Вероятностная динамическая логика мышления♣

Витяев Е.Е.[1,2], Перловский Л.И.[3],
Ковалерчук Б.Я.[4], Сперанский С.О.[2]

[1]Институт математики им. С.Л. Соболева СО РАН, vityaev@math.nsc.ru
[2]Новосибирский государственный университет, netid@ya.ru
[3]Harvard University, Air Force Research Laboratory, US, leonid@seas.harvard.edu
[4]Central Washington University, Ellensburg, WA 98926-7520, US, borisk@cwu.edu

Аннотация. Одним из авторов статьи (Л.И. Перловским) разработан оригинальный подход к моделированию мышления, основанный на теории нейронных моделирующих полей и динамической логике. Этот подход основан на детальном анализе проблем моделирования мышления в искусственном интеллекте – недостаточности 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. Практическое применение разработанной вероятностной динамической логики иллюстрируется на примере моделирования экспертного правила принятия решений в диагностике рака груди.

1. Введение

Ранее был разработан оригинальный подход к моделированию мышления, основанный на теории нейронных моделирующих полей и динамической логике [9,25-27]. Этот подход, с одной стороны, основан на детальном анализе проблем моделирования мышления в Искусственном Интеллекте – недостаточности формальной логики и проблеме комбинаторной сложности, а, с другой стороны, на данных психологии, философии и когнитивной науки об основных механизмах мышления. Проделанный анализ проблем моделирования мышления имеет, фактически, более широкое значение и преодоление этих проблем может привести к другим формализациям процесса мышления. С этой целью в работе [20] получено обобщение теории нейронных моделирующих полей и динамической логики в виде динамической логики мышления и когнитивной динамической логики. Эти логики сформулированы в достаточно 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 [3,13,32-33]. Полученная в результате Вероятностная Динамическая Логика Мышления принадлежит уже к области вероятностных моделей мышления [28,30]. Мы показываем, что данная логика также решает проблемы моделирования мышления – недостаточности

♣ Работа поддержана грантом РФФИ 08-07-00272-а; интеграционными проектами СО РАН № 47, 115, 119, а также работа выполнена при финансовой поддержке Совета по грантам Президента РФ и государственной поддержке ведущих научных школ (проект НШ-3606.2010.1.)

формальной логики и проблему комбинаторной сложности. Таким образом, через обобщение, полученное в работе [20], мы расширяем интерпретацию теории нейронных моделирующих полей и динамической логики на вероятностные модели мышления. Вероятностная динамическая логика была ранее использована для моделирования работы мозга и когнитивных процессов [3-5,7].

2. Проблемы моделирования мышления с вероятностной точки зрения

Повторим и дополним изложение проблем моделирования мышления в Искусственном Интеллекте, изложенных в [9,20]. 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 [9]. Но применение логики в Искусственном интеллекте не дало ожидаемых результатов. Нужно четко различать теоретический и эмпирический подходы. В теории, использующей идеализированные знания, например, в физике, геометрии, химии и некоторых других науках логика и логический вывод оправданы и прекрасно работают. В основе же интеллектуальных систем лежит процесс обучения, который является эмпирическим, а получаемые в результате обучения знания является индуктивными. Для индуктивно выведенных знаний логический вывод работает плохо.

Мозг – не логическое, а предсказывающее устройство, действующее в соответствии с принципом опережающего отражения действительности П.К. Анохина [1]. Но подходящего определения предсказания для индуктивно выведенных знаний в настоящее время 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. Этот же факт отмечается в работе [8]: «Между дедуктивными и индуктивными рассуждениями существует принципиальное различие, которое заключается в следующем. В дедуктивном умозаключении истинность исходных посылок гарантирует истинность финального вывода, а в случае индукции такой гарантии нет».

Эта проблема в настоящее время осознана как проблема синтеза логики и вероятности. Прошло уже 4 симпозиума в 2002-2009 годах под общим названием Projic (Probability+Logic). 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 получено, т.к. от логического вывода в этих исследованиях не отказывается, а в этом случае, с нашей точки зрения, адекватного определения предсказания получить нельзя. Например, в работе [2] рассмотрено множество разнообразных логик, но в их основе всё равно лежат правила логического вывода.

Нами введено новое понятие предсказания, изложенное в работах [3,32-33] и полученное путем отказа от логического вывода и значений истинности «истина» и «ложь». Интересное определение предсказания, не использующее логический вывод, можно также найти в работе Турчина [31]. Вместо логического вывода нами определён семантический вероятностный вывод, используемый далее для определения оператора обучения. Новое определение предсказания принципиально отличается от определения предсказания по К. Попперу – предсказание некоторого факта осуществляется не логическим выводом этого факта из уже имеющихся фактов и теории, а прямым индуктивным выводом правила, предсказывающего интересующий нас факт. При этом логический вывод и правило «modus ponens» не используются и, соответственно, не возникает проблемы уменьшения оценок вероятности в процессе вывода. Наоборот, оценки вероятности строго возрастают в процессе семантического вероятностного вывода. Как отмечено в [8]: «… дедуктивное применение правила modus ponens математиком и индуктивный «вывод», который делает собака, явно различаются». С нашей точки зрения дедуктивный вывод делается по правилу modus ponens, а предсказание, которое делает собака на основании выработанного условного рефлекса, делается с использованием семантического вероятностного вывода.

Формализация предсказания, основанная на семантическом вероятностном выводе, положена нами в основу формальной модели работы нейрона, принципа опережающего отражения действительности П.К. Анохина и формализации Теории функциональных систем П.К. Анохина [4]. На основе этих формализаций разработана логическая модель адаптивной системы управления анимата, решающего задачу фуражирования [7]. Ниже в разделе 4 приведена формальная модель нейрона и группы нейронов, иллюстрирующих Вероятностную динамическую логику на нейронном уровне.

Другой проблемой моделирования мышления в Искусственном Интеллекте является проблема Комбинаторной Сложности (КС) [9,27]. 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 [27].

Последовавшие исследования обнаружили связь КС с логикой в различных алгоритмах [27]. Логика рассматривает каждое даже небольшое изменение в данных или моделях, как новое высказывание. Приписывание значений истинности «истина» и «ложь» не позволяет сравнивать высказывания. Это приводит к КС. В работе [18] доказывается, что даже простейшая задача нахождения совокупности высказываний, описывающей решающие деревья, NP-трудна.

