

# СИНТЕЗ ЛОГИКИ, ВЕРОЯТНОСТИ И ОБУЧЕНИЯ В СЕМАНТИЧЕСКОМ ВЕРОЯТНОСТНОМ ВЫВОДЕ<sup>1</sup>

Витяев Е.Е.,

*к.т.н., с.н.с.*

*Институт математики им. С. Л. Соболева СО РАН,*

*тел.: 336-13-93*

*факс: 333-25-98*

[vityaev@math.nsc.ru](mailto:vityaev@math.nsc.ru)

**Введение.** Предсказание является одним из важнейших понятий в науке. В настоящее время все больше возникает необходимость осуществлять предсказания на основании индуктивно выведенных знаний. Обнаружением таких знаний занимаются в интенсивно развиваемых направлениях Machine Learning и Knowledge discovery in Data Bases and Data Mining (KDD&DM). Частным случаем индуктивно выведенных знаний являются вероятностные законы. Предсказания, получаемые по вероятностным законам, описываются индуктивно-статистическим выводом I-S-выводом (Inductive-Statistical inference).

Гемпелем [1] отмечал, что предсказания, получаемые по I-S выводу статистически двусмысленны. Во избежание двусмысленности он ввел требование максимальной специфичности (Requirement of Maximum Specificity). Однако, это требование не решает полностью проблемы статистической двусмысленности. Статистическая двусмысленность приводит к необходимости разрабатывать логики, работающие с противоречивыми знаниями, например, логики по умолчанию (default logic [2]) и паранепротиворечивые логики.

Другой проблемой, существующей для I-S выводов, является несогласованность вероятностных оценок с логическим выводом. Известно, что вероятностные оценки высказываний резко падают в процессе логического вывода и их нельзя улучшить. Вычислению оценок посвящены работы по вероятностной логике [3–8]. Есть работы, в которых вероятность рассматри-

---

<sup>1</sup> Интеграционными проектами СО РАН №1, №115, Программой президента Российской Федерации поддержки научных школ 4413.2006.1

вается как значение истинности утверждений, а процесс логического вывода обобщается до так называемых “количественных дедукций” [3, 8].

В настоящее время эта проблема рассматривается как проблема синтеза логики и вероятности. В 2002 в королевском колледже Лондона был проведен workshop “Combining Probability and Logic”. В аннотации к нему говорится: “Artificial intelligence is one key discipline in which probability theory competes with other logics for application. It is becoming vitally important to evaluate and integrate systems that are based on very different approaches to reasoning, and there is strong demand for theoretical understanding of the relationships between these approaches”. Во введении к спецвыпуску «Special issue on Combining Probability and Logic» журнала *Journal of Applied Logic* (2003), посвященному этому совещанию, редакторами Jon Williamson и Dov Gabbay отмечается, что существующие подходы к синтезу логики и вероятности можно разбить на две части: “One approach is to argue that *probability is logic*, which requires showing that probability is a determinate relation between statements. Kyburg, Howson and Paris and Vencovská appeal to the concepts of frequency, consistency and entropy respectively to determine this relation. Alternatively one can explore other formalisms which *interface between probability and logic*: argumentation in the case of Fox and Kohlas; default reasoning in the case of Bourne and Weydert.” Однако, эти исследования, с нашей точки зрения не решают проблему синтеза логики и вероятности.

Первый шаг к синтезу логики и вероятности был сделан в “количественных дедукциях”, где значения истинности были обобщены до значений вероятности. Но в количественных дедукциях сохраняется очевидное несоответствие: при обобщении значений истинности не обобщается правило вывода. Правило вывода предназначено для сохранения значений истинности, но, если значения истинности обобщены, то и правило должно быть обобщено, чтобы сохранить эти обобщенные значения.

Рассмотрим процесс вывода с точки зрения “семантического” подхода к программированию [9]. Идея семантического программирования состоит в том, чтобы процесс вывода рассматривать с семантической точки зрения как вычисление истинности утверждений на модели. При таком взгляде на вывод его можно обобщить, определяя новые взаимоотношения высказываний и модели. Можно рассмотреть вывод не только как проверку истинности на модели, но и как поиск фактов в модели, предсказывающих интересное нас высказывание с максимальной вероятностью, или как поиск

наиболее подтверждающих фактов, и т.д. Такой вывод будем называть семантическим. Если значения истинности имеет континуум значений, то поиск наиболее вероятного, достоверного и т.д. утверждения имеет определенный смысл.

