

**Е.Е. ВИТЯЕВ^{1,2}, Л.И. ПЕРЛОВСКИЙ³,
Б.Я. КОВАЛЕРЧУК⁴, С.О. СПЕРАНСКИЙ²**

¹*Институт математики им. С.Л. Соболева СО РАН, Новосибирск
vityaev@math.nsc.ru*

²*Новосибирский государственный университет
netid@ya.ru*

³*Harvard University, Air Force Research Laboratory, USA
leonid@seas.harvard.edu*

⁴*Central Washington University, Ellensburg, USA
borisk@cwu.edu*

ВЕРОЯТНОСТНАЯ ДИНАМИЧЕСКАЯ ЛОГИКА МЫШЛЕНИЯ И ПОЗНАНИЯ

Ранее был разработан оригинальный подход к моделированию мышления, основанный на теории нейронных моделирующих полей и динамической логики. Этот подход основан на анализе неадекватности формальной логики и проблеме комбинаторной сложности. В данной работе мы интерпретируем теорию нейронных моделирующих полей и динамическую логику в логико-вероятностных терминах и показываем, как в этом случае формулируются и решаются проблемы моделирования мышления в искусственном интеллекте (ИИ).

Ключевые слова: логика, мышление, познание, обучение, модель нейрона, вероятностное прогнозирование, предсказание

Введение

Л.И. Перловским разработан оригинальный подход к моделированию мышления, основанный на теории нейронных моделирующих полей и динамической логики [5,14,15]. Этот подход, с одной стороны, основан на детальном анализе проблем моделирования мышления в Искусственном Интеллекте – неадекватности формальной логики и проблеме комбинаторной сложности, а, с другой стороны, на данных психологии, философии и когнитивной науки об основных механизмах мышления. Прделанный анализ проблем моделирования мышления имеет, на самом деле, более широкое значение и преодоление этих проблем может привести к другим формализациям процесса мышления. С этой целью в работе [11] получено обобщение теории нейронных моделирующих полей и динамиче-

ской логики в виде динамической логики мышления и когнитивной динамической логики. Эти логики сформулированы в наиболее общих терминах: *отношения общности, неопределенности, простоты; проблема максимизации сходства с эмпирическим содержанием; метод обучения*.

В данной работе мы интерпретируем эти понятия в терминах логики и вероятности: неопределенность мы интерпретируем как вероятность, а процесс обучения как семантический вероятностный вывод [2, 7, 20, 21]. Полученная в результате Вероятностная Динамическая Логика Мышления и Познания принадлежит уже к области вероятностных моделей мышления и познания [17, 19]. Мы показываем, что данная логика также, но по-своему, решает проблемы моделирования мышления – неадекватности формальной логики и проблему комбинаторной сложности. Таким образом, через обобщение, полученное в работе [11], мы расширяем интерпретацию теории нейронных моделирующих полей и динамической логики на вероятностные модели мышления и познания. Вероятностная динамическая логика уже была ранее применена для моделирования работы мозга и когнитивных процессов [2, 3, 4].

Проблемы моделирования мышления с вероятностной точки зрения

Повторим и дополним изложение проблем моделирования мышления в Искусственном Интеллекте, изложенную в [11]. Основатели искусственного интеллекта в 1950-х и 1960-х годах верили, что, основываясь на правилах логики, они вскоре создадут компьютеры, чей интеллект намного превзойдет человеческий [11].

Но вскоре стало ясно, что логика не работает. На самом деле мозг – не логическое, а предсказывающее устройство, действующее в соответствии с принципом опережающего отражения действительности П.К. Анохина [1]. Но, как выясняется, подходящего определения предсказания для индуктивно выведенных знаний до настоящего времени не существует.

Общепринятое определение предсказания принадлежит Карлу Попперу и состоит в том, что для предсказания некоторого факта необходимо логически вывести его из имеющихся фактов и теории (дедуктивно-номологический вывод). Но это определение не работает для индуктивно выведенных знаний, имеющих некоторую оценку вероятности, подтвержденности и т.д. При логическом выводе предсказаний необходимо, в то же время, получить оценку вероятности, подтвержденности и т.д. полученного предсказания. В случае вероятностей этим занимается вероятностная логика. Но оказывается, что оценки предсказаний резко падают в

процессе их вычисления вслед за логическим выводом и оценка предсказания может оказаться нулевой. Предсказания с нулевыми оценками не являются предсказаниями.

