УДК 519.68:007.5

ПРАВДОПОДОБНЫЕ ВЫВОДЫ И ПРАВДОПОДОБНЫЕ РАССУЖДЕНИЯ

В. К. Финн

§ 1. Основные определения и классификация типов выводов

Вывод есть точно определенная процедура получения из множества исходных формул, называемых посылками, некоторой формулы, называемой заключением (или следствием). Осуществление этой процедуры требует применения специальных правил вида $\frac{\varphi_1, \ldots, \varphi_n}{\psi}$ (или $\varphi_1, \ldots, \varphi_n \vdash \psi$), которые называются правилами вывода и истолковываются следующим образом: «из формул $\varphi_1, \ldots, \varphi_n$ (посылок) следует («$\vdash$» обозначает отношение выводимости) формула ψ (заключение)».

Вывод некоторой формулы ψ может состоять из последовательного применения различных правил вывода, являющихся шагами единого процесса получения заключения ψ.

Уточним теперь понятие вывода. Пусть Σ_0 есть множество формул, принимаемых в качестве допущений (например, некоторое множество выдвигаемых гипотез), пусть, далее, Σ есть множество истинных формул, принимаемых в качестве аксиом, характеризующих предметную область, а $\mathfrak{R}$ множество правил вывода, тогда конечную последовательность формул $\psi_1, \ldots, \psi_S$ (*) будем называть выводом формулы ψ из множества формул Σ_0 (символически: $\Sigma_0 \vdash \psi$, где «$\vdash$» отношение выводимости) если:

1. ψ_S есть ψ,

2. ψ_i, где $i = 1, \ldots, S$, есть либо аксиома из Σ, либо формула из Σ_0, либо ψ_i получена из предыдущих формул последовательности (*) посредством правил из $\mathfrak{R}$.

Пусть $\overline{V} = V \cup \{t, f\}$, где t, f, соответственно, истинностные значения «истина» и «ложь», а V — множество промежуточных истинностных значений между t, f (отметим, что V — может быть и пустым множеством: $V = \varnothing$).

Пусть, далее, задан язык логики предикатов первого порядка или некоторая его модификация. Рассмотрим функцию оценки val, отображающую множество всех замкнутых формул Φ языка L в $\overline{V}$, т. е. val : $\Phi \to \overline{V}$ [Мендельсон, 1976].

Рассмотрим некоторое подмножество $V_d \subset \overline{V}$ такое, что $V_d = V_d' \cup \{t\}$, где $V_d' \subseteq \overline{V}$. V_d будем называть множеством выделенных истинностных значений[1].

Правило вывода $\varphi_1, \ldots, \varphi_S \vdash \psi$ будем называть *достоверным,* если для любой val из val $(\varphi_i) \in V_d$, $i=1, \ldots, S$, следует val $(\psi) \in V_d$.

Правило вывода назовем *недостоверным,* если приведенное выше условие не всегда выполнимо: из выделенности всех посылок выделенность следствия вытекает не всегда.

Вывод формулы ψ из множества допущений Σ_0 будем называть недостоверным, если при построении этого вывода было применено по крайней мере одно недостоверное правило вывода.

Рассмотрим теперь различение между *правилами достоверного вывода* (ПДВ) и *правилами недостоверного вывода* (ПНВ) с несколько иной (неформальной) точки зрения. Пусть F есть множество условий, посредством которых формула φ может быть сделана ложной (или общее: принимающей невыделенное истинностное значение), тогда, согласно Попперу, F назовем множеством *фальсификаторов* формулы φ [31]. Тогда правило вывода формулы φ назовем достоверным, если при его применении гарантируется, что F — пусто для φ, т. е. что у φ не существует фальсификаторов. Правило недостоверного вывода формулы φ назовем правдоподобным, если для известного (к настоящему моменту состояний знаний о мире) множества возможных фальсификаторов F формулы φ ни один из них не делает φ ложной (или общее: принимающей невыделенное истинностное значение). Во втором случае существенным обстоятельством является тот факт, что множество всех фальсификаторов φ не является заданным, следовательно, мы имеем ситуацию с неполной информацией: множество возможных фальсификаторов формулы φ открыто (ср. с идеей открытости интеллектуальных систем [42]). Открытость этого множества F связана с динамикой используемых знаний (т. е. с возможными изменениями базы знаний компьютерной системы; последнее обстоятельство характерно для так называемых *немонотонных рассуждений,* введенных в проблематику искусственного интеллекта Маккарти [62].

Проиллюстрируем теперь различие между достоверным выводом и недостоверным с помощью следующих примеров.

---

[1] Выделенные истинностные значения суть приемлемые с некоторой точки зрения оценки формулы, например, являющиеся степенями правдоподобия или степенями истинности и т. п.

(1) Он или благородный человек, или ребенок. Но он не благородный человек, следовательно, он ребенок.

(2) Он или благородный человек, или ребенок. Но он благородный человек, следовательно, он не ребенок.

Пусть p=он благородный человек, а q=он ребенок; тогда рассуждениям (1) и (2) отвечают, соответственно, правила

(I) $(p\vee q),\ \neg q\vdash p$,

(II) $(p\vee q),\ p\vdash\neg q$.

Очевидно, что (I) — правило достоверного вывода, а (II) — правило недостоверного вывода. В самом деле, (I) и (II) будут ПДВ тогда и только тогда, когда формулы $(((p\vee q)\&\neg q)\supset p)$ и $(((p\vee q)\&p)\supset\neg q)$, соответственно, будут тавтологиями. Легко видеть, что $(((p\vee q)\&\neg q)\supset p)$ — тавтология. В самом деле, пусть $p=0$ (0 — «ложь»), тогда $(0\vee q)$ эквивалентно q, но $q\&\neg q$ — противоречие, следовательно, $(q\&\neg q)\supset 0$ — тавтология. Следовательно, $(((p\vee q)\&\neg q)\supset p)$ не может быть ложной, так как если заключение (p) ложно, то и посылка $(((p\vee q)\&\neg q))$ ложна.

Легко видеть, что $(((p\vee q)\&p)\supset\neg q)$ не является тавтологией. В самом деле, пусть $\neg q=0$, тогда $q=1$ (1 — «истина»), тогда $(p\vee 1)$ эквивалентна 1, и, следовательно, получаем формулу $(1\&p)\supset 0$, которая ложна при $p=1$.

Таким образом, у формулы $(((p\vee q)\&p)\supset\neg q)$ существует единственный фальсификатор, которым является оценка Val: $p=1$, $q=1$. Существенно отметить, что у данной формулы (и, следовательно, у соответствующего ПДВ) известны все возможные фальсификаторы. И, следовательно, известен случай ложности указанной формулы.

Рассмотрим теперь некоторое условие, заданное формулой логики предикатов первого порядка [25] $\varphi(x)$, зависящей только от одной свободной переменной x. Пусть проведено n наблюдений, описаниями которых будут высказывания $\varphi(a_1),\ldots,\varphi(a_n)$, где $a_1,\ldots,a_n$ — некоторые индивидные константы, характеризуемые посредством $\varphi(a_i)$, $i=1,\ldots,n$. Пусть, далее, не известен случай ложности условия $\varphi(x)$, т. е. не установлено, что $\varphi(b)$ ложно, где $b\neq a_i$, $i=1,\ldots,n$, тогда вывод $\varphi(a_1),\ldots,\varphi(a_n)\vdash\forall x\varphi(x)$ ($\forall x$ — квантор общности: «для всех x...»), см., например, [43]), будет недостоверным выводом, однако, не имеющим какого-либо обнаруженного фальсификатора. Заметим, что утверждать, что множество F фальсификаторов формулы $\forall x\varphi(x)$ пусто, мы не имеем оснований. Иначе говоря, $\forall x\varphi(x)$ не следует логически из посылок $\varphi(a_1),\ldots,\varphi(a_n)$, но ее опровержения (фальсификаторы) нам неизвестны (подобная схема рассуждения называется перечислительной, или популярной, индукцией).

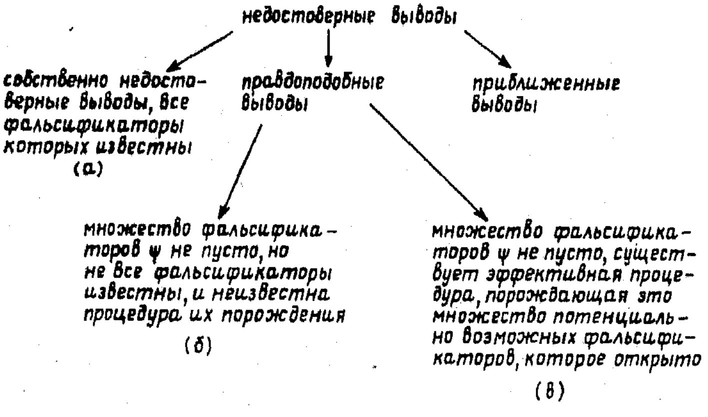
Вывод формулы ψ из посылок $\varphi_1,\ldots,\varphi_n$ будем называть правдоподобным, если он не является достоверным, но множество его фальсификаторов полностью неизвестно. Таким образом, приведенный выше абстрактный пример недостоверного вывода является правдоподобным выводом. Понятие правдоподобного вывода может быть конкретизировано в следующем смысле. Рассмотрим два типа множеств фальсификаторов F_1 и F_2. Посредством F_1 обозначим множество возможных фальсификаторов формулы φ, т. е. такое множество, относительно которого нельзя утверждать его пустоту, но относительно элементов $f_i \in F_1$ неизвестна эффективная процедура, их порождающая. Посредством F_2 обозначим множество потенциальных фальсификаторов для заданного множества формул Φ такое, что существует эффективная процедура, порождающая для $\varphi \in \Phi$ потенциальные фальсификаторы посредством обучения на множестве заданных примеров и контрпримеров. Таким образом, вообще говоря

$$F_2 = \bigcup_{l=1}^{\infty} F_2^{(l)}, \quad \text{где} \quad F_2^{(1)} \subset F_2^{(2)} \subset \ldots \subset F_2^{(l)} \subset F_2^{(l+1)} \subset \ldots$$

в соответствии с расширением обучающей выборки, состоящей из положительных и отрицательных примеров (контрпримеров).

Прежде чем рассмотреть классификацию недостоверных выводов формулы ψ из посылок $\varphi_1, \ldots, \varphi_n$, введем еще один вид недостоверных выводов — приближенные выводы. Под *приближенными выводами* мы будем понимать выводы такие, что оценка заключения есть степень истинностного значения t («истина»). Примерами приближенных выводов являются выводы в вероятностной логике [71], а также выводы в системах формализации логической вероятности в смысле Карнапа и Хинтикки [22]. По-видимому, к приближенным выводам следует отнести различные формализации нечетких выводов [58] (см. в связи с понятием приближенных выводов [57]).

Рассмотрим теперь следующую классификацию недосоверных выводов формулы ψ из посылок $\varphi_1, \ldots, \varphi_n$:



Пример недостоверного вывода типа (а) был приведен выше: $(p \vee q),\ p \vdash \neg q$. Пример правдоподобного вывода типа (б) также был приведен выше: $\varphi(a_1), \ldots, \varphi(a_n) \vdash \forall x \varphi(x)$ (популярная индукция).

Следующий пример также представляет правдоподобный вывод типа (б): $\varphi(x) \vdash \psi(x)$ для всех x за исключением $x \in F$, где F — открытое множество возможных фальсификаторов (например, если x — птица, то x летает для всех x за исключением страусов; F — открытое множество хотя бы уже потому, что в процессе эволюции могли быть уже вымершие и неизвестные птицы, которые не летали). Этот пример связан с проблемой немонотонных рассуждений [62], [72], [63], которые будут охарактеризованы ниже.

Основанием классификации недостоверных выводов, которую мы рассмотрели, была характеризация множества фальсификаторов в смысле (а), (б) и (в).

Рассмотрим теперь классификацию правдоподобных выводов, основанием которой является характеризация **общности** посылок и заключения:

правдоподобные выводы

индукция	*абдукция*
(*вывод* $\dfrac{\text{от частного}}{\text{к общему}}$)	(*вывод* $\dfrac{\text{от частного}}{\text{к частному}}$).

Примером *индукции* является вывод $\varphi(a_1), \ldots, \varphi(a_n) \vdash \forall x \varphi(x)$, если, разумеется, рассматриваемый универсум U содержит $\{a_1, \ldots, a_n\}$, но не равен этому множеству $(\{a_1, \ldots, a_n\} \subset U)$. Ибо в случае $\{a_1, \ldots, a_n\} = U$ мы получим достоверный вывод: $\forall x \varphi(x) \equiv (\varphi(a_1) \& \ldots \& \varphi(a_n))$[1] и $F = \varnothing$ (множество фальсификаторов F пусто).

Приведем теперь примеры *абдукции*.

(1) Пусть a, b — индивидные константы, обозначающие объекты некоторого рода, $\varphi(x)$ — формула логики предикатов 1-го порядка, имеющая вхождение только одной свободной переменной x; тогда абдукцией будет следующий вывод:

$$\frac{a \text{ сходна с } b \qquad \varphi(a)}{\varphi(b)}.$$

Интуитивно ясно, что существуют фальсификаторы для $\varphi(x)$ и что если предположить, что неизвестен способ их порождения, то рассматриваемый вывод есть правдоподобный вы-

[1] Напомним, что «$\equiv$» логическая связка эквиваленции: $(p \equiv q) \rightleftharpoons (p \supset q) \& (q \supset p)$, где «$\rightleftharpoons$» — равенство по определению, «$\supset$» — импликация, «&» — конъюнкция.

вод типа (б). Очевидно, что если бы не были даны истинные посылки «a сходна с b» и $\varphi(a)$, то правдоподобие заключения $\varphi(b)$ было бы меньше. Поэтому, имея в виду повышение степени правдоподобия $\varphi(b)$ при условии задания указанных посылок, Пойа [30] так уточняет приведенную выше абдукцию:

$$\frac{\begin{array}{c} a \text{ сходна с } b \\ \varphi(a) \end{array}}{\varphi(b) \text{ более правдоподобно}}$$

Как мы увидим впоследствии, формализация заключения в виде «$\varphi(b)$ более правдоподобно» потребует использования средств многозначной логики [37], [39].

(2) Обозначим посредством $U^{(1)}=\{d_1, \ldots, d_S\}$ некоторое множество элементов, из которых строятся объекты $B\subseteq U^{(1)}$, т. е. объекты $B\in 2^{U^{(1)}}$, где $2^{U^{(2)}}$ — множеств свойств, присущих объектам из $2^{U^{(1)}}$. Таким образом, $A=\{a_{i_1}, \ldots, a_{i_m}\}\in 2^{U^{(2)}}$, где $1\leqslant m\leqslant r$.

Обозначим посредством $\Rightarrow_1{}^*$ и $\Rightarrow_2{}^*$ бинарные отношения «объект обладает множеством свойств» и «подобъект (часть объекта) есть причина наличия множества свойств». Пусть бинарные предикаты $x\Rightarrow_1 y$ и $x\Rightarrow_2 y$ отвечают, соответственно, отношениям $\Rightarrow_1{}^*$ и $\Rightarrow_2{}^*$. Тогда высказывания вида $C\Rightarrow_1 A$ и $c'\Rightarrow_2 A$, где $c'\subset c$, будут описывать явления, соответствующие отношениям $\Rightarrow_1{}^*$ и $\Rightarrow_2{}^*$, характеризующим реальный мир («наивную семантику» [55]).

Следующий правдоподобный вывод является абдукцией вида

$$\frac{\begin{array}{c} \{d_1, d_2, d_3, d_4\}\Rightarrow_1\{a_1, a_2\} \\ \{d_1, d_2, d_5, d_6\}\Rightarrow_1\{a_1, a_2\} \\ \{d_1, d_2, d_7, d_8\}\Rightarrow_1\{a_1, a_2\} \end{array}}{\{d_1, d_2\}\Rightarrow_2\{a_1, a_2\}}.$$

Пользуясь терминологией Пойа (30), заключение данной абдукции следовало бы записать так «$\{d_1, d_2\}\Rightarrow_2\{a_1, a_2\}$ правдоподобно», т. е. правдоподобно, что подобъект $\{d_1, d_2\}$ есть причина появления $\{a_1, a_2\}$. Рассмотренная абдукция есть запись известного, так называемого индуктивного метода сходства английского логика и философа Джона Стюарта Милля (26)[1].

Легко понять, что мощность множества возможных явлений вида $C\Rightarrow_1 A$, т. е. $|\{<x, y>|x\Rightarrow_1 y$, где $x\in 2^{U^{(1)}}$, $y\in 2^{U^{(2)}}\}|\leqslant\leqslant 2^S\cdot 2^r$, однако для некоторых из пар $<C, A>$ отношение $\Rightarrow_1{}^*$ может быть недоопределено (аналогичное имеет место и

[1] Поскольку метод сходства Д. С. Милля есть вывод «от частного — к частному», то точнее его отнести к абдуктивным выводам, а не к индуктивным, как это традиционно делалось.

для $\Rightarrow_2^*$). Таким образом, $\Rightarrow_1^*$ и $\Rightarrow_2^*$ — частично определенные отношения. Все возможные абдуктивные выводы, полученные из начальной выборки, представляющей отношение $\Rightarrow_1^*$, конечно, являются лишь гипотезами, так как могут быть фальсифицированы либо при доопределении $\Rightarrow_1^*$ как частично определенного отношения, либо при расширении $\Rightarrow_1^*$ посредством добавления новых примеров вида $C\Rightarrow_1 A$. Следовательно, естественно предположить, что множество фальсификаторов F непусто, но не задано исчерпывающим образом (F — открытое множество). Следовательно, приведенный выше пример недостоверного вывода есть правдоподобный вывод (абдукция).

Следующий пример абдукции является записью так называемого *индуктивного метода различия* Д. С. Милля [26]:

$$\frac{\begin{array}{c}\{d_1,\ d_2,\ d_3,\ d_4\}\Rightarrow_1\{a_1,\ a_2\}\\ \{d_1,\ d_2,\ d_3,\ d_5\}\Rightarrow_1\{a_1,\ a_2\}\\ \{d_1,\ d_2,\ d_4,\ d_6\}\Rightarrow_1\{a_3\}\end{array}}{\{d_1,\ d_2,\ d_3\}\Rightarrow_2\{a_1,\ a_2\}}.$$

Рассмотрим также некоторый аналог *метода сопутствующих изменений Д. С. Милля* [26] (подробное обсуждение методов Д. С. Милля читатель найдет в [32]), являющийся примером индукции.

Пусть заданы множества $\mathrm{U}^{(1)}$, B и $\mathrm{U}^{(2)}$, A, из которых, соответственно, строятся объекты и множества свойств, которыми указанные объекты обладают, причем объекты имеют вид $D_i=\{b_i\}\cup B_0^{(i)}$, где $b_i\in B$, $B_0^{(i)}\in 2^{\mathrm{U}^{(1)}}$, а соответствующие множества свойств имеют вид $C_i=\{a_i\}\cup A_0^{(i)}$, где $a_i\in A$, $A_0^{(i)}\in 2^{\mathrm{U}^{(2)}}$.

Аналогично случаем методов сходства и различия введем предикаты $X\Rightarrow_1 Y$ и $Z\Rightarrow_2 Y$, соответственно интерпретируемые как «объект X обладает множеством свойств Y» и «подобъект Z есть причина наличия множества свойств Y». Соответственно будем рассматривать атомарные высказывания $D_i\Rightarrow_1 C_i$ и $D_j\Rightarrow_2 C_j$. Предположим, что B и A — упорядоченные (например, частично упорядоченные) множества изменяющихся параметров, на которых, соответственно, заданы отношения порядка $\prec_1$ и $\prec_2$. $B_0^{(i)}$ ($i=1, 2, \ldots$) суть множества, задающие структуру объекта D_i, b_i есть некоторый параметр, характеризующий D_i; соответственно, $A_0^{(i)}$ — множество свойств, характеризующих объект со структурой $B_0^{(i)}$ и параметром $b_i\in B$, а A есть множество изменяющихся значений свойства некоторого типа. Например, $B_0^{(i)}$ — химическое соединение, b_i — его доза, $A_0^{(i)}$ — множество биологических активностей $B_0^{(i)}$, а a_i — степень канцерогенности этого соединения.