В работах [10-12] определен *семантический вероятностный вывод*, основанный на идее семантического подхода к программированию. В семантическом вероятностном выводе вероятностные оценки высказываний не падают в процессе вывода, а строго возрастают.

Проиллюстрируем основную идею определения семантического вероятностного вывода. Для этого рассмотрим сначала процесс получения предсказания с помощью I-S вывода в рамках логического программирования. В логическом программировании вывод предсказания можно рассматривать как вычисление. Предсказание в логическом программировании формулируется как запрос  $G$  к множеству законов  $L_1, \dots, L_m$  вида  $A_1^i \& \dots \& A_{k_i}^i \Rightarrow G$  и фактам  $C_1, \dots, C_n$  представленных правилами  $(\Rightarrow C_1), \dots, (\Rightarrow C_n)$ .

В процессе вычисления ответа на запрос  $G(x_1, \dots, x_n)$  находятся:

1. вывод  $\{L_1, \dots, L_m, C_1, \dots, C_n\} \vdash \exists x_1, \dots, x_n G$ ;
2. набор термов  $t_1, \dots, t_n \{L_1, \dots, L_m, C_1, \dots, C_n\} \vdash G[x_1/t_1, \dots, x_n/t_n]$ .

Процесс вычисления (вывода) предсказания для хорновых формул можно представить в виде дерева рис. 1. В нём запрос  $G$  унифицируется с законами  $L_1, L_2$  так, чтобы заключение правила совпало с запросом  $G$ . В этом случае запрос выводится из истинности атомов посылки  $A_1^1 \& \dots \& A_i \dots \& A_{k_1}^1$  или посылки  $A_1^2 \& \dots \& A_j \dots \& A_{k_2}^2$ . Если среди законов  $L_1, \dots, L_m$  есть правила  $L_3, L_4, L_5$ , которые унифицируются с некоторыми атомами  $A_i$  или  $A_j$  правил  $L_1, L_2$ , то посылки этих правил –  $B_1^1 \& \dots \& B_{n_1}^1 / B_1^2 \& \dots \& B_{n_2}^2 / B_1^3 \& \dots \& B_{n_3}^3$  подставляются вместо соответствующих атомов  $A_i / A_j$ . Если какие-то правила  $L_3, L_4$  или  $L_5$  являются фактами вида  $\Rightarrow A_i$ , то соответствующий атом после унификации удаляется из посылки правила. Вывод (вычисление) заканчивается, когда найдена такая ветвь дерева вывода, которая содержит правило  $(G \Leftarrow)$ .

Семантический подход к логическому программированию состоит в рассмотрении теоретико-модельной семантики логических программ, когда факты являются высказываниями некоторой модели. В этом случае вывод (вычисление) предсказания состоит в нахождении фактов в модели предсказывающих выводимое высказывание. Такой процесс вывода можно рас-

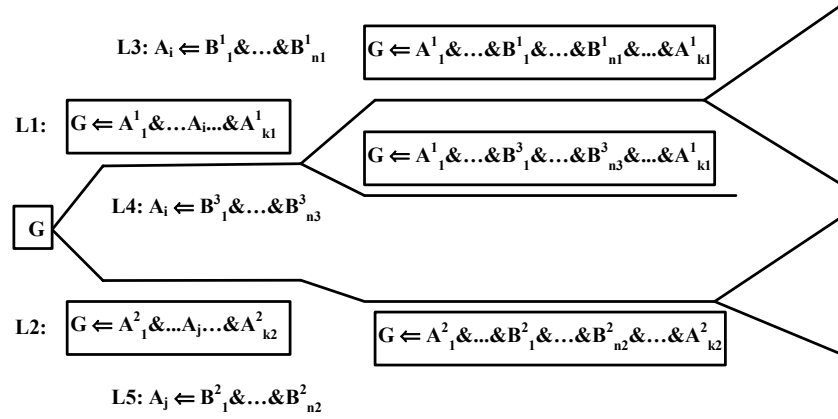


Рис. 1

смаивать как вычисление истинности предсказываемого факта  $G$  на эмпирической системе.