В данной работе мы используем два отношения упорядочения на высказываниях: отношение общности и сравнение по условной вероятности, которые используются 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 [9]. Мы принимаем, что основные механизмы мышления включают инстинкты, концепции, эмоции и поведение.

Из механизмов мышления концепции-модели наиболее доступны сознанию. Рэй Джакендофф [19] 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, точки) до концепций-моделей объектов, отношений между объектами и 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 [20]. 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.

Эксперименты, подтверждающие связь эмоций с инстинктом к знанию можно найти в Информационной теории эмоций П.В.Симонова [12]: «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. Витяев), в то время как радость (эмоция – Е.Е. Витяев) есть ожидание удовольствия в связи с растущей вероятностью удовлетворения потребности».

Следующий эксперимент показывает, что инстинкт к знанию вызывает положительные эмоции [12]: «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 гудок не вызывали КГР (кожногальванический рефлекс, индикатор эмоций – Л.И. Перловский)... 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. Витяев), 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 афферентную 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, вызывающее положительную эмоцию, увеличивает её вероятность и, следовательно, близость концепции-модели нашему миру (проявление инстинкта к знанию). 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.

3. Данные, модели, отношение общности, близость модели к данным

Определим основные понятия вероятностной динамической логики в простейшем – пропозициональном случае. Развернутые определения в языке логики первого порядка (в несколько различных модификациях) можно найти в [7, 13,32-33].

Под данными будем понимать стандартную матрицу объект-признак, в которой на множестве объектов $A = \{a_1, ..., a_m\}$ задано множество признаков $x_1(a), ..., x_n(a)$, где a – переменная по объектам. Определим для каждого значения признака x_i атомарное высказывание $P_j^i(a) = (x_i(a) = x_{ij})$, $j = 1, ..., n_i$, где $x_{i1}, ..., x_{in_i}$ – все значения признака x_i, $i = 1, ..., n$. Множество всех атомарных высказываний (атомов) обозначим через At. Будем обозначать атомы как булевы переменные a, b, c, … . Литералами будем называть множество, состоящее из атомарных высказываний, либо их отрицаний, которые также будем обозначать как булевы переменные a, b, c, … (возможно с индексами). Множество всех литералов обозначим через L.

Будем предполагать, что *данные* представлены *эмпирической системой* [20] (алгебраической системой)

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

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

Под *моделью* будем понимать Булеву функцию Φ, в которой вместо булевых переменных подставляются атомы из At. Известно, что любая Булева функция может быть представлена множеством правил {R} вида

$$(a_1 \& ... \& a_k \Rightarrow b),\ a_1, ..., a_k, b \in L, \qquad (1)$$

поэтому под моделью Φ будем понимать множество правил {R}. Множество всех правил вида (1) обозначим через Pr. Будем говорить, что *правило* (1) *применимо к данным* Data, если посылка $a_1, ..., a_k$ правила истинная в Data и что *модель* $\Phi = \{R\}$ *применима к данным* Data, если каждое правило модели применимо к данным Data.

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

Определение 1. Правило $R_1 = (a_1^1 \& a_2^1 \& ... \& a_{k_1}^1 \Rightarrow c)$ будем называть *более общим*, чем правило $R_2 = (b_1^2 \& b_2^2 \& ... \& b_{k_2}^2 \Rightarrow c)$, обозначим это как $R_1 \succ R_2$, тогда и только тогда, когда $\{a_1^1, a_2^1, ..., a_{k_1}^1\} \subset \{b_1^2, b_2^2, ..., b_{k_2}^2\}$, $k_1 < k_2$ и *не менее общим* $R_1 \gtrsim R_2$, если $k_1 \le k_2$.

Следствие 1. $R_1 \gtrsim R_2 \Rightarrow R_1 \vdash R_2$ и $R_1 \succ R_2 \Rightarrow R_1 \vdash R_2$, где $\vdash$ – доказуемость в исчислении высказываний (если воспринимать атомы как пропозициональные символы).

Таким образом, не менее общие (и более общее) высказывания логически сильнее. Кроме того, более общие правила проще, так как содержит меньшее число литер в посылке правила, поэтому отношение $\succ$ можно воспринимать как *отношение простоты* в смысле [20].

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

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.

Определим множество предложений F, как множество высказываний, полученных из атомов At замыканием относительно логических операций $\wedge, \vee, \neg$.

Определение 3. *Вероятностью* на множестве предложений F назовем отображение $\mu : F \mapsto [0,1]$, удовлетворяющее следующим условиям [15]:

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

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

$$\mu(R) = \mu(c / a_1 \& \ldots \& a_k) = \frac{\mu(a_1 \& \ldots \& a_k \& c)}{\mu(a_1 \& \ldots \& a_k)}, \text{ если } \mu(a_1 \& \ldots \& a_k) > 0 .$$

Если $\mu(a_1 \& \ldots \& a_k) = 0$, то условная вероятность не определена. Множество всех правил из Pr, для которых вероятность определена, обозначим через Pr_0. Будем предполагать, что вероятность μ задаёт вероятности событий, представленных высказываниями об эмпирической системе Data.

Определение 4. *Вероятностным законом* будем называть такое правило $R \in Pr_0$, которое нельзя обобщить (логически усилить), не уменьшив его условную вероятность, т.е. для любого $R' \in Pr_0$, если $R' \succ R$, то $\mu(R') < \mu(R)$.

Вероятностные законы – это наиболее общие, простые и логически сильные правила, среди правил с не большей условной вероятностью. Обозначим множество всех вероятностных законов через PL (Probabilistic Laws). Любое правило можно обобщить (упростить и логически усилить) до вероятностного закона, не уменьшая его условную вероятность.

Лемма 1. Для любого правила $R \in Pr_0$, либо оно является вероятностным законом, либо существует вероятностный закон $R' \in PL$ такой, что $R' \succ R$ и $\mu(R') \geq \mu(R)$.

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

Лемма 2. Для любой модели Φ, либо она является вероятностной закономерной моделью, либо существует не менее общая Φ' “ Φ, общая вероятностная закономерная модель Φ'.

