какого-либо атрибута. Обозначим получающиеся множества атомов, литер, формул и предложений как $At(K)$, $Lit(K)$, $For(K)$ и $Sen(K)$ соответственно.

Множество моделей возникает на контексте естественным образом. Несущее множество $\mathcal{D} = \{g\}$ вместе с подстановкой $\mu(x) = g$ образует модель K_g. Факт истиности формулы Φ на модели объекта g мы обозначим следующим образом: $g \vDash \Phi \Leftrightarrow K_g \vDash \Phi$. $G_\Phi \subseteq G = \{g \in G \mid g \vDash \Phi\}$ будем называть носителем Φ. Если $G_\Phi = G$, то Φ - контекстная тавтология.

Следствие 2. *Заметим, что:*

1. $G_{\neg\Phi} = G \setminus G_\Phi$*;*

2. $G_{\Phi\&\Psi} = G_\Phi \cap G_\Psi$*;*

3. $G_{\Phi\vee\Psi} = G_\Phi \cup G_\Psi$*;*

5 Вероятность формул на контекстах

Последнее, что нам понадобится – реализация понятия вероятности бинарного контекста. Здесь мы дадим одну общую идею для введения вероятностной меры на множестве формул бинарного контекста. Идеи синтеза логики и вероятности, а также и некоторые формулировки, взяты из [Speransky S.O., 2013].

Определение 15. *Рассмотрим верооятностную меру* μ *на множестве G в колмогоровском смысле, которое можно трактовать как множество элементарных событий. Введем контекстную вероятностную меру*

$$\nu : For(K) \to [0,1], \nu(\Phi) = \mu(G_\Phi) = \mu(\{g \mid g \vDash \Phi\})$$

Определение 16. *Статистическими незначимыми объектами мы называем $g \in G$ такие, что $\mu(g) = 0$. Соответственно статистически незначимыми подмножествами - $A \subset G$ такие, что $\mu(A) = 0$.*

Определение 17. *Формулу Φ назовем ν-совместной, если $\nu(\Phi) > 0$. Формулу Φ назовем почти тавтологией, если $\nu(\Phi) = 1$.*

Предложение 1. *Контекстная мера ν обладает следующими свойствами:*

1. *Если Φ - тавтология в классическом смысле, то Φ – контекстная тавтология и $\nu(\Phi) = 1$;*
2. *Пусть $\Vdash \neg(\Phi \& \Psi)$. Тогда $\nu(\Phi \vee \Psi) = \nu(\Phi) + \nu(\Psi)$;*
3. *$\nu(\Phi \& \Psi) \leq \nu(\Phi)$.*

■ 1. Φ - общезначима, поэтому истина на любой модели. В частности, $\forall g \in G(K_g \vDash \Phi)$, то есть $\Vdash \Phi$. Поэтому $\nu(\Phi) = \mu(\{g \mid g \vDash \Phi\}) = \mu(G) = 1$.

2. $\Vdash \neg(\Phi \& \Psi)$, поэтому $\forall g \in G(K_g \vDash \neg\Phi \& \Psi) \Rightarrow$ неверно что $K_g \vDash \Phi \& \Psi$ и $\mu(G_{\Phi \& \Psi}) = 0$. Отсюда и из леммы 4 согласно определению вероятностной меры немедленно вытекает, что $\mu(G_{\Phi \vee \Psi}) = \mu(G_\Phi \cup G_\Psi) = \mu(G_\Phi) + \mu(G_\Psi)$.

3. $G_{\Phi \& \Psi} = G_\Phi \cap G_\Psi \subseteq G_\Phi$; в силу аксиом меры: $\nu(\Phi \& \Psi) = \mu(G_{\Phi \& \Psi}) \leq \mu(G_\Phi) = \nu(\Phi)$.
□

6 Правила на контексте

В данном разделе мы всюду полагаем $L = Lit(K)$, K - бинарный контекст и ν - контекстная мера на нем. Мы ориентируемся на схему изложения, предложенную в [Vityaev E.E., 2006; Smerdov S., Vityaev E., 2011].

Определение 18. *По множеству литер $M \subseteq L$ построим его композицию: $\&M = \underset{P \in M}{\&} P$. Для случая $M = \varnothing$ полагаем $\&M = 1$.*

Формулы, имеющие вид простых конъюнкций $F = m_{i_1} \& m_{i_2} ... \& m_{i_k}$ обладают одним свойством, позволяющим связывать формульные конструкции и классические операции из FCA. В действительности носитель G_F совпадает с $\{m_{i_j}\}'$. В этом смысле можно отождествлять множества литер с их представлением в виде набора атрибутов $\{m_{i_j}\}$.

Более того, формула $m_{i_1} \& m_{i_2} ... \& m_{i_k} \to m = \&\{m_{i_j}\} \to m$ описывает тот же самый процесс что и импликация на контексте в классическом смысле, $(\{m_{i_j}\}, \{m\})$. В связи с этим логично выделить класс импликаций, так как это сделано в FCA, однако уже внутри класса формул. Мы называем их правилами.