Определим семантический вероятностный вывод. Будем предполагать, что определена вероятность для высказываний языка первого порядка [6].

**Определение** [10-13]. *Подправилом* некоторого правила  $C = (A_1 \& \dots \& A_k \Rightarrow A_0)$  вида будем называть логически более сильное правило  $(A_{i1} \& \dots \& A_{ih} \Rightarrow A_0)$ ,  $\{A_{i1}, \dots, A_{ih}\} \subset \{A_1, \dots, A_k\}$ ,  $0 \leq h < k$ .

**Определение** [10-13]. *Вероятностным законом* будем называть правило  $(A_1 \& \dots \& A_k \Rightarrow A_0)$  условная вероятность  $\eta(A_0 \& A_1 \& \dots \& A_k) / \eta(A_1 \& \dots \& A_k)$  которого определена ( $\eta(A_1 \& \dots \& A_k) > 0$ ) и строго больше условных вероятностей каждого из его подправил.

**Определение** [10-12]. *Семантическим вероятностным выводом* вероятностного закона  $C_n$  будем называть такую последовательность вероятностных законов  $C_1 \sqsubset C_2 \sqsubset \dots \sqsubset C_n$ ,  $C_i = (A_1^i \& \dots \& A_{k_i}^i \Rightarrow G)$ , что правило  $C_i$  является подправилом правила  $C_{i+1}$  и  $\eta(C_i) < \eta(C_{i+1})$ ,  $i = 1, 2, \dots, n-1$ ,  $C_n$  – правило не являющееся подправилом никакого другого вероятностного закона.

Рассмотрим множество всех семантических вероятностных выводов факта  $G$ . Его можно представить семантическим вероятностным деревом вывода Рис 2. Сравнение рис. 1 и рис. 2 показывает, что по структуре семантический вероятностный вывод аналогичен выводу предсказания в ло-

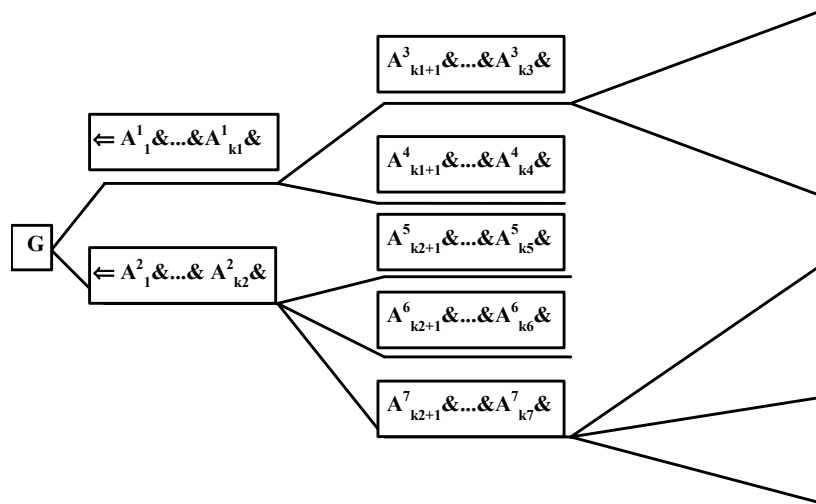


Рис 2. Дерево семантического вероятностного вывода  
гическом программировании за исключением того, что для семантического вероятностного вывода не нужны правила L1, L2, L3, L4, L5 и, значит, не нужен логический вывод, представленный на рис. 1. Единственно, что нужно – это уточнение посылки правил путем добавления дополнительных условий в посылку так, чтобы оценка вероятности предсказания атома G строго увеличивалась. Поскольку вероятность является числом, то для увеличения оценки вероятности не нужен логический вывод, достаточно обнаруживать факты, строго увеличивающие вероятность предсказания атома G.