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 $R_1 \succsim R_2$ и $\mu(R_1) \leq \mu(R_2)$. Если оба неравенства строгие, то отношение вероятностного вывода будем называть строгим *отношением вероятностного вывода* $R_1 \sqsubset R_2 \Leftrightarrow R_1 \succ R_2 \,\&\, \mu(R_1) < \mu(R_2)$.

Определение 7. Под *семантическим вероятностным выводом* [3,13,32-33] будем называть максимальную (которую нельзя продолжить) последовательность вероятностных законов находящихся в отношении строгого вероятностного вывода $R_1 \sqsubset R_2 \sqsubset \ldots \sqsubset R_k$. Последний вероятностный закон R_k в этом выводе будем называть *максимально специфическим*.

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

Определение 8. Вероятностная закономерная модель $\Phi_2 = \{R_j^2\}$ *ближе к данным*, чем вероятностная закономерная модель $\Phi_1 = \{R_i^1\}$, обозначим это как $\Phi_1 \lhd \Phi_2$, тогда и только тогда, когда $\Phi_1 \succ \Phi_2$ и для любого вероятностного закона $R^2 \in \Phi_2$ существует вероятностный закон $R^1 \in \Phi_1$ такой, что $R_1 \sqsubseteq R_2$ и хотя бы для одного вероятностного закона $R^2 \in \Phi_2$, существует вероятностный закон $R^1 \in \Phi_1 \setminus \Phi_2$, 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 к 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* $\Phi = \{R\}$ *к данным* назовём совокупность $\{\mu(R), R \in \Phi\}$ условных вероятностей правил модели.

Следствие 3. Если $\Phi_1 \lhd \Phi_2$, $\Phi_1 = \{R_i^1\}$, $\Phi_2 = \{R_j^2\}$, 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 $\mu(R_2) \in \{\mu(R_2), R_2 \in \Phi_2\}$ существует такая же $\mu(R_1) \le \mu(R_2)$, либо строго меньшая вероятность $\mu(R_1) < \mu(R_2)$, $\mu(R_1) \in \{\mu(R_1), R_1 \in \Phi_1\}$.

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

Определение 10. *Обучающим оператором* $L : \Phi_1 \to \Phi_2$ назовем такое преобразование модели Φ_1 в модель Φ_2, при котором обе модели применимы к данным и близость модели к данным становится выше $\Phi_1 \lhd \Phi_2$, а также все максимально специфические законы модели Φ_1 переходят в модель Φ_2.

Нами разработана программная система Discovery, которая в точности реализует данный обучающий оператор по следующим шагам:

1) начальное множество правил $\Phi_0 = \{R\}$ состоит из правил вида (1) с количеством предикатов в посылке k не более чем некоторое число N, задаваемое пользователем;
2) из множества Φ_0 отбираются только те правила, которые являются вероятностными законами $\Phi_1 = \{R \mid R \in \Phi_0, R \in PL\}$;
3) далее применяется обучающий оператор $L : \Phi_i \to \Phi_{i+1}, i = 1, 2, \ldots$, пока это возможно. Если усиление модели уже невозможно, то процесс обучения заканчивается. Строгое увеличение вероятности в процессе вероятност-

ного вывода проверяется точным критерием независимости Фишера для таблиц сопряженности.

В работе [23] приведен псевдокод данного алгоритма. Эта программа успешно применялась для решения целого ряда прикладных задач в различных предметных областях [3,23-24,29].

4. Нейронная организация вероятностной динамической логики

Поскольку семантический вероятностный вывод может рассматриваться как формальная модель нейрона [4], то мы можем проиллюстрировать введенные в предыдущем параграфе определения на нейронном уровне.

Приведём кратко формальную модель нейрона из [4]. Под *информацией*, поступающей на «вход» мозга, мы будем понимать всю воспринимаемую мозгом афферентацию: мотивационную, обстановочную, пусковую, обратную, санкционирующую, афферентацию об осуществленных действиях, поступающую по коллатералям на «вход» и т. д. Из экологической теории восприятия Дж. Гибсона [6] следует, что под информацией можно понимать любую характеристику энергетического потока света, звука и т. д., поступающую на вход мозга.

Определим информацию, передаваемую по некоторому нервному волокну на дендриты нейрона, одноместными предикатами $P_j^i(a) = (x_i(a) = x_{ij})$, $j = 1,...,n_i$, где $x_i(a)$ – некоторая характеристика энергетического потока, x_{ij} – значение этой характеристики в ситуации (на объекте) a. Если это возбуждение предается на возбуждающий синапс, то оно воспринимается нейроном как информация об истинности предиката $P_j^i(a)$, если на тормозной синапс, то как отрицание предиката $\neg P_j^i(a)$.

Возбуждение нейрона в ситуации (на объекте) a и передачу этого возбуждения на его аксон также определим одноместным предикатом $P_0(a)$. Если нейрон тормозится в ситуации a и не передаёт возбуждение по своему аксону, то определим эту ситуацию как отрицание предиката $\neg P_0(a)$. Известно, что каждый нейрон имеет рецептивное поле, стимуляция которого возбуждает его безусловно. Первоначальной (до всякого обучения) семантикой предиката P_0 является информация, извлекаемая им из рецептивного поля. В процессе обучения эта информация обогащается и может дать достаточно специализированный нейрон, например, «нейрон Билла Клинтона».

Мы предполагаем, что формирование условных правил (связей) на уровне нейрона происходит по правилу Хебба [17] и может быть достаточно точно формализовано семантическим вероятностным выводом. Нейрофизиологические подтверждения этого правила можно найти в работе [11].

Предикаты и их отрицания являются литералами, и мы их, как и в предыдущем разделе, будем обозначать булевыми переменными a, b, c, … ∈ L.

