УДК 510.646 + 510.647

С. О. Сперанский

О ЛОГИЧЕСКОЙ НЕПРОТИВОРЕЧИВОСТИ ВЕРОЯТНОСТНЫХ ПРЕДСКАЗАНИЙ*

Обсуждается проблема статистической двусмысленности (SAP), поставленная Карлом Гемпелем в отношении объяснений, базирующихся на индуктивно-статистической аргументации. С целью устранения SAP вводится формализованное (в терминах логики и вероятности) требование максимальной специфичности (RMS), а также соответствующая ему схема осуществления предсказания. Отметим, определенная в итоге совокупность закономерностей, удовлетворяющих RMS, тесно связана с конструкцией семантического вероятностного предсказания, представленной в работах Евгения Витяева и автора статьи. Установлено, что использование rms-правил в конъюнкции со значимыми относительно рассматриваемой вероятностной меры наборами информационных данных не приводит к возникновению логических противоречий. Иными словами, вероятностные теории, полученные индуктивным образом на основании RMS, выступают в качестве непротиворечивых дедуктивных систем.

Ключевые слова: индуктивно-статистические объяснения и предсказания, проблема статистической двусмысленности, требование максимальной специфичности, вероятностная теория, индуктивная логика, непротиворечивость.

1. Экскурс и обоснование актуальности вопроса

Исследование посвящено работе с *индуктивно произведенным знанием*, синтезированным при совместном использовании логико-вероятностного и обучающего аппаратов [1–3], и сопряженным вопросам. Как известно, логика предикатов первой ступени подразумевает абсолютную достоверность *аксиом* и способов обращения с утверждениями (*правил вывода*). Однако в реальной жизни нам часто приходится оперировать различными видами *нечеткости*, а предположение о тождественной истинности наличествующих средств (их абсолютизация) ведет на практике к увеличению аппроксимационной ошибки округления — эта проблема весьма схожа, например, с возникающими в численном анализе [4], когда «теоретически хорошие» алгоритмы выдают значения, существенно отличные от искомых, и оказываются фактически неприложимы. При моделировании машинных вычислений указанные выше трудности связаны со способами представления величин, хранящихся в памяти компьютера, в то время как для физических и др. теорий — со статистической природой знаний и упрощенными исходными допущениями. Помимо прочего, хорошо известно, что даже маленькая ошибка приближения способна породить весьма большие отклонения результата работы алгоритма от наблюдаемых в действительности значений величин, если задействованы довольно длинные и сложные цепочки логических преобразований. Вдобавок логические модели

* Работа выполнена при частичной поддержке Российского фонда фундаментальных исследований (проект РФФИ-08-07-00272-а) и Совета по грантам Президента РФ и государственной поддержке ведущих научных школ (проект НШ-3606.2010.1)

ISSN 1818-7897. Вестник НГУ. Серия: Математика, механика, информатика. 2011. Т. 11, вып. 1. С. 99–115

наблюдаемых феноменов несовершенны, а от выбранных понятий зависит, каково число шагов, требуемых для достижения ответа — доказательства верности или опровержения суждения, интерпретируемого в онтологии предметной области. Отсюда неизбежно вытекает [1–3]: чисто логическая аргументация обуславливает серьезные ограничения, накладываемые на применимость получаемого *знания*. Согласно Карлу Р. Попперу, обращение к логическому выводу будет справедливо при *дедуктивно-номологических*, но не *индуктивно-статистических* объяснениях [5] (терминология позаимствована у Карла Г. Гемпеля и может быть найдена в его работе [6]).

Итак, в повседневной жизни почти всегда найдутся отклонения (более или менее значительные) от идеальных постулатов теории, причем по аналогии с итерационными методами (в прикладной алгебре) логический вывод склонен накапливать ошибку в процессе своей реализации. При дедуктивном подходе к построению цепочки формул мы оперируем, вообще говоря, *закономерностями статистической природы*, но не имеем в распоряжении «корректной программы» для обращения с ними (классическая схема выводимости хоть и сохраняет истинность, но не сохраняет вероятностных значений). Коль скоро *вероятность* являет собой наиболее утвердившуюся и осознанную формализацию *степени доверия*, информация статистического характера описывается при помощи вероятностных распределений, истинностные значения обобщаются до булевозначных [7; 8] (в силу теоремы Стоуна можно ориентироваться на булевы алгебры подмножеств). Здесь вероятностная мера конструируется (объективно) по статистике, а не по субъективным мнениям экспертов. Отметим, многие люди привыкли думать о вероятности как о размытости, которая может быть сведена к чисто логической модели путем введения большего количества специальных понятий (дополнительного расширения сигнатуры), получаемых в ходе дальнейшего изучения; тем не менее целый ряд дискуссий в основаниях исчисления вероятностей и современной квантовой механике (см. лекцию [9]) указывают на фундаментальный характер случайности, не только устанавливая границы для логического позитивизма и легитимизируя философию индетерминизма, но и обращая наше пристальное внимание на природу явления как такового. Невозможность редукции случайных факторов к атрибутам детерминистского толка и сопутствующая эпистемология излагаются в монографиях [10] и [11].

Наша конечная задача (в общей постановке) — это *извлечение знаний из данных* [2; 3; 12], причем вид индуктивно синтезируемых закономерностей должен быть по возможности прост и удобен в обращении. *Вероятностная аргументация* призвана служить инструментом для предсказания и объяснения свойств малоизученных объектов. В настоящей статье мы сфокусируемся на непротиворечивости совместного использования вероятностных правил и доступных нам данных. Оперирование частичным типом знания (статистической природы) обязывает нас прибегать к индуктивного сорта суждениям. Последние образуют своего рода *макроуровень*, в то время как *микробазой* является исходный материал (например, статистики и модели первого порядка, по которым и была сконструирована вероятностная мера). Обобщение одним уровнем другого не означает лишь поглощения множества элементов или значений истинности, суть обобщения заключается в способности к возврату на первоначальный уровень, к выделению обобщаемого понятия из более общего, в чем нередко возникает необходимость, ибо надстройка пусть и полезна подчас, но сама есть абстракция, при реальном применении нуждающаяся в уточнении контекста. Решения, таким образом, принимаются в рамках традиционной логики (событий микромира). Следовательно, с одной стороны, естественно апеллировать именно к вероятностным рассуждениям (тем более, это в немалой степени может облегчить нагрузку в плане количества предсказываемых свойств, потому что вероятностно независимых признаков обычно сравнительно немного); с другой — важно быть в состоянии проинтерпретировать полученные в результате вероятностной аргументации заключения на логическом уровне предметной области. Выводимые на основе вероятностных соображений теории должны быть *статистически не двусмысленны* (логически непротиворечивы), однако совокупности вероятностных правил часто включают противоречия — указанная проблема и была выявлена Гемпелем в отношении знания, опирающегося на индуктивно-статистические объяснения [6; 13; 14]. Благополучный возврат с макро- на микроуровень будет обеспечен, как только правомерный вопрос о непротиворечивости итоговых предсказаний будет решен положительно.

Несмотря на несомненный интерес к поднимаемой теме согласования и взаимоотношения уровней (их обратной связи), само наличие отрицаний игнорируется во многих работах по синтезу логики и вероятности. Отсутствие таких исследований относит мысли по поводу упомянутого синтеза скорее к разряду неких эвристик, нежели к объективно обоснованной методике обращения с данными. Между тем отрицания крайне важны при вероятностной аргументации. Подробнее опишем возникающие трудности. Разумеется, находясь в рамках чисто логической структуры, имея набор тождественно истинных правил и фактов, мы не испытываем затруднений, связанных со степенями доверия: все ложные формулы (отрицания упомянутых выше фактов и правил) можно отбросить, ведь промежуточных значений истинности нет, а посему нет и смысла в сохранении указанных отрицаний. При работе с вероятностной мерой ситуация меняется категорически: отрицание теперь не только фиксирует жесткую семантическую зависимость P_i и $\neg\mathrm{P}_i$, но и подводит нас к идее о невозможности исключения одного из них из рассмотрения: *a priori* при составлении посылок закономерностей нет никаких причин предпочесть предложение Φ с вероятностью, отличной от нуля, предложению $\neg\Phi$, ведь в конъюнкции с другими формулами каждое может оказаться существенно полезно (без заранее известного приоритета) и специфично. Точнее, будучи прикреплено к посылке, любое из них (при этом не важно, выполнено ли $\mu(\Phi) > \mu(\neg\Phi)$ или обратное неравенство) потенциально способно увеличить условную вероятность: в ходе предсказания для замкнутой формулы Ψ (отличной от Φ) вполне может оказаться

$$\mu(\Psi \mid \neg\Phi) > \mu(\Psi) > \mu(\Psi \mid \Phi)$$

или наоборот («$<$» вместо «$>$»), где μ — *вероятность на формулах* [15; 16].