Пусть, далее, $B\cap\mathrm{U}^{(1)}=\varnothing$ и $A\cap\mathrm{U}^{(2)}=\varnothing$, а x и y, соответственно, переменные, областью изменения которых являются B и A;

тогда аналог метода сопутствующих изменений Д. С. Милля [26] (см. также [32]) сформулируем следующим образом:

$$\frac{\begin{array}{l}(\{b_1\}\cup B_0)\Rightarrow_1(\{a_1\}\cup A_0^{(1)})\\ (\{b_2\}\cup B_0)\Rightarrow_1(\{a_2\}\cup A_0^{(2)})\\ \dots\dots\dots\dots\\ (\{b_m\}\cup B_0)\Rightarrow_1(\{a_m\}\cup A_0^{(m)})\\ b_1 \prec_1 b_2 \prec_1 \dots \prec_1 b_{m-1} \prec_1 b_m\\ a_1 \prec_2 a_2 \prec_2 \dots \prec_2 a_{m-1} \prec_2 a_m\end{array}}{\forall x \exists y\,(x \in B \supset (\{x\}\cup B_0)\Rightarrow_2 \{y\} \,\&\, y \in A)},$$

где $A_0^{(i)} \in 2^{U(2)}$.

Приведенный выше метод правдоподобного вывода имеет еще три родственных варианта, соответственно, с посылками:

$$b_1 \,{}_1\!\succ b_2 \,{}_1\!\succ \dots {}_1\!\succ b_m \quad \text{и} \quad a_1 \,{}_2\!\succ a_2 \,{}_2\!\succ \dots {}_2\!\succ a_m;$$

$$b_1 \prec_1 b_2 \prec_1 \dots \prec_1 b_m \quad \text{и} \quad a_1 \,{}_2\!\succ a_2 \,{}_2\!\succ \dots {}_2\!\succ a_m;$$

$$b_1 \,{}_1\!\succ b_2 \,{}_1\!\succ \dots {}_1\!\succ b_m \quad \text{и} \quad a_1 \prec_2 a_2 \prec_2 \dots \prec_2 a_m.$$

Естественным уточнением в стиле Д. Пойа будет указание на то, что заключение метода сопутствующих изменений не достоверно, но правдоподобно.

Обобщим метод сопутствующих изменений Д. С. Милля и сформулируем комбинированный метод сопутствующих изменений+метод сходства.

Пусть $B_0^{(i)} \in 2^{U(1)}$, $i=1, 2, \dots$; а $A_0^{(i)} \in 2^{U(2)}$, $i=1, 2, \dots$; тогда

$$\frac{\begin{array}{l}(\{b_1\}\cup B_0^{(1)})\Rightarrow_1(\{a_1\}\cup A_0^{(1)})\\ (\{b_2\}\cup B_0^{(2)})\Rightarrow_1(\{a_2\}\cup A_0^{(2)})\\ \dots\dots\dots\dots\\ (\{b_m\}\cup B_0^{(m)})\Rightarrow_1(\{a_m\}\cup A_0^{(m)})\\ B_0^{(1)}\cap B_0^{(2)}\cap\dots\cap B_0^{(m)}=D\neq\varnothing\\ A_0^{(1)}\cap A_0^{(2)}\cap\dots\cap A_0^{(m)}=C\neq\varnothing\\ b_1 \prec_1 b_2 \prec_1 \dots \prec_1 b_m\\ a_1 \prec_2 a_2 \prec_2 \dots \prec_2 a_m\end{array}}{\forall x \exists y\,(x \in B \supset ((\{x\}\cup D)\Rightarrow_2(\{y\}\cup C) \,\&\, y \in A)}.$$

Сформулированный метод правдоподобного вывода, как и метод сопутствующих изменений, имеет четыре варианта связи отношений $\prec_1$ и $\prec_2$.

Очевидно, что данный правдоподобный вывод представляет собой соединение абдукции и индукции[1]. Следует отметить, что естественным будет случай такой, что B и A — числовые множества, а $\prec_1$ и $\prec_2$ — отношения линейного порядка ($\leqslant$).

[1] Очевидно, что можно сформулировать комбинированный метод: *метод сопутствующих изменений+метод различия.*

Мы уже отмечали выше, что правдоподобный вывод является логическим средством получения нового знания в условиях неполноты информации, необходимым условием которой является нарушение закона исключенного третьего («каждое высказывание либо истинно, либо ложно, а третьего не дано»). Типичной схемой организации знания, в которой применим правдоподобный вывод, является *квазиаксиоматическая теория* (КАТ) [40], характеризующая ситуации с неполной информацией.

Пусть Σ — множество аксиом, заведомо неполно характеризующих предметную область, пусть Σ' — открытое множество элементарных (протокольных) высказываний, представляющих факты (будем называть их просто фактами), где $\Sigma' = \bigcup\limits_{n=1}^{\infty} \Sigma_n$, где Σ_n — множество фактов, соответствующих n-му состоянию КАТ. Пусть далее, $\mathfrak{R}$ — множество правил вывода такое, что $\mathfrak{R} = \mathfrak{R}' \cup \mathfrak{R}_0$, где $\mathfrak{R}'$ — множество правил правдоподобного вывода (ППВ), а $\mathfrak{R}_0$ — множество правил достоверного (дедуктивного) вывода, тогда $\Gamma' = \langle \Sigma, \Sigma_j, \mathfrak{R} \rangle$, $\Gamma = \langle \Sigma, \mathfrak{R} \rangle$ и $\Gamma_n = \langle \Sigma, \Sigma_n, \mathfrak{R} \rangle$ будем называть КАТ, каркасом КАТ и n-ым состоянием КАТ, соответственно.

Компьютерные аналоги введенных понятий очевидны: Σ_n — состояние базы данных (базы фактов — БФ), Γ — база знаний, содержащая декларативное знание (Σ) и процедурное знание (τ). Сама КАТ есть предельное (идеальное) понятие, что все состояния каркаса Γ достижимы в пределе, т. е., например, имеет место $\Sigma_1 \subseteq \Sigma_2 \subseteq \ldots \subseteq \Sigma_n \subseteq \ldots$

В связи со сказанным отметим, что открытость множества Σ' — характерная черта интеллектуальных систем, реализующих рассуждение в условиях неполноты информации [42]; [3]. Однако это обстоятельство порождает трудную, но важную проблему выбора критерия принятия заключения правдоподобного вывода в КАТ, называемого либо критерием рациональности правдоподобного вывода [10], либо достаточным основанием правдоподобного вывода [69].

Пусть $\varphi \in \Sigma$, $\chi \in \Sigma'$ и дано ППВ $\frac{\varphi, \chi}{\psi}$, тогда ППВ рационально по Гаеку—Гавранеку, если невыделенность заключения $\psi (Val(\psi) \overline{\in} V_d)$ влечет, что φ выделено $(Val(\varphi) \in V_d)$ лишь для достаточно малого числа случаев в соответствующим образом определенной семантике.

Формула ψ выведена из посылок $\varphi_1, \ldots, \varphi_n$ в КАТ на достаточном основании в смысле К. Поппера, если она не опровергается наиболее сильными для нее фальсификаторами из множества возможных фальсификаторов F (разумеется, что предполагается некоторое упорядочение на F).

Конкретизируем теперь определение КАТ $\Gamma' = \langle \Sigma, \Sigma', \mathfrak{R} \rangle$, положив $\Sigma = \Sigma_{pr} \cup \Sigma_{dc}$, где Σ_{pr}, Σ_{dc} — множества процедурных и декларативных аксиом, соответственно (55). Причем $\Sigma_{dc} = \Sigma_{dc,0} \cup \cup \Sigma_{dc,1}$, где $\Sigma_{dc,0}$ и $\Sigma_{dc,1}$ — множества аксиом структуры данных и аксиом связи исходных предикатов, соответственно[1]. Отметим здесь, однако, что в приведенных выше примерах методов сходства и различия Д. С. Милля $\Sigma_{dc,0}$ содержит аксиомы булевых алгебр множеств $\mathscr{B}_1 = \langle 2^{U^{(1)}}, -, \cap, \cup \rangle$ и $\mathscr{B}_2 = \langle 2^{U^{(2)}}, -, \cap \cup \rangle$; а $\Sigma_{dc,1}$ содержит аксиомы связи предикатов $X \Rightarrow_1 Y$ и $Z \Rightarrow_2 Y$.

Говоря точнее, структура данных, используемая для формализации правдоподобных выводов, есть многоосновная алгебраическая система [23] с частично определенными предикатами. Посредством операций соответствующих алгебраических систем строятся термы, представляющие объекты рассматриваемой предметной области (ПО), а посредством отношений между указанными объектами задается наивная семантика, отражающая «внешний мир» (точнее, его фрагмент, отвечающий ПО). Частичная определенность исходных отношений отражает неполноту информации о ПО, которая характеризуется в КАТ. Правдоподобные выводы — средства уменьшения неполноты информации о ПО. Таким образом, правдоподобные выводы полезны для ПО, данные которой хорошо структурированы (т. е. структура данных формально описана в $\Sigma_{dc,0}$), а знания о которой плохо формализованы, т. е. ПО отвечает не *аксиоматическая теория* (АТ), а охарактеризованная выше КАТ. Примерами наук и научных дисциплин, уровень формализации которых достигает КАТ, но не АТ, являются так называемые науки о жизни (биохимия, биология, фармакология, медицина, социология), а также синтетическая органическая химия, геология, техническая диагностика, некоторые дисциплины, относящиеся к управлению (например, ситуационное управление [32]).

Посредством понятия «правдоподобный вывод» определяется важное для теории интеллектуальных систем понятие «правдоподобного рассуждения», для определения которого нам потребуется характеризация понятия «рассуждения».

Под рассуждением в КАТ будем понимать конечную последовательность формул такую, что ее элементами являются или формулы из Σ (аксиомы ПО), или формулы из Σ' (факты), или формулы, полученные из предыдущих формул последовательности посредством правил вывода из $\mathfrak{R}$; результатом (или целью) рассуждения считается формула, стоящая в конце указанной последовательности, причем

(1) существенно, что Σ' — множество допустимых посылок рассуждения — является открытым множеством (новые посылки, релевантные цели рассуждения, привлекаются в процессе рас-

[1] Примеры Σ_{pr} и Σ_{dc} будут приведены впоследствии.

суждения в условиях неполноты информации посредством специально заданных отношений сходства);
(2) имеются специальные металогические средства управления выводом, посредством которых оцениваются промежуточные результаты, осуществляется проверка на непротиворечивость с привлекаемыми в процессе рассуждения посылками, формулируются критерии отбора релевантных посылок в условиях неполноты информации[1].

Примером логической формализации рассуждений являются различные уточнения немонотонных рассуждений [62].

Правдоподобным рассуждением будем называть рассуждение такое, что при построении последовательности, заканчивающейся результатом рассуждения, использовались правила правдоподобного вывода из $\mathfrak{R}'$ и, быть может, правила достоверного вывода из $\mathfrak{R}$.

§ 2. Краткая история вопроса

История попыток формализации индуктивных выводов восходит к Сократу. Описание схемы вывода, называемой популярной индукцией, или индукцией через простое перечисление, принадлежит Аристотелю. Существенный вклад в описание индуктивных рассуждений внес Ф. Бэкон [8], который впервые (1620 г.) попытался формализовать индуктивные выводы посредством таблиц причин. Ф. Бэкон явился родоначальником исследований «эмпирической структурной индукции», целью которой является обнаружение эмпирических зависимостей в виде индуктивных обобщений, полученных на основе сравнения объектов, имеющих структуру и входящих в явления, которые представляются примерами и контрпримерами. Учение Ф. Бэкона об индукции было развито Д. С. Миллем [26], который предложил свои известные методы сходства, различия, остатков и сопутствующих изменений. Впоследствии ряд авторов пытались средствами современной логики построить формальные имитации методов Д. С. Милля ([74], [61] и др.). Однако использование лишь аппарата двузначной логики и сугубо философские цели экспликации милллевских методов структурной индукции не дали каких-либо результатов, повлиявших на создание программных систем, реализующих правдоподобные рассуждения.

Абдуктивные выводы были предложены одним из создателей математической логики Ч. Пирсом [67]. Интересно отметить, что абдуктивные выводы характерны для распознавания образов [12].

---

[1] Отметим, что вывод в стандартном логическом смысле есть один из видов рассуждений, в которых множество посылок не является открытым и не допускается применение метасредств типа проверок на невыводимость, непротиворечивость и т. п.

Термин «правдоподобное рассуждение» принадлежит Пойа [30], примерами правдоподобных рассуждений в смысле Д. Пойа являются индукция через простое перечисление, аналогия и различные схемы недостоверных (в двузначной логике высказываний) выводов.

Д. Пойа сформулировал два важных принципа *вывод по аналогии* [30]:

1 принцип: «предположение становится более правдоподобным, когда оказывается истинным аналогичное предположение»;

2 принцип: «предположение становится несколько более правдоподобным, когда становится более правдоподобным аналогичное предположение».

Пусть φ и ψ — упомянутые выше предположения, тогда принципам Пойа отвечают, соответственно, следующие схемы правдоподобных выводов:

$$\frac{\varphi \text{ аналогично } \psi \quad \psi \text{ истинно}}{\varphi \text{ более правдоподобно}}$$

$$\frac{\varphi \text{ аналогично } \psi \quad \psi \text{ более правдоподобно}}{\varphi \text{ несколько более правдоподобно}}$$ [1].

Формальные уточнения схем таких правдоподобных выводов связаны, во-первых, с формализацией аналогичности (сходства) φ и ψ, использующей средства описания структуры данных предметной области; и, во-вторых, с формализацией «степени правдоподобия» φ. Существенность сходства структур в выводах по аналогии была отмечена еще И. Г. Лейбницем.

Существенным сдвигом в исследовании правдоподобных выводов и правдоподобных рассуждений явилась потребность имитировать их на ЭВМ. Важным импульсом в этом направлении явилась работа [68], а также [66], в которых индуктивное рассуждение имитировалось средствами дедуктивной логики предикатов первого порядка.

На создание компьютерных систем, реализующих индуктивные обобщения средствами логики предикатов первого порядка, существенное влияние оказали идеи Поппера об индукции [69], [31]. Значительной работой по применению индуктивных рассуждений, корректирующих посредством базы данных некоторую теорию, является [73] (индуктивные выводы в этой работе реализованы посредством языка ПРОЛОГ).

Интересные результаты в области применения правил индуктивного вывода получены в [5]. В этой работе предложены

---

[1] φ несколько более правдоподобно в том смысле, что без информации о правдоподобности ψ, φ было бы менее правдоподобным.

правила индуктивного вывода, которые достаточно часто (рационально в упомянутом выше смысле) позволяют по описанию работы алгоритма на конкретных примерах получить описание алгоритма для общего случая или по формальному доказательству для частного случая получить доказательство для общего случая. Подход к формализации индуктивного обобщения, предложенный в упомянутой работе, применен к проблеме синтеза программ (автором доказаны теоремы об асимптотической полноте системы правил вывода для графических программ). Данное исследование, по-видимому, надо отнести к направлению «структурной индукции» (точнее, к ее разделу — «алгоритмической структурной индукции»)[1].

Фактически, к направлению «структурной индукции» относится и работа [36], в которой осуществляется индуктивное обобщение посредством некоторого алгоритма, который строит «структурные описания» по примерам. К направлению «эмпирической структурной индукции» следует отнести INDUCE — метод [65], который порождает полное и непротиворечивое описание некоторого класса объектов по множеству примеров и контрпримеров соответствующего класса объектов (см. в связи с этим обзор [20]).

Важным этапом в развитии теории правдоподобных рассуждений стало созданное Д. Маккарти направление исследований по немонотонным рассуждениям, которые характерны как логические средства баз данных (БД) и баз знаний (БЗ) с неполной информацией [7]. Правила немонотонного вывода, согласно предложенной выше классификации недостоверных выводов, относятся к правдоподобным выводам типа (б), у которых множество фальсификаторов заключений не пусто, но не все фальсификаторы известны и не известна процедура, их порождающая.

При неполной информации в базах данных можно применять нестандартные процедуры логической обработки данных, имитирующие правдоподобные рассуждения. В этих условиях возникает так называемый эффект немонотонности, который проявляется, в частности, в изменении отношения выводимости (ранее выводимые утверждения могут оказаться невыводимыми) или в изменении истинностных значений исходных утверждений о внешнем мире.

Специальные логические процедуры преобразуют исходную внешнюю информацию: система активно воздействует на входной поток данных. Это приводит к тому, что часть исходной информации становится невыводимой (системный эффект неполноты, ср. рисунок). С другой стороны, система становится более интеллектуальной, реализуя сложные процедуры вывода ответа на запрос.

---

[1] Результаты упомянутого направления исследований изложены также в [50], [46].

Большинство известных в литературе так называемых немонотонных логических систем, по существу, выражают различные подходы к имитации тех или иных рациональных (разумных, человекоподобных и т. п.) способов рассуждения в условиях неполноты информации. При этом «эффекты немонотонности» связаны с тем, что при получении дополнительной информации может потребоваться пересмотреть некоторые ранее сделанные заключения, которые окажутся несовместимыми с новым, более полным описанием ситуации. А поскольку явления такого рода распространены достаточно широко, то при моделировании их возникают самые разные, подчас значительно удаленные друг от друга, теоретические и прикладные системы с немонотонной логикой.

Один из таких подходов представлен в работах группы Дж. Маккарти[1].

При этом, в отличие от классической логики, которую (очень нестрого) можно было бы назвать «логикой достоверного заключения», здесь вводятся, так сказать, системы «правдоподобного предположения». Иначе говоря, в отличие от обычного отношения логического следования $\Gamma \vdash A$ (замкнутая формула A выводима в классическом исчислении предикатов [25, гл. 2, § 33] из множества замкнутых формул Γ) определяется немонотонное отношение $\Gamma \overset{*}{\vdash} A$ (формула A «допустима как предположение», исходя из заданного набора сведений Γ); при этом немонотонность понимается в том смысле, что, вообще говоря, $\Gamma \overset{*}{\vdash} A$ не влечет $(\Gamma \cup \Gamma') \overset{*}{\vdash} A$ (допустимое предположение A может оказаться недопустимым при расширении набора исходных сведений).

Ситуации, моделируемые в таких системах, можно проиллюстрировать следующим простым примером [75]. Представим себе робота-шофера, умеющего ходить и водить машину, которому поручили что-то куда-то отвезти. Человек на его месте, естественно, прежде всего отправился бы искать машину туда, где он ее оставил. Но робот, рассчитанный на работу в условиях полной информации, должен был бы сперва доказать, что машина находится именно там. А поскольку непосредственно такой информации в его памяти нет, то ему пришлось бы последовательно исключать все возможные причины, по которым она могла бы оттуда исчезнуть (например, попробовать позвонить в милицию, нет ли там сведений об угоне; потом подумать, а могли ли туда не успеть сообщить, и т. д. без ограничения). И в результате отвозить груз пришлось бы человеку, привыкшему действовать в условиях неполной информации и готовому предположить, что машина скорее всего там, где ее

---

[1] См., например, сборник статей по немонотонной логике «Artif. Intell.», 1980, *13*, № 1-2 и другие публикации тех же авторов.

оставили, — ну а уж если ее там не окажется, тогда и будем думать об этом, смотря по ситуации.

Разумеется, этот пример вполне игрушечный и может показаться надуманным, но фактически он иллюстрирует важную в приложениях идею «заключения по умолчанию». Для выражения этой идеи Дж. Маккарти предложил понятие «минимальной модели»: предполагается существование только тех объектов, которые с необходимостью вытекают из условий задачи, и определил «минимальное следование» (minimal entailement) — $\Gamma \underset{m}{\vdash} A \rightleftharpoons$ «A истинно во всякой минимальной модели множества формул Γ» (немонотонность очевидна: минимальная модель $(\Gamma \cup \Gamma')$, хотя и является моделью множества формул Γ, но может не быть его минимальной моделью). При этом, например, для известного множества аксиом арифметики Пеано S [25, гл. 3, § 1] единственной минимальной моделью является стандартный натуральный ряд — это означает, что $(S \underset{m}{\vdash} A) \Leftrightarrow$ $\Leftrightarrow$ «A истинно на натуральном ряду». Таким образом, отошение минимального следования не может быть эффективно (рекурсивно) аксиоматизировано (более того, оно даже не может быть выражено обычными средствами арифметики).