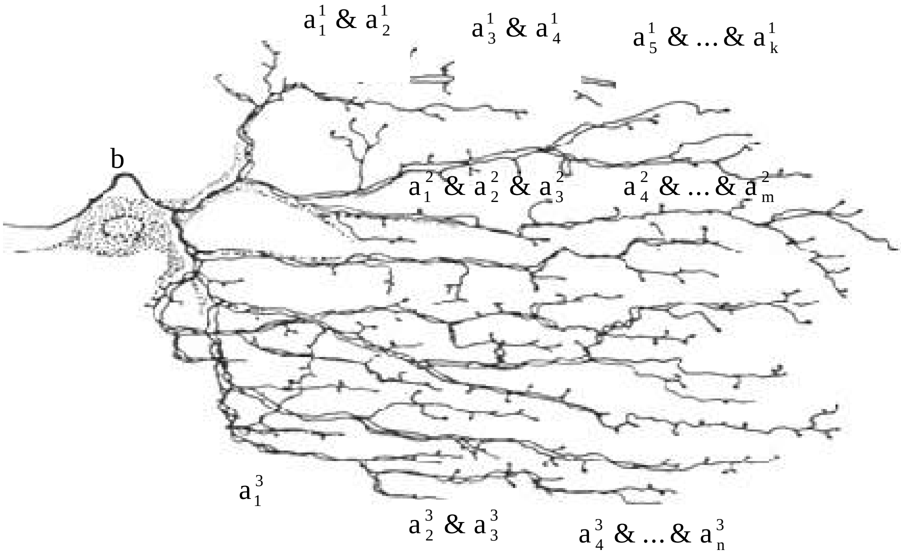

Рис. 1

Нейрон определим как множество правил {R} вида (1):

$$(a_1 \& \ldots \& a_k \Rightarrow b),\ a_1,\ldots,a_k,b \in L,$$

где $a_1,\ldots,a_k$ – значения (возбуждающие/тормозные) предикатов (стимулов), приходящих на вход нейрона, а b – значение предиката $P_0(a)$, обозначающего выход (аксон) нейрона. По этому определению нейрон является частным случаем модели, т.к. заключение правил нейрона фиксировано и состоит в возбуждении нейрона $P_0(a)$.

Определим метод вычисления условных вероятностей правил $(a_1 \& \ldots \& a_k \Rightarrow b)$ нейрона. Подсчитаем число случаев $n(a_1,\ldots,a_k,b)$, когда произошло событие $\langle a_1,\ldots,a_k,b\rangle$ – одновременное возбуждение/торможение входов $\langle a_1,\ldots,a_k\rangle$ нейрона и самого нейрона непосредственно перед действием подкрепления. Подкрепление может быть как положительным, так и отрицательным и осуществляться как санкционирующей афферентацией, так и эмоциями [4,5].

Среди случаев $n(a_1,\ldots,a_k,b)$ подсчитаем случаи $n^+(a_1,\ldots,a_k,b)$, когда подкрепление было положительным, а также число случаев $n^-(a_1,\ldots,a_k,b)$, когда подкрепление было отрицательным. Условная вероятность (эмпирическая) правила $(a_1 \& \ldots \& a_k \Rightarrow b)$ нейрона тогда вычисляется следующим образом:

$$\mu(b / a_1,\ldots,a_k) = \frac{n^+(a_1,\ldots,a_k,b) - n^-(a_1,\ldots,a_k,b)}{n(a_1,\ldots,a_k,b)}.$$

В процессе выработки условного рефлекса условные сигналы связываются с результатом – подкреплением. На уровне нейронов условный рефлекс проявляется в виде пластичности нейронов [11]. Сигнальные стимулы связываются со свойством (характеристикой), за которое отвечает нейрон так, что появление одного только стимула может вызвать возбуждение нейрона.

Кроме того, если обнаруживаются новые стимулы, которые в ещё большей степени (с большей вероятностью) позволяют предсказывать возбуждение нейрона, то они присоединяются к данной условной связи. Формально присоединение новых стимулов к имеющейся связи определяется вероятностным выводом (определение 6), который собственно и означает, что новые стимулы добавляются к условной связи, если они увеличивают условную вероятность предсказания возбуждения нейрона.

Формализация процесса замыкания условных связей на уровне нейрона дается семантическим вероятностным выводом (определение 7). На рис. 1 схематически показано несколько семантических вероятностных выводов, осуществляемых нейроном:

1. $(b \Leftarrow a_1^1 \,\&\, a_2^1) \sqsubset (b \Leftarrow a_1^1 \,\&\, a_2^1 \,\&\, a_3^1 \,\&\, a_4^1) \sqsubset$
$\sqsubset (b \Leftarrow a_1^1 \,\&\, a_2^1 \,\&\, a_3^1 \,\&\, a_4^1 \,\&\, a_5^1 \,\&\ldots\&\, a_k^1)$
2. $(b \Leftarrow a_1^2 \,\&\, a_2^2 \,\&\, a_3^2) \sqsubset (b \Leftarrow a_1^2 \,\&\, a_2^2 \,\&\, a_3^2 \,\&\, a_4^2 \,\&\ldots\&\, a_m^2)$
3. $(b \Leftarrow a_1^3) \sqsubset (b \Leftarrow a_1^3 \,\&\, a_2^3 \,\&\, a_3^3) \sqsubset (b \Leftarrow a_1^3 \,\&\, a_2^3 \,\&\, a_3^3 \,\&\, a_4^3 \,\&\ldots\&\, a_n^3)$

По определению семантического вероятностного вывода правила, входящие в него, должны быть вероятностными законами (определение 4). Это означает, что в правило не включаются стимулы, которые не являются сигнальными, т.е. каждый стимул должен увеличивать вероятность (предсказание) возбуждения нейрона, что достаточно очевидно для условной связи.

Ещё одним свойством семантического вероятностного вывода является требование увеличения вероятности правил в процессе вероятностного вывода $R_1 \sqsubset R_2$ (определение 6), т.е. $(b \Leftarrow a_1^1 \,\&\, a_2^1) \sqsubset (b \Leftarrow a_1^1 \,\&\, a_2^1 \,\&\, a_3^1 \,\&\, a_4^1)$. Этот означает, что условная связь $(b \Leftarrow a_1^1 \,\&\, a_2^1)$ усиливается новыми стимулами $a_3^1 \,\&\, a_4^1$ до связи $(b \Leftarrow a_1^1 \,\&\, a_2^1 \,\&\, a_3^1 \,\&\, a_4^1)$, если они увеличивают условную вероятность предсказания возбуждения нейрона b.