Эта проблема в настоящее время осознана как проблема синтеза логики и вероятности. Прошло уже 4 симпозиума под общим названием Projic (Probability+Logic), в 2002–2009 гг.. Во введении к симпозиуму 2002 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". Однако решение проблемы до сих пор не найдено, т.к. от логического вывода никто не отказывается, а в этом случае, с нашей точки зрения, адекватного определения предсказания получить нельзя.

Нами введено новое понятие предсказания, изложенное в работах [2, 20, 21], полученное путем отказа от логического вывода и значений истинности "истина" и "ложь". Вместо логического вывода нами определен семантический вероятностный вывод, используемый далее для определения оператора обучения.

Другой проблемой моделирования мышления в Искусственном Интеллекте является проблема Комбинаторной Сложности (КС) [11]. В процессе восприятия мышление ассоциирует подмножества сигналов, соответствующих объектам, с представлениями об этих объектах. Математическое описание этого, казалось бы, простого шага – процесса ассоциации-распознавания-понимания оказалось далеко не простым делом, и это связано с понятием комбинаторной сложности (КС) [16].

Последовавшие исследования обнаружили связь КС с логикой в различных алгоритмах [16]. Логика рассматривает каждое даже небольшое изменение в данных или моделях, как новое высказывание (гипотезу). Это приводит к КС. В работе [9] доказывается, что даже нахождение простейшей совокупности высказываний, которой описываются решающие деревья, NP-трудна.

Приписывание значений истинности "истина" и "ложь" не позволяет сравнивать высказывания. В данной работе мы используем два отношения упорядочения на высказываниях: отношение общности и сравнение по условной вероятности, которые используются далее в семантическом вероятностном выводе. Это принципиально сокращает перебор и, наряду с применением статистических оценок, делает его практически приемлемым и решает проблему КС.

Напомним и дополним основные определения, связанные с моделированием мышления [11]. Мы принимаем, что основные механизмы мышления включают инстинкты, концепции, эмоции и поведение.

Из механизмов мышления концепции-модели наиболее доступны сознанию. Рэй Джакендофф [10] считает, что наиболее адекватный термин для механизма концепций – это модель, или внутренняя модель мышления. Концепции – это модели в буквальном смысле. Они моделируют в нашем мышлении объекты и ситуации в мире. Мышление включает иерархию многих уровней концепций-моделей, от простейших элементов восприятия (линии, движущиеся точки) до концепций-моделей объектов, отношений между объектами, до сложных ситуаций.

Фундаментальная роль эмоций в мышлении состоит в том, что они связаны с инстинктом к знанию – максимизацией меры близости между концепциями-моделями и миром [11]. Этот эмоциональный механизм оказался принципиально важен для того, чтобы «разорвать замкнутый круг» комбинаторной сложности. В процессе обучения и понимания входных сигналов, модели адаптируются так, чтобы лучше представлять входные сигналы, и чтобы схожесть между ними увеличивалась. Это увеличение схожести удовлетворяет инстинкт к знанию и ощущается как эстетическая эмоция.

Экспериментальные подтверждения связи эмоций с инстинктом к знанию можно найти в Информационной теории эмоций П.В.Симонова [6]: "Суммируя результаты собственных опытов и данные литературы, мы пришли ... к выводу о том, что эмоция есть отражение мозгом человека и животных какой-либо актуальной потребности (её качества и величины) и вероятности (возможности) её удовлетворения, которую мозг оценивает на основе генетического и ранее приобретенного индивидуального опыта...". "Удовольствие всегда есть результат уже происходящего (контактного) взаимодействия (удовлетворения потребности – Е.Е.), в то время как радость (эмоция – Е.Е.) есть ожидание удовольствия в связи с растущей вероятностью удовлетворения потребности".