В представленном вниманию тексте вводятся конструкции, формализующие *требование максимальной специфичности* (RMS), накладываемое на правила и развивающее предложенные Гемпелем концепции по устранению статистической двусмысленности. Интуитивно: применяемые закономерности должны обладать всей специфичной информацией относительно предсказываемых по ним свойств; такая информация оптимизирует правдоподобность и объяснительную силу рассматриваемой закономерности. Хотя идеология, стоящая за требованием, и была представлена в работах Гемпеля [13; 14], в литературе RMS обычно фигурирует в довольно туманной форме либо дается при весьма серьезных ограничениях (скажем, для основных предложений языка монадических предикатов, причем вероятности входящих в аргументацию правил полагаются приблизительно равными единице), от которых в общем случае хотелось бы избавиться. Наше *формальное* RMS, главным образом, связано с парадигмой *семантического вероятностного предсказания* [1–3], а серия доказанных в статье теорем (см. п. 4–6) решает означенную проблему двусмысленности для различных теорий, удовлетворяющих вариантам RMS (нами будет определено несколько разновидностей RMS-требования), и тем самым снимает вопрос о логической непротиворечивости соответствующих данным вариантам вероятностных теорий (см. также [23]).

2. Вероятностная мера над языком 1-й ступени

В первую очередь зафиксируем язык $\mathfrak{L}$ не более чем счетной сигнатуры

$$\Sigma = \{\mathrm{P}_1, \mathrm{P}_2, \dots; f_1, f_2, \dots; \mathrm{c}_1, \mathrm{c}_2, \dots\};$$

с операциями $\wedge$, $\vee$ и $\neg$. Пусть X — счетное множество переменных. Обозначим: $\mathbf{T}_{\mathfrak{L}}$ — множество всех термов языка; $\mathbf{T}^0_{\mathfrak{L}}$ — множество основных термов (термов, не содержащих переменных); $\mathbf{A}_{\mathfrak{L}}$ — множество всех атомов; $\mathbf{A}^0_{\mathfrak{L}}$ — множество основных атомов (атомов, не содержащих переменных); $\mathbf{F}_{\mathfrak{L}}$ — множество всех формул; $\mathbf{S}_{\mathfrak{L}}$ — множество предложений; $\mathbf{S}^0_{\mathfrak{L}}$ — множество основных предложений (бескванторных формул без свободных переменных). Отметим, что $\mathbf{S}^0_{\mathfrak{L}}$ совпадет с замыканием $\mathbf{A}^0_{\mathfrak{L}}$ относительно рассматриваемых логических операций. Затем Θ — совокупность всевозможных подстановок, а $\Theta^0 \leftrightharpoons \{\theta \mid \theta : \mathrm{X} \to \mathbf{T}^0_{\mathfrak{L}}\}$ — подстановки основных термов.

Определение 1. Назовем *литералом* всякий атом, либо его отрицание. Обозначим $\mathbf{L}_{\mathfrak{L}}$ — множество всех литералов; $\mathbf{L}^0_{\mathfrak{L}} \rightleftharpoons \mathbf{L}_{\mathfrak{L}} \cap \mathbf{S}^0_{\mathfrak{L}}$ — основные литералы.

Под *отрицанием* литерала A будем понимать непосредственно $\neg$A, если $\mathrm{A} \in \mathbf{A}_{\mathfrak{L}}$, и B, если $\mathrm{A} \equiv \neg\mathrm{B}$, где $\mathrm{B} \in \mathbf{A}_{\mathfrak{L}}$. На письме используется обозначение «$\neg$A», т. е. двойное отрицание отождествляется с отсутствием таковых. Разрешим присутствие литералов (не только атомов) в качестве структурных элементов *правил* — традиционных объектов логического программирования [17].

Определение 2. *Правило* — это запись вида $\mathrm{C} \leftrightharpoons \tilde{\forall}(\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \dots \wedge \mathrm{B}_n)$, где A, $\mathrm{B}_1, \dots, \mathrm{B}_n$ суть литералы. *Вариантом правила* C называется любое $\mathrm{C}\theta \rightleftharpoons \tilde{\forall}(\mathrm{A}\theta \leftarrow \mathrm{B}_1\theta \wedge \dots \wedge \mathrm{B}_n\theta)$, где θ — произвольная биекция множества переменных на себя. Далее обозначим $\mathtt{Rule}_{\mathfrak{L}}$ — множество всех правил языка $\mathfrak{L}$.

Замечание. Отношение следования «$\leftarrow$» чуть ниже будет нами пояснено и сопровождено вероятностной семантикой; при рассмотрении же вопроса о непротиворечивости связка «$\leftarrow$» в правилах логически интерпретируется как классическая импликация.

Опираясь на анализ, проведенный в работах [15; 16], дадим следующее определение.

Определение 3. *Вероятность* на множестве $F \subseteq \mathbf{S}^0_{\mathfrak{L}}$, замкнутом относительно $\wedge$, $\vee$ и $\neg$, — это функция $\mu : F \to [0, 1]$, обладающая следующими свойствами:

(1) если $\vdash \Phi$ («Φ — тавтология»), то $\mu(\Phi) = 1$;
(2) если $\vdash \neg(\Phi \wedge \Psi)$, то $\mu(\Phi \vee \Psi) = \mu(\Phi) + \mu(\Psi)$.

Отныне считем, что μ — вероятность на основных предложениях. В п. 3 мы также определим μ через *вероятностную меру по Колмогорову* [18], что потенциально позволяет оперировать счетными дизъюнкциями и конъюнкциями, финитно приближать вероятности последних, в пределе получая точное значение.

«Материальность» классической импликации служит аргументом в пользу отказа от введения меры правила как значения вероятности соответствующей замкнутой формулы; вместо этого продуктивным оказывается распространить вероятность на основные экземпляры правил как условную (см. также [2; 16]), а именно

$$\mu(\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \leftrightharpoons \mu(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) = \mu(\mathrm{A}) / \mu(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n).$$

Такое определение более адекватно улавливает причинно-следственную связь между посылкой и заключением. Сопутствующие обозначения:

$$\mathtt{Rule}^0_{\mathfrak{L}} \leftrightharpoons \left\{\mathrm{C}\theta \mid \mathrm{C} \in \mathtt{Rule}_{\mathfrak{L}} \text{ и } \theta \in \Theta^0\right\} \text{ —}$$

основные частные случаи правил рассматриваемого языка $\mathfrak{L}$;

$$\mathtt{Rule}^{0,\mu}_{\mathfrak{L}} \leftrightharpoons \left\{\mathrm{C} \mid \mathrm{C} \leftrightharpoons (\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathtt{Rule}^0_{\mathfrak{L}} \text{ и } \mu(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \neq 0\right\} \text{ —}$$

те из элементов $\mathtt{Rule}^0_{\mathfrak{L}}$, чья условная вероятность определена.

Определение 4. Пусть $\mathrm{C}_1 \leftrightharpoons \tilde{\forall}(\mathrm{A}_1 \leftarrow \bigwedge_{i=1}^{m_1} \mathrm{B}_i)$ и $\mathrm{C}_2 \leftrightharpoons \tilde{\forall}(\mathrm{A}_2 \leftarrow \bigwedge_{k=1}^{m_2} \mathrm{D}_k)$ — некоторые правила (выберем варианты с непересекающимися множествами входящих переменных). Отношение $\mathrm{C}_1 \succ \mathrm{C}_2$ (*«быть более общим»*) имеет место тогда и только тогда, когда можно подобрать такую $\theta \in \Theta$, что $\mathrm{A}_1\theta \equiv \mathrm{A}_2$, $\{\mathrm{B}_1\theta, \ldots, \mathrm{B}_{m_1}\theta\} \subseteq \{\mathrm{D}_1, \ldots, \mathrm{D}_{m_2}\}$, причем $m_1 \leqslant m_2$ и C_1 не является вариантом C_2. Легко видеть, второе правило будет логическим следствием первого (если воспринимать «$\leftarrow$» как материальную импликацию). Обозначим $\succcurlyeq$ — рефлексивное замыкание для $\succ$.