Для возможных приложений понятие «минимального следования» несколько обобщается: допускается минимальность не обязательно по множеству индивидов (элементов «универсума» — предметной области), а по некоторому свойству (в рассмотренном примере — по множеству возможных препятствий для нахождения машины на том месте, где она была оставлена). Такой подход ограничен из-за недостаточной определенности тех свойств, по которым предполагается минимальность. Например, если бы робот минимизировал не препятствия, а причины нахождения машины на месте, то потребность в доказательстве их существования оказалась бы логически безупречной (хотя, как уже отмечалось, совершенно непродуктивной). Мотивы, которыми руководствуется реальный человек при выборе «минимизируемых» свойств, могут быть достаточно сложны, и включать различные субъективные оценки «весовых» характеристик типа «значимости», «вероятности», «опасности» и т. п. Чтобы не анализировать подобные мотивы, в ряде работ из упомянутого выше сборника предлагается рассматривать немонотонные теории с заранее заданным набором правил «заключения по умолчанию». При этом моделируются схемы рассуждения по правилам типа «Птицы обычно летают», представленным в виде:

$$\frac{(x\text{ — птица}),\ M\,(x\text{ — летает})}{(x\text{ — летает})}, \qquad (*)$$

где выражение MX означает, что в системе нет сведений, противоречащих предположению о том, что выполняется x (иначе говоря: в системе не доказуемо $\neg X$). Ясно, что определение доказуемости в системах такого рода наталкивается на «порочный круг» — доказуемость определяется через недоказуемость. Для того, чтобы избежать этого, приходится применять технику «неподвижных точек», после чего определение доказуемости становится неконструктивным и не допускает эффективной аксиоматизации.

Например, Рейтер [72] рассматривает теории вида $\Delta = (D, W)$, где W — множество замкнутых предикатных формул (предложений), а D — множество правил *вывода «по умолчанию»* (default-rule) вида

$$\left(\frac{A;\ MB_1,\ \ldots,\ MB_n}{C}\right), \qquad (**)$$

где A, $B_1, \ldots, B_n$, C — предложения, а символ M понимается как в $(*)$. При этом определяется оператор $\Gamma = \Gamma_\Delta$, а именно: если E — произвольное множество предложений, то $\Gamma_\Delta(E)$ — наименьшее (по включению) множество предложений, удовлетворяющее условиям

$\Delta 1$: $W \subseteq \Gamma_\Delta(E)$ (аксиомы теории Δ принадлежат $\Gamma_\Delta(E)$);

$\Delta 2$: $\Gamma_\Delta(E) \vdash X \Rightarrow X \in \Gamma_\Delta(E)$ (замкнутость относительно классической выводимости);

$\Delta 3$:
$$\left(\frac{A;\ MB_i : i=1,\ \ldots,\ n}{C}\right) \in \Delta, \quad \underline{A \in \Gamma_\Delta(E)}$$
$$(\neg B_i) \notin E\ (i=1,\ \ldots,\ n) \Rightarrow \underline{C \in \Gamma_\Delta(E)}$$

(«размыкание» порочного круга в понимании символа M).

Теперь расширением теории Δ называется неподвижная точка оператора Γ_Δ, т. е. такое множество предложений E, что $E = \Gamma_\Delta(E)$ (так сказать, «множество гипотез, допустимое в теории Δ» — при этом, как видно, условие $\Delta 3$ в точности выражает понимание MB_i как «невыполнение $\neg B_i$»). В частности, если $D = \varnothing$ (т. е. правило $\Delta 3$ «не работает»), то теория $\Delta = W$ имеет единственное расширение $Th(W) = \{X \mid W \vdash X\}$ (при отсутствии правил вывода «по умолчанию» получаем обычную классическую выводимость). В общем случае, у теории может быть несколько расширений или не быть ни одного. Например, теория $\Delta_1 = \left(\left\{\left(\frac{MX}{X}\right), \left(\frac{M\neg X}{\neg X}\right)\right\}, \varnothing\right)$ имеет два расширения $Th(X)$ и $Th(\neg X)$, а теория $\Delta_2 = \left(\left\{\left(\frac{MX}{\neg X}\right)\right\}, \varnothing\right)$ — ни одного. Это позволяет по-разному определять множество теорем. Например, Рейтер предлагает $Th_r(\Delta) = \{X \mid X \in E\}$ для некоторого расширения E теории Δ (формула X «приемлема как допущение в теории Δ») — разумеется, это не единственно возможное определение. Эффекты «немонотонности» проявляются, например, в том, что при расширении теории $\Delta_1 = (D_1, \varnothing)$ до $\Delta_3 = (D_1, \{X\})$ расширение $Th(\neg X)$ пропадает и остается единственное расширение $Th(X)$ (и значит, $Th_r(\Delta_3) \subsetneqq Th_r(\Delta_1)$).

В [72] специально изучаются теории с правилами вида $\left(\frac{A,\ MC}{C}\right)$ (правила типа (*) — так называемые нормальные теории). Доказывается, что у такой теории всегда имеется расширение, и любые два расширения несовместимы (т. е. $(E_1 \neq E_2) \Rightarrow$ $\Rightarrow (Th(E_1 \cup E_2))$ есть противоречивая теория)). Для них выполняется так называемая «полумонотонность» (монотонность «по правилам»): $Th_r(D, W) \subseteq Th_r(D \cup D', W)$ Это позволяет развить некоторый аппарат теории доказательств. При этом устанавливается, что отношение доказумости $\Delta \underset{r}{\vdash} X \rightleftharpoons (X \in Th_r(\Delta)$ (для конечных теорий Δ) не является рекурсивно перечислимым. Это связано с тем, что, проверяя доказуемость формулы X в теории Δ, на некотором шаге приходится проверять классическую совместимость (непротиворечивость) множества формул. Поэтому, в частности, для пропозициональных теорий или для разрешимых классов предикатных формул (например, для логики одноместных предикатов) отношение оказывается разрешимым. Также, для всякой конкретной конечной нормальной теории Δ множество $Th_r(\Delta)$ — рекурсивно перечислимо (хотя эффективного общего способа аксиоматизации всех таких теорий не существует). Кроме того, в [72] рассмотрены теории с правилами типа (**), где формулы A и C не обязательно замкнуты, и для этих теорий также описан некоторый аппарат выводимости, основанный на технике метода резолюций и связанный со сколемовским построением универсума из индивидов—констант.

Альтернативный вариант построения немонотонной логики рассмотрели Макдермотт и Дойл в [63].

При этом символ M («совместимость») понимается не как внешний «метасимвол» при формулировке правил, а вводится в язык — добавляется к языку логики предикатов в качестве дополнительного модального оператора. Интуитивный смысл символа M выражается определением расширения теории T как неподвижной точки оператора $NM_T(E) = Th(T \cup As(E))$, где $As(E) = \{Mx \mid (\neg x) \notin E\}$ (здесь T — множество аксиом теории, т. е. предложений языка с символом M). Как и в [72], теория T может иметь одно или несколько расширений или не иметь ни одного (в последнем случае она называется противоречивой, и все формулы считаются ее теоремами). Например, непротиворечивая теория $T = \{(MA \rightarrow \neg A), \neg A\}$ (равносильная $\{\neg A\}$) содержит противоречивую подтеорию $T' = \{(MA \rightarrow \neg(A)\}$ (пример немонотонности). Ситуация, изучаемая в [63], представляется несколько более общей, чем в [72], а именно: правилам вида (**) здесь соответствуют аксиомы $(A \& MB_1 \& \ldots \& MB_n \rightarrow C)$. Однако поскольку расширения определяются как неподвижные точки разных операторов, то и результаты для соответствующих теорий могут оказаться различными (например, в [72] утверждается, что теория $\Delta=(D, W)$ с $D=\left(\frac{A, MB}{B}\right), \left(\frac{C, MD}{D}\right)$, $W=\{(A\vee C)\}$ имеет единственное расширение $Th(A\vee C)$, а расширение соответствующей теории $T=\{(A\&MB\rightarrow B), (C\&MD\rightarrow\rightarrow D), (A\vee C)\}$ содержит $(A\vee C)$ и $(B\vee D)$. Еще одно отличие состоит в том, что Макдермотт и Дойл определяют множество теорем $Th_d(T)$ не как объединение, а как пересечение всех неподвижных точек, т. е. изучают, так сказать, «надежные» допущения, приемлемые в любом расширении теории T. При этом в [63] формулируется некоторый аппарат семантических таблиц и с его помощью устанавливается разрешимость отношения $T \vdash_d X \rightleftarrows (X\in Th_d(T))$ в случае пропозициональной логики.

Дальнейшее развитие системы Макдермотта и Дойла связано с тем, что в качестве базисной логики при определении оператора NM_T вместо чистого классического исчисления предикатов используется та или иная модальная система, постулирующая определенные свойства оператора M. Исследования в этом направлении см., например, в [64], где применяется известный аппарат моделей Крипке для модальной логики, в которых истинность формул вида MX определяется с учетом истинности X в «возможных мирах». С этим связан также подход к построению немонотонной системы на базе интуиционисткой логики, предложенный Д. Габбаем [56]. А именно: рассматриваются модели Крипке для интуиционистской логики (где возможные миры связаны между собой отношением достижимости, а истинность формул сохраняется при переходе вверх по этому отношению) и добавляется оператор M, понимаемый как «возможность»: MX истинно в некотором мире X, если X истинно в некотором мире Y, достижимом из X. При этом «немонотонность» проявляется при переходе вверх по отношению достижимости (что позволяет определить некоторым образом отношение «немонотонной выводимости» $\Gamma \vdash_h X$).

Машинно-ориентированная формализация правдоподобных выводов, которыми являются методы сходства и различия Д. С. Милля, средствами многозначных логик предикатов с кванторами по кортежам переменной длины была предложена в [37], [38], [39]. Компьютерная реализация правдоподобных рассуждений в стиле Д. С. Милля была осуществлена в [16], [17], [27]. Имитация этих правдоподобных рассуждений на ЭВМ образует ДСМ-метод автоматического порождения гипотез. Исследование средств формализации процедур правдоподобных рассуждений, сформулированных в указанных работах, содержатся в [34], где изучены выразительные средства языка логики предикатов с кванторами по кортежам переменной длины, а также в [1], [3], [4]. Предложенные в этих работах правдоподобные выводы относятся к типу (в), т. е. являются правдоподобными выводами, использующими эффективно порождаемые фальсификаторы: заключение ДСМ-метода автоматического порождения гипотез принимается в качестве гипотезы, если оно не имеет в данном состоянии КАТ фальсификаторов, имеющихся в этом состоянии.

Ввиду важности для приложений в интеллектуальных системах правдоподобных выводов типа (в) в основной части этого раздела мы изложим подробно ДСМ-метод автоматического порождения гипотез, реализующий правдоподобные рассуждения, содержащие правила правдоподобного вывода типа (в).

§ 3. Исходные понятия теории правдоподобных выводов

Исходные принципы формализации правдоподобных выводов и, следовательно, правдоподобных рассуждений сводятся к следующему.

I. Теория правдоподобных рассуждений предполагает принятие некоторой *наивной семантики*, которая является непосредственной символизацией изучаемого «внешнего мира» и содержит *частично* определенные отношения.

II. Частичная определенность отношений наивной семантики определяет ситуацию с *неполной информацией*, которая с необходимостью требует применения правдоподобных рассуждений, являющихся формальным средством порождения гипотез (компьютерная реализация правдоподобных рассуждений дает возможность формулировать различные методы *автоматического порождения гипотез* [10], [16], [17], [65]).

III. Механизмом получения правдоподобных выводов является поиск *сходства* [13], [14], [15] и *различия* рассматриваемых объектов, представимых как структурированные конструктивные объекты (конечные множества, кортежи, слова, графы), причем найденные сходства и различия на заданных выборках объектов должны коррелировать с наличием (или отсутствием) множества свойств, характеризующих эти объекты. Таким образом, целью рассматриваемых нами правдоподобных рассуждений является порождение заключений типа «структура объекта — множество свойств» и, в частности, заключений типа «подструктура объекта — причина наличия (отсутствия) множества свойств».

IV. В связи со сказанным выше рассматриваемый подход к формализации правдоподобных рассуждений и выводов является формализацией эвристических методов Ф. Бэкона — Д. С. Милля и будет систематически изложен в настоящем подразделе.

В настоящем разделе предпринята попытка построения (при определенных эпистемологических допущениях) понятийного каркаса для теории правдоподобных рассуждений. Термин «правдоподобное рассуждение» принадлежит Пойа [30], а его актуальность, по-видимому, связана с тем, что теория эмпирической индукции стимулирована исследованиями по искусственному интеллекту, связанными с имитацией рассуждений на компьютерах. Основные черты нашей точки зрения на теорию правдоподобных рассуждений сводятся к следующему (и частично были затронуты в предыдущих подразделах).

I. Теория правдоподобных рассуждений начинается с принятия некоторой *наивной семантики*, которая является непосредственной символизацией изучаемой действительности и содержит *частично* определенные отношения.

II. Частичная определенность отношений наивной семантики определяет ситуацию с *неполной информацией*, которая с необходимостью требует применения правдоподобных рассуждений, являющихся формальным средством порождения гипотез[1].

III. Механизмом получения правдоподобных заключений является поиск *сходства* и *различия* рассматриваемых объектов, представимых как структурированные конструктивные объекты (конечные множества, кортежи, слова, графы), причем найденные сходства и различия на заданных выборках объектов должны коррелировать с наличием (или отсутствием) множеств свойств, характеризующих эти объекты. Таким образом, целью рассматриваемых нами правдоподобных рассуждений является порождение заключений типа «структура объекта — множество свойств» и, в частности, заключений типа «подструктура объекта — причина наличия (отсутствия) множества свойств».

IV. В связи со сказанным наш подход к формализации правдоподобных рассуждений есть попытка формализации эвристики в стиле Ф. Бэкона — Д. С. Милля, в основу которой положены следующие принципы [26]:

(1) для порождения правдоподобного заключения необходимо рассматривать не только *примеры* явления, но и контрпримеры[2];

(2) в силу предполагаемого единообразия природы сходство в структуре объектов связано с распознаваемой эмпирической зависимостью.

Из сказанного следует, что:

1. так называемые индуктивные методы Д. С. Милля являются «вызовом» формальной теории правдоподобных рассуждений: она должна быть способной их формализовать;

2. частотность повторения в рассматриваемой выборке появления некоторого множества свойств и объектов или их ча-

[1] Компьютерная реализация правдоподобных рассуждений дает возможность формулировать различные методы *автоматического порождения гипотез*.

[2] Сравните этот принцип с техникой распознавания образов.

стей (подобъектов) не является достаточным основанием для порождения гипотез.

V. Достаточное основание для принятия правдоподобного заключения (гипотезы) зависит от способа представления знаний, которым является *квазиаксиоматическая теория* (КАТ); КАТ

$\Gamma=\langle\Sigma, \Sigma', \mathfrak{R}\rangle$, где Σ — множество аксиом, заведомо неполно описывающее предметную область (ПО), Σ' — множество эмпирических элементарных предложений (протокольные предложения), которое является открытым (оно может пополняться посредством проведения новых экспериментов, наблюдений и т. п.); $\mathfrak{R}$ — множество правил вывода, где $\mathfrak{R}=\mathfrak{R}'\cup\mathfrak{R}^0$, $\mathfrak{R}'$ — множество правил правдоподобного вывода ($\mathfrak{R}'\neq\varnothing$), а $\mathfrak{R}^0$ — множество правил достоверного вывода[1].

Каркасом КАТ Γ является $F=\langle\Sigma, \mathfrak{R}\rangle$, множество состояний Γ есть $\Gamma_0, \Gamma_1, \ldots, \Gamma_n, \ldots$, где $\Gamma_n=\langle\Sigma, \Sigma_n', \mathfrak{R}\rangle$ $(n=1, 2, \ldots)$.

Посредством упорядочения (по степеням достоверности заключений множества правил из $\mathfrak{R}'$, проверок сохранения истинностных значений в Γ_n и проверок на непротиворечивость $\Gamma_n\cup\{\varphi\}$, где φ — гипотеза, полученная посредством применения правил из $\mathfrak{R}'$, строится обоснование гипотезы φ, порожденной средствами КАТ[2].

VI. В КАТ реализуются правдоподобные рассуждения, под которыми мы понимаем построение последовательности формул $\psi_1, \ldots, \psi_m$ такой, что ψ_m есть φ, где φ — цель рассуждения, а ψ_i либо суть формулы из $\Sigma\cup\Sigma_n'$ (Σ_n' — открытое множество), либо получены из предыдущих формул посредством правил из $\mathfrak{R}$ так, что использовались правила из $\mathfrak{R}'$; причем в процессе построения $\psi_1, \ldots, \psi_m$ использовались метасредства (например, проверки на непротиворечивость и невыводимость, критерии выбора формул из Σ_n', релевантных ψ_m, и т. п.).

VII. Рассматриваемые нами правдоподобные рассуждения (мы называем их ДСМ-рассуждениями в честь Д. С. Милля) характеризуются следующими условиями:

заданы (посредством Σ') частично определенные отношения $\Rightarrow_1^*$ и $\Rightarrow_2^*$ («объект X обладает множеством свойств Y» и «подобъект X есть причина наличия (отсутствия) множества свойств Y», соответственно), образующие наивную семантику; причем предполагается, что существуют как причины, вызывающие нали-

---

[1] КАТ — широко распространенная форма представления знаний в естественных науках и инженерных дисциплинах (например, в науках о жизни: биохимии, фармакологии, медицине, социологии), уровень формализации которых не достигает аксиоматической теории (АТ).

[2] Очевидно, что в намеченном выше смысле обоснование гипотезы φ слабее, чем ее объяснение.

чие свойств у объектов, так и причины, вызывающие отсутствие этих свойств[1].

VIII. Важным допущением, которое мы принимаем, является допущение о том, что наивная семантика описывается на специальном языке, который мы называем *внутренним языком*. Во внутреннем языке строятся лишь фактоподобные формулы. Факт в нашем понимании есть «фактоподобная формула+ +оценка». Во внутреннем языке описывается структура возможного факта, который может быть построен после характеризации фактоподобной формулы посредством *оценки*. Итак, во внутреннем языке нет средств выражения для оценок и рассуждения. Средства для выражения оценок и формализации рассуждения задаются во *внешнем языке*. В этом смысле внешний язык является метаязыком для внутреннего языка. Идея различения внешнего и внутреннего языка принадлежит Бочвару [6] (в связи с анализом логических и семантических парадоксов) и Рамсею [70] (в связи с логическим анализом научных теорий).

В настоящей работе мы рассмотрим следующую иерархию формальных языков

$$\mathcal{L}_i^0 \subset \mathcal{L}_i \subset \mathcal{L}_e \subset \mathcal{L}'_e \subset \mathcal{L}''_e,$$

где $\mathcal{L}_i^0$, $\mathcal{L}_i$ — внутренние языки, а $\mathcal{L}_e$, $\mathcal{L}'_e$, $\mathcal{L}''_e$ — внешние языки. $\mathcal{L}_i^0$ — язык нулевого порядка, в котором выражаются фактоподобные формулы, $\mathcal{L}_i$ — расширение $\mathcal{L}_i^0$, содержащее индивидные переменные и кванторы; $\mathcal{L}_e$ — расширение $\mathcal{L}_i$, в котором формулы характеризуются некоторыми оценками («истинностными значениями» многозначной логики), в $\mathcal{L}_e$ представимы факты; $\mathcal{L}'_e$ — расширение $\mathcal{L}_e$ такое, что в нем выразимы правила правдоподобного вывода; $\mathcal{L}''_e$ — расширение $\mathcal{L}'_e$ такое, что в нем представимо правдоподобное рассуждение, предствляющее некоторую *стратегию* применения правил правдоподобного вывода, использующую упорядочение $\mathfrak{R}'$.

IX. Задание оценки формул внешних языков $\mathcal{L}_e$ и $\mathcal{L}'_e$ производится в соответствии с принятой наивной семантикой и некоторым способом оценки формул внутреннего языка, соответствующим выбранной *эпистемологии*. Ниже мы опишем один из возможных способов построения оценок формул, полученных посредством правил правдоподобного вывода.

3.1. Внутренние и внешние языки: $\mathcal{L}_i^0$, $\mathcal{L}_i$, $\mathcal{L}_e$.

Пусть дана КАТ $\Gamma = \langle \Sigma, \Sigma', \mathfrak{R} \rangle$, где Σ, Σ' и $\mathfrak{R}$, соответственно, множество аксиом предметной области, открытое множество атомарных (протокольных) высказываний и множество правил

[1] Сравните это требование с принципом обязательного рассмотрения как примеров, так и контрпримеров для построения правдоподобных рассуждений в стиле Бэкона — Милля.

вывода, содержащее помимо дедуктивных правил правила, представляющие правдоподобные заключения (например, правила индуктивного вывода или абдуктивного вывода).