Следующий эксперимент показывает, что инстинкт к знанию вызывает положительные эмоции [6]: «В наших опытах на экране, установленном перед испытуемым, проецировались наборы из пяти цифр – единиц и нулей. Испытуемого предупреждали, что некоторые из кадров, содержащие общий признак (например, два нуля подряд 00), будут сопровождаться гудком. Задача испытуемого состояла в обнаружении этого общего признака. До возникновения первой (как правило, ошибочной, например 01) гипотезы относительно подкрепляемого признака ни новые кадры, ни гу-

док не вызывали КГР (кожногальванический рефлекс, индикатор эмоций – Е.Е.). Возникновение гипотезы сопровождается КГР. После формирования гипотезы возможны две ситуации, которые мы рассматриваем в качестве экспериментальных моделей отрицательной и положительной эмоциональных реакций. Гипотеза не верна, и кадр, содержащий подкрепляемый признак (00 и, следовательно, не подтверждающий гипотезу 01 – Е.Е.), не вызывает КГР. Когда же гудок показывает испытуемому, что он ошибся, регистрируется КГР как результат рассогласования гипотезы с наличным раздражителем. Испытуемый несколько раз меняет гипотезу, и в какой-то момент она начинает соответствовать действительности. Теперь уже само появление подкрепляемого кадра вызывает КГР, а его подкрепление гудком приводит к ещё более сильным кожногальваническим сдвигам. Как понять этот эффект? Ведь в данном случае произошло полное совпадение гипотезы с наличным стимулом. Отсутствие рассогласования должно было бы повлечь за собой отсутствие КГР. На самом деле в последнем случае мы также встречаемся с рассогласованием, но рассогласованием иного рода, чем при проверке ложной гипотезы. Формирующийся в процессе повторных сочетаний прогноз содержит не только дифференциальную модель цели, но и вероятность достижения этой цели. В момент подкрепления кадра гудком прогнозируемая *вероятность решения задачи (правильность гипотезы)* резко возросла, и это рассогласование прогноза с поступившей информацией привело к сильной КГР».

Таким образом, подтверждение гипотезы, увеличивающее её вероятность и, следовательно, близость концепции-модели нашему миру (проявление инстинкта к знанию), вызывают положительную эмоцию. Весь процесс обучения, когда человек добивается всё более точных и правильных действий в реальном мире, поддерживается эмоциями – положительные эмоции подкрепляют правильные действия (и соответствующие правильные предсказания, увеличивая их вероятность), а отрицательные эмоции корректируют рассогласования модели и мира (и соответствующие неправильные предсказания, уменьшая их вероятность).

Близость концепций-моделей нашему миру, контролируемая эмоциями, в нашем случае оценивается вероятностью предсказаний. Семантический вероятностный вывод, лежащий в основе оператора обучения, осуществляет направленный поиск всё более вероятных правил путём добавления в условие правил таких дополнительных свойств мира, которые позволяют увеличивать условную вероятность прогноза и, следовательно, обеспечивают большую адекватность и близость миру. Такой направленный поиск снимает проблему КС.

Данные, модели, отношение общности, близость модели к данным

Определим основные понятия вероятностной динамической логики простейшим образом. Развернутые определения в языке логики первого порядка определены в [20, 21, 22].

Под данными будем понимать стандартную матрицу объект-признак, в которой на множестве объектов $A = \{a_1, \dots, a_m\}$ задано множество признаков $x_1(a), \dots, x_n(a)$, где a – переменная по объектам. Определим для каждого значения признака атомарное высказывание $P_i^j(a) = (x_i(a) = x_{ij})$, где x_{ij} – некоторое значение признака x_i . Множество всех атомарных высказываний обозначим через At . Литералами будем называть множество атомарных высказываний, либо их отрицаний. Будем обозначать литералы как булевы переменные a, b, c, \dots , а множество всех литералов обозначим через L .

Будем предполагать, что данные представлены эмпирической системой [11] (являющейся алгебраической системой)

$$Data = \langle A, \{P_1^1, \dots, P_{n_1}^1, P_1^2, \dots, P_{n_2}^2, \dots, P_1^k, \dots, P_{n_k}^k\} \rangle,$$

в которой заданы значения истинности всех атомарных высказываний на множестве объектов A .