Совокупность семантических вероятностных выводов, которые обнаруживает нейрон в процессе обучения, составляет его вероятностную закономерную модель (определение 5), предсказывающую возбуждение $P_0(a)$ нейрона.

Вероятностная закономерная модель нейрона становится ближе к данным (определение 8), если хотя бы одна из его условных связей (один из вероятностных законов) стала сильнее – добавились новые стимулы с увеличением условной вероятности (осуществился вероятностный вывод – определение 6). Данные в этом случае представляют собой запись опыта обучения.

Близость к данным означает, что нейрон более точно реагирует на условные стимулы – с большей условной вероятностью. Мерой близости вероятностной закономерной модели к данным (определение 9) называется совокупность условных вероятностей предсказания возбуждения нейрона по всем его условным связям.

Обучение нейрона (обучающим оператором – определение 10) называется такое преобразование закономерной модели нейрона, при котором в результате получения опыта, взятого из данных, происходит строгое усиление хотя бы одной из его условных связей. Таким образом, оператор обучения постоянно усиливает условные связи нейрона на основании опыта – происходит дифференциация опыта, развитие навыков и интуиции.

Известно, что в процессе выработки условных связей, а также при замыкании условных связей на уровне отдельного нейрона, скорость проведения импульса от условного раздражителя к аксону нейрона, т. е. скорость ответа нейрона на условный сигнал, тем выше, чем выше вероятность условной связи. Это подтверждает, что мозг реагирует, прежде всего, на высоковероятные прогнозы и нейроны срабатывают на самые сильные закономерности с максимальными оценками условных вероятностей.

Моделью группы нейронов, представляющих собой, например, интуицию эксперта, как в задаче аппроксимации экспертной модели принятия решений в диагностике рака груди, рассмотренной в следующем разделе, является совокупность закономерных моделей нейронов.

Если стоит задача моделирования некоторой Булевой функции Φ , как описано в начале раздела 3, или промоделировать Булеву функцию экспертной модели принятия решений, как описано в разделе 5, то необходимо рассмотреть все закономерные модели нейронов, предсказывающих булевы переменные этих функций.

Теория нейронных моделирующих полей и динамическая логика [9,25-27] также могут быть представлены на нейронном уровне. Для определения силы условной связи на уровне нейрона для этого используются веса, приписываемые условному стимулу и определяющие силу его воздействия на возбуждение нейрона. В процессе обучения эти веса – как у нас вероятности – модифицируются.

Основное сходство динамической логики и вероятностной динамической логики в их динамичности:

- в теории нейронных моделирующих полей она заключается в том, что мера сходства меняется постепенно и, сначала, находятся приближенные решения и модели, а затем мера сходства уточняется, и находятся более точные модели;
- в случае семантического вероятностного вывода – сначала для предсказания используются простые правила, с одним или двумя условными стимулами в посылке, дающих не очень хорошую условную вероятность (меру близости), а затем правила наращиваются путём добавления новых уточняющих условий так, чтобы усилить эту условную вероятность.

5. Экстрагирование экспертной модели принятия решений в диагностике рака груди

Мы применили разработанный оператор обучения для аппроксимации экспертной модели принятия решений в диагностике рака груди [24]. Сначала мы экстрагировали эту модель из эксперта радиолога J. Ruiz, используя специальную процедуру извлечения экспертных знаний, основанную на монотонных булевых функциях [21,24]. Затем мы применили систему ‘Discovery’ [3,23-24] и, реализованный в ней обучающий оператор, для аппроксимации этой модели.

1. Иерархический подход. Сначала мы попросили эксперта описать конкретные случаи, используя бинарные признаки. Затем, попросили эксперта принять решение относительно каждого случая. Типичный запрос к эксперту имел следующий вид: «Если признак 1 имеет значение v_1, признак 2 – значение v_2, ..., признак n – значение v_n, то является ли этот случай подозрительным на рак или нет?». Каждое множество значений $v_1, v_2, ..., v_n$ представляет собой возможный клинический случай. Практически невозможно попросить радиолога дать диагнозы для тысяч возможных случаев. Мы использовали иерархический подход вместе со свойством монотонности, что бы свести число вопросов к эксперту к нескольким десяткам. Сначала мы построили иерархию клинических свойств. В низу иерархии располагались 11 клинических бинарных признаков $w_1, w_2, w_3, y_1, y_2, y_3, y_4, y_5, x_3, x_4, x_5$ (со значениями: 1 – «подозрение на рак», 0 – «доброкачественный случай»). Эксперт установил, что эти 11 признаков могут быть организованы в иерархию путем введения двух обобщенных признаков: x_1 – «количество и объем кальциноза», зависящего от признаков w_1, w_2, w_3:

w_1 – число кальциноза/ cm^3,

w_2 – объем кальциноза/ cm^3,

w_3 – общее число кальциноза.

и x_2 – «форма и плотность кальциноза», зависящего от признаков y_1, y_2, y_3, y_4, y_5:

y_1 – ‘нерегулярность в форме отдельных кальцинозов’

y_2 – ‘вариации в форме кальциноза’

y_3 – ‘вариации в размере кальциноза’

y_4 – ‘вариации в плотности кальциноза’

y_5 – ‘плотность кальциноза’.

Мы будем рассматривать признак x_1 как функцию $g(w_1, w_2, w_3)$ и признак и x_2 как функцию $x_2 = h(y_1, y_2, y_3, y_4, y_5)$, которые требуется найти. В результате мы получаем декомпозицию задачи в виде следующей функции: $f(x_1, x_2, x_3, x_4, x_5) = f(g(w_1, w_2, w_3), h(y_1, y_2, y_3, y_4, y_5), x_3, x_4, x_5)$.

2. **Монотонность.** В терминах введенных признаков и функций мы можем представить клинические случаи в терминах векторов с пятью обобщенными переменными $(x_1, x_2, x_3, x_4, x_5)$. Рассмотрим два клинических случая представленных векторами: (10110) и (10100). Если радиолог правильно диагностирует случай (10100) как подозрительный на рак, то, используя свойство монотонности, мы можем заключить, что случай (10110) также должен быть подозрительным на рак, так как признаков со значением 1, способствующих раку, в этом случае больше. Эксперт согласен с предположением о монотонности функций $f(x_1, x_2, x_3, x_4, x_5)$ и $h(y_1, y_2, y_3, y_4, y_5)$.