Естественно, что КАТ — форма организации знаний о предметных областях таких, что знания о них хорошо структурированы, но слабо формализованы. Слабая формализованность знаний характеризуется тремя аспектами:

(1) множество аксиом Σ заведомо *неполно* описывает предметную область, представляя лишь фрагментарную ее аксиоматизацию;

(2) в $\mathfrak{R}$ имеются правдоподобные правила вывода, не гарантирующие истинности заключений, но частично аппроксимирующие неполноту описания предметной области посредством аксиом;

(3) множество атомарных высказываний Σ', описывающих изменяющийся мир экспериментов, наблюдений и т. п., *открыто*;

(4) Σ' представляет систему отношений, содержащую частично определенные отношения.

Эти аспекты (1) — (4) характеризуют рассматриваемую эпистемологическую ситуацию как ситуацию с неполной информацией (т. е. ситуацию с заведомо незавершенным и лишь частично заданным знанием о предметной области).

С КАТ посредством Σ' связана наивная семантика, порожденная непосредственно реальным миром — например, некоторой экспериментальной средой или средой наблюдений и т. п. Наивная семантика предполагает как интерпретацию термов и предикатных символов из Σ' в экспериментальной среде (или среде наблюдений), так и задание частично определенных отношений, принадлежащих этой экспериментальной среде и соответствующих высказываниям из Σ'.

Задание указанных частично определенных отношений должно быть осуществлено в специальном языке (впоследствии обозначенном посредством $\mathcal{L}_i^0$). Понятно, что наивная семантика не является формальной семантикой в стандартном логическом смысле.

Эпистемология, соответствующая КАТ, в которой представлены знания с неполной информацией о предметной области, характеризуется определенными связями между наивной семантикой и *оценками* формул из Σ'. Эти оценки задаются в метаязыке языка представления отношений наивной семантики, который мы обозначали посредством $\mathcal{L}_i^0$ [1]). Указанный метаязык языка $\mathcal{L}_i^0$, обозначаемый посредством $\mathcal{L}_e$, мы рассмотрим ниже.

[1]) В связи с этим отметим, что с нашей точки зрения теория правдоподобных рассуждений (и, в частности, теория эмпирической индукции) вряд ли возможна без допущения некоторой наивной семантики. Иллюстрацией этого тезиса является распознавание образов — аппарат прикладной абдукции [67] и индукции [12], использующий в качестве наивной семантики обучающие выборки, заданные как частично определенные отношения.

Задание оценок формул, интерпретированных в наивной семантике, образуют *эпистемологию*, отвечающую КАТ.

Отметим одну важную черту наивной семантики — ее фрагментарность (сравните это обстоятельство с тем, что Σ' — *открытое* множество!). А именно: формула φ выводима в КАТ, заданной посредством Σ, Σ' и $\mathfrak{R}$, может оказаться ложной для другого состояния КАТ, определенного посредством Σ, Σ'' и $\mathfrak{R}$, где $\Sigma' \subset \Sigma''$.

В силу сказанного выше, выводы в КАТ являются *правдоподобными* выводами относительно рассматриваемой наивной семантики.

Упомянутая связь между наивной семантикой и способом оценки формул КАТ может быть задана различными способами, т. е. различными формальными эпистемологиями. В настоящей работе мы пользуемся эпистемологическим принципом, который можно назвать юмовским[1]. Этот принцип состоит в следующем: оценка формулы φ, выводимой в КАТ с использованием правил из $\mathfrak{R}'$, зависит от оценок атомарных формул из Σ' так, что φ не принимает истинностных значений двузначной логики «истина» или «ложь», а имеет аппроксимирующие их (но отличные от них) промежуточные оценки[2].

Строение КАТ предполагает наличие двух языков для формализации правдоподобных выводов: *внутреннего* языка [6], [70] (языка фактоподобных формул), который является языком-посредником между миром экспериментов (наблюдений и т. п.) и миром рассуждений; и внешнего языка [6], [70] — языка рассуждений о фактах, в котором факты анализируются, устанавливается их сходство и различия, формулируются обобщения.

Предлагаемая формализация правдоподобных рассуждений, предполагающая наивную семантику и соответствующую ей эпистемологию юмовского типа, использует иерархию внутренних и внешних языков

$$\mathscr{L}_i^0 \subset \mathscr{L}_i \subset \mathscr{L}_e \subset \mathscr{L}_e' \subset \mathscr{L}_e'',$$

где: $\mathscr{L}_i^0$ — внутренний язык фактоподобных формул, $\mathscr{L}_i$ — внутренний язык, $\mathscr{L}_e$ — внешний язык представления фактов (язык теории), $\mathscr{L}_e'$ — внешний язык формулирования правил правдоподобных выводов, $\mathscr{L}_e''$ — внешний язык формулирования правдоподобных рассуждений (т. е. стратегий, организующих заданное множество правил правдоподобных выводов).

Каркас аппарата исследований правдоподобных рассуждений описывается следующим образом.

[1] В силу его аналогичности тезису Юма [59] об основаниях индукции.

[2] Существует очевидное соответствие этого принципа с юмовским тезисом о том, что универсальные высказывания, полученные в результате индукции, не могут быть верифицированы.

1. Отношения, образующие наивную семантику, формализуются посредством фактоподобных формул, которыми мы называем атомарные формулы $\mathscr{L}_i^0$ и формулы, полученные из них посредством & и V; соответственно, будем говорить об элементарных и неэлементарных фактоподобных формулах. Таким образом, $\mathscr{L}_i^0$ является *минимальным позитивным языком.*

2. $\mathscr{L}_i^0$ погрузим в $\mathscr{L}_e$, задав оценку фактоподобных формул (из $\mathscr{L}_i^0$) на основе наивной семантики. Переводами формул $\mathscr{L}_i^0$ в $\mathscr{L}_e$ будут, таким образом, формулы $\mathscr{L}_e$, имеющие оценки, взятые из наивной семантики и соответствующие принятой эпистемологии.

Итак, в нашем случае погружение внутреннего языка во внешний осуществляется в соответствии с принятой юмовской эпистемологией и отражает заданную наивную семантику. В результате в порождаются *факты,* которые есть

«фактоподобные формулы+оценка».

Следовательно, факты представимы только в теории (в языке $\mathscr{L}_e$).

3.2. Язык $\mathscr{L}_i$.

X, Z, V (быть может, с нижними индексами) — переменные сорта 1 для множеств, представляющих объекты и подобъекты;

Y, U, V (быть может, с нижними индексами) — переменные сорта 2 для множеств свойств;

двухместными предикатными символами являются $\Rightarrow_1$, $\Rightarrow_2$, соответствующие предикатам $X \Rightarrow_1 Y$, $X \Rightarrow_2 Y$;

индивидные константы сорта 1: C, C_1, ..., C_i, ...;

индивидные константы сорта 2: A, A_1, ..., A_j, ...;

логические связки: &, V;

кванторы: $\forall$, $\exists$ (соответственно, для переменных сорта 1 и 2).

Определение терма $\mathscr{L}_i$:

1°. Переменная сорта 1 или индивидная константа сорта 1 есть терм сорта 1;

2°. Переменная сорта 2 или индивидная константа сорта 2 есть терм сорта 2;

3°. T есть терм $\mathscr{L}_i$, если T есть терм сорта 1 или сорта 2.

Определение формулы $\mathscr{L}_i$:

1°. Если T_1 и T_2 — термы сорта 1 и 2, соответственно, то $T_1 \Rightarrow_1 T_2$ и $T_1 \Rightarrow_2 T_2$ — формулы;

2°. если φ и ψ — формулы, то $(\varphi \& \psi)$ и $(\varphi \vee \psi)$ — формулы;

3°. если $\varphi(X)$ и $\psi(Y)$ — формулы, а X и Y входят свободно, соответственно, в φ и ψ, то $\forall X\varphi(X)$, $\exists X\varphi(X)$ и $\forall Y\psi(Y)$, $\exists Y\psi(Y)$ — формулы.

3.3. Язык $\mathscr{L}_i^0$—подъязык $\mathscr{L}_i$.

1°. Если T_1, T_2, соответственно, индивидные константы сорта 1 и 2, то $T_1 \Rightarrow_1 T_2$ и $T_1 \Rightarrow_2 T_2$ — формулы $\mathscr{L}_i^0$;

2°. если φ и ψ — формулы $\mathcal{L}_i^0$, то $(\varphi \& \psi)$ и $(\varphi \vee \psi)$ — формулы $\mathcal{L}_i^0$.

В наивной семантике интерпретируются термы $\mathcal{L}_i$ и предикатные символы $\Rightarrow_1$ и $\Rightarrow_2$, соответствующие частично определенным отношениям $\Rightarrow_1^*$ и $\Rightarrow_2^*$, которые понимаются следующим образом: $X \Rightarrow_1 Y$ и $X \Rightarrow_2 Y$ означают, соответственно, что «объект X обладает (не обладает) множеством свойств Y» и «подобъект X есть причина наличия (отсутствия) множества свойств Y». Атомарные формулы вида $C \Rightarrow_1 A$ и $C \Rightarrow_2 A$ будем называть *фактоподобными*, т. е. формулами способными представлять факты [1].

Обозначим посредством Φ_i и Φ_i^0 — множества формул языков $\mathcal{L}_i$ и $\mathcal{L}_i^0$, соответственно. Очевидно, что $\Phi_i^0 \subset \Phi_i$. Формулы, принадлежащие Φ_i, будем называть *внутренними*. Φ_i^0 будем называть множеством фактоподобных формул. Частичная определенность отношений $\Rightarrow_1^*$ и $\Rightarrow_2^*$, отражающая неполноту информации рассматриваемой предметной области, характеризуемой наивной семантикой, будет формализована средствами специально построенных многозначных логик [2].

3.4. Язык $\mathcal{L}_e$.

Пусть $V_e = V_i \cup \{t, f\}$ — множество истинностных значений (оценок) формул из Φ_i, где t, f, соответственно, «истина» и «ложь» двузначной логики, $t, f \notin V_i$, $V_i \neq \varnothing$, а V_i — множество «промежуточных» истинностных значений, смысл которых будет уточнен ниже.

Словарь внешнего языка $\mathcal{L}_e$ содержит словарь внутреннего языка $\mathcal{L}_i$, а также имеет следующие дополнительные средства:

1. Переменные сорта 3 — i, j, k, l, m, n (быть может, с нижними индексами) для натуральных чисел;

2. Индивидные константы сорта 3 — 0, 1, 2, ..., принадлежащие множеству натуральных чисел $\mathbf{N}$;

3. Множество унарных логических связок $\{J_{\bar{v}} \mid \bar{v} \in V_e\}$ [3], где

$$J_{\bar{v}}\varphi = \begin{cases} t, & \text{если } v(\varphi) = \bar{v} \\ f, & \text{если } v(\varphi) \neq \bar{v} \end{cases}$$

$v(\varphi) \in V_i \cup \{t, f\}$ (v — функция оценки такая, что $v: \Phi_e \to V_e$, где Φ_e — множество формул $\mathcal{L}_e$, определяемое ниже);

---

[1] Ниже будет показано, что эти формулы в языке $\mathcal{L}_e'$ могут представлять как факты, соответсствующие наивной семантике, так и полуфакты, являющиеся гипотезами, порожденными посредством правдоподобных рассуждений.

[2] В [2] сформулированы условия, которым должны удовлетворять многозначные логики, пригодные для формализации правдоподобных рассуждений.

[3] $J_{\bar{v}}$ — характеристические логические связки Россера — Тюркетта.

4. $\cap$, $\cup$, $\setminus$ — функциональные символы (соответствующие теоретико-множественным операциям пересечения, объединения и разности);

5. Предикатные символы: $=$ (предикаты равенства для термов сорта 1 и сорта 2, соответственно), $\leqslant$ (для отношения «меньше или равно» на $\mathbf{N}$).

Определение терма $\mathscr{L}_e$:

1°. Переменные и индивидные константы сортов 1—3 суть термы $\mathscr{L}_e$ сортов 1—3;

2°. Если T_1 и T_2 — термы сорта i $(i=1,2)$, то $T_1 \cap T_2$, $T_1 \cup T_2$, $T_1 \setminus T_2$ — термы сорта i;

3°. Термом $\mathscr{L}_e$ является любой терм сорта i $(i=1,2,3)$. Определение квази-формулы $\mathscr{L}_e$:

1°. Если T_1 и T_2 — термы сортов 1 и 2, соответственно, то $T_1 \Rightarrow_1 T_2$ и $T_1 \Rightarrow_2 T_2$ — квази-формулы;

2°. Если φ и ψ — квази-формулы, то $(\varphi \& \psi)$ и $(\varphi \vee \psi)$ — квази-формулы;

3°. Если $\varphi(X)$ и $\psi(Y)$ — квази-формулы такие, что X и Y входят свободно в φ и ψ, соответственно, то $\forall X \varphi(X)$, $\exists X \varphi(X)$ и $\forall Y \psi(Y)$, $\exists Y \psi(Y)$ — квази-формулы[1]. Определение формулы $\mathscr{L}_e$;

1°. Если T_1 и T_2 — термы сорта i $(i=1,2)$, то $T_1 = T_2$ — формула $\mathscr{L}_e$;

2°. Если T_1 и T_2 — термы сорта 3, то $T_1 \leqslant T_2$ — формула $\mathscr{L}_e$;

3°. Если φ — формула $\mathscr{L}_e$ или квази-формула $\mathscr{L}_e$, то $J_{\bar{\nu}}\varphi$ — формула $\mathscr{L}_e$;

4°. Если φ и ψ — формулы $\mathscr{L}_e$, то $(\varphi \& \psi)$, $(\varphi \vee \psi)$ — формулы $\mathscr{L}_e$;

5°. Если $\varphi(X)$ и $\psi(Y)$ — формулы $\mathscr{L}_e$, а X и Y входят свободно, соответственно, в φ и ψ, то $\forall X \varphi(X)$, $\exists X \varphi(X)$ и $\forall Y \psi(Y)$, $\exists Y \psi(Y)$ — формулы $\mathscr{L}_e$.

Введем теперь следующие определения:

$$T_1 \subseteq T_2 \leftrightharpoons T_1 \cap T_2 = T_1,$$

$$\varphi \to \psi \leftrightharpoons J_f (J_t \varphi) \vee J_t \psi,$$

$$\varphi \leftrightarrow \psi \leftrightharpoons (\varphi \to \psi) \& (\psi \to \varphi).$$

Квази-формулы $\mathscr{L}_i$ будем называть *внутренними* формулами; формулы $\mathscr{L}_e$ будем называть *внешними* формулами (очевидно, что во внешней формуле всякая квази-формула, являющаяся ее подформулой, находится в сфере действия некоторой связки J_ν);

$$\neg \varphi \leftrightharpoons J_f \varphi, \text{ если } \varphi \text{ — внешняя формула.}$$

Уточним теперь понятие «наивной семантики» представимой в $\mathscr{L}_i^0$. Пусть заданы исходные множества $\mathbf{U}^{(1)}$ и $\mathbf{U}^{(2)}$, $\mathbf{U}^{(j)} \neq \varnothing$ $(j=1,2)$. $2^{\mathbf{U}^{(1)}}$ будем называть универсумом объектов. X есть

[1] Очевидно, что каждая формула из Φ_i есть квази-формула $\mathscr{L}_e$.

объект, если $X \subset U^{(1)}$. $2^{U^{(2)}}$ будем называть универсумом множеств свойств, Y — множество свойств, если $Y \subseteq U^{(2)}$.

Рассмотрим множество истинностных значений

$$V_l^{(0)} = \{ \langle 1, 0 \rangle, \langle -1, 0 \rangle, \langle \tau, 0 \rangle \}, \quad V_l^{(0)} \subset V_l,$$

где 1, —1 и τ обозначают, соответственно, типы оценок «фактическая истина», «фактическая ложь» и «недоопределенность». Ради унификации логической терминологии будем эти типы оценок называть типами «истинностных значений». Оценка $\langle \nu, 0 \rangle$ есть результат приписывания типа истинностного значения ν в начальном состоянии КАТ: $n=0$. Тогда $\bar{\nu} = \langle \nu, 0 \rangle$, где $\nu \in \{1, -1, \tau\}$, будем называть *истинностным значением* (оценкой фактоподобной формулы) в *начальном состоянии*, где «0» — указатель начального состояния КАТ. Например, в качестве объектов реального мира можно взять химические соединения, тогда соответствующие им термы будут множества фрагментов структуры этих химических соединений. В качестве же множества свойств можно взять множество биологических активностей химических соединений.

Аналогично: в качестве объектов мы можем взять описание генотипа, а в качестве множества свойств — описание фенотипа; однако в этом случае потребуется использовать другую структуру данных — объекты будут представлены словами[1].

Итак, если в реальном мире проинтерпретированы термы и описаны частично определенные отношения между рассматриваемыми объектами, то наивная семантика является заданной.

В нашем случае мы рассмотрим два частично определенных отношения $\Rightarrow_1^*$ и $\Rightarrow_2^*$: «объект обладает множеством свойств» и «подобъект есть причина наличия (отсутствия) множества свойств», соответственно.

$$\Rightarrow_1^* = \langle \Omega'_{1,0}; \Omega'_{-1,0}; \Omega'_{\tau,0} \rangle,$$
$$\Rightarrow_2^* = \langle \Omega''_{1,0}; \Omega''_{-1,0}; \Omega''_{\tau,0} \rangle,$$

где $\Omega'_{1,0}$ — множество всех пар $\langle C_i, A_j \rangle$ таких, что «C_i обладает A_j» (в начальном состоянии), где $C_i \in 2^{U^{(1)}}$, $A_j \in 2^{U^{(2)}}$; $\Omega'_{-1,0}$ — множество всех пар $\langle C_i, A_j \rangle$ таких, что «C_i не обладает A_j» (в начальном состоянии); $\Omega'_{\tau,0}$ — множество всех пар $\langle C_i, A_j \rangle$ таких, что $\Rightarrow_1^*$ для них недоопределено. Аналогичен случай $\Rightarrow_2^*$.

Для развиваемой нами концепции правдоподобных рассуждений характерно получение истинностных значений как вычислимых оценок, являющихся результатом применения соот-

[1] Сохранение структуры правил правдоподобного вывода при вариации структур данных — необходимое условие разумной теории правдоподобных рассуждений.

ветствующих процедур. В соответствии с этим обстоятельством, а также в связи с необходимостью формализации ситуаций с неполной информацией, в которой в начальном состоянии имеются частично определенные отношения, уточняемые посредством применения правил правдоподобных рассуждений, вводятся 6 типов истинностных значений, а также степени (для 4 из указанных типов). Типы истинностных значений:

(1) t — истина,
(2) f — ложь,
(3) 1 — фактическая истина,
(4) —1 — фактическая ложь,
(5) 0 — фактическая противоречивость,
(6) τ — недоопределенность.

Истинностные значения из V_l имеют вид: $\bar{\nu} = \langle \nu, n \rangle$, где $\nu \in \{-1, 0, 1\}$, $n \in \bar{\mathrm{N}}$, $\bar{\mathrm{N}}$ — множество натуральных чисел, т. е. $\bar{\mathrm{N}} = \mathrm{N} \cup \{\omega\}$ или $\bar{\nu} = (\tau, n)$, для $n < \omega$ имеет место

$$(\tau, n) = \{\langle 1, n+1 \rangle, \langle -1, n+1 \rangle, \langle 0, n+1 \rangle\} \cup (\tau, n+1).$$

Пусть $V = \{1, 0, -1, \tau\}$, $V_\nu = \{\langle \nu, n \rangle \mid n \in \bar{\mathrm{N}}\}$, тогда $V_l = (\bigcup_{\nu \in V} V_\nu)$, где $\bar{\mathrm{N}} = \{0, 1, 2, \ldots, \omega\}$. Множество выделенных значений $\tilde{V}_d$ может задаваться различным способом: $\tilde{V}_d = \{t\} \cup V_d$, где $V_d = (\bigcup_{\nu \in V'_d} V_\nu)$ для некоторого $V'_d \subseteq \{1, 0, -1, \tau\}$. Например, $V'_d = \{1, -1\}$ или $\{1\}$ и т. д.[1]

В связи с обсуждением понятия наивной семантики мы рассмотрели $V_l^{(0)} = \{\langle 1, 0 \rangle, \langle -1, 0 \rangle, \langle 0, 0 \rangle, \langle \tau, 0 \rangle\}$, где $\langle \tau, 0 \rangle = (\tau, 0)$. Истинностные значения из $V_l^{(0)}$ приписываются фактоподобным формулам в силу наивной семантики, тогда как истинностные значения из $V_l \setminus V_l^{(0)}$ приписываются фактоподобным формулам из $\mathscr{L}_l$ посредством процедур правдоподобного вывода, конструктивно сопоставляя этим формулам переводы, соответствующие им в $\mathscr{L}_e$.