Определение 19.

1. *Правило - формула вида $R = (H_1 \& H_2 ... \& H_k \Rightarrow T)$, где $H_i \in L$, $T \notin \{H_1, H_2, ... H_k\}$.*
2. *Для правила R под* head(R) *мы имеем в виду множество $\{H_1, H_2 ..., H_k\}$, а* tail$(R) = T$.
3. *Длина правила есть мощность его посылки:* len$(R) = |$head$(R)|$.

Определение 20. *Вероятностью правила R, определенного выше, называется величина*

$$\eta(R) = \nu(\text{tail}(R)|\text{head}(R)) = \frac{\nu(\&\text{head}(R)\&\text{tail}(R))}{\nu(\&\text{head}(R))}$$

Правило назовем нетривиальным, если выражение в знаменателе меньше единицы. Если выражение в знаменателе равно нулю, вероятность правила остается неопределенной.

Определение 21. *Правило R_1 - подправило правила R_2, если* $\text{head}(R_1) \subseteq \text{head}(R_2)$ *и* $\text{tail}(R_1) = \text{tail}(R_2)$. *Сей факт обозначим как $R_1 \succ R_2$.*

Определение 22. $R_1 = R_2$, *когда* $\text{head}(R_1) = \text{head}(R_2)$ *и* $\text{tail}(R_1) = \text{tail}(R_2)$.

Определение 23. *R_1 - обобщение правила R_2, то есть $R_1 \succeq R_2$, когда $R_1 \succ R_2$ или $R_1 = R_2$.*

Определение 24. *R_1 - уточнение правила R_2, $R_1 > R_2$, если $R_2 \succ R_1$ и $\eta(R_1) > \eta(R_2)$.*

Теорема 3. *Пусть R - нетривиальное правило на контексте K с мерой ν.*

1. *Вероятность R меньше или равна вероятности соответствующей импликации:*

$$\eta(R) \leq \nu(\text{head}(R) \to \text{tail}(R))$$

2. *R – почти тавтология $\Leftrightarrow \eta(R) = \nu(R) = 1$.*

■ Положим $H = \&\text{head}(R)$, $T = \text{tail}(R)$ и рассмотрим разность (1): $\nu(H)(\eta(R) - \nu(H \to T))$. Заметим, что $H \to T = T \vee \neg H = (T\&H) \vee \neg H$, при этом $(T\&H)\&\neg H = 0$. Значит, по лемме 4, $\nu(H \to T) = \nu(T\&H) + \nu(\neg H)$. Отсюда разность (1) можно преобразовать как

$$\nu(H)(\eta(R) - \nu(H \to T)) = \nu(H\&T) - \nu(H\&T)\nu(H) - \nu(\neg H)\nu(H) =$$
$$\nu(H\&T)\nu(\neg H) - \nu(H)\nu(\neg H) = -\nu(H\&\neg T)\nu(\neg H) \leq 0$$

Далее, равенство 0 достигается только в случае $\nu(H\&\neg T) = 0$. Однако это равносильно тому что $\nu(H\&T) = \nu(H) - \nu(H\&\neg T) = \nu(H)$ и $\eta(R) = \frac{\nu(H\&T)}{\nu(H)} = 1$. Отсюда с учетом выкладок выше заключаем что R – почти тавтология. □

Следствие 1. *Если мера μ не допускает незначимых объектов, то множество почти тавтологий превращается в множество тавтологий, а $\eta(R) = 1 \Leftrightarrow R$ - контекстная тавтология.*

Определение 25. *R – вероятностный закон, если он является уточнением любого своего подправила, т.е. $(R' \succ R) \Rightarrow (R > R')$.*

Докажем несколько технических вещей, которые нам понадобятся в дальнейшей работе с правилами.

Следствие 3. *Если добавление литеры H в посылку R уменьшает его вероятность, $\eta(\&\text{head}(R)\ \&\ H \Rightarrow \text{tail}(R)) < \eta(R)$, то $\neg H$ ее увеличивает [Vityaev E.E., 1992].*

Следствие 4. *Для любого правила R существует его обобщение R' такое, что:*

1. *R' – вероятностный закон;*

2. *$\nu(R') \geq \nu(R)$.*

■ Рассмотрим множество $\Pi = \{A \mid \nu(A) \geq \nu(R),\ A \succeq R\}$. $R \in \Pi$, поэтому $\Pi \neq \varnothing$. Значит, существует минимальный в смысле отношения $\succeq$ элемент, назовем его $S = \min \Pi$. Условие 2 леммы выполнено для S по построению Π.