Введем нижнюю оценку для значений вероятности основных случаев правила с переменными:

$$\mathtt{Rule}^{\mu}_{\mathfrak{L}} \rightleftharpoons \left\{\mathrm{C} \in \mathtt{Rule}_{\mathfrak{L}} \mid \mathrm{C}\theta \in \mathtt{Rule}^{0,\mu}_{\mathfrak{L}} \text{ для какого-нибудь } \theta \in \Theta^0\right\};$$

$$\underline{\mu}(\mathrm{C}) \rightleftharpoons \inf\left\{\mu(\mathrm{C}\theta) \mid \theta \in \Theta^0 \text{ и } \mathrm{C}\theta \in \mathtt{Rule}^{0,\mu}_{\mathfrak{L}}\right\}, \quad \text{где } \mathrm{C} \in \mathtt{Rule}^{\mu}_{\mathfrak{L}}.$$

Определение 5. Отношение $\mathrm{C}_1 \sqsubset \mathrm{C}_2$ (*«вероятностное следование»*) для $\{\mathrm{C}_1, \mathrm{C}_2\} \subseteq \mathtt{Rule}^{\mu}_{\mathfrak{L}}$ эквивалентно $\mathrm{C}_1 \succ \mathrm{C}_2$ вместе с $\underline{\mu}(\mathrm{C}_1) < \underline{\mu}(\mathrm{C}_2)$.

Определение 6. Назовем μ-*законом* всякое C из $\mathtt{Rule}^{\mu}_{\mathfrak{L}}$, для которого $\mathrm{C}' \succ \mathrm{C}$ влечет $\mathrm{C}' \sqsubset \mathrm{C}$ при любом $\mathrm{C}' \in \mathtt{Rule}^{\mu}_{\mathfrak{L}}$; обозначим $\mathtt{GLaw}^{\mu}_{\mathfrak{L}}$ — совокупность всех μ-законов.

3. Семантическое μ-предсказание на данных

Для начала обоснуем задачу предсказания свойств, выраженных формулами в языке $\mathfrak{L}$, а затем перейдем к ее формализации и постановке вопроса непротиворечивости.

Чаще всего *случайной величиной* называется измеримая функция из пространства элементарных исходов в $\mathbb{R}$ (см. [18]), тогда статистические выборки представляют собой конечные последовательности вещественных чисел. Из желания подчеркнуть, что всякий исход может иметь довольно произвольную природу, вместо $\mathbb{R}$ (или $\mathbb{Q}$) мы воспользуемся некоторым классом моделей первого порядка. Подобный переход, разумеется, не позволяет непосредственно производить алгебраические операции и сравнивать значения случайных величин на вещественной прямой, поэтому подразумевается наличие взаимно-однозначного соответствия (своего рода шкалирования) между рассматриваемым классом и поднабором из $\mathbb{R}$ (при надобности отождествляя одно с другим).

Пусть $\mathfrak{G}^*$ — некоторый счетный класс алгебраических систем первого порядка, наблюдаемых в изучаемом нами явлении; иными словами, любой исход эксперимента трактуется как некий элемент из $\mathfrak{G}^*$ (однако нам не обязательно известно, какой конкретно). Непосредственно осуществлению предсказания предшествует ряд этапов:

- построение вероятностной меры μ по репрезентативной выборке $\mathfrak{G} \subset \mathfrak{G}^*$, состоящей из хорошо исследованных структур, чьи свойства нами полностью идентифицированы;
- индуктивный синтез предсказывающих закономерностей, а вместе с тем и подбор статистически значимых данных для изучаемой нами системы $\mathfrak{B} \in \mathfrak{G}^*$ — в соответствии с ними будут приниматься / отклоняться гипотезы относительно свойств последней.

Остановимся подробнее на описании указанных выше этапов. Первый из них сопровождается применением аппарата математической статистики, в результате чего получаем вероятностное распределение P, действующее на пространстве элементарных исходов $\Omega = \mathfrak{G}^*$ (при обобщении здесь потребуется колмогоровская аксиоматизация); в том или ином смысле оно должно аппроксимировать «настоящие» значения *эмпирической вероятности* P^* (подробнее о теории интервального и точечного оценивания можно прочесть в монографии А. А. Боровкова [19], для нас же это сейчас не столь важно). Качество *репрезентативности* выборки из генеральной совокупности обеспечивает нашу возможность говорить о вероятностной мере в целом, заменяя в рассуждениях P^* на P. Имея P, для произвольного предложения Φ определяем

$$\mu(\Phi) \leftrightharpoons \mathsf{P}(\llbracket\Phi\rrbracket) = \sum_{\omega\in\llbracket\Phi\rrbracket} \mathsf{P}(\omega), \quad \text{где} \quad \llbracket\Phi\rrbracket \leftrightharpoons \{\mathfrak{A} \in \mathfrak{G}^* \mid \mathfrak{A} \models \Phi\}.$$

Нам понадобится только μ на основных предложениях. Нетрудно понять, все необходимые свойства вероятности на $\mathbf{S}^0_{\mathfrak{L}}$ (см. определение 3) выполнены. В приведенной схеме происходит естественный переход от двузначной истинности к булевозначной (ибо $2^{\mathfrak{G}^*}$ есть булева алгебра), которую, в свою очередь, можно «померить» при помощи P.

Далее берется интересующая нас $\mathfrak{B} \in \mathfrak{G}^*$ (знание относительно ее свойств частично, т. е. наблюдается некоторая модель, однако все ее характеристики нам не известны) и ставится задача предсказания, точнее — для произвольного основного литерала H выяснить, что более достоверно: $\mathfrak{B} \vDash \mathrm{H}$, $\mathfrak{B} \nvDash \mathrm{H}$ или у нас недостаточно информации для какого-либо из представленных выводов; причем в первых двух случаях приводится *вероятностная оценка* степени правдоподобия. Ниже будут определены правила и наборы данных, по которым эта задача решается в *семантическом* μ-*предсказании* [1–3].

Зафиксируем $\mathtt{Fact}_{\mathrm{o}} \subset \mathbf{A}^0_{\mathfrak{L}}$ — те из основных атомов, что позволяют непосредственную проверку в любой реально возникающей модели $\mathfrak{B} \in \mathfrak{G}^*$; иначе говоря, постулируется наличие $\zeta : \mathfrak{G}^* \times \mathtt{Fact}_{\mathrm{o}} \to \{0,1\}$ (считаем, что задано произвольное такое отображение). Нами также используется аббревиатура $\zeta_{\mathfrak{B}}(\cdot)$ для одноместной функции $\zeta(\mathfrak{B},\cdot)$.

Определим *полный набор альтернатив*, допускающих проверку, как

$$\mathtt{Fact}^*_{\mathrm{o}} \leftrightharpoons \mathtt{Fact}_{\mathrm{o}} \cup \mathtt{Fact}^{\neg}_{\mathrm{o}}, \quad \text{где} \quad \mathtt{Fact}^{\neg}_{\mathrm{o}} \rightleftharpoons \{\neg \mathrm{A} \mid \mathrm{A} \in \mathtt{Fact}_{\mathrm{o}}\}\,.$$

Теперь положим (не накладывая дополнительных ограничений на тип функции $\zeta_{\mathfrak{B}}$)

$$\mathtt{Data}[\mathfrak{B}] \rightleftharpoons \{\mathrm{A} \mid \mathrm{A} \in \mathtt{Fact}_{\mathrm{o}} \ \textit{и}\ \zeta_{\mathfrak{B}}(\mathrm{A}) = 1\} \cup \{\neg\mathrm{A} \mid \mathrm{A} \in \mathtt{Fact}_{\mathrm{o}} \ \textit{и}\ \zeta_{\mathfrak{B}}(\mathrm{A}) = 0\}\,.$$

Итак, *данные* являют собой не что иное, как максимальное непротиворечивое подмножество полного набора альтернатив.

Определение 7. Подмножество $S \subseteq \mathbf{L}^0_{\mathfrak{L}}$ называется μ-*совместным* или *статистически значимым относительно* μ (для вероятности μ, построенной по распределению P) в том и только том случае, когда $\mathsf{P}(\{\mathfrak{A} \mid \mathfrak{A} \vDash S\}) \neq 0$.

Определение 8. *Максимально специфичным правилом (ms-правилом) для предсказания литерала* H называется $\mathrm{C} \leftrightharpoons \tilde{\forall}(\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathtt{Rule}_{\mathfrak{L}}$, обладающее следующими характеристиками:

(1) найдется подстановка $\theta \in \Theta^0$ такая, что $\mathrm{A}\theta \equiv \mathrm{H}\theta$, $\{\mathrm{B}_1\theta, \ldots, \mathrm{B}_n\theta\} \subseteq \mathtt{Fact}^*_{\mathrm{o}}$, $\mu(\mathrm{B}_1\theta \wedge \cdots \wedge \mathrm{B}_n\theta) \neq 0$ и $\mu(\mathrm{C}\theta) > \mu(\mathrm{H}\theta)$;

(2) не существует $\mathrm{C}' \in \mathtt{Rule}^{\mu}_{\mathfrak{L}}$, для которого верно (1) и $\mathrm{C} \sqsubset \mathrm{C}'$;

(3) его нельзя обобщить до некоторого $\mathrm{C}' \in \mathtt{Rule}^{\mu}_{\mathfrak{L}}$, не потеряв при этом в значении $\underline{\mu}(\cdot)$ и одновременно сохранив свойства (1–2).