Итак, если φ — атомарная формула $\mathscr{L}_l^{(0)}$ (т. е. фактоподобная формула) такая, что ее оценка $v(\varphi) = \langle \nu, 0 \rangle$, где $\nu \in \{-1, 0, 1, \tau\}$, то формулу $J_{\bar{\nu}}\varphi$, где $\bar{\nu} = \langle \nu, 0 \rangle$ мы называем *элементарным фактом* (очевидно, что $J_{\bar{\nu}}\varphi$ — элементарный факт, если $J_{\bar{\nu}}\varphi$ — истинна). Естественным образом определяется язык $\mathscr{L}_e^0 \subset \mathscr{L}_e$ такой, что множество формул Φ_e^0, соответствующее ему, состоит из суперпозиций формул вида $J_{\bar{\nu}}\varphi$, где φ — атомарная формула, построенных посредством & и V. Указанные суперпозиции формул вида $J_{\bar{\nu}}\varphi$ (где φ — атомарная формула) будем называть

[1] Подробное рассмотрение ∞-значной логики, используемой для формализации правдоподобных рассуждений, см. в [1], [3], [49].

неэлементарными фактами, если эти суперпозиции неэквивалентны элементарным фактам и являются истинными формулами

Таким образом, в $\mathscr{L}_e^0$ представлена наивная семантика, заданная частично определенными отношениями $\Rightarrow_1^*$ и $\Rightarrow_2^*$. Причем неполнота информации в рассматриваемой экспериментальной среде формализуется посредством введения четырех типов фактических оценок (-1, 1, 0, τ), рассматриваемых как *внутренние* истинностные значения, отличные от *внешних* истинностных значений t и f. Типы истинностных значений -1, 0, 1 используются для *определенных* оценок, а тип оценок τ — для *недоопределенных*.

Выше мы говорили, что истинностные значения имеют вид $\bar{\nu} = \langle \nu, n \rangle$ ($\nu \in \{-1, 0, 1\}$) или $\bar{\nu} = (\tau, n)$. При $n=0$ истинностные значения относятся к наивной семантике, а $J_{\bar{\nu}}(c \Rightarrow_1 A)$ и $J_{\bar{\nu}}(c \Rightarrow_2 A)$ суть элементарные факты. Напомним, что начальное состояние КАТ с каркасом $\langle \Sigma, \mathfrak{R} \rangle$ имеет вид $T_0 = \langle \Sigma, \Sigma_0', \mathfrak{R} \rangle$.

Конкретизируем теперь T_0: $\Sigma_0' \subset \Phi_e^0$, где Σ_0' — множество всех атомарных формул $\mathscr{L}_e^0$, т. е. множество фактов рассматриваемой наивной семантики. Пусть

$$\Sigma_{\nu,1}^{(0)} = \{J_{\bar{\nu}}(C \Rightarrow_1 A) \mid J_{\bar{\nu}}(C \Rightarrow_1 A) \in \Sigma_0'\},$$

$$\Sigma_{\nu,2}^{(0)} = \{J_{\bar{\nu}}(C \Rightarrow_2 A) \mid J_{\bar{\nu}}(C \Rightarrow_2 A) \in \Sigma_0'\},$$

где $\nu \in \{1, 0, \tau, -1\}$.

Предположим в целях дальнейших построений $\Sigma_{\nu,1}^{(0)} \neq \varnothing$, где $\nu \in \{1, \tau, -1\}$, $\Sigma_{0,1}^{(0)} = \varnothing$, $\Sigma_{\nu,2}^{(0)} = \varnothing$, где $\nu \in \{1, 0, -1\}$, а $\Sigma_{\tau,2}^{(0)} \neq \varnothing$. Соответствия $\Omega_{\nu,0}'$ и $\Sigma_{\nu,1}^{(0)}$, $\Omega_{\nu,0}''$ и $\Sigma_{\nu,2}^{(0)}$ очевидны. Состояние КАТ $T_n = \langle \Sigma, \Sigma_n', \mathfrak{R} \rangle$, где Σ_n' — множество формул вида $J_{\bar{\nu}}(C \Rightarrow_1 A)$ и $J_{\bar{\nu}}(C \Rightarrow_2 A)$, где $\bar{\nu} = \langle \nu, n \rangle$ ($\nu \in \{-1, 0, 1\}$) или $\bar{\nu} = (\tau, n)$, а $n > 0$, охарактеризуем следующим образом: $\Sigma_n' \subset \Phi_e$, где Φ_e — множество формул $\mathscr{L}_l$, а для каждого $m \in \mathrm{N}$ (*) $\Sigma_{\nu,j}^{(m-1)} \subseteq \Sigma_{\nu,j}^{(m)}$, где $m \geqslant 1$, $j = 1, 2$, $\nu \in \{-1, 0, 1\}$. Очевидно, что $\Sigma_{\nu,j}^{(0)} \subset \Sigma_{\nu,j}^{(m)}$ для всех $m \in \mathrm{N}$ и $\nu \in \{-1, 0, 1\}$.

Сформулируем условие *монотонного* преобразования матриц, соответствующих частично определенным отношениям $\Rightarrow_1^*$ и $\Rightarrow_2^*$ при условии расширений $\mathrm{U}^{(1)}$. Указанное условие состоит в следующем: $\tilde{\Sigma}_{\nu,n}' \subseteq \tilde{\Sigma}_{\nu,n+1}'$, где

$$\tilde{\Sigma}_{\nu,m}' = \{C \Rightarrow_1 A \mid J_{\langle \nu,m \rangle}(C \Rightarrow_1 A)\} \cup \{C' \Rightarrow_2 A \mid J_{\langle \nu,m \rangle}(C' \Rightarrow_2 A)\},$$

где

$$\nu \in \{-1, 0, 1\}, \ m \in \mathrm{N}.$$

Если это условие не выполняется, то рассматриваемое преобразование указанных матриц называется немонотонным.

В случае монотонного преобразования этих матриц имеет место следующее утверждение: $\tilde{\Sigma}'_{\tau,n+1} \subseteq \tilde{\Sigma}'_{\tau,n}$, где

$$\tilde{\Sigma}'_{\tau,n} = \{C \Rightarrow_1 A \mid J_{(\tau,m)}(C \Rightarrow_1 A)\} \cup \{C \Rightarrow_2 A \mid J_{(\tau,m)}(C \Rightarrow_2 A)\} \quad (m \in \mathbf{N}).$$

Это утверждение неверно в случае немонотонных преобразований матриц $\Rightarrow_1^*$ и $\Rightarrow_2^*$.

Пусть φ — атомарная формула такая, что $\varphi \in \Phi_i^0$ (т. е. является фактоподобной внутренней формулой $\mathscr{L}_i^0$) и $J_{\bar{\nu}}\varphi$ истинна ($v(\varphi) = \bar{\nu}$), где $\bar{\nu} = \langle \nu, n \rangle$, $\nu \in \{-1, 0, 1\}$ или $\bar{\nu} = (\tau, n)$, где $n > 0$, тогда $J_{\bar{\nu}}\varphi$ будем называть элементарным *псевдофактом.*

Псевдофактом будем называть любую их суперпозицию такую, что она истинна относительно рассматриваемой наивной семантики. *Полуфактом* будем называть факт или псевдофакт. Заметим, что если φ — полуфакт, то φ зависит от n, входящего в соответствующую связку $J_{\bar{\nu}}$, где $\bar{\nu} = \langle \nu, n \rangle$ или $\bar{\nu} = (\tau, n)$.

Таким образом, индуктивное обобщение, полученное посредством правдоподобных выводов, будет обобщением полуфактов, имеющим вид $\forall n \varphi(n)$, где n будет выражать неустранимость фактических оценок, порожденных наивной семантикой. Очевидно также, что если ψ — внутренняя формула, то $J_t\psi$ и $J_f\psi$ — ложны. Оба указанных выше обстоятельства специфицируют логическими средствами юмовский эпистемологический принцип, согласно которому индуктивные обобщения не могут быть верифицированы.

Из построений, сформулированных ниже, следует, что псевдофакты суть заключения правдоподобных выводов, т. е. *гипотезы*. Интуитивно ясно, что степень правдоподобия гипотезы $J_{\langle 1, n \rangle}(c \Rightarrow_i A)$ больше степени правдоподобия гипотезы $J_{\langle 1, m \rangle}(c \Rightarrow_i A)$, если $n < m$ (где $m > 0$, $n > 0$, а $i = 1,2$) и т. п. Естественно, что истинностные значения $\langle \nu, n \rangle$ понимаются как степени или приближения к истинностному значению $\langle \nu, 0 \rangle$, где $\nu \in \{-1, 0, 1\}$.

Одной из основных задач нашей работы является построение правил правдоподобного вывода (абдукций), посредством которых порождаются гипотезы и уменьшается, сколь это возможно, область недоопределенности $\Sigma_\tau^{(0)} = \Sigma_{\tau,1}^{(0)} \cup \Sigma_{\tau,2}^{(0)}$, специфичная для ситуаций с неполной информацией.

Важно отметить, что в рассматриваемых логиках правдоподобных рассуждений можно различным образом задавать типы *выделенных* истинностных значений. Укажем здесь на две интересные возможности, каждая из которых соответствует некоторому эпистемологическому допущению, характеризуемому посредством таблицы

	Выделенные истинностные значения	Невыделенные истинностные значения
(1)	1, −1	0, τ
(2)	1, 0, −1	τ

(1) характеризуется условием принятия высказывания, когда имеются либо аргументы «за» него, либо аргументы «против» него; (2) характеризуется условием неприятия высказывания, когда нет как аргументов «за», так и аргументов «против» (ср. эти условия принятия с глубокой идеей Нильса Бора о взаимоисключающих (дополнительных) описаниях одного и того же объекта различными экспериментальными устройствами [51].

Сделаем теперь некоторые замечания о природе КАТ. Информативность КАТ не исчерпывается, так сказать, «случайностями» и фрагментарностью наблюдаемого мира (экспериментальной средой, т. е. начальным состоянием КАТ), относящимся к наивной семантике, так как в КАТ содержатся утверждения с кванторами общности (в том числе чисто универсальные утверждения о связи полуфактов, описываемых в КАТ). Эти утверждения либо принадлежат Σ, либо выводятся из $\Sigma \cup \Sigma_n'$ посредством $\mathfrak{R}$. Упомянутые утверждения о связи полуфактов в КАТ содержат элементы сигнатуры (т. е. некоторые символы словаря языка), отражающие *лишь существование* наивной семантики, а не специфичность конкретной выборки фактоподобных высказываний, переводимых в $\mathcal{L}_e$ в качестве фактов. Таким образом, в КАТ могут появиться индуктивные обобщения, не содержащие индивидных констант, однако истинностные оценки подформул этих утверждений зависят от принятой связи наивной семантики и эпистемологии (т. е. способа оценки формул). Это обстоятельство выражается посредством $J_{\bar{v}}$ где $v \in V_{\bar{t}}$. Иными словами, замыканию КАТ могут принадлежать формулы с кванторами общности, содержащие $\overline{J_v}$, где $\overline{v} \in V_t$; причем либо эти формулы выводимы из Σ, либо получены с использованием правил правдоподобного вывода, применимым к формулам из Σ_n'.

Обычно дедуктивная теория имеет дело с классом моделей, КАТ же имеет дело с единственной порожденной моделью, соответствующей наивной семантике. Наивная семантика переводится в модель, порожденную пополнением сигнатуры соответствующими индивидными константами (именами) и, возможно, некоторыми отношениями. Приписывание истинностных значений из V_t (т. е. оценок) связано только с моделью, соответствующей наивной семантике; саму же теорию мы интерпретируем в этой модели.

Существенно, что все процедуры правдоподобного рассуждения осуществляются в упомянутой модели, а полученные в результате правдоподобного рассуждения следствия (даже те, которые не имеют индивидных констант) интерпретируются только в модели, соответствующей наивной семантике. Таким образом, формулы с кванторами, не содержащие индивидных констант, являются обобщениями над универсумом упомянутой модели.

Приведенные выше рассмотрения являются эскизом варианта теории правдоподобных рассуждений с юмовским эпистемологическим принципом (в нашей формулировке) и правилами правдоподобных выводов, эксплицирующих схемы рассуждений в стиле Бэкона [8] и Д. С. Милля [26][1]. Однако изложенные выше соображения оснований теории правдоподобных рассуждений могут быть распространены и на другие схемы правдоподобных рассуждений, так что схемы Бэкона — Милля можно рассматривать лишь как важную иллюстрацию предлагаемого подхода к формализации правдоподобных рассуждений.

Второй юмовский эпистемологический тезис о том, что причинно-следственные отношения не являются внутренне присущими объектам реального мира, а устанавливаются при повторяющихся наблюдениях в силу «ассоциации идей», противоположен философским установкам Бэкона и Д. С. Милля о единообразии природы и о существовании в силу этого причинно-следственных отношений, внутренне присущих объектам реального мира. Эти причинно-следственные отношения устанавливаются посредством методов индуктивного рассуждения, порождающих утверждения о сходстве и различии объектов. Эта точка зрения Бэкона — Милля об *эвристической* и продуктивной роли правдоподобных рассуждений есть необходимая метатеоретическая предпосылка наших построений.

§ 4. Внешний язык $\mathscr{L}_e'$ и правила правдоподобных выводов

Для формализации правдоподобных выводов необходимо использовать некоторое расширение языка логики предикатов 1-го порядка [1], [39], [34]. С этой целью мы ниже сформулируем язык $\mathscr{L}_e'$, содержащий подъязык $\mathscr{L}_e$.

При поиске эмпирических зависимостей требуется установить сходство или различие на конечном, но заранее *неопределенном* множестве подтверждающих или опровергающих гипотезу примеров. Число этих примеров k, следовательно, является переменной величиной (k мы называем параметром эмпирической индукции). Это обстоятельство требует расширить язык логики предикатов 1-го порядка, допустив «формулы перемен-

[1] В логико-философской литературе они известны как «таблицы Бэкона» и «Индуктивные методы Д. С. Милля».

ной длины»[1], содержащие кванторы по кортежам переменной длины. Ниже для формализации правил правдоподобного вывода нам понадобится использовать выражения вида

$$\exists k \exists X_0 \exists X_1 \ldots \exists X_{k-1} \exists Y_0 \ldots \exists Y_{k-1} \left[\ldots \underset{l=0}{\overset{k-1}{\&}} (X_l \Rightarrow_r Y_l) \& \ldots\right],$$

где $r=1,2$. Для представления таких «кванторов переменной длины» можно ввести просто обычные кванторы по натуральным числам и функциям, определенным на конечном начальном отрезке натурального ряда (34, 3). При этом кортежу $\langle X_0, X_1, \ldots, X_{k-1}\rangle$ сопоставляется функция $f(i)=X_i$ (для $i<k=l(f)$, где $l(f)$ — длина кортежа f), и обратно, функции f с областью определения $\{0, 1, \ldots, k-1\}$ сопоставляется кортеж $\langle f(0), f(1), \ldots, f(k-1)\rangle$. Тогда, например, формулы вида

$$\exists k \exists X_0 \ldots \exists X_{k-1} \exists Y_0 \ldots \exists Y_{k-1} \left[\ldots \underset{l=0}{\overset{k-1}{\&}} (X_l \Rightarrow_r Y_l) \& \ldots\right]$$

представляются в виде:

$$\exists f \exists g\, [l(f)=l(g) \& (\forall i < l(f)\, (f(i) \Rightarrow_r g(i))) \& \ldots],$$

где $l(f)=k \Leftrightarrow [(\forall i<k) \exists X\, (f(i)=X) \& \neg \exists Y\, (f(k)=Y)]$ ($l(g)$ определяется аналогично).

Мы будем пользоваться более экономной записью формул с кванторами по кортежам переменной длины, считая, что формальное уточнение этой записи может быть сделано указанным выше способом.

Ниже определяется понятие формулы $\mathscr{L}'_e$.

Определение терма $\mathscr{L}'_e$:

1. термы $\mathscr{L}_e$ суть термы $\mathscr{L}'_e$;

2. если $T_1, \ldots, T_k$ — термы сорта 1, k — переменная сорта 3, то $T_1 \cup \ldots \cup T_k$ и $T_1 \cap \ldots \cap T_k$ — термы $\mathscr{L}'_e$.

1. Если T и T' — термы $\mathscr{L}'_e$ сорта i ($i=1, 2$), то $T=T'$ — формула $\mathscr{L}'_e$;

2. если T и T' термы сорта 3, то $T \leqslant T'$ — формула $\mathscr{L}'_e$;

3. если T и T' — термы $\mathscr{L}'_e$ сорта 1 и сорта 2 соответственно, то $T \Rightarrow_1 T'$ и $T \Rightarrow_2 T'$ — формулы $\mathscr{L}'_e$;

4. если $T, T_1, \ldots, T_k$ — термы сорта 1, k — переменная сорта 3, то $(T=T_1 \vee \ldots T=T_k)$ — формула $\mathscr{L}'_e$;

5. если φ, ψ — формулы $\mathscr{L}'_e$, то $(\varphi \& \psi)$, $(\varphi \vee \psi)$ — формулы $\mathscr{L}'_e$;

6. если φ — формула $\mathscr{L}'_e$, то $J_{\bar{\nu}}\varphi$ — формула $\mathscr{L}'_e$, где $\bar{\nu} \in V_e$;

7. если $\varphi(X)$, $\psi(Y)$ — формулы $\mathscr{L}'_e$, а X, Y входят свободно

[1] Таким образом, имеется три вида логических формул: формулы конечной и постоянной длины, формулы конечной, но переменной длины и формулы бесконечной длины.

в φ и ψ соответственно, то $\forall X\varphi(X)$, $\exists X\varphi(X)$, $\forall Y\psi(Y)$, $\exists Y\psi(Y)$ — формулы $\mathscr{L}'_e$ [1];

8. если $T_1, \ldots, T_k$ и $T'_1, \ldots, T'_k$ — термы сорта 1 и сорта 2 соответственно, то $(T_1 \Rightarrow_j T'_1) \& \ldots \& (T_k \Rightarrow_j T'_k)$, $(T_1 \Rightarrow_j T'_1) \vee \ldots \ldots \vee (T_k \Rightarrow_j T'_k)$ — формулы $\mathscr{L}'_e$ $(j=1, 2)$;

9. если φ — формула $\mathscr{L}'_e$, содержащая вхождение $J_{\langle \nu, n \rangle}$ $(\nu \in \{-1, 0, 1\})$ или $J_{(\tau, n)}$, где n входит свободно, то $\exists n\varphi$ и $\forall n\varphi$ — формулы $\mathscr{L}'_e$;

10. если φ — формула $\mathscr{L}'_e$, в которую входят свободно переменные $X_1, \ldots, X_k$, где k — переменная сорта 3, то $\mathsf{Y}_1 X_1 \ldots \ldots \mathsf{Y}_k X_k \varphi$ — формула $\mathscr{L}'_e$, где $\mathsf{Y}_i = \forall, \exists$, $i = 1, \ldots, k$;

11. если φ — формула, в которую входят свободно переменные $Y_1, \ldots, Y_k$, где k — переменная сорта 3, то $\mathsf{Y}_1 Y_1 \ldots \mathsf{Y}_k Y_k \varphi$ — формула $\mathscr{L}'_e$, где $\mathsf{Y}_i = \forall, \exists$, $i = 1, \ldots, k$;

12. если φ — формула, в которую входит свободно переменная k сорта 3, то $\forall k\varphi$, $\exists k\varphi$ — формулы $\mathscr{L}'_e$.

Сделаем теперь следующие замечания.

Замечание 1. Многоточие мы используем в определении терма $\mathscr{L}_e$: для $T_1 \cup \ldots \cup T_k$, $T_1 \cap \ldots \cap T_k$, где k — переменная сорта 3 (близкое по смыслу понятие многоточия было введено в (5)).