Пусть S не является законом, т.е. найдется подправило S', такое что $\nu(S') \geq \nu(S)$ и $S' \succ S$, с учетом $S \succeq R$ заключаем, что $S' \succ R$. С другой стороны, $\nu(S') \geq \nu(S) \geq \nu(R)$, откуда вытекает, что $S' \in \Pi$, противоречащее минимальности S. □

7 Теорема об уточнении

Важным шагом будет применение предложенной в [Vityaev E.E., 1992] техники. Дадим необходимые определения полностью аналогичные определениям для класса правил.

Определение 26.

1. *Псевдоправило - формула вида $R = ((P_1 \& ... \& P_k)\ \&\ \neg(N_1 \& ... \& N_s) \Rightarrow T)$;*

2. *Для псевдоправила R, $\text{head}(R) = (P_1 \& ... \& P_k)\ \&\ \neg(N_1 \& ... \& N_s)$ и $\text{tail}(R) = T$;*

3. *Литеры P_i мы называем позитивной частью посылки, а литеры N_j – негативной;*

4. *Вероятностью псевдоправила R называется величина*

$$\eta(R) = \nu(\text{tail}(R) | \text{head}(R)) = \frac{\nu(\&\text{head}(R)\ \&\ \text{tail}(R))}{\nu(\&\text{head}(R))}$$

Теорема 4. *(Об уточнении) Пусть S – псевдоправило $((\&A)\&\neg(\&B)) \Rightarrow G), R$ - соответствующее правило без негативной части $\&A \Rightarrow G$, притом $\eta(S) > \eta(R)$. Тогда для R существует уточнение с помощью негативной части псевдоправила S в виде правила $R' > R$.*

■ Обозначим для краткости $\overline{A} = \&A$, $\overline{B} = \&B$. Выпишем вероятность псевдоправила S:

$$\eta(S) = \nu(G \mid \overline{A}\ \&\ \neg\overline{B}) = \nu(G \mid \overline{A}\ \&\ (\neg B_1 \vee ... \vee \neg B_m)) \tag{1}$$

Представим дизъюнкцию как дизъюнкцию конъюнкций:

$$\neg B_1 \vee ... \vee \neg B_m = \bigvee_{i=(0,...,0)}^{i=(1,...,1,0)} (B_1^{i_1}\ \&\ ...\ \&\ B_m^{i_m})$$

где 0 в мультииндексе обозначает наличие отрицания, а 1 - его отсутствие. Включены все мультииндексы за исключением $(1,...1)$, который соответствует конъюнкции B_1 & ... & B_m.

Тогда условная вероятность (1) перепишется как

$$\eta(S) = \nu(G \mid \bigvee_{i=(0,...,0)}^{i=1,...,1,0)} (\overline{A}\ \&\ B_1^{i_1}\ \&\ ...\ \&\ B_m^{i_m})) \tag{2}$$

Теперь пусть утверждение теоремы неверно, то есть одновременно выполнены все неравенства $\nu(G \mid \overline{A}$ & $B_1^{i_1}$ & ... & $B_m^{i_m}) \leq \nu(G \mid \overline{A})$, если конечно соответствующие вероятности определены. Поскольку $\nu(\overline{A}$ & $\neg\overline{B}) \neq 0)$, то найдется хотя бы один мультииндекс, для которого это так. Тогда

$$\nu(G \,\&\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m}) \leq \nu(G \mid \overline{A})\nu(\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m});$$

$$\nu(G \mid \bigvee_{i=(0,...,0)}^{i=1,...,1,0)} (\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})) = \frac{\nu(\vee\, G \,\&\, \overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})}{\nu(\vee\, \overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})} =$$

$$\frac{\sum \nu(G \,\&\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})}{\sum \nu(\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})} \leq \frac{\nu(G \mid \overline{A}) \sum \nu(\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})}{\sum \nu(\overline{A} \,\&\, B_1^{i_1} \,\&\, ... \,\&\, B_m^{i_m})} = \nu(G \mid \overline{A});$$

Последнее согласно (2) означает, что $\eta(S) \leq \eta(R)$ – противоречие с условием. Значит, наше предположение неверно и для одного из правил $\nu(G \mid \overline{A}$ & $B_1^{i_1}$ & ... & $B_m^{i_m}) > \nu(G \mid \overline{A})$. □

8 Семантический вероятностный вывод

Определим одно из ключевых для данной работы понятий, отношение семантической выводимости на множестве правил.

Определение 27. *Правило R – семантически выводимо из правила R', если:*

1. *R, R' - вероятностные законы;*
2. len$(R) =$ len$(R') + 1$*;*
3. *$R > R'$.*

Отношение выводимости есть $\rangle = \{(R', R) \mid R-$ семантически выводимо из $R'\}$.

Определение 28. *Вероятностный закон R – сильнейший, если $\forall R' \neg(R \rangle R')$.*

Определение 29. *Семантическим вероятностным выводом (СВВ) называется последовательность $R_0 \rangle R_1 \rangle R_2 ... \rangle R_m$, в которой:*

1. len$(R_0) = 0$*;*
2. *R_m – сильнейший вероятностный закон.*

Иначе говоря, СВВ требует процедуру вывода от начала и до конца.