Иногда, говоря о введенных нами условиях, мы будем употреблять аббревиатуру «ms.i» для обозначения i-го пункта ms-определения (когда из контекста ясно, о предсказании какого именно литерала идет речь).

Интуиция, стоящая за пунктами 1–3 определения, такова: ms.1 отвечает за соответствие предсказываемому свойству и проверяемость посылки правила на возникающих данных; ms.2 — непосредственно за максимальную специфичность в классе правил, чьи посылки и заголовки валидны по п.1; ms.3 — за простоту (общность) закономерности. Кроме того, во избежание деления на ноль возникает техническое требование «$\mu(\mathrm{B}_1\theta \wedge \cdots \wedge \mathrm{B}_n\theta) > 0$», а также «$\mu(\mathrm{C}\theta) > \mu(\mathrm{H}\theta)$», символизирующее производительность набора причин $\{\mathrm{B}_1\theta, \ldots, \mathrm{B}_n\theta\ \}$ и называемое условием *мотивированности.* Последнее можно переписать как $\mu(\mathrm{A}\theta \mid \mathrm{B}_1\theta \wedge \cdots \wedge \mathrm{B}_n\theta) > \mu(\mathrm{H}\theta \mid \top)$, где $\top$ — символ тавтологии (пустой конъюнкции). Конечно же, нет смысла предсказывать литерал по правилу, оценка вероятности которого хуже, чем безусловная вероятность самого литерала. Мы еще вернемся к анализу $\mathsf{ms.1}$–$\mathsf{3}$ в некоторых модификациях, пока же заметим, что условия мотивации, сходные с упомянутым выше, встречаются и в вероятностной теории причинности Суппеса [20], и в индуктивной логике Карнапа [21].

Предположим, нас интересует правдоподобность $\mathrm{H} \in \mathbf{L}^0_{\mathfrak{L}}$. Сперва мы находим ms-правило $\mathrm{C} \leftrightharpoons \tilde{\forall}(\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n)$ для предсказания H, причем должна существовать $\theta \in \Theta^0$ такая, что $\{\mathrm{B}_1\theta, \ldots, \mathrm{B}_n\theta\}$ содержится в $\mathtt{Data}\,[\mathfrak{B}]$. Тогда свойство H подтверждается в модели $\mathfrak{B}$ с оценкой $\gamma\,(\mathrm{H}) \leftrightharpoons \underline{\mu}\,(\mathrm{C})$. Последняя может служить критерием принятия или отклонения гипотезы об истинности запроса: например, доверяем лишь прогнозам с оценкой не меньше $1 - \varepsilon$, где ε — *доверительная граница*. Аналогично поступаем с предсказанием $\neg\mathrm{H}$. При наличии нескольких ms-правил естественно выбрать то, которое доставляет максимум величине $\mu\,(\cdot)$; отметим, на практике все множество ms-правил у нас обычно отсутствует (их нахождение наталкивается на проблемы вычислительной сложности), поэтому часто приходится оперировать поднаборами. Разумеется, особый интерес вызывает предвосхищение свойств, не лежащих в $\mathtt{Fact}^*_{\mathrm{o}}$.

Для каждого ms-правила C, предсказывающего некоторый литерал H, рассмотрим всевозможные экземпляры $\mathrm{C}\theta$, где θ принадлежит к числу подстановок, удовлетворяющих п.1 ms-определения для C и H. Объединение всех таких основных случаев (по всем правилам и литералам) обозначим $\mathsf{MSR}^{\mu}_{\theta}$. Множество $\mathsf{MSR}^{\mu}_{\theta}$ можно считать итогом проведения так называемого *семантического* μ*-вывода*, идея которого была предложена в работах Е. Е. Витяева [1–3] и основывается на *требовании максимальной специфичности*, выдвинутом К. Г. Гемпелем для устранения *проблемы статистической двусмысленности* при обращении с индуктивно произведенным знанием [6; 13; 14]. Цель последующих разделов состоит в доказательстве логической непротиворечивости $\mathsf{MSR}^{\mu}_{\theta} \cup \mathtt{Data}\,[\mathfrak{B}]$, где $\mathtt{Data}\,[\mathfrak{B}]$ статистически значимо. Отметим, что в разных вариациях непротиворечивость индуктивных схем постоянно фигурирует в философии науки: у Гемпеля — как необходимое условие при формировании рациональной системы убеждений [6]; у М. Родерика и Р. Мартина — как адаптация эпистемологических принципов, также принятых в концепции знания Я. Хинтикки [22]. Кроме того, в пятом параграфе конструкция ms-правила будет обобщена для введения переменных (ибо, хотя ms-правила и включают переменные, но в действительности закономерности присутствуют лишь в виде $\mathsf{MSR}^{\mu}_{\theta}$ — соответствующих основных экземпляров); наконец, в шестом параграфе мы обсудим слабомотивированные предсказания с доверительным уровнем $1-\varepsilon$ и установим, что для гарантии совместности достаточно взять любое $\varepsilon < 0{,}5$.

4. Доказательство непротиворечивости

Утверждение 1. *Допустим,* $\mathrm{C} \leftrightharpoons (\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) \in \mathtt{Rule}^{0,\mu}_{\mathfrak{L}}$, *причем имеет место строгое неравенство*

$$\mu\,(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) < \mu\,(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n)\,.$$

Тогда правило $\mathrm{C}' \leftrightharpoons (\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D})$ *лежит в* $\mathtt{Rule}^{0,\mu}_{\mathfrak{L}}$, *а* $\neg\mathrm{D}$ *уже увеличивает условную вероятность*

$$\mu\,(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D}) > \mu\,(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n)\,.$$

ДОКАЗАТЕЛЬСТВО. Если бы $\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D})$ равнялось нулю, то имеем

$$\mu\,(\mathrm{A} \wedge \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) = \mu\,(\mathrm{A} \wedge \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) \; + \\ + \; \mu\,(\mathrm{A} \wedge \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D}) = \mu\,(\mathrm{A} \wedge \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n),$$

а также

$$\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) = \mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}) \; + \\ + \; \mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D}) = \mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n),$$

откуда (с учетом условий)

$$\mu\,(\mathrm{C}) = \tfrac{\mu(\mathrm{A}\wedge\mathrm{B}_1\wedge\cdots\wedge\mathrm{B}_n\wedge\mathrm{D})}{\mu(\mathrm{B}_1\wedge\cdots\wedge\mathrm{B}_n\wedge\mathrm{D})} = \tfrac{\mu(\mathrm{A}\wedge\mathrm{B}_1\wedge\cdots\wedge\mathrm{B}_n)}{\mu(\mathrm{B}_1\wedge\cdots\wedge\mathrm{B}_n)} = \mu\,(\mathrm{A} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) > \mu\,(\mathrm{C}).$$

Возникает противоречие. Значит, $\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D}) > 0$, т. е. и условная вероятность для C' определена. В результате дальнейших преобразований получим:

$$\begin{gathered}
\mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D}\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D}\right) = \\
= \mu\left(\mathrm{A} \wedge \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D}\right) = \mu\left(\mathrm{A} \wedge \bigwedge_{i=1}^{n} \mathrm{B}_i\right) - \mu\left(\mathrm{A} \wedge \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}\right) = \\
= \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i\right) - \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}\right) > \\
> \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i\right) - \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}\right) = \\
= \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) \times \left(\mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i\right) - \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}\right)\right) = \\
= \mu\left(\mathrm{A} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) \times \mu\left(\bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D}\right).
\end{gathered}$$

И, таким образом, вслед за сокращением $\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \neg\mathrm{D}) \neq 0$ требуемое нам неравенство будет установлено. □

Лемма 1 (об альтернативах). *Допустим,* $\mathrm{C}_{\mathrm{pos}} \leftrightharpoons (\mathrm{H} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathsf{MSR}^{\mu}_{\theta}$ *и* $\mathrm{C}_{\mathrm{neg}} \leftrightharpoons (\neg\mathrm{H} \leftarrow \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) \in \mathsf{MSR}^{\mu}_{\theta}$ *(заключения ms-экземпляров противоречат друг другу). Тогда с необходимостью* $\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) = 0$.