6. Извлечение экспертного правила принятия решений

Опишем интервью с экспертом, используя минимальные последовательности вопросов, для полного восстановления функций f и h. Эти последовательности базируются на фундаментальной лемме Ханселя [16,21]. Мы опустим описание математических деталей, которые можно найти в [21].

Рассмотрим таблицу 1. Колонки 2 и 3 представляют собой значения функций f и h, которые должны быть восстановлены. Мы опускаем восстановление функции $g(w_1, w_2, w_3)$, потому что требуется всего несколько вопросов для её восстановления. Все 32 возможных случая для пяти булевых переменных $\langle x_1, x_2, x_3, x_4, x_5 \rangle$ представлены в колонке 1. Они сгруппированы в цепи Ханселя [21,24]. Последовательность цепей начинается с короткой цепи 1 – $(01100) < (11100)$. Наибольшая цепь 10 содержит 6 упорядоченных случая: $(00000) < (00001) < (00011) < (00111) < (01111) < (11111)$. Цепи пронумерованы от 1 до 10, и каждый случай имеет свой собственный номер в цепи, например, 1.2 – второй случай в первой цепи. Звездочки в колонках 2 и 3 означают ответы, полученные от эксперта, например, 1* для случая (01100) в колонке 3 означает, что эксперт сказал ‘Да’ («подозрение на рак») относительно этого случая. Ответы для некоторых других случаев в колонках 2 и 3 автоматически получаются на основании свойства монотонности. Значение $f(01100) = 1$ для случая 1.1 может быть распространено на случаи 1.2, 6.3. и 7.3 вследствие монотонности. Аналогично вычисляются значения монотонной булевой функции h, только атрибуты, например, в последовательности (10010) интерпретируются как признаки y_1, y_2, y_3, y_4, y_5. Цепи Ханселя остаются теми же самыми, если число переменных не изменяется.

Колонки 4 и 5 перечисляют случаи, на которые можно распространить значения функций f и h без опроса эксперта. Колонка 4 для распространения значения 1 функций и колонка 5 для распространения значения 0 функций. Если эксперт даст ответ $f(01100) = 0$ для случая (01100), то это значение должно быть распространено в колонке 2 на случаи 7.1 (00100) и 8.1 (01000), записанные в колонке 5. Поэтому не надо спрашивать эксперта о случаях 7.1 и 8.1, т.к. они следуют из монотонности. Ответ $f(01100) = 0$ не может быть распространен на случай 1.2 со значением $f(11100)$, поэтому эксперт должен дать ответ относительно случая $f(11100)$. Если его ответ будет $f(11100) = 0$, то это значение должно быть распространено на случаи 5.1. и 3.1, записанные в колонке 5.

Общее число вопросов, отмеченных звездочкой (*) в колонках 2 и 3 равно 13 и 12 соответственно. Это число показывает, что 13 вопросов нужны для восстановления функции $f(x_1, x_2, x_3, x_4, x_5)$ и 12 вопросов для восстановления функции $h(y_1, y_2, y_3, y_4, y_5)$. Это только 37.5% из 32 возможных вопросов. Полное число вопросов, требуемых для восстановления этих функций без условия монотонности и иерархии равно $2 * 2^{11} = 4096$.

Таблица 1. Динамическая последовательность вопросов эксперту

Случай	f	h	Распространение по монотонности		Цепь	Случай
1	**2**	**3**	**4:** $1 \mapsto 1$	**5:** $0 \mapsto 0$	**6**	**7**
(01100)	1*	1*	1.2, 6.3, 7.3	7.1, 8.1	Цепь 1	1.1
(11100)	1	1	6.4, 7.4	5.1, 3.1		1.2
(01010)	0*	1*	2.2, 6.3, 8.3	6.1, 8.1	Цепь 2	2.1
(11010)	1*	1	6.4, 8.4	3.1, 6.1		2.2
(11000)	1*	1*	3.2	8.1, 9.1	Цепь 3	3.1
(11001)	1	1	7.4, 8.4	8.2, 9.2		3.2
(10010)	0*	1*	4.2, 9.3	6.1, 9.1	Цепь 4	4.1
(10110)	1*	1	6.4, 9.4	6.2, 5.1		4.2
(10100)	1*	1*	5.2	7.1, 9.1	Цепь 5	5.1
(10101)	1	1	7.4, 9.4	7.2, 9.2		5.2
(00010)	0	0*	6.2, 10.3	10.1	Цепь 6	6.1
(00110)	1*	0*	6.3, 10.4	7.1		6.2
(01110)	1	1	6.4, 10.5			6.3
(11110)	1	1	10.6			6.4
(00100)	1*	0*	7.2, 10.4	10.1	Цепь 7	7.1
(00101)	1	0*	7.3, 10.4	10.2		7.2
(01101)	1	1*	7.4, 10.5	8.2, 10.2		7.3
(11101)	1	1	5.6			7.4
(01000)	0	1*	8.2	10.1	Цепь 8	8.1
(01001)	1*	1	8.3	10.2		8.2
(01011)	1	1	8.4	10.3		8.3
(11011)	1	1	10.6	9.3		8.4
(10000)	0	1*	9.2	10.1	Цепь 9	9.1
(10001)	1*	1	9.3	10.2		9.2
(10011)	1	1	9.4	10.3		9.3
(10111)	1	1	10.6	10.4		9.4
(00000)	0	0	10.2		Цепь 10	10.1
(00001)	0*	0	10.3			10.2
(00011)	1*	0	10.4			10.3
(00111)	1	1*	10.5			10.4
(01111)	1	1	10.6			10.5
(11111)	1	1				10.6
Вопросы	13	12				

7. Решающие правила (модель) извлеченные из эксперта

Найдем Булевы функции $f(x_1, x_2, x_3, x_4, x_5)$ и $h(y_1, y_2, y_3, y_4, y_5)$ по таблице 1 следующим образом:

1. надо найти все максимальные нижние единицы для всех цепей и представить их в виде конъюнкций;
2. взять дизъюнкцию полученных конъюнкций;
3. исключить лишние (выводящиеся из других) конъюнкции.