Определение 30. *Максимально специфическим законом для предиката T назовем тот сильнейший вероятностный закон, который имеет максимальную условную вероятность среди всех сильнейших вероятностных законов с заключением T.*

Множество максимально специфических законов на контексте K мы будем обозначать как MSR$_K$ или MSR, если не возникает неоднозначностей. MSR(G) есть подмножество MSR, в котором все правила имеют в заключении G.

Следствие 5. *Для любого правила R,* tail$(R) = G$*, вероятность которого определена, всегда существует максимально специфический закон W с тем же самым заключением G, такой что $\eta(R') \geq \eta(R)$.*

■ Согласно лемме 9 для R существует обобщение R' являющееся вероятностным законом. Но для R' существует сильнейший вероятностный закон R'' т.ч. $\eta(R'') \geq \eta(R')$. Для последнего существует максимум множества сильнейших вероятностных законов, т.е. максимально специфический R''' и $\eta(R''') \geq \eta(R'') \geq \eta(R') \geq \eta(R)$. $W = R'''$ - искомый. □

9 Классы правил

В работе [Speransky S.O., 2013] представлены классы правил, используемые для обоснования корректности семантического вероятностного вывода. Немного видоизменив эти определения, мы дадим в некотором смысле аналогичный результат.

Определение 31. $R \in M_1(G) \Leftrightarrow ((\varnothing \Rightarrow G) \succ R \;\Rightarrow\; R > (\varnothing \Rightarrow G))$

Определение 32. $R \in M_2(G) \Leftrightarrow R \in M_1(G)$ *и* $(\forall R' \in M_1(G))[R \succ R' \Rightarrow \eta(R') \leq \eta(R)]$

Фактически, класс M_1 обязывает правила быть значащими, таким образом чтобы их рассмотрение имело смысл по сравнению с безусловным утверждением G. Класс M_2 требует от правил принципиальной неуточняемости (как бы мы ни расширяли правило R, мы никогда не получим улучшения оценки его вероятности). Имеет место следующая взаимосвязь:

Предложение 2. MSR $\subset M_2 \subset M_1$.

■ Второе включение очевидно. Пусть $R \in$ MSR.

1. Для R существует некоторый СВВ, согласно п.1 определения 29 начинающийся с безусловного правила $R' = \varnothing \Rightarrow$ tail(R). Если посылка R вдруг непуста, то $\varnothing \Rightarrow$ tail$(R) \succ R$ и из цепочки отношений семантических выводимостей следует что $R > R'$ и $R \in M_1$. Если посылка R пуста, то последнее выполнено автоматически.

2. Рассмотрим $R \succ R' \in M_1$ и предположим что $\eta(R') > \eta(R)$. Из леммы 11 следует, что найдется $S \in$ MSR : $\eta(S) \geq \eta(R') > \eta(R)$. Последнее противоречит максимальной специфичности R и поэтому $\eta(R') \leq \eta(R)$. Отсюда $R \in M_2$. □

Определение 33. *Набором правил будем называть* $\Pi \subseteq M_2$.

В духе подхода, обозначенного в работе [Demin A., Ponomaryov D., Vityaev E., 2012], для исследования формальных понятий на бинарном контексте K, нам достаточно понять структуру неподвижных точек соответствующего оператора предсказания на клоне K^*. Здесь мы ставим цель исследования неподвижных точек для вероятностностного оператора предсказания на K, с учетом наличия отрицаний в форме логических отрицаний на формулах специального вида. Мы так долго шли к понятиям:

Определение 34. *Оператор непосредственного предсказания на наборе правил* Π *работает следующим образом:*

$$\text{Pr}_\Pi(L) = L \cup \{G \mid \exists R \in \Pi : \text{head}(R) \subseteq L, \text{tail}(R) = G\}$$

Т.е. Pr$_\Pi$ *добавляет операнду заключения всех импликаций, посылка которых уже содержится в* L *и в некотором смысле выполнена на нем.*

Определение 35. *Замыкание набора литер* L *– это наименьшая неподвижная точка оператора непосредственного предсказания:* PR$_\Pi(L) = \text{Pr}_\Pi^\infty(L)$.

10 Теорема совместности

Предложенная конструкция нуждается в обосновании ее корректности, аналогично [Speransky S.O., 2013]. Корректность здесь понимается в двух смыслах: в вероятностном и в логическом. Покажем, что оба эти требования выполнены.

Определение 36. *Набор литер L назовем совместным, если &L – ν-совместна на K.*

По сути, набор является совместным, когда находится набор статистически значимых объектов из G_K, на которых выполнена формула &L. На «нормальных» вероятностых контекстах K (с не более чем счетным множеством объектов G, каждый из которых является статистически значимым), совместными окажутся такие множества литер L (а вместе с ними и соответствующие множества атрибутов L из M_{K^*}), для которых L$' \neq \varnothing$.