**Определение** [11]. *Максимально специфическим законом*  $MC3(G)$  для вывода факта G будем называть вероятностный закон, принадлежащий семантическому вероятностному дереву вывода факта G, имеющий максимальное значение условной вероятности среди всех правил дерева. Множество всех максимально специфических законов  $MC3(G)$  для всех атомов G обозначим через MC3.

**Теорема** [11]. I-S вывод непротиворечив для любой теории  $T \subset MC3$ .

Данная теорема решает проблему статистической двусмысленности и позволяет осуществлять предсказания без противоречий, используя максимально специфические правила.

Семантический вероятностный вывод решает также проблему синтеза логики, вероятности и обучения, обсуждаемую в широко цитируемой работе L.De Raedt and K.Kersting «Probabilistic logic learning» [14]. В ней говорится, что «One of the central open questions in data mining and artificial intelligence, concerns probabilistic logic learning, i.e. the integration of relational or logical representations, probabilistic reasoning mechanisms with machine learning and data mining principals». Синтез логики, вероятности и обучения в семантическом вероятностном выводе осуществляется следующим образом. Процесс вывода-вычисления позволяет обнаружить максимально специфические правила, содержащие максимум информации для получения предсказаний. Процесс вывода-вычисления совпадает с целью индуктивного вывода знаний – получения максимально точных обобщений фактов.

Семантический вероятностный вывод позволил решить не только перечисленные выше проблемы, но также позволил получить следующие результаты.

1. Нами разработан реляционный подход к методам KDD&DM, опубликованный в монографии [15], в главах в монографиях [16-17] и в серии статей (более подробная информация представлена на сайте [18]). В этом подходе на основании семантического вероятностного вывода была разработана программная система Discovery, обнаруживающая множество законов, вероятностных законов и максимально специфических правил для некоторой эмпирической системы М. Система Discovery успешно применялась для решения большого числа практических задач и давала лучшие результаты, чем другие методы (подробности сравнения приведены на сайте [18]).

2. Семантический вероятностный вывод был положен в основу формальной модели нейрона и некоторых моделей когнитивных процессов. С точки зрения когнитивных процессов мозг – это предсказывающее устройство, а не логическое. Работа мозга представлена в работах [19-22] как непрерывное во времени предсказание мозгом событий окружающей среды с одновременным контролем акцептором результатов действия правильности полученных предсказаний. В этих же работах приведено объяснение теории функциональных систем работы мозга П.К.Анохина [23] и информационной теории эмоций П.В.Симонова [24]. В работах [25-27] представлены схемы, описывающие эти теории, а также теорию движений Н.А.Берштейна [28] и принятия решений. В работе [29] проведены машинные эксперименты по моделированию анимата (искусственного автономного интеллектуального

агента) на основе разработанных моделей. Проведено сравнение предложенной модели с существующими подходами. Показано, что анимат, основанный на семантическом вероятностном выводе, обучается намного эффективнее, чем аниматы, основанные на нейронных сетях и Reinforcement Learning.