ДОКАЗАТЕЛЬСТВО. Ясно, что $\mathrm{C}_{\mathrm{pos}}$ и $\mathrm{C}_{\mathrm{neg}}$ суть основные случаи некоторых ms-правил C_1 и C_2, причем для каждого выполнен соответствующий ему п. 1 определения ms-правила. Очевидно, $\mathrm{C}_1 \succcurlyeq \mathrm{C}_{\mathrm{pos}}$. Отсюда (в силу ms.2) имеем $\underline{\mu}\,(\mathrm{C}_1) \geqslant \mu\,(\mathrm{C}_{\mathrm{pos}})$; с другой стороны, $\underline{\mu}\,(\mathrm{C}_1) \leqslant \mu\,(\mathrm{C}_{\mathrm{pos}})$ — всегда будет выполнено для нижней границы $\underline{\mu}\,(\cdot)$. Таким образом, $\mu\,(\mathrm{C}_{\mathrm{pos}}) = \underline{\mu}\,(\mathrm{C}_1)$. Тем же методом получим и $\mu\,(\mathrm{C}_{\mathrm{neg}}) = \underline{\mu}\,(\mathrm{C}_2)$.

«От противного». Пусть $\mu\,(\mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) \neq 0$. Следовательно, и для любого подмножества $S \subseteq \{\mathrm{B}_1, \ldots, \mathrm{B}_n, \mathrm{D}_1, \ldots, \mathrm{D}_m\}$ имеет место $\mu\left(\bigwedge_{\mathrm{A}\in S}\mathrm{A}\right) \neq 0$. Далее зафиксируем какое-нибудь $\mathrm{D} \in \{\mathrm{D}_1, \ldots, \mathrm{D}_m\}$.

Этап 1. Сначала докажем равенство $\mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i) = \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D})$. В самом деле, $\mathrm{C} \leftrightharpoons (\mathrm{H} \leftarrow \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D})$ принадлежит $\mathtt{Rule}_{\mathfrak{L}}^{0,\mu}$, к тому же верно $\mathrm{C}_1 \succcurlyeq \mathrm{C}_{\mathrm{pos}} \succ \mathrm{C}$. Легко понять, это C полностью удовлетворяет условиям п. 1 определения ms-правила (для литерала, предсказываемого с помощью C_1) за исключением, быть может, «$\mu\,(\mathrm{C}) > \mu\,(\mathrm{H})$». Рассмотрим возможные альтернативы.

1. $\underline{\mu}\,(\mathrm{C}_1) = \mu\,(\mathrm{C}_{\mathrm{pos}}) = \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i) \leqslant \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D})$. Из п. 2 определения ms-правила, однако, вытекает обратное неравенство (ибо появилось недостающее условие «мотивации» $\mu\,(\mathrm{C}) \geqslant \mu\,(\mathrm{C}_{\mathrm{pos}}) > \mu\,(\mathrm{H})$); в итоге установим нужное равенство.

2. $\underline{\mu}\,(\mathrm{C}_1) = \mu\,(\mathrm{C}_{\mathrm{pos}}) = \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i) > \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D})$. Исходя из утверждения 1, $\mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i) < \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D})$, где $\mathrm{C}' \leftrightharpoons (\mathrm{H} \leftarrow \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D}) \in \mathtt{Rule}_{\mathfrak{L}}^{0,\mu}$. Повторяя предыдущие рассуждения (теперь с заменой «D» на «$\neg$D»), приходим к противоречащему допущениям тождеству $\mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i) = \mu\,(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \neg\mathrm{D})$.

Этап 2. Ежели нами уже добавлена часть литералов $\{\mathrm{D}_1, \dots, \mathrm{D}_m\}$, берем очередное (по порядку) и еще не задействованное D_k. Индукционное предположение гласит

$$\underline{\mu}\,(\mathrm{C}_1) = \mu\left(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i\right) = \mu\left(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \mathrm{D}_1\right) = \dots = \mu\left(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \bigwedge_{j=1}^{k-1} \mathrm{D}_j\right).$$

Для последнего из выражений (внутри скобок) и D_k выкладки будут без изменений по отношению к проделанным ранее вплоть до момента наступления разобранных выше альтернатив (1–2) и не используют никаких дополнительных свойств. В (1–2) необходимо аккуратно применить индукционную гипотезу, после чего действия становятся тривиальны. Итак,

$$\underline{\mu}\,(\mathrm{C}_1) = \mu\left(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \bigwedge_{j=1}^{m} \mathrm{D}_j\right).$$

Аналогично

$$\underline{\mu}\,(\mathrm{C}_1) = \mu\left(\neg\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \bigwedge_{j=1}^{m} \mathrm{D}_j\right).$$

В результате нетрудно извлечь (с привлечением одной из составляющих п. 1 определения ms-правила для предсказания) противоречие:

$$1 = \mu\,(\neg\mathrm{H}) + \mu\,(\mathrm{H}) < \mu\,(\mathrm{C}_{\mathrm{pos}}) + \mu\,(\mathrm{C}_{\mathrm{neg}}) = \\ = \mu\left(\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \bigwedge_{j=1}^{m} \mathrm{D}_j\right) + \mu\left(\neg\mathrm{H} \mid \bigwedge_{i=1}^{n} \mathrm{B}_i \wedge \bigwedge_{j=1}^{m} \mathrm{D}_j\right) = 1.$$

□

Замечание. Из доказательства леммы видно, что нами установлен следующий факт. Пусть $\mathrm{C} \leftrightharpoons (\mathrm{H} \leftarrow \mathrm{B}_1 \wedge \dots \wedge \mathrm{B}_n) \in \mathsf{MSR}_{\theta}^{\mu}$ и $\{\mathrm{B}_1 \cdots \mathrm{B}_n\} \subseteq S \subseteq \mathtt{Fact}_{\mathrm{o}}^{*}$, где S конечно и μ-совместно. Тогда окажется выполненым равенство $\mu\,(\mathrm{C}) = \mu\left(\mathrm{H} \mid \bigwedge_{\mathrm{B} \in S} \mathrm{B}\right)$.

$$\mathsf{Prdct}\,[\mathfrak{B}]_{\theta}^{\mu} \rightleftharpoons \mathsf{MSR}_{\theta}^{\mu} \cup \mathtt{Data}\,[\mathfrak{B}].$$

Теорема 1 (о непротиворечивости ms-теории). *Теория* $\mathsf{Prdct}\,[\mathfrak{B}]_{\theta}^{\mu}$ *логически непротиворечива для всякого* μ*-совместного множества* $\mathtt{Data}\,[\mathfrak{B}]$.

ДОКАЗАТЕЛЬСТВО. Для выполнимости $\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta}$ достаточно показать выполнимость множества

$$\begin{aligned}&\Gamma \ \leftrightharpoons \mathsf{Pr} \cup \mathsf{Dt}_1, \quad \text{где} \quad \mathsf{Dt}_1 \leftrightharpoons \big\{\,\mathrm{D} \mid \mathrm{D} \in \mathtt{Data}\,[\mathfrak{B}] \ \text{ и } \ \mu\,(\mathrm{D}) = 1\,\big\},\\ &\mathsf{Pr} \leftrightharpoons \big\{\,\mathrm{A} \mid (\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathsf{MSR}^{\mu}_{\theta}, \quad \{\mathrm{B}_1, \ldots, \mathrm{B}_n\} \subseteq \mathtt{Data}\,[\mathfrak{B}]\,\big\},\end{aligned}$$

ибо $\mathsf{Dt}_{<} \leftrightharpoons \{\,\mathrm{D} \mid \mathrm{D} \in \mathtt{Data}\,[\mathfrak{B}],\ \mu\,(\mathrm{D}) < 1\} \subseteq \mathsf{Pr}$. В самом деле, для каждого $\mathrm{D} \in \mathtt{Data}\,[\mathfrak{B}]$ команда $\mathrm{C} \leftrightharpoons (\mathrm{D} \leftarrow \mathrm{D})$ лежит в $\mathtt{Rule}^{0,\mu}_{\mathfrak{L}}$ (ввиду μ-совместности данных). Далее указанное правило обобщается до некоторого $\mathrm{C}' \in \mathtt{GLaw}^{\mu}_{\mathfrak{L}}$ без потери в условной вероятности: $\underline{\mu}\,(\mathrm{C}') = \mu\,(\mathrm{D} \mid \mathrm{D}) = 1 > \mu\,(\mathrm{D})$ (это всегда можно сделать). Оценка полученного C' не может быть улучшена, т. е. получено ms-правило, причем его частный случай C предсказывает D. Кроме того, в модели для Γ окажутся реализованными все данные и заключения импликаций $\mathsf{MSR}^{\mu}_{\theta}$ с посылками из $\mathtt{Data}\,[\mathfrak{B}]$ (именно они участвовали в формировании Pr). Посылка всякого правила $\mathsf{MSR}^{\mu}_{\theta}$ лежит в $\mathtt{Fact}^{*}_{\mathrm{o}}$. Стало быть, если она не включена в данные полностью, то найдутся литералы дополнения: такие литералы, что их отрицания принадлежат $\mathtt{Data}\,[\mathfrak{B}]$, — отсюда посылка рассматриваемого правила будет ложна, а сама импликация из $\mathsf{MSR}^{\mu}_{\theta}$, разумеется, истинна. Поэтому эрбранова модель для Γ будет моделью исходного $\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta}$.