Специфика построенного бинарного контекста также подразумевает для каждого объекта либо наличие некоторого атрибута $m \in M_{K^*}$, либо его отсутствие, что равноценно обладанию атрибутом $\overline{m} \in M_{K^*}$.

Определение 37. *Набор литер L – непротиворечив, если он не содержит одновременно какой-либо атом G вместе с его отрицанием $\neg G$.*

Предложение 3. *Если L – совместно, то L – непротиворечиво.*

■ В противном случае противоречивость L означает что найдется G, такой что $G \in L$ и $\neg G \in L$, поэтому $\nu(\&L) \leq \nu(G \;\&\; \neg G) = 0$. □

Покажем сначала, что непосредственное предсказание сохраняет свойство совместности.

Теорема 5. *(Совместности) Если L – совместно, то* Pr(L) *– также совместно.*

■ Доказательство нетрудно получить, взглянув на теорему об уточнении. Рассмотрим все правила, которые вносят вклад в формирование образа L: $T = \{R \in \Pi \mid \text{head}(R) \subseteq L\}$. Занумеруем все элементы T произвольным образом, $T = \{T_1, ...T_m\}$, и рассмотрим последовательность множеств $U_i = U_{i-1} \cup \{\text{tail}(T_i)\}, U_0 = L$. Покажем что каждое U_i - совместно.

База очевидна из посылки теоремы.

Пусть U_i - совместно. Для краткости обозначим $U = U_i, W = U_{i+1}, R = R_{i+1}$ и $G = \text{tail}(R), H = \text{head}(R), N = U \setminus H$. Предположим, что W - несовместно, то есть $\nu(\&W) = 0$. Аналогично теореме об уточнении рассмотрим псевдоправило $F = (\&H \;\&\; \neg(\&N)) \Rightarrow G$. Рассмотрим 2 случая:

1. $\nu(\&\text{head}(F)) \neq 0$. Тогда вероятность F определена и

$$\eta(F) = \frac{\nu(\&H \;\&\; \neg(\&N) \;\&\; G)}{\nu(\&H \;\&\; \neg(\&N))} = \frac{\nu(\&H \;\&\; G) - \nu(\&H \;\&\; (\&N) \;\&\; G)}{\nu(\&H) - \nu(\&H \;\&\; (\&N))} =$$
$$\frac{\nu(\&H \;\&\; G) - \nu(\&W)}{\nu(\&H) - \nu(\&U)} = \frac{\nu(\&H \;\&\; G)}{\nu(\&H) - \nu(\&U)} > \frac{\nu(\&H \;\&\; G)}{\nu(\&H)} = \eta(R) > 0.$$

   По теореме об уточнении, найдется правило S такое, что $S > R$, что противоречит неуточняемости правила R (то есть тому, что $R \in M_2$).

2. $\nu(\&\text{head}(F)) = 0$. Тогда

$$\nu(\&\text{head}(F)) = \nu(\&H \;\&\; \neg(\&N)) = 0 \Rightarrow \nu(\&H \;\&\; \neg(\&N) \;\&\; G) = 0;$$
$$0 = \nu(\&H \;\&\; (\&N) \;\&\; G) = \nu(\&H \;\&\; G) - \nu(\&H \;\&\; \neg(\&N) \;\&\; G) = \nu(\&H \;\&\; G).$$

Последнее означает $\eta(R) = 0$, противоречащее $R \in M_1$ ($0 = \eta(R) > \eta(\varnothing \Rightarrow G) \geq 0$). □

Следствие 2. *Если L – совместно, то* PR(L) *– тоже совместно.*

Следствие 3. *Если L – совместно, то* PR(L) *– непротиворечиво.*

11 О несовместных наборах

С совместными наборами ситуация достаточно ясная. Непосредственное предсказание по совместному набору L и замыкание этого набора всегда гарантированно совместно и непротиворечиво.

Попробуем разобраться со структурой несовместных наборов L. Начнем с достаточно тривиального утверждения, обратного к теореме о совместности.

Предложение 4. *Если L – несовместно, то* PR(L) *– тоже несовместно.*

■ Предположив совместность PR(L) получим, что любое подмножество, и в частности L, оказывается совместным. □

Несколько труднее оказывается вопрос о непротиворечивости таких замыканий. Для более подробного исследования структуры несовместных систем литер нам понадобится следующее понятие.

Определение 38. $M \underset{\nu}{\subseteq} L$ *(или же M – ν-максимально в L), если когда M – максимальное по включению множество в L, такое что M совместно.*

Определение 39. *Систему правил Π назовем полной, если* MSR $\subset \Pi$.

Далее речь пойдет лишь о полных системах правил. Требование полноты можно немного ослабить, как будет видно из теоремы ниже, но в данной статье мы ограничимся максимально специфичными правилами. Следует отметить, что согласно предложению 2, системы, содержащие M_2, оказываются полными.