Под моделью будем понимать Булеву функцию Φ от булевых переменных из L . Известно, что любая Булева функция может быть представлена множеством правил $\{R\}$ вида

$$(a \& b \& \dots \& c \Rightarrow d), a, b, c, d \in L,$$

поэтому под моделью Φ будем понимать множество правил $\{R\}$.

Для моделей, определенных как совокупность правил, возникает проблема комбинаторной сложности. Чтобы избежать этой проблемы, определим упорядочения на множествах правил и моделей, а также меру близости между моделью и данными.

Определим множество предложений F как множество высказываний, получающихся из литер L замыканием относительно логических операций \wedge, \vee и \neg .

Определение 1. Правило $R_1 = (a_1^1 \& a_2^1 \& \dots \& a_{k_1}^1 \Rightarrow c)$ называется строго более общим $R_1 \succ R_2$, чем правило $R_2 = (b_1^2 \& b_2^2 \& \dots \& b_{k_2}^2 \Rightarrow c)$, тогда и только тогда, когда $\{a_1^1, a_2^1, \dots, a_{k_1}^1\} \subset \{b_1^2, b_2^2, \dots, b_{k_2}^2\}$, $k_1 < k_2$ и более общим $R_1 \succeq R_2$, если $k_1 \leq k_2$.

Следствие 1. $R_1 \succ R_2 \Rightarrow R_1 \vdash R_2$, где \vdash – доказуемость в исчислении высказываний.

Таким образом, более общее высказывание логически сильнее. Кроме того, более общее правило проще, так как содержит меньшее число литер в посылке правила, поэтому отношение общности \succ также является *отношением простоты*.

Определение 2. Модель $\Phi_1 = \{R_i^1\}$ будем называть *более общей* $\Phi_1 \succ \Phi_2$, чем модель $\Phi_2 = \{R_j^2\}$, тогда и только тогда, когда для любого правила $R^2 \in \Phi_2$ существует более общее правило $R^1 \in \Phi_1$, $R_1 \succeq R_2$ и хотя бы для одного правила $R^1 \in \Phi_1$ отношение общности строгое $R_1 \succ R_2$.

Следствие 2. $\Phi_1 \succ \Phi_2 \Rightarrow \Phi_1 \vdash \Phi_2$.

Из следствия 2 следует, что более общая модель логически сильнее и одновременно проще.

Определение 3. *Вероятностью* на множестве предложений F назовем отображение $\mu : F \mapsto [0,1]$, удовлетворяющее следующим условиям [8]:

1. Если $\vdash \Phi$, то $\mu(\Phi) = 1$;
2. Если $\vdash \neg(\Phi \wedge \Psi)$, то $\mu(\Phi \vee \Psi) = \mu(\Phi) + \mu(\Psi)$.

Определим условную вероятность правила $R = (a_1 \& \dots \& a_k \Rightarrow c)$ как

$$\mu(R) = \mu(c / a_1 \& \dots \& a_k) = \frac{\mu(a_1 \& \dots \& a_k \& c)}{\mu(a_1 \& \dots \& a_k)}.$$

Мы предполагаем, что вероятность μ отражает вероятности событий, представленных в данных Data.

Определение 4. *Вероятностным законом* будем называть такое правило R , которое нельзя обобщить (логически усилить), не уменьшая его условную вероятность, т.е. для любого R' , если $R' \succ R$, то $\mu(R') < \mu(R)$.

Вероятностные законы – это наиболее общие, простые и логически сильные правила, среди более общих правил с той же условной вероятностью. Обозначим множество всех вероятностных законов через PL (Probabilistic Laws). Любое правило можно обобщить (упростить и логически усилить) до вероятностного закона с сохранением условной вероятности.

Лемма 1. Для любого правила R , либо оно является вероятностным законом, либо существует вероятностный закон $R' \in PL$ такой, что $R' \succ R$ и $\mu(R') = \mu(R)$.

Определение 5. Под *вероятностной закономерной моделью* будем понимать модель $\Phi = \{R\}, R \in PL$.

Лемма 2. Для любой модели Φ существует более общая вероятностная закономерная модель $\Phi' \succ \Phi$.

Определим отношение упорядочения на множестве вероятностных законов PL .