Замечание 2. Расширение языка логики предикатов 1-ого порядка посредством введения кванторов по кортежам переменной длины позволяет выразить транзитивное замыкание бинарного отношения R:

$$XR^{\otimes}Z \leftrightarrows \exists k \exists X_0 \ldots \exists X_{k-1} \left[\mathop{\&}_{i=1}^{k-1} (X_{i-1} R X_i) \& (X = X_0) \& (Z = X_{k-1}) \right].$$

(Известно, что транзитивное замыкание не представимо обычными средствами логики предикатов 1-го порядка.)

Замечание 3. Положим, что $|\mathbf{U}^{(1)}| \leqslant \aleph_0$, $|\mathbf{U}^{(2)}| < \aleph_0$. Для случая конечного $\mathbf{U}^{(1)}$ зададим частично определенные отношения $\Rightarrow^*_1$ и $\Rightarrow^*_2$ посредством матриц $M_1^{(n)}$ и $M_2^{(n)}$ со значениями в V_l:

$$M_1^{(n)} \quad \begin{array}{c|ccccc} \Rightarrow^*_1 & A_1 & \ldots & A_j & \ldots & A_{s_2} \\ \hline C_1 & \bar{\nu}_{11} & \ldots & \bar{\nu}_{1j} & \ldots & \bar{\nu}_{1s_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_i & \bar{\nu}_{i1} & \ldots & \bar{\nu}_{ij} & \ldots & \bar{\nu}_{is_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C_{s_1} & \bar{\nu}_{s_1 1} & \ldots & \bar{\nu}_{s_1 j} & \ldots & \bar{\nu}_{s_1 s_2} \end{array} \qquad M_2^{(n)} \quad \begin{array}{c|ccccc} \Rightarrow^*_2 & A_1 & \ldots & A_j & \ldots & A_{s_2} \\ \hline C'_1 & \bar{\mu}_{11} & \ldots & \bar{\mu}_{1j} & \ldots & \bar{\mu}_{1s_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C'_l & \bar{\mu}_{l1} & \ldots & \bar{\mu}_{lj} & \ldots & \bar{\mu}_{ls_2} \\ \vdots & \vdots & & \vdots & & \vdots \\ C'_r & \bar{\mu}_{r1} & \ldots & \bar{\mu}_{rj} & \ldots & \bar{\mu}_{rs_2} \end{array}$$

где $\bar{\nu}_{ij} = \langle \nu_{ij}, n \rangle$ или $\bar{\nu}_{ij} = (\tau, n)$, $\nu_{ij} \in \{-1, 0, 1\}$; $\bar{\mu}_{lj} = \langle \mu_{lj}, n \rangle$ или

$$\bar{\mu}_{lj} = (\tau, n), \quad \mu_{lj} \in \{-1, 0, 1\}, \quad n \in \mathbf{N};$$

[1] Напомним, что X и Y — переменные сорта 1 и сорта 2, соответственно.

$$U^{(1)}=\{d_1, \ldots, d_{r_1}\}, \quad U^{(2)}=\{a_1, \ldots, a_{r_2}\},$$

$$C_i, \ C'_l \in 2^{U^{(1)}}, \quad (i=1, \ldots, s_1;\ l=1, \ldots, r;\ s_1<2^{r_1});$$

$$A_j \in 2^{U^{(2)}} \ (j=1, \ldots, s_2;\ s_2=2^{r_2}).$$

Начальное состояние матрицы $M_1^{(0)}=\|\bar{\nu}_{ij}\|$ представляет бинарное отношение $\Rightarrow_1^*$, заданное на декартовом произведении $\mathscr{X}_1\times 2^{U^{(2)}}$, где $\mathscr{X}_1=\{C_1, \ldots, C_{s_1}\}$, $2^{U^{(2)}}=\{A_1, \ldots, A_{s_2}\}$, причем $\bar{\nu}_{ij}$ для $M_1^{(0)}$ есть либо $\langle 1, 0\rangle$, либо $\langle -1, 0\rangle$, либо $(\tau, 0)$. $M_2^{(0)}=\|\bar{\mu}_{ij}\|$, представляющая бинарное отношение $\Rightarrow_2^*$ в начальном состоянии, задается на декартовом произведении $\mathscr{X}_2\times 2^{U^{(2)}}$, где $\mathscr{X}_2=2^{U^{(1)}}\setminus\mathscr{X}_1$, для $M_2^{(0)}$ $\bar{\mu}_{ij}=(\tau, 0)$, $1\leqslant i\leqslant r$, $r=2^{r_1}-s_1$, $1\leqslant j\leqslant 2^{r_2}$. Преобразования матриц $M_e^{(n)}$ $(e=1, 2)$, означающие доопределения (τ, n)-случаев для $\Rightarrow_1^*$ и $\Rightarrow_2^*$, происходят следующим образом. К матрице $M_1^{(0)}$ применяются п. п. в. I-го рода, а в результате происходит для некоторых атомарных формул $C_i\Rightarrow_2 A_j$ замена оценки $(\tau, 0)$ на $\langle \mu_{ij}, 1\rangle$, т. е. получаем матрицу $M_2^{(1)}$. Затем к $M_2^{(1)}$ применяются п. п. в. II-го рода для доопределения случаев $C_i\Rightarrow_1 A_j$ с оценкой $(\tau, 0)$. Таким образом, посредством $M_1^{(n)}$ преобразуется $M_2^{(n+1)}$, а посредством $M_2^{(n+1)}$ преобразуется $M_1^{(n)}$ в $M_1^{(n+1)}$ и т. д. Этот процесс продолжается пока, не наступает стабилизация, т. е. $(\tau, n+1)=(\tau, n)$ для всех $\bar{\nu}_{ij}$ и $\bar{\mu}_{ij}$. Таким образом, истинностные значения $\bar{\nu}_{ij}$ и $\bar{\mu}_{ij}$ порождаются *конструктивно* посредством п. п. в. I-го и II-го родов.

Замечание 4. В [1], [3], [49] был определен класс бесконечнозначных логик, пригодных для формализации правдоподобных рассуждений в рассматриваемом нами смысле. Рассмотрим следующие примеры определения связок &, $\vee$ из указанного класса.

&	f	-1	0	1	τ	t
f	f	f	f	f	f	f
-1	f	-1	0	0	τ	-1
0	f	0	0	0	0	0
1	f	0	0	1	τ	1
τ	f	τ	0	τ	τ	τ
t	f	-1	0	1	τ	t

$\vee$	f	-1	0	1	τ	t
f	f	-1	0	1	τ	t
-1	-1	-1	-1	τ	τ	t
0	0	-1	0	1	τ	t
1	1	τ	1	1	τ	t
τ	τ	τ	τ	τ	τ	t
t	t	t	t	t	t	t

(матрицы для типов истинностных значений).

Положим

$$\langle \nu, n\rangle \& \langle \nu', m\rangle=\langle \nu \& \nu', \max(n, m)\rangle,$$
$$\langle \nu, n\rangle \vee \langle \nu', m\rangle=\langle \nu \vee \nu', \min(n, m)\rangle.$$

Можно также рассмотреть вместо & связку $\&_1$, определяемую аналогично &, но с заменой условия $\tau\&0=0\&\tau=0$ на $\tau\&_1 0= =0\&_1\tau=\tau$ (т. е. фактически $\tau\&_1\nu=\nu\&_1\tau=\tau$ для всех $\nu\in V$).

Введем теперь следующие определения.

$$V_{(\nu,n)}=\{\langle \nu, m \rangle \mid m\leqslant n\} \text{ для } n\in\overline{\mathbf{N}},\quad \nu\in\{-1, 0, 1\}$$

$$J_{(\nu,n)}\varphi=\begin{cases} t, & \text{если } v(\varphi)\in V_{(\nu,n)},\\ f, & \text{если } v(\varphi)\notin V_{(\nu,n)}.\end{cases}$$

Иначе: $J_{(\nu,n)}\varphi \leftrightharpoons \bigvee\limits_{i=0}^{n} J_{\langle \nu,i \rangle}\varphi$;

$$J_e\varphi \leftrightharpoons J_t\varphi\vee J_f\varphi=\begin{cases} t, & \text{если } v(\varphi)\in\{t, f\},\\ f, & \text{если } v(\varphi)\notin\{t, f\};\end{cases}$$

$$J_i\varphi \leftrightharpoons \neg J_e\varphi=\begin{cases} t, & \text{если } v(\varphi)\in V_i,\\ f, & \text{если } v(\varphi)\notin V_i,\end{cases}$$

$$J_{(\tau,\omega)}\varphi=J_{\langle \tau,\omega \rangle}\varphi=J_i\varphi\&\left(\underset{\nu\in V'}{\&}\neg J_{(\nu,\omega)}\varphi\right)=$$

$$=\begin{cases} t, & \text{если } v(\varphi)=\langle \tau, \omega \rangle\\ f, & \text{если } v(\varphi)\neq\langle \pi, \omega \rangle\end{cases}, \text{ где } V'=\{-1, 0, 1\};$$

$$J_{(\tau,n)}\varphi \leftrightharpoons \neg J_{(0,n)}\varphi\&\neg J_{(-1,n)}\varphi\&\neg J_{(1,n)}\varphi,$$

$$J_{(d,n)}\varphi \leftrightharpoons J_t\varphi\vee\left(\bigvee_{\nu\in V'_d} J_{(\nu,n)}\varphi\right).$$

Ниже мы определим формальные уточнения известных методов правдоподобного вывода Д. С. Милля [26] (методы сходства и различия), которые будут конкретными примерами изучаемых нами правдоподобных выводов.

Предварительно введем неформальные допущения о применимости предлагаемых правил правдоподобного вывода для соответствующих предметных областей (ПО).

I. Для ПО можно определить процедуры установления сходства и различия объектов, представимых как конструктивные объекты (конечные множества, кортежи, слова, графы, пространственные графы и т. п.).

II. В ПО существуют эмпирические зависимости (ЭЗ), представляющие причинно-следственные связи между объектами. Причем *существуют как причины* (являющиеся собственными частями объектов, т. е. подобъектами) *наличия* множеств свойств у объектов из ПО, так и *причины отсутствия* множеств свойств у объектов из ПО. Первые причины будем называть (+)-*причинами*, вторые — (—)-*причинами* (отсутствие свойств у объекта означает или наличие у объекта (—)-причин, или отсутствие (+)-причин, или отсутствие какой-либо определенной информации в рассматриваемой ситуации).

III. Исходные отношения $\Rightarrow_1^*$ и $\Rightarrow_2^*$, выражающие наивную семантику, частично определены, что обуславливает *неполноту информации* о ПО.

IV. В КАТ имеется $\mathfrak{R}'\subset\mathfrak{R}$ такое, что $\mathfrak{R}'\neq\varnothing$, где $\mathfrak{R}'$ — множество правил правдоподобного вывода; причем

$$\mathfrak{R}'=\mathfrak{R}'_{\mathrm{I}}\cup\mathfrak{R}'_{\mathrm{II}},\quad \mathfrak{R}'_{\mathrm{I}}\cap\mathfrak{R}'_{\mathrm{II}}=\varnothing,$$

где $\mathfrak{R}'_{\mathrm{I}}$ — правила, порождающие гипотезы вида $J_{\langle\nu,n\rangle}(C\Rightarrow_2 A)$ или $J_{(\tau,n)}(C\Rightarrow_2 A)$ (C причина A с оценкой правдоподобия $\langle\nu,n\rangle$, где $\nu\in\{-1,0,1\}$, или, соответственно, с оценкой (τ,n)); а $\mathfrak{R}'_{\mathrm{II}}$ — правила, порождающие гипотезы вида $J_{\langle\nu,n\rangle}(C\Rightarrow_1 A)$ ($\nu\in\{1,-1,0\}$) или $J_{(\tau,n)}(C\Rightarrow_1 A)$. Соответственно, правила правдоподобных выводов (п. п. в.) из $\mathfrak{R}'_{\mathrm{I}}$ и $\mathfrak{R}'_{\mathrm{II}}$ будем называть п. п. в. I-го и II-го рода. П. п. в. II-го рода используют гипотезы, полученные как следствия п. п. в. I-го рода для доопределения формул вида $J_{(\tau,n)}(C\Rightarrow_1 A)$, т. е. получения формул (с уточненной оценкой) $J_{\langle\nu,n+1\rangle}(C\Rightarrow_1 A)$ ($\nu\in\{-1,0,1\}$) или $J_{(\tau,n+1)}(C\Rightarrow_1 A)$, где n — номер шага вычислений (п. п. в. II-го рода есть в некотором расширенном смысле правила вывода по аналогии). Таким образом, последовательное применение п. п. в. I-го и II-го родов образуют некоторый способ правдоподобных рассуждений.

Пусть $\mathcal{M}^{+}_{x,n}(V,W)$ и $\mathcal{M}^{-}_{x,n}(V,W)$ — условия, характеризующие наличие (+)-причины V для множества свойств W и наличие (—)-причины V для множества свойств W, соответственно, выразимые в $\mathcal{L}_e'$, где $x\in\mathcal{I}$, $n\in\mathrm{N}$ (n — номер шага вычислений), а $\mathcal{I}$ — множество индексов (имен) соответствующих условий ($\mathcal{I}$ — конечное множество). Тогда п. п. в. I-го рода имеют вид:

$$(I^{(+)}):\frac{J_{(\tau,n)}(V\Rightarrow_2 W),\ \mathcal{M}^{+}_{x,n}(V,W)\,\&\,\neg\mathcal{M}^{-}_{x,n}(V,W)}{J_{\langle 1,n+1\rangle}(V\Rightarrow_2 W)},$$

$$(I^{(-)}):\frac{J_{(\tau,n)}(V\Rightarrow_2 W),\ \mathcal{M}^{-}_{x,n}(V,W)\,\&\,\neg\mathcal{M}^{+}_{x,n}(V,W)}{J_{\langle -1,n+1\rangle}(V\Rightarrow_2 W)},$$

$$(I^{(0)}):\frac{J_{(\tau,n)}(V\Rightarrow_2 W),\ \mathcal{M}^{+}_{x,n}(V,W)\,\&\,\mathcal{M}^{-}_{x,n}(V,W)}{J_{\langle 0,n+1\rangle}(V\Rightarrow_2 W)},$$

$$(I^{(\tau)}):\frac{J_{(\tau,n)}(V\Rightarrow_2 W),\ \neg\mathcal{M}^{+}_{x,n}(V,W)\,\&\,\neg\mathcal{M}^{-}_{x,n}(V,W)}{J_{(\tau,n+1)}(V\Rightarrow_2 W)}.$$

Пусть $\mathfrak{M}=\mathfrak{M}^{+}\cup\mathfrak{M}^{-}$ — множество $\mathcal{M}^{\sigma}_{x,n}$ — условий ($\sigma=+,-$), тогда для приводимого набора правил

$$\mathfrak{M}^{+}=\{\mathcal{M}^{+}_{x,n}(V,W)\}_{n\in\mathrm{N}},\quad \mathfrak{M}^{-}=\{\mathcal{M}^{-}_{x,n}(V,W)\}_{n\in\mathrm{N}},$$

$$\mathfrak{R}'_{\mathrm{I}}=\{(I^{(+)}),\ (I^{(-)}),\ (I^{(0)}),\ (I^{(\tau)})\},$$

где $\mathfrak{R}'_{\mathrm{I}}\subset\mathfrak{R}'$, а $\mathfrak{R}'$ — множество п. п. в. КАТ.

Из определений ($I^{(\sigma)}$), $\sigma \in \{+, -, 0, \tau\}$ ясен конструктивный способ определения истинностных значений из V_i, а также смысл типов истинностных значений —1, 1, 0, и смысл самих истинностных значений $\langle \nu, n \rangle$, где $\nu \in \{-1, 0, 1\}$, и (τ, n). Очевидно, что $\langle 1, n \rangle$ и $\langle -1, n \rangle$ означают установление степени наличия корреляции (ее отсутствия) у V и W (на n-ом шаге), интерпретируемые как «фактическая истина» и «фактическая ложь». Соответственно, $\langle 0, n \rangle$ и $\langle \tau, n \rangle$ означают степени «фактического противоречия» и «недоопределенности» (на n-ом шаге).

Пусть $\Pi_n^{(+)}(V, W)$, $\Pi_n^{(-)}(V, W)$ и $\Pi_n^{(0)}(V, W)$ — условия, характеризующие тот факт, что объект V обладает множеством свойств W, не обладает множеством свойств W, фактическую противоречивость утверждения, что V обладает W (на n-ом шаге вычислений). Тогда п. п. в. II-го рода имеют вид:

$$(II^{(+)}): \frac{J_{(\tau, n)}(V \Rightarrow_1 W),\ \Pi_n^{(+)}(V, W)}{J_{\langle 1, n+1 \rangle}(V \Rightarrow_1 W)},$$

$$(II^{(-)}): \frac{J_{(\tau, n)}(V \Rightarrow_1 W),\ \Pi_n^{(-)}(V, W)}{J_{\langle -1, n+1 \rangle}(V \Rightarrow_1 W)},$$

$$(II^{(0)}): \frac{J_{(\tau, n)}(V \Rightarrow_1 W),\ \Pi_n^{(0)}(V, W)}{J_{\langle 0, n+1 \rangle}(V \Rightarrow_1 W)},$$

$$(II^{(\tau)}): \frac{J_{(\tau, n)}(V \Rightarrow_1 W),\ \Pi_n^{(\tau)}(V, W)}{J_{(\tau, n+1)}(V \Rightarrow_1 W)},$$

где $\Pi_n^{(\tau)}(V, W) \rightleftharpoons \neg \Pi_n^{(+)}(V, W) \& \neg \Pi_n^{(-)}(V, W) \& \neg \Pi_n^{(0)}(V, W)$. $\mathfrak{R}'_{\mathrm{II}} = \{(II^{(+)}), (II^{(-)}), (II^{(0)}), (II^{(\tau)})\}$, $\mathfrak{R}' = \mathfrak{R}'_{\mathrm{I}} \cup \mathfrak{R}'_{\mathrm{II}}$.

Ниже мы сформулируем примеры $\mathscr{M}^{\sigma}_{x,n}$-предикатов и $\Pi_n^{(\sigma)}$-предикатов.

Прямой положительный метод сходства $(x = a)$ Д. С. Милля: $\mathscr{M}^{+}_{a,n}(V, W) \rightleftharpoons \exists k \tilde{\mathscr{M}}^{+}_{a,n}(V, W, k)$, где

$$\begin{aligned} &\tilde{\mathscr{M}}^{+}_{a,n}(V, W, k) \rightleftharpoons \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots \exists U_k (J_{(1,n)}(Z_1 \Rightarrow_1 U_1) \& \\ &\& \forall U (J_{(1,n)}(Z_1 \Rightarrow_1 U) \to U \subseteq U_1)) \& \ldots \& (J_{(1,n)}(Z_k \Rightarrow_1 U_k) \& \\ &\& \forall U (J_{(1,n)}(Z_k \Rightarrow_1 U) \to U \subseteq U_k)) \& (Z_1 \cap \ldots \cap Z_k = V) \& V \neq \varnothing \& \ldots \\ &\& \forall i \forall j ((i \neq j \& 1 \leqslant i,\ j \leqslant k) \to Z_i \neq Z_j) \& \\ &\& \forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \& \forall U (J_{(1,n)}(X \Rightarrow_1 U) \to U \subseteq Y) \& V \subset X) \to \\ &\to (W \neq \varnothing \& W \subseteq Y \& (\bigvee_{i=1}^{k} (X = Z_i)))) \& k \geqslant 2). \end{aligned}$$

Комментарии.

(1) Мы определили $(+)$-предикат, выражающий локальное сходство на $(+)$-примерах $J_{(1,n)}(Z_i \Rightarrow_1 U_i)$, $i = 1, \ldots, k$, где k —

переменная (в этом смысле мы говорили выше, что п. п. в. 1-го рода применяются к $M_1^{(n)}$ для $\overset{*}{\Rightarrow}_1$).

(2) Локальное сходство для рассматриваемой нами булевой структуры данных с $\mathscr{B}_1 = \langle 2^{U^{(1)}}, U^{(1)}, \varnothing, \setminus, \cap, \cup \rangle$, $\mathscr{B}_2 = \langle 2^{U^{(2)}}, U^{(2)}, \varnothing, \setminus, \cap, \cup \rangle$ выражается посредством подформулы переменной длины

$$(Z_1 \cap \ldots \cap Z_k) = V \,\&\, V \neq \varnothing.$$

(3) Подформулы $\forall U\,(J_{(1,n)}(Z_i \Rightarrow_1 U) \rightarrow U \subseteq U_i)$, $i = 1, \ldots, k$, выражают тот факт, что в примере $J_{(1,n)}(Z_i \Rightarrow_1 U_i)$ объект Z_i обладает максимально возможным множеством свойств ($i = 1, \ldots, k$).