Теорема 6. *Пусть $M \underset{\nu}{\subseteq} L$. Тогда $M \cup \neg(L \setminus M) \subseteq$* PR$(M)$.

■ Положим x принадлежащим левой части выражения. Случай $x \in M$ очевиден, тогда $x \in$ PR(M) по определению замыкания.

Теперь пусть $x \in L \setminus M$. По определению ν-максимального подмножества, множество $M \cup \{x\}$ оказывается несовместным (иначе получим новое максимальное по включению множество). Последнее означает, что

$$\nu(\&M \;\&\; x) = 0;$$
$$\nu(\&M \;\&\; \neg x) = \nu(\&M) - \nu(\&M \;\&\; x) = \nu(\&M);$$

Пусть $R = (\&M \Rightarrow \neg x)$. Из соотношений выше нетрудно вычислить вероятность правила R:

$$\eta(R) = \frac{\nu(\&M \;\&\; \neg x)}{\nu(\&M)} = 1;$$

Лемма 5 утверждает, что найдется правило $S \in$ MSR $\subset \Pi$, являющееся МСЗ, имеющее в заключении $\neg x$, притом для S выполнено $\nu(S) \geq 1$. Таким образом, правило S неизбежно добавит $\neg x$ в непосредственное предсказание Pr(L). □

Теорема 7. *Рассмотрим $M \underset{\nu}{\subseteq} L$, $N \underset{\nu}{\subseteq} L$ и $M \neq N$. Тогда*

1. $\exists x \;:\; x \in \text{PR}(M)$ *и* $\neg x \in \text{PR}(N)$*;*

2. $\text{PR}(M) \supseteq \text{PR}(M \cap N) \subseteq \text{PR}(N)$ *и* $\text{PR}(M) \neq \text{PR}(N)$*.*

■ 1. $M \neq N$ означает, что $\exists x \in M \setminus N$ (действительно, $M \subset N$ было бы противоречащим максимальности M). $x \in M \Rightarrow \text{PR}(M)$ и аналогично теореме 18, $\neg x \in \text{PR}(N)$.

2. $\text{PR}(N)$ – совместно и непротиворечиво, а $\neg x \in \text{PR}(N)$; значит $x \notin \text{PR}(N)$ и $x \in \text{PR}(M) \setminus \text{PR}(N)$. Далее, $M \cap N \subset M$, поэтому $\text{PR}(M \cap N) \subseteq \text{PR}(M)$. □

Последние две теоремы заключают, что существует взаимно-однозначное соответствие между ν-максимальными подмножествами в L и неподвижными точками, полностью покрывающими весь набор атомов из L (содержащие их или их отрицания).

Для совместных множеств непротиворечивость и совместность неподвижных точек доказана разделом выше. Для несовместных множеств ответ дает следующая теорема.

Теорема 8. *Если L – несовместно, то* $\text{PR}(L)$ *– противоречиво.*

■ Найдем ν-максимальное подмножество в L и обозначим его M. $M \neq L$, иначе L оказалось бы совместным. Поэтому найдется $x \in L \setminus M$. Множество $\{x\}$ расширим до максимального совместного $N \underset{\nu}{\subseteq} L$. По построению $x \in N \setminus M \Rightarrow M \neq N$. По теореме 19, найдется y такое что $y \in \text{PR}(M)$ и $\neg y \in \text{PR}(N)$:

$$\left.\begin{matrix} M \subseteq L \\ N \subseteq L \end{matrix}\right\} \Rightarrow \left.\begin{matrix} y \in \text{PR}(M) \subseteq \text{PR}(L) \\ \neg y \in \text{PR}(N) \subseteq \text{PR}(L) \end{matrix}\right\} \Rightarrow \text{PR}(L) \text{ - противоречиво. } \square$$

12 Вероятностные формальные понятия

Неподвижные точки оператора PR оказываются во многом замечательными. Однако целью их рассмотрения служил мотив введения аналогов формальных понятий, но уже в вероятностном смысле. Пользуясь идеей теоремы 3, нетрудно предложить в качестве кандидатов [Demin A., Ponomaryov D., Vityaev E., 2012] на роль содержания таких понятий как раз неподвижные точки PR.

Выбор кандидатов на включение в объем чуть более сложен. Но поскольку все множества литер, такие что $\text{PR}(M) = B$, имеют реальное отношение к замыканию, логично предложить собрать все попадающие под них объекты. То есть:

Определение 40. *Вероятностным формальным понятием на K назовем (A, B), такую что:*

$$\text{PR}(B) = B,\ A = \bigcup_{\text{PR}(C) = B} G_C$$

Чтобы отличать вероятностные понятия от обычных в смысле контекста K^*, последние мы будем называть строгими формальными понятиями. Наш выбор оправдывает следующее утверждение, связывающее вероятностные и строгие формальные понятия на одном и том же контексте.