Определение 6. Под отношением *вероятностного вывода* $R_1 \sqsubseteq R_2$, $R_1, R_2 \in PL$ для вероятностных законов будем понимать одновременное выполнение двух неравенств $R_1 \succeq R_2$ и $\mu(R_1) \leq \mu(R_2)$. Если оба неравенства строгие, то отношение вероятностного вывода будем называть строгим $R_1 \sqsubset R_2 \Leftrightarrow R_1 \succ R_2 \ \& \ \mu(R_1) < \mu(R_2)$.

Определение 7. *Семантическим вероятностным выводом* [2, 20, 21] будем называть максимальную (которую нельзя продолжить) последовательность вероятностных законов находящихся в отношении вероятностного вывода $R_1 \sqsubset R_2 \sqsubset \dots \sqsubset R_k$. Последний вероятностный закон R_k в этом выводе будем называть *максимально специфическим*.

Расширим определение семантического вероятностного вывода и определим отношение близости на вероятностных закономерных моделях.

Определение 8. Вероятностная закономерная модель $\Phi_2 = \{R_j^2\}$, $R_j^2 \in PL$ *ближе к данным*, чем вероятностная закономерная модель $\Phi_1 = \{R_i^1\}$, $R_i^1 \in PL$ обозначим $\Phi_1 \triangleleft \Phi_2$, тогда и только тогда, когда $\Phi_1 \succ \Phi_2$ и для любого вероятностного закона $R^2 \in \Phi_2$ существует вероятностный закон $R^1 \in \Phi_1$, $R_1 \sqsubseteq R_2$ и для хотя бы одного вероятностного закона $R^2 \in \Phi_2$ существует вероятностный закон $R^1 \in \Phi_1$, со строгим отношением вероятностного вывода $R_1 \sqsubset R_2$.

Это определение означает, что при переходе от вероятностной закономерной модели Φ_1 к модели Φ_2 происходит такое наращивание посылок правил, которое (строго) увеличивает условную вероятность этих правил при минимальной их общности и простоте. Увеличение условных вероятностей правил модели означает увеличение предсказательной способности модели и её близость нашим данным.

Как говорилось во введении, инстинкт к знанию состоит в «максимизации меры близости между концепциями-моделями и миром». В нашем определении мера близости определяется через совокупность условных вероятностей правил модели, т.е. через совокупную точность предсказаний модели.

Определение 9. *Мерой близости вероятностной закономерной модели $\Phi = \{R\}$, $R \in PL$ к данным назовём совокупность условных вероятностей правил модели $\{\mu(R), R \in \Phi\}$.*

Следствие 3. Если $\Phi_1 \triangleleft \Phi_2$, $\Phi_1 = \{R_1^1\}$, $\Phi_2 = \{R_j^2\}$, $R_1^1, R_j^2 \in PL$, то мера близости $\{\mu(R), R \in \Phi_2\}$ модели Φ_2 аппроксимирует меру близости $\{\mu(R), R \in \Phi_1\}$ модели Φ_1 в том смысле, что для любой вероятности $\mu(R_1) \in \{\mu(R_1), R_1 \in \Phi_1\}$ существует большая вероятность $\mu(R_1) \leq \mu(R_2)$ в мере близости $\mu(R_2) \in \{\mu(R_2), R_2 \in \Phi_2\}$ модели Φ_2 .

Инстинкт к знанию – это процесс, который проявляется динамически – путём последовательного приближения к данным.

Определение 10. *Обучающим оператором $L: \Phi_1 \rightarrow \Phi_2$ назовем такое преобразование одной модели Φ_1 в модель Φ_2 , при котором близость модели к данным становится выше $\Phi_1 \triangleleft \Phi_2$ и все максимально специфические законы модели Φ_1 переходят в модель Φ_2 .*

Нами разработана программная система Discovery, которая реализует данный обучающий оператор. Эта программа успешно применялась для решения целого ряда практических задач [12,13,18].

Практическая применимость данного подхода и, в частности, обучающего оператора продемонстрирована нами на примере аппроксимации экспертной модели принятия решений в диагностике рака груди. Краткое описание этого применения в терминах вероятностной динамической логики приведено в [22]. Полное описание этого применения приведено в работах [13].