(4) Формула $\forall X \forall Y\,((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, \forall U\,(J_{(1,n)}(X \Rightarrow_1 U) \rightarrow U \subseteq Y) \,\&\, V \subset X) \rightarrow (W \neq \varnothing \,\&\, W \subseteq Y))$ следует из $\mathscr{M}^+_{a,n}(V, W)$ и выражает *эмпирическую зависимость* (ЭЗ) между V и W, *прогнозируемую* как причинное отношение («V причина W»).

(5) Подформула $\bigvee\limits_{i=1}^{k}(X = Z_i)$ выражает так называемое условие исчерпываемости — требование рассмотрения всех имеющихся в модели примеров, включающих подходящее V.

(6) Условие $k \geqslant 2$ очевидно.

(7) Уточнением (+)-метода сходства Д. С. Милля будет, таким образом, п. п. в.:

$$(I^{(+)}): \frac{J_{(\tau,n)}(V \Rightarrow_2 W),\ \mathscr{M}^+_{a,n}(V, W) \,\&\, \neg \mathscr{M}^-_{a;n}(V, W)}{J_{\langle 1,n+1 \rangle}(V \Rightarrow_2 W)},$$

где $\mathscr{M}^-_{a,n}(V, W) \rightleftharpoons \exists k \tilde{\mathscr{M}}^-_{a,n}(V, W)$, а $\tilde{\mathscr{M}}^-_{a,n}(V, W)$ получается из $\tilde{\mathscr{M}}^+_{a,n}(V, W)$ заменой $J_{(1,n)}$ на $J_{(-1,n)}$, $J_{\langle 1,n+1 \rangle}$ на $J_{\langle -1,n+1 \rangle}$.

Аналогично формулируется (—)-метод сходства как п. п. в. $(I^{(-)})$. Наша формализация требует также введения п. п. в. $(I^{(0)})$ и $(I^{(\tau)})$. Таким образом, строго говоря, уточнением прямого метода сходства Д. С. Милля с ЭЗ типа *«от причины — к следствию»* (см. (4)) является множество п. п. в. $\{(I^{(+)}), (I^{(-)}), (I^{(0)}), (I^{(\tau)})\}$. (8) Легко видеть, что п. п. в. $(I^{(\sigma)})$; $\sigma \in \{+, -, 0, \tau\}$, являются *абдуктивными* [55] п. п. в., а не индуктивными (как их традиционно называют), так как если C' и A, соответственно, значения V и W, то выводимые формулы

$$J_{\langle \nu,n+1 \rangle}(C' \Rightarrow_2 A) \text{ или } J_{(\tau,n+1)}(C' \Rightarrow_2 A),$$

где $\nu \in \{1, 0, -1\}$, суть элементарные или сингулярные формулы $\mathscr{L}_e$, и мы получаем вывод *«от частного — к частному»*, т. е. абдукцию в смысле Пирса [67].

Расширим теперь множества $\mathfrak{M}^+$ и $\mathfrak{M}^-$, а следовательно и $\mathfrak{R}'_e$, следующим образом. Рассмотрим следующие формулы $(1)^+$—$(6)^+$, которые добавим к $\mathscr{M}^+_{a,n}(V, W)$ и получим

$\mathscr{M}^+_{x,n}(V, W)$, где $\mathscr{M}^+_{x,n}(V, W) \leftrightharpoons \mathscr{M}^+_{a,n}(V; W) \& (m)^{(+)}$, $m=1, \ldots$ $\ldots, 6$; положим $x=(am)^+$.

$(1)^+$ $\forall X \forall Y ((V \subset X \& W \subseteq Y) \to (J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(0,n)}(X \Rightarrow_1 Y) \vee$ $\vee J_{(\tau,n)}(X \Rightarrow_1 Y)))$,

$(2)^+$ $\exists Z \exists U (\neg (V \subset Z) \& J_{(-1,n)}(Z \Rightarrow_1 U) \& W \subset U)$,

$(3)^+$ $\exists Z \exists U (\neg (V \subset Z) \& (J_{(-1,n)}(Z \Rightarrow_1 U) \vee J_{(\tau,n)}(Z \Rightarrow_1 U)) \& W \subseteq U)$,

$(4)^+$ $\forall X \forall Y (V \subset X \& W \subseteq Y) \to (J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)))$,

$(5)^+$ $\forall Z \forall U ((\neg (V \subset Z) \& W \subseteq U) \to (J_{(-1,n)}(Z \Rightarrow_1 U) \vee$ $\vee J_{(\tau,n)}(Z \Rightarrow_1 U)))$,

$(6)^+$ $\forall X \forall Y \exists Z \exists V_0 \exists W_0 ((J_{(1,n)}(X \Rightarrow_1 Y) \& W \subseteq Y \& V \subseteq X) \to$ $\to (Z = ((X \setminus V) \cup V_0) \& \neg (V \subset V_0) \& V_0 \neq \varnothing \& (J_{(-1,n)}(Z \Rightarrow_1 W \cup W_0) \vee$ $\vee J_{(\tau,n)}(Z \Rightarrow_1 W \cup W_0))))$.

Легко показать, что логически зависимыми являются лишь $(4)^+$ и $(1)^+$, $(2)^+$ и $(3)^+$, а именно: $(4)^+ \vdash (1)^+$ и $(2)^+ \vdash (3)^+$. $(4)^+ \& (2)^+ \nvdash (6)^+$, $(4)^+ \& (2)^+ \nvdash (5)^+$ и т. д.

В особенности интересны $(4)^+$, $(5)^+$ и $(6)^+$. $(4)^+$ выражает следующее условие «запрета на контрпримеры»: $(+)$-причина V множества свойств W *не* должна включаться в $(-)$-*примеры для* $\Rightarrow_1$ [28].

$(5)^+$ выражает условие *единственности* $(+)$-причины V множества свойств W [28].

$(6)^+$ выражает дополнительное условие для $(+)$-метода сходства Д. С. *Милля*, превращающее последний метод в $(+)$-*метод различия*.

Естественным образом, могут быть получены $(-)$-аналоги условий $(1)^+$—$(6)^+$: $(1)^-$—$(6)^-$. Например:

$(4)^-$ $\forall X \forall Y ((V \subset X \& W \subseteq Y) \to (J_{(-1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)))$.

Пусть $\mathfrak{M}^+$ — множество всех $\mathscr{M}^+_{x,n}$-предикатов, полученных посредством выбора всех непустых подмножеств из $\Delta^+ = \{(1)^+, \ldots, (6)^+\}$. Используя $p \& q \to p$ и $(4)^+ \vdash (1)^+$, $(2)^+ \vdash (3)^+$, определим отношение частичного порядка на $\mathfrak{M}^+$: $\mathscr{M}^+_{x,n} \sqsupseteq \mathscr{M}^+_{y,n}$ («$\mathscr{M}^+_{y,n}$ следует из $\mathscr{M}^+_{x,n}$»), где x и y отвечает объединение $\{a\}$ и некоторого подмножества из Δ^+ (например, $\{a, (1)^+, (3)^+, (6)^+\}$, тогда $x = a(1)^+(3)^+(6)^+$ запишем в естественном порядке элементов из Δ^+). Так как имеет место

$$\mathscr{M}^+_{x,n} \sqsupseteq \mathscr{M}^+_{y,n} \& \mathscr{M}^+_{y,n} \sqsupseteq \mathscr{M}^+_{x,n} \to \mathscr{M}^+_{x,n} = \mathscr{M}^+_{y,n},$$

то $\sqsupseteq$ — частичный порядок. Более того, $\mathfrak{M}^+$ — решетка с наименьшим элементом $\mathscr{M}^+_{a,n}$. Аналогично определим решетку $\mathfrak{M}^-$ с наименьшим элементом $\mathscr{M}^-_{a,n}$. Ради удобства обозначений будем рассматривать множество имен $\mathcal{J}^+$ и $\mathcal{J}^-$ элементов решеток $\mathfrak{M}^+$ и $\mathfrak{M}^-$. Пусть $\mathrm{I}^+ = \langle \mathcal{J}^+, \sqsupseteq \rangle$ и $\mathrm{I}^- = \langle \mathcal{J}^-, \sqsupseteq \rangle$. Очевидно, что I^+ и I^- — изоморфны. Пусть h — изоморфизм I^+ на I^-

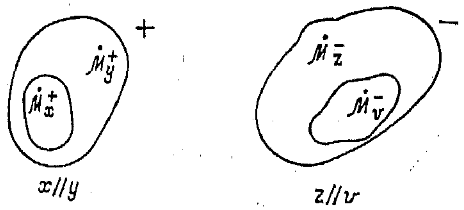
$$h: \mathrm{I}^+ \to \mathrm{I}^-.$$

Рассмотрим $\mathcal{I}=\mathcal{I}^+\cup\mathcal{I}^-$. На $\mathcal{I}$ определим предпорядок $\succeq$ следующим образом[1]. $x\succeq y$ для $x, y\in\mathcal{I}$, если:

1°. $x, y\in\mathcal{I}^+ (x, y\in\mathcal{I}^-)$ и $x\sqsupseteq y$ в $\mathrm{I}^+ (\mathrm{I}^-)$; или

2°. $x\in\mathcal{I}^+$, $y\in\mathcal{I}^-$, $h(x)=y'$ и $y'\sqsupseteq y$ в I^-.

Определим отношение $\parallel$ несравнимости x и y на $\mathcal{I}$: $x\parallel y\leftrightharpoons\neg(x\succeq y\vee y\succeq x)$.



Теперь для построенного $\mathfrak{M}=\mathfrak{M}^+\cup\mathfrak{M}^-$ (соответственно; $\mathcal{I}=\mathcal{I}^+\cup\mathcal{I}^-$) сформулируем п. п. в. I-го рода:

$$(I^{(+)}):\ \frac{J_{(\tau,n)}(V, W),\ \mathcal{M}^+_{x,n}(V, W)\,\&\,\forall y\,(y\succeq x\to\neg\mathcal{M}^-_{y,n}(V, W))}{J_{\langle 1,n+1\rangle}(V\Rightarrow_2 W)},$$

$$(I^{(-)}):\ \frac{J_{(\tau,n)}(V, W),\ \mathcal{M}^-_{x,n}(V, W)\,\&\,\forall y\,(y\succeq x\to\neg\mathcal{M}^+_{y,n}(V, W))}{J_{\langle -1,n+1\rangle}(V\Rightarrow_2 W)},$$

$$(I^{(0)}):\ \frac{J_{(\tau,n)}(V, W),\ \mathcal{M}^+_{x,n}(V, W)\,\&\,\mathcal{M}^-_{y,n}(V, W)\,\&\,(x\succeq y\,\&\,y\succeq x)}{J_{\langle 0,n+1\rangle}(V\Rightarrow_2 W)},$$

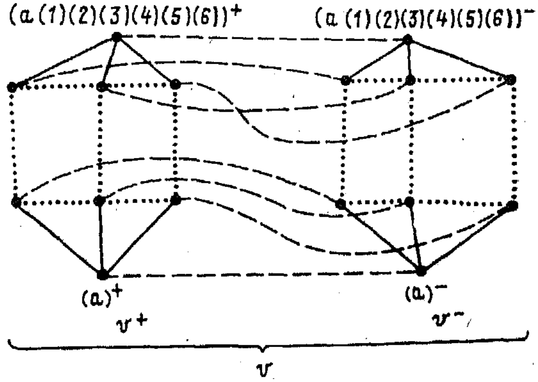
$$(I_1^{(\tau)}):\ \frac{J_{(\tau,n)}(V, W),\ \neg\mathcal{M}^+_{a,n}(V, W)\,\&\,\neg\mathcal{M}^-_{a,n}(V, W)}{J_{(\tau,n+1)}(V\Rightarrow_2 W)},$$

$$(I_2^{(\tau)}):\ \frac{J_{(\tau,n)}(V\Rightarrow_2 W),\ \mathcal{M}^+_{x,n}(V, W)\,\&\,\mathcal{M}^-_{y,n}(V, W)\,\&\,(x\parallel y)}{J_{(\tau,n+1)}(V\Rightarrow_2 W)}.$$

З а м е ч а н и е 5. Пусть v^+ и v^- — диаграммы решеток I^+ и I^-, соответственно,

---

[1] Из приводимого ниже определения $\succ$ следует, что для $\succ$ не имеет место антисимметричность.



а v — граф предпорядка $\mathrm{I}=\langle I, \succ\!\!-\rangle$, определенного выше (вершины $x\,v^{+}$ соединяем с вершинами $y\,v^{-}$ тогда и только тогда, когда $y=h(x)$ для всех $x, y: x\in I^{+}, y\in I^{-}$).

Очевидно, что антисимметричность в I не имеет места, так как из $x\succ\!\!-y\,\&\,y\succ\!\!-x$ не следует $x=y$ в силу $\mathscr{M}^{+}_{x,n}\neq\mathscr{M}^{-}_{y,n}$.

Замечание 6. Введение отношений $\sqsupseteq$ и $\succ\!\!-$ означает расширение $\mathscr{L}'_{e}$ до $\mathscr{L}''_{e}$, содержащего множество имен $\mathcal{I}=\mathcal{I}^{+}\cup\mathcal{I}^{-}$, соответствующие переменные $x, y, z, \ldots$ сорта 4 и предикаты, представляющие $\sqsupseteq$ и $\succ\!\!-$[1].

Замечание 7. Использование $\mathrm{I}=\langle \mathcal{I}, \succ\!\!-\rangle$ делает возможным ввести нелинейные оценки формул внутреннего языка $\mathscr{L}_{i}$ в зависимости от $\mathscr{M}^{\sigma}_{x,n}$, используемого в применяемом п. п. в. ($\sigma\in\{+, -\}$), что означает двупараметричность оценки

$$\tilde{v}=\langle \bar{v}, \mathscr{M}^{\sigma}_{x,n}\rangle=\langle\langle v, n+1\rangle, \mathscr{M}^{\sigma}_{x,n}\rangle.$$

Таким образом, множество оценок $\tilde{v}$ является *вычислимым* и *нелинейным*. Развиваемая нами теория правдоподобных рассуждений основана на следующем тезисе: предпорядок с вычислимыми (посредством п. п. в.) истинностными значениями противопоставлен линейным оценкам из $[0, 1]$ для нечетких множеств, вводимым посредством выбираемым «внешним» образом функций принадлежности $\mu(x)$. В нашем случае истинностные значения порождаются «снизу» из обрабатываемой модели мира (связанной с наивной семантикой)[2], а не наклады-

[1] В настоящей работе мы не рассматриваем конструкции языка $\mathscr{L}''_{e}$, который является метаязыком языка $\mathscr{L}'_{e}$ ($\mathscr{L}'_{e}\subset\mathscr{L}''_{e}$); в $\mathscr{L}''_{e}$ формализуются стратегии правдоподобных выводов, образующие правдоподобные рассуждения.

[2] Причем, учитываются аргументы «за гипотезу» и «против нее».

ваются «сверху»; $+1$, -1, 0, τ — ранжирование оценок и последующая их градация по степеням правдоподобия, зависящим от шага вычисления n и от вида правил (т. е. предпорядка, учитывающего их силу), нам представляется более жизненным логическим аппаратом, имитирующим нюансы человеческих рассуждений.

Замечание 8. Организацию, заданную на множестве процедур, определяющую последовательность их применения и условие окончания их работы, будем называть *стратегией* (STR). В нашем случае процедуры — это п. п. в., а STR образует структуру рассуждения, приводящего к порождению **гипотез**. Таким образом, метод автоматического порождения гипотез реализует некоторое множество γ. В настоящей работе мы не рассматриваем вопросы, связанные с точным определением STR и упорядочением их на γ (см. в связи с этим [28]). Однако отметим следующие обстоятельства.

1. Определив STR формально как некоторое алгоритмическое управление п. п. в. и задав на γ *упорядочение*, мы можем использовать более тонкие оценки, претендующие на формализацию *рефлексии* как оценки способа получения гипотезы. Таким образом, теперь оценка может быть трехпараметрической $\nu^* = \langle \tilde{\nu}, \mathrm{STR} \rangle = \langle \langle \nu, n+1 \rangle, \mathscr{M}^{\sigma}_{x,n}, \mathrm{STR} \rangle$.

2. Формализация STR требует введения двух уровней внешнего языка: уровня описания п. п. в. (язык $\mathscr{L}'_e$) и уровня описания STR (язык $\mathscr{L}''_e$). Таким образом, мы ищем последовательность следующих языков, $\mathscr{L}^0_i$, $\mathscr{L}_i$, $\mathscr{L}_e$, $\mathscr{L}'_e$, $\mathscr{L}''_e$, где $\mathscr{L}^0_i$ — язык фактоподобных формул, $\mathscr{L}_i$ — внутренний язык ($\mathscr{L}^0_i \subset \mathscr{L}_i$); $\mathscr{L}_e$ — язык представления полуфактов (т. е. фактов и гипотез), $\mathscr{L}'_e$ — язык описания п. п. в., $\mathscr{L}''_e$ — язык описания STR, $\mathscr{L}''_e$ — метаязык $\mathscr{L}'_e$. $\mathscr{L}_e$ — метаязык $\mathscr{L}_i$. Как мы уже отмечали, фармализация правдоподобных рассуждений дается не дешево.

Замечание 9. Вернемся к простому случаю $\mathfrak{M}^+ = \{\mathscr{M}^+_{a,n}\}_{n \in N}$ и $\mathfrak{M}^- = \{\mathscr{M}^-_{a,n}\}_{n \in N}$. Расширим $\mathfrak{M}^+$ и $\mathfrak{M}^-$, например, взяв $\mathscr{Y} = \{a, (1), \ldots, (6)\}$, но не определяя предпорядка на $\mathscr{Y}$. Тогда получим STR типа «мешок» с п. п. в. I-го рода вида

$$(I^{(\sigma)}): \ \frac{J_{(\tau,n)}(V \Rightarrow_2 W),\ \mathscr{M}^{\sigma}_{x,n}(V, W),\ \{\neg \mathscr{M}^{\bar{\sigma}}_{y,n}(V, W) \mid y \in \mathscr{Y}\}}{J_{\langle \nu, n+1 \rangle}(V \Rightarrow_2 W)},$$

где $\sigma \in \{+, -\}$, $\bar{\sigma} = \begin{cases} +, & \text{если } \sigma = - \\ -, & \text{если } \sigma = + \end{cases}$,

$$\nu = \begin{cases} 1, & \text{если } \sigma = + \\ -1, & \text{если } \sigma = - \end{cases}.$$

$(I^{(0)})$ и $(I^{(\tau)})$ определяются аналогично. Очевидно, что STR «мешок» менее информативна, чем STR с предпорядком на $\mathcal{Y}$.

Мы рассмотрели прямые п. п. в. I-го рода, распознающие ЭЗ типа «от причины — к следствию». Ниже мы рассмотрим пример п. п. в. I-го рода, распознающего ЭЗ типа *«от следствия — к причине»*.

$$\overline{\mathcal{M}}^{+}_{a(s)\overset{+}{0},n}(V, W) \Leftrightarrow \exists k \overline{\tilde{\mathcal{M}}}^{+}_{a(s)\overset{+}{0},n}(W, V, k),,$$

где $(5)^{\circ}$ — ЭЗ, являющаяся аналогом (5) (для случая «от следствия — к причине»).

$$\overline{\tilde{\mathcal{M}}}^{+}_{a(s)\overset{+}{0},n}(W, V, k) \Leftrightarrow \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots$$

$$\exists U_k \left(\mathop{\&}_{i=1}^{k} (J_{(1,n)}(Z_i \Rightarrow_1 U_i) \,\&\, \forall U (J_{(1,n)}(Z_i \Rightarrow_1 U) \to U \subseteq U_i)) \,\&\right.$$

$$\& \left(W = \bigcap_{i=1}^{k} U_i \right) \& \forall i \forall j\, ((i \neq j) \,\&\, 1 \leqslant i,\ j \leqslant k) \to Z_i \neq Z_j \Big) \,\&$$

$$\& W \neq \varnothing \,\&\, \forall X \forall Y\, ((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, \forall U\, (J_{(1,n)}(X \Rightarrow_1 U) \to U \subseteq Y) \,\&$$

$$\left.\& W \subseteq Y) \to \left(V \subset X \,\&\, V \neq \varnothing \,\&\left(\bigvee_{i=1}^{k} (X = Z_i) \right)\right)\right) \,\&\, k \geqslant 2 \Big).$$

Отметим, что Д. С. Милль в [26] предусмотрел в формулировке метода сходства две возможности: «от причины — к следствию» и «от следствия — к причине». Он также рассматривал вариант метода сходства, в котором прогнозируемая причина *единственна* (метод единственного сходства). Приведенный выше предикат $\overline{\mathcal{M}}^{+}_{a(s)\overset{+}{0},n}(W, V)$ и есть предикат положительного обратного метода единственного сходства. Для формулировки соответствующего п. п. в. аналогично определим $\overline{\mathcal{M}}^{-}_{a(s)\overset{-}{0},n}(W, V)$ заменой в $\overline{\mathcal{M}}^{+}_{a(s)\overset{+}{0},n}(W, V)$ $J_{(1,n)}$ на $J_{(-1,n)}$) и введем новое отношение $\overset{*}{_3}\Leftarrow$ («W следствие V») и соответствующий предикат $W \overset{*}{_3}\Leftarrow V$ (очевидно, что $\overset{*}{_3}\Leftarrow$ не является обратным отношением для $\Rightarrow^{*}_{2}$).