Отметим, что $\mathsf{Pr} \cap \mathsf{Dt}_1 = \varnothing$. Действительно, любое D из данных, обладающее свойством $\mu\,(\mathrm{D}) = 1$, нельзя семантически μ-предсказать никаким ms-правилом (при сильной мотивации), ведь не существует правила с условной вероятностью большей единицы. Одновременно, коль скоро значение $\mu\,(\neg\mathrm{D}) = 1 - \mu\,(\mathrm{D}) = 0$ также нельзя увеличить (для произвольного основного экземпляра $(\neg\mathrm{D} \leftarrow \mathrm{D}_1, \ldots, \mathrm{D}_m) \in \mathtt{Rule}^{0,\mu}_{\mathfrak{L}}$ будем иметь $\mu\,(\neg\mathrm{D} \mid \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) = \mu\,(\neg\mathrm{D} \wedge \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m)\,/\,\mu\,(\mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) = 0$), то и отрицание D не может попасть в Pr.

Из сказанного вытекает: Γ будет логически совместно, как только совместны (в отдельности) Dt_1 и Pr. Проверим это. Множество Dt_1 непротиворечиво по условию (иначе бы вероятность данных равнялась нулю). Pr же непротиворечиво в силу леммы 1 и статистической значимости набора $\mathtt{Data}\,[\mathfrak{B}]$. Теперь зададим интерпретацию $v : \mathbf{A}^{0}_{\mathfrak{L}} \to \{0, 1\}$ согласно инструкции

$$v\,(\mathrm{A}) \rightleftharpoons \begin{cases} 1, \ \text{если верно } \mathrm{A} \in \mathsf{Pr} \text{ либо } \mathrm{A} \in \mathsf{Dt}_1 \\ 0, \ \text{если верно } \neg\mathrm{A} \in \mathsf{Pr} \text{ либо } \neg\mathrm{A} \in \mathsf{Dt}_1 \\ \text{произвольно, на оставшихся атомах.} \end{cases}$$

Значит, существует эрбранова модель для Γ. В итоге множество $\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta}$ выполнимо. Следовательно, и непротиворечиво. □

Замечание. В ходе доказательства леммы и теоремы нам фактически не понадобился п. 3 определения ms-правила, отвечающий за общность. Более того, ежели мы рассмотрим совокупность правил, удовлетворяющих только первым двум пунктам ms-определения (для каких-нибудь предсказываемых литералов), а затем соберем все их основные экземпляры по соответствующим п. 1 подстановкам (см. аналогию с образованием $\mathsf{MSR}^{\mu}_{\theta}$) в одно множество, то последнее будет логически следовать из $\mathsf{MSR}^{\mu}_{\theta}$. Убедимся в этом. Пусть $\mathrm{C} \leftrightharpoons \tilde{\forall}\,(\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n)$ обладает ms.1–2 для $\mathrm{H} \in \mathbf{L}_{\mathfrak{L}}$. Тогда оно может быть обобщено (без потери в нижней границе $\underline{\mu}\,(\cdot)$ и при сохранении $\mathsf{ms.1\text{–}2}$) до некоторого ms-правила $\mathrm{C}' \leftrightharpoons \tilde{\forall}\,(\mathrm{A}' \leftarrow \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m)$. Далее для θ (из $\mathsf{ms.1}$ для C и H) всегда найдется θ' такая, что $\mathrm{A}\theta \equiv \mathrm{A}'\theta'$ и $\{\mathrm{D}_1\theta', \ldots, \mathrm{D}_m\theta'\} \subseteq \{\mathrm{B}_1\theta, \ldots, \mathrm{B}_n\theta\}$. Учитывая равенство $\underline{\mu}\,(\mathrm{C}) = \mu\,(\mathrm{C}\theta)$ (следует из выбора θ и $\mathsf{ms.2}$), по построению имеем:

$$\mu\left(\mathrm{C}'\theta'\right) \geqslant \underline{\mu}\left(\mathrm{C}'\right) \geqslant \underline{\mu}\,(\mathrm{C}) = \mu\,(\mathrm{C}\theta) > \mu\,(\mathrm{A}\theta) = \mu\left(\mathrm{A}'\theta'\right),$$

откуда $\mathrm{C}'\theta' \in \mathsf{MSR}^{\mu}_{\theta}$, причем $\mathrm{C}'\theta' \succcurlyeq \mathrm{C}\theta$ и выполнено $\mu\,(\mathrm{C}'\theta') = \underline{\mu}\,(\mathrm{C}')$.

○ Дополнительно: схема предсказания в теореме 1.

Статистическая значимость множества $\mathtt{Data}\,[\mathfrak{B}]$ означает ненулевую вероятность конъюнкций посылок ms-экземпляров, предсказывающих на данных, а точнее — правил

$$\mathsf{Pr}' \leftrightharpoons \left\{\mathrm{C} \mid \mathrm{C} \leftrightharpoons (\mathrm{A} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathsf{MSR}^{\mu}_{\theta},\ \{\mathrm{B}_1, \ldots, \mathrm{B}_n\} \subseteq \mathtt{Data}\,[\mathfrak{B}]\right\},$$

так как по монотонности $\mu\left(\bigwedge_{\mathrm{B}\in S'}\mathrm{B}\right) \geqslant \mu\left(\bigwedge_{\mathrm{D}\in S}\mathrm{D}\right)$, где $S' \subseteq S \subseteq \mathbf{L}^0_{\mathfrak{L}}$.

Зафиксируем некоторую нумерацию всех основных атомов сигнатуры Σ и выберем очередное A_i. Возникает одна из следующих альтернатив:

1) атом A_i семантически μ-предсказывается на данных системы $\mathfrak{B}$;

2) литерал $\neg\mathrm{A}_i$ семантически μ-предсказывается на данных системы $\mathfrak{B}$;

3) ни сам рассматриваемый атом, ни его отрицание не могут быть семантически μ-предсказаны средствами $\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta}$.

Случаи (1–3) исчерпывают возможные ситуации и взаимно исключают друг друга. Очевидно, первый из них (по определению) эквивалентен $\mathrm{A}_i \in \mathsf{Pr}$, второй равносилен $\neg\mathrm{A}_i \in \mathsf{Pr}$, а в третьем будут содержаться (в частности) все элементы Dt_1 (обозначения см. в доказательстве теоремы).

5. Универсальная версия ms-теории

Подобно $\mathtt{Fact}_{\mathrm{o}}$, определенному в третьем разделе, зададим $\mathtt{Fact}_v \subset \mathbf{A}_{\mathfrak{L}}$ — атомы (не обязательно основные), позволяющие проверку в изучаемых нами системах $\mathfrak{G}^*$. Естественно считать, что при этом мы в состоянии верифицировать и фальсифицировать истинность любых элементов

$$\mathtt{Fact}_v\Theta^0 \leftrightharpoons \left\{\mathrm{A}\theta \mid \theta \in \Theta^0,\ \mathrm{A} \in \mathtt{Fact}_v\right\}.$$

Как и ранее, мы обозначим $\mathtt{Fact}^*_v \leftrightharpoons \mathtt{Fact}_v \cup \{\neg\mathrm{A} \mid \mathrm{A} \in \mathtt{Fact}_v\}$ — замыкание множества $\mathtt{Fact}_v$ относительно отрицаний.

Пусть $\mathrm{C} \leftrightharpoons \tilde{\forall}\,(\mathrm{A} \leftarrow \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) \in \mathtt{Rule}_{\mathfrak{L}}$. Среди подстановок выделим

$$\mathtt{Sub}\,[\mathrm{C}]^{\mu} \leftrightharpoons \left\{\theta \in \Theta^0 \mid \mu\,(\mathrm{D}_1\theta \wedge \cdots \wedge \mathrm{D}_m\theta) > 0\right\} \ —$$

те из основных θ, при которых условная вероятность экземпляра $\mathrm{C}\theta$ определена.