Список литературы

1. Анохин П.К. Опережающее отражение действительности // Философские аспекты теории функциональных систем. М.: Наука, 1978. С. 7–27.
2. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирский гос. ун-т. Новосибирск, 2006. 293 с.

3. Витяев Е.Е. Принципы работы мозга, содержащиеся в теории функциональных систем П.К. Анохина и теории эмоций П.В. Симонова // Нейроинформатика, 2008, том 3, № 1, стр. 25–78.
4. Демин А.В., Витяев Е.Е. Логическая модель адаптивной системы управления // Нейроинформатика, 2008, том 3, № 1, стр. 79–107.
5. Перловский Л.И. К физической теории мышления: теория нейронных моделирующих полей // Нейроинформатика, 2006, том 1, № 2, pp. 175–196.
6. Симонов П.В. Эмоциональный мозг. М.: Наука, 1981. С. 140.
7. Смердов С.О., Витяев Е.Е. Синтез логики, вероятности и обучения: формализация предсказания // Сибирские Электронные Математические Известия. Т.6, Институт математики им.С.Л. Соболева СО РАН, 2009. С. 340–365.
8. Halpern J.Y. An analysis of first-order logics of probability // Artificial Intelligence, 46, 1990. Pp. 311–350.
9. L. Hyafil, R.L. Rivest, Constructing optimal binary decision trees is NP-Complete. Information Processing Letters 5:1 (1976), Pp. 15–17.
10. Jackendoff, R. Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford Univ. Press, New York, NY, 2002.
11. Kovalerchuk B., Ya., Perlovsky L.I. Dynamic logic of phenomena and cognition // IJCNN, 2008. Pp. 3530–3537.
12. Kovalerchuk B.Ya., Vityaev E.E. Data mining in finance: advances in relational and hybrid methods. Kluwer Academic Publisher. 2000.
13. Kovalerchuk B.Ya., Vityaev E.E., Ruiz J.F. Consistent and complete data and "expert" mining in medicine // Medical data mining and knowledge discovery, Springer, 2001, pp. 238–280.
14. Perlovsky L.I. Toward physics of the mind: concepts, emotions, consciousness, and symbols // Physics of Life Reviews, 3, 2006, pp. 23–55.
15. Perlovsky L.I. Neural Networks, Fuzzy Models and Dynamic Logic // R. Kohler and A. Mehler, eds., Aspects of Automatic Text Analysis (Festschrift in Honor of Burghard Rieger), Springer, Germany, 2007, pp. 363–386.
16. Perlovsky, L.I. (1998). Conundrum of Combinatorial Complexity // IEEE Trans. PAMI, 20(6). Pp.666–70.
17. Probabilistic models of cognition // Special issue of the journal: Trends in cognitive science, v.10, Issue 7. 2006. Pp. 287–344.
18. Scientific Discovery website <http://math.nsc.ru/AP/ScientificDiscovery>
19. The Probabilistic Mind. Prospects for Bayesian cognitive science // Eds. Nick Chater, Mike Oaksford, Oxford University Press. 2008. P. 536.

20. Vityaev E.E. The logic of prediction // Mathematical Logic in Asia 2005, Proceedings of the 9th Asian Logic Conference, eds. Goncharov S.S., Downey R. and Ono.H., August 16-19, Novosibirsk, Russia, World Scientific. 2006. Pp. 263–276.

21. Vityaev E.E., Smerdov S.O. New definition of prediction without logical inference // Proceedings of the IASTED international conference on Computational Intelligence (CI 2009), ed. Kovalerchuk B., August 17–19, Honolulu, Hawaii, USA. Pp. 48–54.

22. Evgenii Vityaev, Boris Kovalerchuk, Leonid Perlovsky, Stanislav Smerdov. Probabilistic Dynamic Logic of Phenomena and Cognition // WCCI 2010 IEEE World Congress on Computational Intelligence July, 18-23, 2010 – CCIB, Barcelona, Spain, IJCNN, IEEE Catalog Number: CFP10US-DVD, ISBN: 978-1-4244-6917-8, pp. 3361–3366.