## ЛИТЕРАТУРА

1. *Hempel, C. G.* 'Maximal Specificity and Lawlikeness in Probabilistic Explanation', *Philosophy of Science* **35**, 1968. – P. 16–33.
2. *Yao-Hua Tan.* Is default logic a reinvention of inductive-statistical reasoning? *Synthese*, 110: 357–379, 1997. Kluwer Academic Publishers.
3. *R.T.Ng, V.S.Subrahmanian.* Probabilistic reasoning in Logic Programming // Proc. 5th Symposium on Methodologies for Intelligent Systems, Knoxville, North-Holland, 1990. – P. 9-16.
4. *H.Gaifman.* Concerning measure in first order calculi // *Israel journal of Math.* v.2, N1, 1964 – P. 1–18.
5. *Nils J. Nilsson.* Probability logic // *Artif. Intell.*, v.28, N1, 1986. – P. 71-87.
6. *Halpern, J.Y.* An analysis of first-order logic of probability. *Artificial Intelligence*. v. 46, 1990. – P. 311-350.
7. *D.S.Scott, P.Krauss.* Assigning Probabilities to Logical Formulas // *Aspects of Inductive Logic* / eds. J.Hintikka, P.Suppes, N.Holland, 1966. – P. 219-264.
8. *M.N.Van Emden.* Quantitative deduction and its fixpoint theory // *J. Logic Programming*, v.3, N.1. 1986, – P. 37-53.
9. *S.S.Goncharov, Yu.L.Ershov, D.I.Sviridenko.* Semantic programming // 10th World Congress Information Processing 86, Dublin, Oct.,1986. - Amsterdam, 1986. – P. 1093-1100.
10. *Вутяев Е.Е.* Семантический подход к созданию баз знаний. Семантический вероятностный вывод наилучших для предсказания ПРОЛОГ-программ по вероятностной модели данных. // *Логика и семантическое программирование* (Выч. сист., вып. 146). Новосибирск, 1992. – С. 19-49.
11. *Evgenii Vityaev.* The logic of prediction. In: *Mathematical Logic in Asia. Proceedings of the 9th Asian Logic Conference* (August 16-19, 2005, Novosibirsk, Russia), World Scientific, Singapore, 2006. – P. 263-276
12. *Вутяев Е.Е.* Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов: Моногр. // НГУ. Новосибирск, 2006. 293 с.
13. *Evgenii Vityaev, Boris Kovalerchuk.* Empirical Theories Discovery based on the Measurement Theory. *Mind and Machine*, v.14, #4, 2004. – P. 551-573
14. *De Raedt L., Kersting K.* Probabilistic logic learning // *ACM-SIGKDD Explorations* / special issue on Multi-Relational Data Mining. Vol. 5(1). 2004. – P. 31–48, July.

15. Kovalerchuk, B., Vityaev, E. Data Mining in finance: Advances in Relational and Hybrid Methods, Kluwer Academic Publishers, 2000. – P. 308.
16. Kovalerchuk, B., Vityaev, E., Ruiz, J.F. Consistent and Complete Data and 'Expert' Mining in Medicine // *Medical Data Mining and Knowledge Discovery*, Springer, 2001. – P. 238-280.
17. Evgenii Vityaev, Boris Kovalerchuk. Data Mining For Financial Applications. In: O. Maimon and L. Rokach (eds.), *Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers*, Springer 2005, – P. 1203-1224.
18. Scientific Discovery. <http://www.math.nsc.ru/AP/ScientificDiscovery>
19. Витяев Е.Е. Принципы работы мозга. 2003.  
<http://wsni2003.narod.ru/Papers/Vityaev.htm>
20. Витяев Е.Е. Целеполагание как принцип работы мозга // *Модели когнитивных процессов* (Выч. системы, 158), Новосибирск, 1997. – С. 9-52.
21. Витяев Е.Е. Вероятностное прогнозирование и предсказание как принцип работы мозга // *Измерение и модели когнитивных процессов* (Выч. системы, 162), Новосибирск, 1998. – С. 14-40.
22. Витяев Е.Е. Формальная модель работы мозга, основанная на принципе предсказания // *Модели Когнитивных Процессов*. (Выч. сист., 164), Новосибирск, 1998. – С. 3-61
23. Анохин П.К. Принципиальные вопросы теории функциональных систем // *Философские аспекты теории функциональных систем*. М. Наука, 1978. с. 49 - 106.
24. Симонов П.В. Эмоциональный мозг. М.: Наука, 1981. – С. 140.
25. Михиенко Е.В., Витяев Е.Е. Моделирование работы функциональной системы, VI Всероссийская научно-техническая конференция «Нейроинформатика-2004». Сборник научных трудов. В 2-х частях. ч.2., М. МИФИ, 2004. с. 124-129.
26. Витяев Е.Е. Объяснение Теории Движений Н.А.Бернштейна. VII Всероссийская научно-техническая конференция «Нейроинформатика-2005», ч.1. Москва, 2005. с. 234-240
27. Витяев Е.Е. Принятие решений. Переключающая и подкрепляющая функции эмоций // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006», Сборник научных трудов, Москва, 2006. – С. 24-30
28. Бернштейн Н.А. Биомеханика и физиология движений. // *Избранные психологические труды*, Москва-Воронеж, 1997. – С. 605
29. Демин А.В., Витяев Е.Е. Реализация модели анимата на основе семантического вероятностного вывода // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006», ч.2, Москва, 2006. – С. 16-24