Определение 9. Привило $\mathrm{C} \leftrightharpoons \tilde{\forall}\,(\mathrm{A} \leftarrow \bigwedge_{i=0}^{n}\mathrm{B}_i)$ называется *максимально специфичным u-правилом (ums-правилом) для предсказания литерала* H в том и лишь том случае, когда выполнены следующие условия:

(1) найдется подстановка θ (не обязательно основная) такая, что $A \equiv H\theta$, $\{B_1, \dots, B_n\} \subseteq \{B\theta \mid B \in \mathtt{Fact}_v^*\}$;

(2) если $\theta_o \in \mathtt{Sub}\,[C]^\mu$, то $\mu\,(C\theta_o) > \mu\,(H\theta_o)$;

(3) не существует $C' \in \mathtt{Rule}_{\mathfrak{L}}^\mu$, для которого верно (1–2) и $C \sqsubset C'$;

(4) правило C нельзя обобщить до некоторого $C' \in \mathtt{Rule}_{\mathfrak{L}}^\mu$, не потеряв при этом в значении $\underline{\mu}\,(\cdot)$ и одновременно сохранив свойства (1–3).

Пусть MSR_u^μ — всевозможные ums-правила (пробегая по предсказываемым литералам); также воспользуемся сокращением «ums.i» для обозначения i-го пункта определения. Теперь $\mathtt{Data}\,[\mathfrak{B}]$ суть максимальное непротиворечивое подмножество полного набора альтернатив для $\mathtt{Fact}_o \rightleftharpoons \mathtt{Fact}_v\Theta^0$.

$$\mathsf{Prdct}\,[\mathfrak{B}]_u^\mu \rightleftharpoons \mathsf{MSR}_u^\mu \cup \mathtt{Data}\,[\mathfrak{B}]\ .$$

Теорема 2 (о непротиворечивости универсальной ms-теории). *Теория* $\mathsf{Prdct}\,[\mathfrak{B}]_u^\mu$ *логически непротиворечива для всякого* μ*-совместного* $\mathtt{Data}\,[\mathfrak{B}]$.

ДОКАЗАТЕЛЬСТВО. Рассмотрим множество основных экземпляров

$$\Delta = \{C\theta_\delta \mid C \in \mathsf{MSR}_u^\mu,\ \theta_\delta \in \mathtt{Sub}\,[C]^\mu\}\,.$$

Убедимся, что все его элементы удовлетворяют пунктам ms.1–2 (например, для предсказания соответствующих заключений), где в качестве $\mathtt{Fact}_o$ используется $\mathtt{Fact}_v\Theta^0$. Действительно, пусть $C_\delta \leftrightharpoons (H \leftarrow B_1 \wedge \dots \wedge B_n) \in \Delta$. Это означает $C_\delta = C\theta_\delta$ (и, в частности, $C \succcurlyeq C_\delta$), где $C \in \mathsf{MSR}_u^\mu$, причем в силу ums.1–2:

$$\{B_1, \dots, B_n\} \subseteq \mathtt{Fact}_o^*, \mu\,(C_\delta) = \mu\,(C\theta_\delta) > \mu\,(H\theta_\delta) = \mu\,(H)\,,$$

т. е. для C_δ выполнено ms.1 при предсказании литерала H. Наличие ms.2 докажем «от противного». Если найдется $C'_\delta \sqsupset C_\delta$, сохраняющее п. 1 определения ms-правила для H, то мы бы уточнили исходное ums-правило C ($C \succcurlyeq C_\delta \succ C'_\delta$) со строгим увеличением нижней оценки: $\underline{\mu}\,(C) \leqslant \mu\,(C_\delta) < \mu\,(C'_\delta)$. Сказанное вступает в противоречие с п. 3 ums-правил, так как требования ums.1–2 совпадут с ms.1 для основного C'_δ.

С учетом замечания к теореме 1 получаем: $\mathsf{MSR}_\theta^\mu \cup \mathtt{Data}\,[\mathfrak{B}] \vDash \Delta'$, где $\Delta' \leftrightharpoons \Delta \cup \cup\, \mathtt{Data}\,[\mathfrak{B}]$. Отметим, в любой модели Δ' реализуются все данные, а значит, и те основные случаи правил MSR_u^μ, для которых условная вероятность не определена (ибо их посылки с необходимостью не могут содержаться полностью в $\mathtt{Data}\,[\mathfrak{B}]$, т. е. ложны в модели). Таким образом, из теоремы 1 вытекает выполнимость

$$\Gamma \leftrightharpoons \left\{C\theta_\delta \mid C \in \mathsf{MSR}_u^\mu,\ \theta_\delta \in \Theta^0\right\} \cup \mathtt{Data}\,[\mathfrak{B}]\,.$$

В итоге существует эрбранова модель для $\mathsf{Prdct}\,[\mathfrak{B}]_u^\mu$, поэтому указанное множество и будет непротиворечиво. $\square$

6. Несколько непротиворечивых модификаций

Уберем требование «$\mu(\mathrm{C}\theta) > \mu(\mathrm{H}\theta)$» из п. 1 определения ms-правила для предсказания $\mathrm{H} \in \mathbf{L}_{\mathfrak{L}}$. Список получившихся характеристик для C и H обозначим как wms.1–3; определенное в результате множество правил есть *слабые ms-правила (wms-правила)*, или *слабо мотивированные ms-правила* для H. Пусть $\mathsf{WMS}^{\mu}_{\theta}$ — соответствующая им совокупность основных экземпляров (по всем литералам; задается по аналогии с $\mathsf{MSR}^{\mu}_{\theta}$).

Лемма 2 (об уровне доверия $1 - \varepsilon$). *Допустим,* $\mathrm{C_{pos}} \leftrightharpoons (\mathrm{H} \leftarrow \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n) \in \mathsf{WMS}^{\mu}_{\theta}$ *и* $\mathrm{C_{neg}} \leftrightharpoons (\neg\mathrm{H} \leftarrow \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) \in \mathsf{WMS}^{\mu}_{\theta}$, *причем вероятность конъюнкции посылок* $\mathrm{C_{pos}}$ *и* $\mathrm{C_{neg}}$ *не равна нулю. Тогда, если в первом случае для оценки справедливо* $\mu(\mathrm{C_{pos}}) \geqslant 1-\varepsilon$, *то во втором случае реализуется неравенство* $\mu(\mathrm{C_{neg}}) \leqslant \varepsilon$.

ДОКАЗАТЕЛЬСТВО. По определению $\mathrm{C_{pos}}$ и $\mathrm{C_{neg}}$ являются основными случаями некоторых wms-правил C_1 и C_2 соответственно (для подстановок, удовлетворяющих wms.1). Как и в лемме 1, нетрудно показать равенства $\mu(\mathrm{C_{pos}}) = \underline{\mu}(\mathrm{C}_1)$ и $\mu(\mathrm{C_{neg}}) = \underline{\mu}(\mathrm{C}_2)$. Заметим, свойства $\mu(\mathrm{C_{pos}}) > \mu(\mathrm{H})$ и $\mu(\mathrm{C_{neg}}) > \mu(\neg\mathrm{H})$, отвечающие за сильную мотивацию, используются в доказательстве леммы 1 исключительно на завершающем этапе (получения противоречия); предшествующие же рассуждения проводятся в актуальном для нас предположении μ-совместности посылок $\mathrm{C_{pos}}$ и $\mathrm{C_{neg}}$. Осталось понять, что

$$\begin{gathered} 1 = \mu(\mathrm{H} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) + \\ + \mu(\neg\mathrm{H} \mid \mathrm{B}_1 \wedge \cdots \wedge \mathrm{B}_n \wedge \mathrm{D}_1 \wedge \cdots \wedge \mathrm{D}_m) = \mu(\mathrm{C_{pos}}) + \mu(\mathrm{C_{neg}}) = \\ = \underline{\mu}(\mathrm{C}_1) + \underline{\mu}(\mathrm{C}_2) \geqslant 1 - \varepsilon + \underline{\mu}(\mathrm{C}_2), \end{gathered}$$

поэтому

$$\mu(\mathrm{C_{neg}}) = \underline{\mu}(\mathrm{C}_2) \leqslant 1 - (1 - \varepsilon) = \varepsilon. \qquad \square$$

$$\varepsilon\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta} \rightleftharpoons \bigl\{\mathrm{C} \mid \mathrm{C} \in \mathsf{WMS}^{\mu}_{\theta},\ \mu(\mathrm{C}) \geqslant 1 - \varepsilon\bigr\} \cup \mathtt{Data}\,[\mathfrak{B}], \quad \text{где } \varepsilon < \tfrac{1}{2}.$$

Теорема 3 (о непротиворечивости wms-теории уровня $1 - \varepsilon$). *Теория* $\varepsilon\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta}$ *логически непротиворечива для всякого* μ*-совместного множества* $\mathtt{Data}\,[\mathfrak{B}]$.