Теорема 9. *Пусть K – бинарный контекст.*

1. *Если (A, B) - строгое понятие на K, то существует вероятностое понятие (N, M) такое, что $A \subseteq N$, а $B \subseteq M$.*

2. *Если (N, M) - вероятностое понятие на K, то найдется набор строгих понятий $\mathcal{C}$, таких что*

$$\forall (A, B) \in \mathcal{C} \ (\mathsf{PR}(B) = M),$$
$$N = \bigcup_{(A,B)\in\mathcal{C}} A.$$

■ Положим $\mathcal{S} = \{S \mid \mathsf{PR}(S) = M\}$.

1. Пусть $M = \mathsf{PR}(B)$. Тогда $B \in S$ и $A = G_B \subseteq \bigcup\limits_{\mathcal{S}} G_S = N$. Поэтому (A, B) – искомое.

2. По $\mathcal{S}$ построим множество строгих понятий $\mathcal{C} = \{(S''', S'') \mid S \in \mathcal{S}\}$. Из леммы 1 нетрудно понять, что $B''' = B'$, то есть $\mathcal{C} = \{(S', S'') \mid S \in \mathcal{S}\}$ и все $(A, B) \in \mathcal{C}$ - строгие понятия. Отсюда

$$N = \bigcup_{\mathcal{S}} S' = \bigcup_{(A,B)\in\mathcal{C}} A$$

Остается добавить, что $M \in \mathcal{S}$ и следовательно $\mathcal{C} \neq \varnothing$. □

Вероятностное понятие являются некоторым кластером, объединяющим набор слабо различимых с точки зрения набора правил Π строгих понятий.

13 Поиск вероятностых понятий

В этом разделе мы ограничимся случаем конечных контекстов K. Из последних не умаляя общности можно выкинуть статистически незначимые объекты.

Предположим, что набор правил Π на контексте K уже найден одним из алгоритмов, например алгоритмами из [Vityaev E.E., 1992; Kovalerchuk B., Vityaev E., 2000]. Из определения вероятностного понятия вытекает следующая процедура поиска.

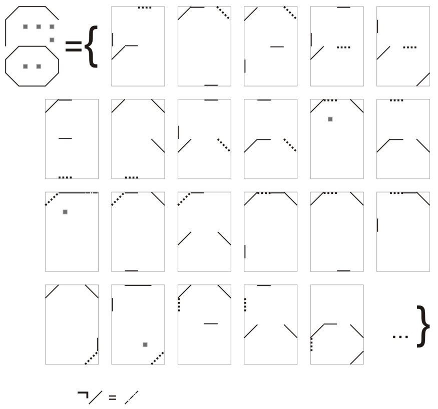
1. На шаге $k = 1$ генерируется множество $C^{(1)} = \{\mathsf{PR}(\mathrm{head}(R)) \mid R \in \Pi\}$.

2. На шаге $k > 1$ в случае если $C^{(k-1)} = \varnothing$ алгоритм завершает свое выполнение и на выход подается список обнаруженных вероятностных понятий.

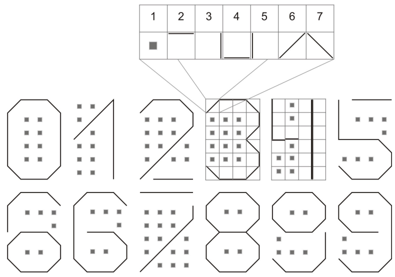
3. Иначе на шаге $k > 1$ для каждого $B \in C^{(k-1)}$ рассматриваем семейство импликаций $L_B = \{A \Rightarrow m \in L \mid A \subseteq B\}$, вычисляем множество $A = \{g \in G \mid \mathsf{PR}(g' \cap B) = B\}$. Если $A \neq \varnothing$, то пара (A, B) добавляется в список найденных понятий.

4. Генерируется множество $C^{(k)} = \{\mathsf{PR}(B \cup C) \mid B, C \in C^{(k-1)}, \mathsf{PR}(B \cup C) \notin C^{(k-1)}\}$

5. Полагаем $k := k + 1$ и переходим к пункту 2.



(a) Неподвижная точка для цифры 6



(b) Кодирование цифр

Приведем один из многочисленных примеров. За дополнительными идеями можно обратиться, например, к [Vityaev E., Lapardin K., Khomicheva I., Proskura A., 2014]. В работе [Neupokoev N., Vityaev E., 2014] в качестве данных были взяты 12 цифр, используемых обозначения индекса на почтовых конвертах (по 2 варианта для «6» и «9»). Контекст основывался на 24 атрибутах, каждый из которых имел по 7 значений (для различных начертаний в соответствующем секторе разбиения), в каждой цифре отсутствовал один из признаков цифры. Множество G состояло из 360 цифр (по 30 копий каждой цифры, в каждой из которых был случайно удален какого-то признак) плюс отрицательная выборка, состоящая из 1050 объектов со случайными признаками. На этих данных было обнаружено 73458 закономерностей. по которым были обнаружены неподвижные точки, которых оказалось 14. Из них 12 цифр – в точности цифры, а для цифр 6 и 9 еще по одной неподвижной точке, содержащих в признаках дополнительный пробел, отличающий 2 варианта цифр.