По колонке 1 и 3 мы находим

$$x_2 = h(y_1, y_2, y_3, y_4, y_5) = y_2y_3 \vee y_2y_4 \vee y_1y_2 \vee y_1y_4 \vee y_1y_3 \vee$$
$$\vee y_2y_3y_5 \vee y_2 \vee y_1 \vee y_3y_4y_5 \equiv y_2 \vee y_1 \vee y_3y_4y_5.$$

Функция $g(w_1, w_2, w_3) = w_2 \vee w_1w_3$ может быть получена заданием $2^3 = 8$ вопросов эксперту. По колонкам 1 и 2 мы находим

$$f(\overline{x}) = x_2x_3 \vee x_1x_2x_4 \vee x_1x_2 \vee x_1x_3x_4 \vee x_1x_3 \vee x_3x_4 \vee x_3$$
$$\vee x_2x_5 \vee x_1x_5 \vee x_4x_5 \equiv x_1x_2 \vee x_3 \vee (x_2x_1x_4)x_5 \equiv$$
$$(w_2 \vee w_1w_3)(y_1 \vee y_2 \vee y_3y_4y_5) \vee x_3 \vee$$
$$\vee (y_1 \vee y_2 \vee y_3y_4y_5)(w_2 \vee w_1w_3)x_4x_5.$$

8. Аппроксимация модели эксперта обучающим оператором

Для аппроксимации экспертной модели принятия решений была использована программная система ‘Discovery’ [3, 23-24,29], реализующая обучающий оператор. Было обнаружено несколько десятков диагностических правил, аппроксимирующие эту экспертную модель. Они были статистически значимы относительно статистического критерия отбора правил для уровней 0.01, 0.05 0.1. Правила были обнаружены на 156 случаях (73 злокачественных, 77 доброкачественных, 2 подозрительных на рак и 4 со смешанным диагнозом) [24].

Совокупность правил была протестирована методом скользящего контроля. Диагноз был получен для 134 случаев (в 22 случаях диагноз не поставлен). Точность диагноза составила 86%. Неправильный диагноз был получен в 19 случаях (14% от всех диагностических случаев). Ошибка первого рода составила 5.2% (7 злокачественных случаев были диагностированы как доброкачественные) и ошибка второго рода была 8.9% (12 доброкачественных случая были диагностированы как злокачественные). Некоторые их этих правил приведены в таблице 2. В таблице приводятся примеры правил вместе с их статистической значимостью по критерию Фишера. В этой таблице: ‘NUM’ – число кальциноза в cm^3; ‘VOL’ – объем in cm^3; ‘TOT’ – общее число кальциноза; ‘DEN’ – плотность кальциноза; ‘VAR’ – вариации в форме кальциноза; ‘SIZE’ – вариации в размере кальциноза; ‘IRR’ – нерегулярность в форме кальциноза; ‘SHAPE’ – форма кальциноза.

Мы рассмотрели три уровня меры близости для обучающего оператора 0.75, 0.85 и 0.95, которые означали, что условные вероятности всех вероятностных

Таблица 2. Примеры обнаруженных диагностических правил

Диагностическое правило	f-критерий		Значение. f-критерия 0.01	0.05	0.1	Точность на контроле
Если $10 < \mathrm{NUM} < 20$	NUM	0.0029	+	+	+	93.3%
и $\mathrm{VOL} > 5$	VOL	0.0040	+	+	+	
то *злокачественный*						
Если $\mathrm{TOT} > 30$	TOT	0.0229	-	+	+	100.0%
и $\mathrm{VOL} > 5$	VOL	0.0124	-	+	+	
и DEN *умеренная*	DEN	0.0325	-	+	+	
то *злокачественный*						
Если VAR *заметная*	VAR	0.0044	+	+	+	100.0%
и $10 < \mathrm{NUM} < 20$	NUM	0.0039	+	+	+	
и IRR *умеренная*	IRR	0.0254	-	+	+	
то *злокачественный*						
Если SIZE *умеренная*	SIZE	0.0150	-	+	+	92.86%
и SHAPE *слабая*	SHAPE	0.0114	-	+	+	
и IRR *слабая*	IRR	0.0878	-	-	+	
то *доброкачественный*						

законов закономерной модели больше либо равны этих значений.

Больший уровень условной вероятности уменьшает число правил и диагностируемых случаев, но увеличивает точность диагноза и близость данным. Было обнаружено 44 статистически значимых диагностических правил с уровнем критерия (f -критерия) 0.05 и условной вероятностью не меньшей 0.75; 30 правил с условной вероятностью не меньшей 0.85; 18 правил с условной вероятностью не меньшей 0.95. Из них 30 правил дали точность 90% и 18 правил дали точность 96.6% с только 3 случаями ошибки второго рода (3.4%).

Как видно из полученных результатов, требуемая близость к данным 0.75, 0.85 и 0.95 меньше, чем получаемая в результате точность 86%, 90% и 96.6% на скользящем контроле. Это является следствием применяемого системой Discovery точного критерия независимости Фишера, используемого для проверки статистической значимости увеличения условных вероятностей в семантическом вероятностном выводе. Это препятствует переобучению системы Discovery. Другие эксперименты [22-23] показывают достаточно высокую устойчивость системы Discovery к переобучению.

В результате обучающему оператору удалось достаточно точно приблизиться к представленным данным. Полученный результат оказался лучше, чем применения нейронных сетей (Brainmaker), которые дали 100% точность на обучении, но на скользящем контроле только 66 %. Решающие деревья (программное обеспечение SIPINA) дали точность 76–82 % на обучении.

9. Сравнение экспертной модели с её аппроксимацией обучающим оператором

Для сравнения модели (правил), полученных системой Discovery с моделью (правилами), извлеченными из эксперта мы попросили эксперта оценить первые из правил. Ниже приводятся некоторые правила, обнаруженные системой

Discovery и комментарии радиолога об их соответствии его модели принятия решений.

ЕСЛИ «общее число кальциноза» больше, чем 30,
и «объем» больше 5 cm^3
и «плотность» кальциноза умеренная,
ТО «подозрение на рак».

Значимость f -критерия – 0.05. Точность диагноза на контроле – 100%. Комментарий радиолога – «это правило перспективное, но может быть рискованным».