ДОКАЗАТЕЛЬСТВО. Находясь в условиях леммы 1: если $\mu(\mathrm{C_{pos}}) \geqslant 1-\varepsilon > \frac{1}{2}$, то верно и $\mu(\mathrm{C_{neg}}) \leqslant \varepsilon < \frac{1}{2}$; таким образом, wms-предсказания H и $\neg$H (для $\mathrm{H} \in \mathbf{L}^0_{\mathfrak{L}}$) при уровне доверия $\varepsilon < \frac{1}{2}$ взаимно исключают друг друга. Дальнейшие рассуждения проводятся в духе доказательства теоремы 1. $\square$

Имея $\mathtt{Fact}_v$, обратимся к ослабленной версии ums-правил (см. п. 5): исключим пункт ums.2 для $\mathrm{H} \in \mathbf{L}_{\mathfrak{L}}$ (точнее говоря, заменим его на условие, тождественно выполненное при любом $\mathrm{C} \in \mathtt{Rule}_{\mathfrak{L}}$). В итоге нами будут определены *слабые ums-правила* для предсказания H. Совокупность последних (объединенную по всевозможным литералам) обозначим WMS^{μ}_{u}.

$$\varepsilon\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{u} \rightleftharpoons \bigl\{\mathrm{C} \mid \mathrm{C} \in \mathsf{WMS}^{\mu}_{u},\ \underline{\mu}(\mathrm{C}) \geqslant 1 - \varepsilon\bigr\} \cup \mathtt{Data}\,[\mathfrak{B}], \quad \text{где } \varepsilon < \tfrac{1}{2}.$$

Теорема 4 (о непротиворечивости wums-теории уровня $1-\varepsilon$). *Теория* $\varepsilon\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{u}$ *логически непротиворечива для всякого* μ*-совместного множества* $\mathtt{Data}\,[\mathfrak{B}]$.

ДОКАЗАТЕЛЬСТВО. Нужно лишь убедиться, что замечание к теореме 1 остается справедливо в ситуации слабого μ-предсказания уровня $1-\varepsilon$ (остальное будет аналогично доказательству теоремы 2). Рассмотрим произвольное правило C, для которого выполнены п. 1–2 wms-определения. Пусть θ — какая-нибудь основная подстановка, удовлетворяющая $\mathsf{wms.1}$ для C и предсказываемого им литерала, причем $\mu\,(C\theta) \geqslant 1-\varepsilon$. Теперь, действуя как и в упомянутом замечании, находим wms-правило C' и $\theta' \in \Theta^0$ такие, что $C'\theta' \in \mathsf{WMS}^{\mu}_{\theta}$, а $C'\theta' \succcurlyeq C\theta$. При этом $\mu\,(C'\theta') \geqslant \mu\,(C\theta) \geqslant 1-\varepsilon$; следовательно, $C'\theta'$ лежит в множестве $\varepsilon\mathsf{Prdct}\,[\mathfrak{B}]^{\mu}_{\theta} \setminus \mathtt{Data}\,[\mathfrak{B}]$. $\square$

Замечание. Пусть $\mathbf{T}_{\mathfrak{L},\leqslant d} \subseteq \mathbf{T}_{\mathfrak{L}}$ — термы, в записи которых встречается не более $d \in \mathbb{N}$ вхождений функциональных символов Σ. Далее можно определить множества подстановок, формул и правил, построенных при участии лишь элементов $\mathbf{T}_{\mathfrak{L},\leqslant d}$. Кроме того, все использованные в работе конструкции преобразуются естественным образом; в частности, оценка $\underline{\mu}\,(\cdot)$ для правила будет вычисляться лишь относительно конечного числа допустимых подстановок основных термов [2]. Нетрудно видеть, мы не выйдем за пределы некоторой фиксированной глубины d в ходе размышлений. Иными словами, все приведенные нами выше леммы и теоремы можно переформулировать и доказать для ситуации с глубиной, ограниченной сверху числом d.

Заключение

Индукция — крайне важный аспект научного объяснения того или иного эмпирически наблюдаемого феномена; между тем статистика обеспечивает необходимый мост между экспериментальными данными и их теоретической (математизированной) спецификацией. Приведенное в статье требование максимальной специфичности и его модификации позволяют избежать проблемы статистической двусмысленности при работе с индуктивно произведенным знанием, причем сказанное касается как теорий пропозиционального сорта, так и теорий универсальных формул. Сопутствующие теоремы не только демонстрируют решение немаловажного философского вопроса с помощью современного логико-вероятностного аппарата, но и обосновывают, в каком ключе будет корректно использовать индуктивно-статистическую аргументацию на практике при формировании синтаксических систем для наблюдаемых физических, биологических, социологических и прочих явлений.

Список литературы

1. *Vityaev E. E.* The Logic of Prediction // Mathematical Logic in Asia 2005: Proc. of the 9th Asian Logic Conference. August 16–19, Novosibirsk, Russia / Eds. S. S. Goncharov, R. Downey, H. Ono. World Scientific Publisher, 2006. P. 263–276.

2. *Смердов С. О., Витяев Е. Е.* Синтез логики, вероятности и обучения: формализация предсказания // Сибирские электронные математические известия. 2009. Т. 9. С. 340–365.

3. *Vityaev E. E., Smerdov S. O.* New Definition of Prediction without Logical Inference // Proc. of the IASTED International Conference on Computational Intelligence (CI 2009) / Ed. by B. Ya. Kovalerchuk. Honolulu, Hawaii, USA, 2009. P. 48–54.

4. *Годунов С. К.* Современные аспекты линейной алгебры. Новосибирск: Научная книга, 1997.

5. *Popper K. R.* The Logic of Scientific Discovery. Routledge publishers, 1959.

6. *Hempel C. G.* Deductive-Nomological versus Inductive-Statistical Explanation // Minnesota Studies in the Philosophy of Science / Eds. H. Feigl, G. Maxwell. Minneapolis, Minnesota: University of Minnesote Press, 1962. Vol. 3. P. 98–169.

7. *Bell J. L.* Boolean-Valued Models and Independence Proofs in Set Theory. Oxford: Clarendon Press, 2005.

8. *Кусраев А. Г., Кутателадзе С. С.* Введение в булевозначный анализ. М.: Наука, 2005.

9. *Popper K. R.* Of Clouds and Clocks. An Aproach to the Problem of Rationality and the Freedom of Man // Objective Knowledge. An Evolutionary Approach. Oxford: Clarendon Press, 1979. Ch. 6. P. 206–255.

10. *Suppes P.* Probabilistic Metaphysics. Basil Blackwell, Oxford, 1984.

11. *Carnap R.* Logical Foundations of Probability. University of Chicago Press, 1950.

12. *Vityaev E. E., Kovalerchuk B. Ya.* Relational Methodology for Data Mining and Knowledge Discovery // Intelligent Data Analysis, Special Issue on Philosophies and Methodologies for Knowledge Discovery and Intelligent Data Analysis / Eds. K. Rennolls, E. Vityaev. 2008. Vol. 12. No. 2. P. 189–210.

13. *Hempel C. G.* Aspects of Scientific Explanation and Other Essays in the Philosophy of Science. N. Y.: Free Press, 1966.

14. *Hempel C. G.* Maximal Specificity and Lawlikeness in Probabilistic Explanation // Philosophy of Science. 1968. Vol. 35. No. 2. P. 116–133.

15. *Scott D., Krauss P.* Assigning Probabilities to Logical Formulas // Aspects of Inductive Logic / Eds. J. Hintikka, P. Suppes. Amsterdam: North-Holland Publishing Company, 1966. P. 219–264.

16. *Halpern J. Y.* An Analysis of First-Order Logics of Probability // Artificial Intelligence. 1990. Vol. 46. P. 311–350.

17. *Lloyd J. W.* Foundations of Logic Programming. Springer-Verlag, 1987.

18. *Borovkov A. A.* Probability Theory. Gordon and Breach, 1998.

19. *Borovkov A. A.* Mathematical Statistics. Gordon and Breach, 1998.

20. *Suppes P.* A Probabilistic Theory of Causality. Amsterdam: North-Holland Publishing Company, 1970.

21. *Carnap R.* A Basic System of Inductive Logic. Part 1 // Studies in Inductive Logic and Probability / Eds. R. Carnap, R. C. Jeffrey. Univericy of California Press, 1971. Vol. 1. P. 33–165.

22. *Hintikka J., Hilpinen R.* Knowledge, Acceptance and Inductive Logic / Eds. J. Hintikka, P. Suppes // Aspects of Inductive Logic. Amsterdam: North-Holland Publishing Company, 1966. P. 1–20.

23. *Smerdov S. O.* On the Question of Consistence of the Semantic μ-prediction // Abstracts, Logic Colloquium 2009, ASL European Summer Meeting, July 31–August 5, Sofia, Bulgaria. P. 82–83.

Материал поступил в редколлегию 21.11.2009

Адрес автора

СПЕРАНСКИЙ Станислав Олегович

Новосибирский государственный университет

ул. Пирогова, 2, Новосибирск, 630090, Россия

e-mail: katze.tail@gmail.com