14 Заключение

Введя в формальный контекст отрицания (и вообще значения для атрибутов) мы получаем гораздо более выразительную систему описаний. Заметим, что это обеспечивает предлагаемому метода наличие (в некотором смысле) таких свойств как корректность и полнота (теоремы 14 и 18). Более того, оказывается возможной классическая логика предикатов. Это направление, обращенное на изучение синтеза вероятности, логики и контекста, заслуживает дополнительного внимания.

Другой достаточно интересной и слабо изученной конструкцией являются правила на контекстах. Правила способны хранить в себе знания о взаимосвязях объектов и признаках, что делает их пригодными для наглядного описания любых процессов, формализуемых в виде матриц объект-признак. Концепция неподвижных точек предсказания оказывается вполне естественной попыткой развить формалную среду для проведения рассуждений в стиле классификации объектов [Vityaev E., Lapardin K., Khomicheva I., Proskura A., 2014; Neupokoev N., Vityaev E., 2014]. Поэтому область применения предлагаемого метода не заканчивается на приведенных примерах: здесь мы говорим о причастности метода к направлению Data Mining.

Алгоритм позволяет находить вероятностные понятия, извлекая универсальные знания, и в то же время не теряя строгих понятий. Наложение двоичных шумов на значения атрибутов оставляет набор понятий неизменным [Demin A., Ponomaryov D., Vityaev E., 2012; Neupokoev N., Vityaev E., 2014].

15 Благодарности

Эта работа выполнена при финансовой поддержке Российского Гуманитарного научного фонда (проект 12-01-12026), интеграционными проектами СО РАН 3, 87, 136 и Советом по грантам Президента РФ для государственной поддержки ведущих научных школ (проект НШ-860.2014.1).

Литература

Ganter B. Formal Concept Analysis: Methods, and Applications in Computer Science. – TU Dresden, 2003.

Ganter B., Wille R. Formal concept analysis – Mathematical Foundations. – Berlin-Heidelberg-New York: Springer, 1999.

Demin A., Ponomaryov D., Vityaev E. Probabilistic concepts in formal contexts. // Proc. ARCOE Workshop at IJCAI'11. – Barselona, 2011: 31-35.

Demin A.V., Ponomaryov D. K., Vityaev E.E. Probabilistic Generalization of Formal Concepts. // Programming and Computer Software. – 2012 – 38(5): 219-230.

Vityaev E. Probabilistic fix-points determination. // International conference on Mal'tsev readings (12-16 november, 2012). – Novosibirsk (2012): 21 / `http://www.math.nsc.ru/conference/malmeet/12/malmeet_2012.pdf`

Ganter B., Obiedkov S. Implications in Triadic Formal Contexts. – TU Dresden: Springer, 2004.

Missaoui R., Kwuida L. Implications in Triadic Formal Contexts. // 9th International Conference, ICFCA 2011. – Nicosia, Cyprus: Springer, 2011.

Kovalerchuk B., Vityaev E. Empirical Theories Discovery based on the Measurement Theory. // Mind and Machine. – 2004 – 14(4): 551-573

Vityaev E.E. The logic of prediction. In: Mathematical Logic in Asia. // Proceedings of the 9th Asian Logic Conference, World Scientific. – Singapore, 2006: 263-276.

Smerdov S., Vityaev E. New defention of prediction without logical inference. // Proceedings of the IASTED International Conference, CI 2009, USA, 2009.

Smerdov S., Vityaev E. On the Problem of Prediction. // KONT/KPP 2007, LNAI 6581 Heidelberg: Springer, 2011: 280-296.

Vityaev E.E. Semantic approach to knowledge base development: Semantic probabilistic inference. // Comp. Syst. 146, 1992: 19-49 (in Russian).

Kovalerchuk B., Vityaev E. Data Mining in Finance: Advances in Relational and Hybrid methods. // Kluwer Academic Publishers, 2000.

Speransky S.O. On the logical consistency of probabilistic predictions // Messenger of NGU: mathematics, mechanics, informatics. – 2011 – 11(1): 99-115 (in Russian).

Speransky S.O. Logic of probability and probability logic. – Novosibirsk State University, Novosibirsk, PhD thesis, 2013 (in Russian).

Vityaev E., Lapardin K., Khomicheva I., Proskura A. Transcription factor binding site recognition by regularity matrices based on the natural classification method. // Machine learning and bioinformatics, IOS Press, 2008. – v. 12(5): 495-512.

Neupokoev N.V., Vityaev E.E. Formal model of perception and pattern as fix point of anticipations // Approaches to the thinking modeling. – Moscow: URSS Editorial, 2014: 155-172 (in Russian).