ЕСЛИ «вариации в форме кальциноза» заметные
и «число кальциноза» между 10 и 20
и «нерегулярность в форме кальциноза» умеренная,
ТО «подозрение на рак».

Значимость f -критерия – 0.05. Точность диагноза на контроле – 100%. Комментарий радиолога – «я бы доверял этому правилу».

ЕСЛИ «вариации в размере кальциноза» умеренные
и «вариации в форме кальциноза» слабые
и «нерегулярность в форме кальциноза» слабая
ТО «доброкачественный».

Значимость f -критерия – 0.05. Точность диагноза на контроле – 92.86%. Комментарий радиолога – «я бы доверял этому правилу».

Таким образом, обучающий оператор обнаружил правила, которые достаточно хорошо соответствовали интуиции эксперта. Более подробное сравнение обнаруженных правил и экспертных правил приведено в [24].

10. Заключение

В результате нами получена интерпретация теории нейронных моделирующих полей и динамической логики в логико-вероятностных терминах. В обоих подходах удалось найти схожее решение двух коренных проблем искусственного интеллекта – недостаточности формальной логики и проблемы комбинаторной сложности. В обоих подходах, кроме того, используется динамическая логика процесса обучения – приближение к данным путем движения от простого к сложному, а также динамическая оценки качества приближения.

Список литературы

1. Анохин П. К. Опережающее отражение действительности // Философские аспекты теории функциональных систем. М.: Наука, 1978. С. 7–27.
2. Вагин В.Н., Головина Е.Ю., Загорянская А.А., Фомина М.В. Достоверный и правдоподобный вывод в интеллектуальных системах. М., Физматлит., 2004, 704с.
3. Витяев Е.Е. Извлечение знаний из данных. Компьютерное познание. Модели когнитивных процессов. Новосибирский гос. ун-т. Новосибирск, 2006. 293 с.
4. 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.
   URL: http://www.niisi.ru/iont/ni/Journal/V3/N1/Vityaev.pdf
5. Витяев Е.Е. Принятие решений. Переключающая и подкрепляющая функции эмо-

ций // VIII Всероссийская научно-техническая конференция «Нейроинформатика-2006», Сборник научных трудов, Москва, 2006, с.24-30
6. Гибсон Дж. Экологический подход к зрительному восприятию. М.: Прогресс, 1988. С. 462.
7. Демин А.В., Витяев Е.Е., Логическая модель адаптивной системы управления // Нейроинформатика, 2008, том 3, № 1, стр. 79-107. URL: http://www.niisi.ru/iont/ni/Journal/V3/N1/DeminVityaev.pdf
8. Новиков Н.Б. Происхождение человеческой логики // Нейроинформатика, 2010, том 4, № 1, стр. 1-30. URL: http://www.niisi.ru/iont/ni/Journal/V4/N1/Novikov.pdf
9. Перловский Л.И. К физической теории мышления: теория нейронных моделирующих полей // Нейроинформатика, 2006, том 1, № 2, pp. 175-196. URL: http://www.niisi.ru/iont/ni/Journal/N2/Perlovsky.pdf
10. Редько В.Г. Как исследовать происхождение человеческой логики // Нейроинформатика, 2010, том 4, № 1, стр. 31-36. URL: http://www.niisi.ru/iont/ni/Journal/V4/N1/Redko.pdf
11. Русинова Е.В. Пластические перестройки нейронной активности сенсомоторной коры во время выработки клеточного аналога условного рефлекса // Журн. высш. нервн. деят. им. И.П. Павлова. 1977. Т. 27. С. 941–948.
12. Симонов П. В. Эмоциональный мозг. М.: Наука, 1981. С. 140.
13. 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, стр. 340-365.
14. Сперанский С.О. О логической непротиворечивости вероятностных предсказаний // Вестник НГУ. Серия: механика, математика, информатика. Т.11, Вып. 1 (2011), (в печати).
15. Halpern J.Y. An analysis of first-order logics of probability // Artificial Intelligence, 46, 1990, pp. 311-350.
16. G. Hansel. Sur le nombre des fonctions Booleenes monotones de n variables. C.R. Acad. Sci., Paris (in French), 262(20):1088-1090, 1966.
17. Hebb D.O. The organization of behavior. A neurophysiological theory. NY, 1949. 335 p.
18. L. Hyafil, R.L. Rivest, Constructing optimal binary decision trees is NP-Complete. Information Processing Letters 5:1 (1976), 15-17.
19. Jackendoff, R. Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford Univ. Press, New York, NY, 2002.
20. Kovalerchuk B.,Ya., Perlovsky L.I. Dynamic logic of phenomena and cognition // IJCNN, 2008, pp. 3530-3537.
21. Kovalerchuk B, Triantaphyllou E, Despande A, Vityaev E (1996): Interactive Learning of Monotone Boolean Function. Information Sciences 94 (1-4):87-118.
22. Kovalerchuk, B., Vityaev E. Discovering Lawlike Regularities in Financial Time Series, Journal of Computational Intelligence in Finance, v.6, No.3, 1998 pp.12-26.
23. Kovalerchuk B.Ya., Vityaev E.E. Data mining in finance: advances in relational and hybrid methods. Kluwer Academic Publisher, 2000.
24. 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.
25. Perlovsky L.I. Toward physics of the mind: concepts, emotions, consciousness, and symbols // Physics of Life Reviews, 3, 2006, pp. 23-55.
26. 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.
27. Perlovsky, L.I. (1998). Conundrum of Combinatorial Complexity. *IEEE Trans. PAMI*, 20(6)

p.666-70.
28. Probabilistic models of cognition // Special issue of the journal: Trends in cognitive science, v.10, Issue 7, 2006, pp. 287-344
29. Scientific Discovery website \\ http://math.nsc.ru/AP/ScientificDiscovery
30. The Probabilistic Mind. Prospects for Bayesian cognitive science // Eds. Nick Chater, Mike Oaksford, Oxfor University Press, 2008, pp.536
31. Turchin V.F. A constructive interpretation of the full set theory // Journal of Symbolic Logic, 1987. V. 52. No. 1. PP. 172 -201
32. 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.
33. 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.
34. 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: CFP1OUS-DVD, ISBN: 978-1-4244-6917-8, pp. 3361-3366