Получим следующее п. п. в.:

$$\overline{(I^{(+)})}: \frac{J_{(\tau,n)}(W_3 \Leftarrow V),\ \overline{\mathcal{M}}^{+}_{a(s)\overset{+}{0},n}(W, V) \,\&\, \neg \overline{\mathcal{M}}^{-}_{a(s)\overset{-}{0},n}(W, V)}{J_{\langle 1,n+1 \rangle}(W_3 \Leftarrow V)}.$$

Аналогично определим $\overline{(I^{(\sigma)})}$, $\sigma \in \{0, -, \tau\}$. Применение обратных п. п. в. естественно для предметных областей с богатой феноменологией, когда свойств много и они существенно характеризуют объекты, структура которых не представлена детально. Грубо говоря, в этих предметных областях действует принцип: *множество свойств, присущих объекту, идентифицирует объект*. Понятно, что обратные п. п. в. существенно связаны с диагностическими задачами.

Отметим, что мы не приводим здесь общий случай определения обратного п. п. в. для неединственного сходства ввиду его громоздкости. Однако обратные аналоги ЭЗ прямых п. п. в. a, (1), ..., (6) могут быть сформулированы (наш пример был с аналогом ЭЗ (5)).

Обратимся теперь к правилам II-го рода, порождающим утверждения относительно наличия или отсутствия множества свойств у объекта V, т. е. рассмотрим п. п. в., позволяющие преобразовывать матрицу $M_1^{(n)}$ (отношения $\Rightarrow_1^*$ (уменьшать, если возможно, число τ-случаев отношения $\Rightarrow_1^*$).

$$\widetilde{\Pi}_n^+(V, W, k) \leftrightharpoons \exists Y_1 \ldots \exists Y_k \Bigg(\Big(\mathop{\&}_{i=1}^{k} \exists X_i \, (J_{(1,n)}(X_i \Rightarrow_2 Y_i) \,\&\, X_i \subset V) \,\&$$
$$\& \Big(\bigcup_{i=1}^{k} Y_i = W \Big) \& \forall Y \Big(\exists X (J_{(1,n)}(X \Rightarrow_2 Y) \,\&\, X \subset V) \to \Big(\bigvee_{k=1}^{k} (Y = Y_l) \Big) \Big) \&$$
$$\& \forall u \, ((u \subseteq W \,\&\, u \neq \varnothing) \to \neg \exists Z \, (J_{(-1,n)} \, (Z \Rightarrow_2 U) \,\&\, Z \subseteq V)) \Bigg).$$

Комментарии.

Подформула $\mathop{\&}_{i=1}^{k} (J_{(1,n)} \, (X_i \Rightarrow_2 Y_i) \,\&\, X_i \subset V) \,\&\, \Big(\bigcup_{i=1}^{k} Y_i = W \Big)$ выражает утверждение, что V содержит положительные причины $X_1, \ldots, X_k$ наличия свойств $Y_1, \ldots, Y_k$ соответственно, причем W полностью покрывается этими множествами свойств.

Подформула

$$\forall Y \Big(\exists X \, (J_{(1,n)} \, (X \Rightarrow_2 Y) \,\&\, X \subset V) \to \Big(\bigvee_{l=1}^{k} (Y = Y_l) \Big) \Big)$$

выражает условие исчерпываемости: мы должны рассматривать все входящие в V причины появления множества свойств W.

Подформула

$$\forall u \, ((u \subseteq W \,\&\, u \neq \varnothing) \to \neg \exists Z \, (J_{(-1,n)} \, (Z \to_2 u) \,\&\, Z \subset V)$$

говорит о том, что ни для каких подмножеств W, V не содержит отрицательных причин Z.

Аналогичным образом определяется $\Pi_n^-(V, W)$ — достаточно в $\Pi_n^+(V, W)$ заменить $J_{(1,n)}$, всюду на $J_{(-1,n)}$, и наоборот.

Предикат, порождающий гипотезы о $\Rightarrow_1^*$ с оценкой $\langle 0, n+1 \rangle$ (фактическое противоречие), определим следующим образом:

$$\Pi_n^0(V, W) \leftrightharpoons \exists X_1 \exists Y_1 \exists X_2 \exists Y_2 \, (J_{(1,n)} \, (X_1 \Rightarrow_2 Y_1) \,\&$$

$$\& J_{(-1,n)}(X_2 \Rightarrow_2 Y_2) \& Y_1 \cap Y_2 = \varnothing \& X_1 \subset V \& \\ \& X_2 \subset V \& Y_1 \subseteq W \& Y_2 \subseteq W).$$

Сформулируем теперь снова п. п. в. II-го рода:

$$(II^{(\sigma)}): \frac{J_{(\tau,n)}(V \Rightarrow_1 W), \Pi_n^\sigma(V, W)}{J_{\langle \nu, n+1 \rangle}(V \Rightarrow_1 W)},$$

где $\sigma \in \{+, -\}$, $\nu = \begin{cases} 1, & \text{если } \sigma = + \\ -1, & \text{если } \sigma = -; \end{cases}$

$$(II^{(0)}): \frac{J_{(\tau,n)}(V \Rightarrow_1 W), \ \Pi_n^0(V, W)}{J_{\langle 0, n+1 \rangle}(V \Rightarrow_1 W)},$$

$$(II^{(\tau)}): \frac{J_{(\tau,n)}(V \Rightarrow_1 W), \ \neg \Pi_n^+(V, W) \& \neg \Pi_n^-(V, W) \& \neg \Pi_n^0(V, W)}{J_{(\tau,n+1)}(V \Rightarrow_1 W)},$$

а $\Pi_n^\sigma(V, W) \rightleftharpoons \exists k \Pi_n^\sigma(V, W, k)$, где $\sigma \in \{+, 0, -\}$.

Очевидно, что эти п. п. в. являются абдукциями.

Если в качестве $\mathfrak{R}'$[1] для соответствующей КАТ выбрать п. п. в. I-го рода в первоначальной формулировке (без предпорядка $\succ$), то может быть реализована STR типа «мешок», упомянутая выше, комбинирующая применения п. п. в. I-го и II-го родов следующим образом.

1°. Применяем п. п. в. I-го рода к $M_1^{(0)}$, порождаем множество гипотез $\Gamma_{\rm I}^{(1)} = \Gamma_{{\rm I},1}^{(1)} \cup \Gamma_{{\rm I},-1}^{(1)} \cup \Gamma_{{\rm I},0}^{(1)} \cup \Gamma_{{\rm I},\tau}^{(1)}$, где $J_{\langle \nu, 1 \rangle}(C' \Rightarrow_2 A) \in \Gamma_{{\rm I},\nu}^{(1)}$, $\nu = \{-1, 0, 1\}$, которому отвечает $M_2^{(1)}$ (для $\Rightarrow_2^*$); если $\Gamma_{\rm I}^{(1)} = \Gamma_{{\rm I},\tau}^{(1)}$, процесс заканчивается; если $\Gamma_{\rm I}^{(1)} \neq \Gamma_{{\rm I},\tau}^{(1)}$, то переходим к 2°;

2°. Применяем п. п. в. II-го рода к $M_1^{(0)}$ и $M_2^{(1)}$, порождая множество гипотез

$$\Gamma_{\rm II}^{(1)} = \Gamma_{{\rm II},1}^{(1)} \cup \Gamma_{{\rm II},-1}^{(1)} \cup \Gamma_{{\rm II},0}^{(1)} \cup \Gamma_{{\rm II},\tau}^{(1)},$$

которому отвечает $M_1^{(1)}$ (для $\Rightarrow_1^*$); если $\Gamma_{\rm II}^{(1)} = \Gamma_{{\rm II},\tau}^{(1)}$, то процесс заканчивается, если $\Gamma_{\rm II}^{(1)} \neq \Gamma_{{\rm II},\tau}^{(1)}$, то переходим к 3° и т. д. Процесс заканчивается, когда найдется такое n, что $\Gamma_{\rm II}^{(n)} = \Gamma_{\rm II}^{(n+1)}$, что означает, что $(\tau, n) = (\tau, n+1)$ для всех τ-случаев отношения $\Rightarrow_1^*$.

STR «мешок» реализует правдоподобное рассуждение в КАТ такой, что $\Sigma = \varnothing$, Σ_0' состоит из $\Rightarrow_1$-фактов и $\Rightarrow_2$-фактов (точнее: из фактов, т. е. формул с оценками $\langle \nu, 0 \rangle$, и гипотез о фактах с оценками $\langle \nu, n \rangle$, где $n > 0$), а $\mathfrak{R}' = \mathfrak{R}_{\rm I}' \cup \mathfrak{R}_{\rm II}'$, где $\mathfrak{R}_{\rm I}'$, $\mathfrak{R}_{\rm II}'$, соответственно, п. п. в. I-го и II-го родов. Определим теперь цель рассуждения φ следующим образом.

[1] Напомним, что $\mathfrak{R}' \subset \mathfrak{R}$ — множество всех п. п. в. КАТ.

Пусть $s = s_1 \cdot s_2$, $r_1 = |\mathrm{U}^{(1)}|$[1], $s_2 = |\mathrm{U}^{(2)}|$, $r = 2^{r_1} - s_1$, $\overset{*}{\Rightarrow}_1$ задано на $\mathscr{X}_1 \times 2^{\mathrm{U}^{(2)}}$, где $|\mathscr{X}_1| = s_1$, тогда:

$$s_\tau = s_{\tau,\nu} + s_{\tau,\tau}, \quad s_\tau < s,$$

где $s_{\tau,\nu}$ и $s_{\tau,\tau}$ — числа гипотез вида $J_{\langle \nu, n \rangle}(C \Rightarrow_1 A)$ и $J_{(\tau,n)}(C \Rightarrow_1 A)$ $(\nu \in \{1, 0, -1\})$, соответственно;

$$r' = r \cdot s_2, \quad r_\tau = r_{\tau,\mu} + r_{\tau,\tau}, \quad r_\tau \leqslant r',$$

где $r_{\tau,\mu}$ и $r_{\tau,\tau}$ — числа гипотез вида $J_{\langle \mu, n \rangle}(C' \Rightarrow_2 A')$ $(\mu \in \{-1, 0, 1\})$ и $J_{\langle \tau, n \rangle}(C' \Rightarrow_2 A')$ $\varphi \rightleftharpoons \psi \,\&\, \chi$, где

$$\psi = \psi_1 \,\&\, \psi_{-1} \,\&\, \psi_0 \,\&\, \psi_\tau,$$

$$\chi = \chi_1 \,\&\, \chi_{-1} \,\&\, \chi_0 \,\&\, \chi_\tau \; ^{2)},$$

$$\psi_\nu = \mathop{\&}_{l=1}^{s_{\tau,\nu}} J_{\langle \nu, n \rangle}(C_l \Rightarrow_1 A_l), \quad \nu \in \{-1, 0, 1\},$$

$$\psi_\tau = \mathop{\&}_{l=s_{\tau,\nu}+1}^{s_{\tau,\nu}+s_{\tau,\tau}} J_{(\tau,n)}(C_l \Rightarrow_1 A_l),$$

$$\chi_\mu = \mathop{\&}_{j=1}^{r_{\tau,\mu}} J_{\langle \mu, n \rangle}(C'_j \Rightarrow_2 A'_j), \quad \mu \in \{-1, 0, 1\},$$

$$\chi_\tau = \mathop{\&}_{j=r_{\tau,\mu}+1}^{r_{\tau,\mu}+r_{\tau,\tau}} J_{(\tau,n)}(C'_j \Rightarrow_2 A'_j).$$

Очевидно, что «качество достижения цели» зависит от оценок вида $\langle \overline{\nu}, \mathscr{M}^{\sigma}_{x,n}, \mathrm{STR} \rangle$ и $\langle \overline{\mu}, \Pi^{\sigma}_{n}, \mathrm{STR} \rangle$ для порожденных гипотез. В связи с этим необходимо определить некоторое упорядочение[3] как на множестве п. п. в. (на $\mathfrak{M}$ или I соответственно), так и на множестве стратегий γ.

Неформальный смысл сформулированного выше правдоподобного рассуждения (будем называть его ДСМ-рассуждением) состоит в следующем.

Посредством п. п. в. I-го рода порождаются гипотезы о причинах наличия или отсутствия множеств свойств у объектов ((+)- или (—)-причин). Далее, обнаруженные гипотезы о (+)- или (—)-причинах используются в п. п. в. II-го рода для порождения гипотез о том, обладают или нет некоторые объекты соответствующими множествами свойств. Итерация указанной комбинации п. п. в. приводит к стабилизации ча-

[1] В данном примере мы рассматриваем конечный случай, т. е. $|\mathrm{U}^{(1)}| < \aleph_0$.

[2] Возможно, что некоторые из ψ_ν, ψ_τ, χ_μ, χ_τ, $(\nu, \mu \in \{-1, 0, 1\})$ отсутствуют.

[3] Это было сделано для I, что породило STR, отличные от STR типа «мешок».

стично определенных отношений $\Rightarrow_1^*$ и $\Rightarrow_2^*$: число τ-случаев больше не уменьшается.

Таким образом, описанный вариант ДСМ-рассуждения есть рассуждение о фактах, выражающих наличие или отсутствие причинного отношения. Естественно, что ДСМ-рассуждение может быть усложнено посредством добавления п. п. в., выражающих обратные ЭЗ «от следствия — к причине», а также посредством использования декларативных аксиом Σ из КАТ (например, осуществляя проверку на непротиворечивость полученных гипотез и Σ).

§ 5. Процедурные и декларативные аксиомы

Ниже мы сформулируем пример КАТ $\Gamma = \langle \Sigma, \Sigma', \mathfrak{R} \rangle$ такой, что Σ содержит лишь так называемые процедурные аксиомы, представляющие п. п. в. в виде импликаций, а $\mathfrak{R}$ содержит лишь правила достоверного вывода (т. е. $\mathfrak{R}' = \varnothing$).

Пусть $\mathfrak{M}_{\mathrm{I}} = \{\mathscr{M}^+_{a,n}(V, W), \mathscr{M}^-_{a,n}(V, W)\}$, а

$$\mathfrak{M}_{\mathrm{II}} = \{\Pi^+_n(V, W), \Pi^-_n(V, W), \Pi^0_n(V, W)\},$$

тогда получим следующую систему процедурных аксиом:

1. $(J_{(\tau,n)}(V \Rightarrow_2 W) \,\&\, \mathscr{M}^+_{a,n}(V, W) \,\&\, \neg \mathscr{M}^-_{a,n}(V, W)) \to$
$\to J_{\langle 1,n+1 \rangle}(V \Rightarrow_2 W)$;

2. $(J_{(\tau,n)}(V \Rightarrow_2 W) \,\&\, \neg \mathscr{M}^+_{a,n}(V, W) \,\&\, \mathscr{M}^-_{a,n}(V, W)) \to$
$\to J_{\langle -1,n+1 \rangle}(V \Rightarrow_2 W)$;

3. $(J_{(\tau,n)}(V \Rightarrow_2 W) \,\&\, \mathscr{M}^+_{a,n}(V, W) \,\&\, \mathscr{M}^-_{a,n}(V, W)) \to$
$\to J_{\langle 0,n+1 \rangle}(V \Rightarrow_2 W)$;

4. $(J_{(\tau,n)}(V \Rightarrow_2 W) \,\&\, \neg \mathscr{M}^+_{a,n}(V, W) \,\&\, \neg \mathscr{M}^-_{a,n}(V, W)) \to$
$\to J_{(\tau,n+1)}(V \Rightarrow_2 W)$;

5. $(J_{(\tau,n)}(V \Rightarrow_1 W) \,\&\, \Pi^+_n(V, W)) \to J_{\langle 1,n+1 \rangle}(V \Rightarrow_1 W)$;

6. $(J_{(\tau,n)}(V \Rightarrow_1 W) \,\&\, \Pi^-_n(V, W)) \to J_{\langle -1,n+1 \rangle}(V \Rightarrow_1 W)$;

7. $(J_{(\tau,n)}(V \Rightarrow_1 W) \,\&\, \Pi^0_n(V, W)) \to J_{\langle 0,n+1 \rangle}(V \Rightarrow_1 W)$;

8. $(J_{(\tau,n)}(V \Rightarrow_1 W) \,\&\, (\mathop{\&}\limits_{\sigma \in \{+,0,-\}} \neg \Pi^\sigma_n(V, W))) \to J_{(\tau,n+1)}(V \Rightarrow_1 W)$;

9. $\mathop{\&}\limits_{h=1}^{2} [\forall X \forall Y (J_{(\tau,n)}(X \Rightarrow_h Y) \to J_{(\tau,n+1)}(X \Rightarrow_h Y))] \,\&$
$\& J_{(\tau,n)}(V \Rightarrow_l W) \to J_{\langle \tau,\omega \rangle}(V \Rightarrow_l W), \quad l = 1, 2.$

Аксиомы 1—9 образуют теорию $\mathrm{T}(\mathfrak{M})$, где $\mathfrak{M} = \mathfrak{M}_{\mathrm{I}} \cup \mathfrak{M}_{\mathrm{II}}$.

Теорема. Теория $\mathrm{T}(\mathfrak{M})$ непротиворечива, если все $\mathscr{M}^{\pm}_{a,n}$ и Π^{σ}_{n} невырождены:

$$\left[\forall X \forall Y \mathop{\&}_{l=1}^{2} J_{\langle \tau, \omega \rangle} (X \Rightarrow_l Y)\right] \to [\forall X \forall Y \neg \mathscr{M}^{\pm}_{a,n}(X, Y)]$$

и аналогично для Π^{σ}_{n} («из полной недоопределенности ничего не выводится»).

Пусть даны конечные множества $\mathrm{U}^{(j)}$ $(j=1, 2)$ и булевы алгебры $\mathscr{B}^{(j)}=2^{\mathrm{U}^{(j)}}$, а язык расширен константами для всех элементов $\mathscr{B}^{(j)}$. Диаграмма:

$$D(\mathrm{U}^{(j)})=D'(\mathrm{U}^{(j)})+\left[\forall X^{(j)}\left(\bigvee_{B\in\mathscr{B}^{(j)}}(X^{(j)}=B)\right)\right],$$

где $D'(\mathrm{U}^{(j)})$ — множество всех замкнутых истинных равенств булевой алгебры $\mathscr{B}^{(j)}$ $(j=1, 2)$, а $X^{(j)}$ есть либо X, либо Y (т. е. переменные сорта 1 или 2). Начальное распределение ρ — набор функций ρ_r:

$$2^{\mathrm{U}^{(1)}}\times 2^{\mathrm{U}^{(2)}} \to \{\langle 1, 0\rangle, \langle -1, 0\rangle, \langle 0, 0\rangle, \langle \tau, 0\rangle\},$$

где $r=1, 2$, определяющих исходные матрицы $M_1^{(0)}$, $M_2^{(0)}$ для отношений $\Rightarrow_r^*$.

Теория

$$\mathrm{T}_\rho(\mathfrak{M})=[\mathrm{T}(\mathfrak{M})+D(\mathrm{U}^{(j)})_{(j=1,2)}+D(\rho_r)_{(r=1,2)}],$$

где

$$D(\rho_r)=\{J_{\langle \nu, 0\rangle}(C\Rightarrow_r A)/\rho_r(C, A)=\nu,\ \nu\in\{1, 0, -1, \tau\}\}$$

(«диаграмма» начальных состояний матриц $\Rightarrow_r^*$).

Теорема. При определенных условиях на связки ∞-значной логики теория $\Gamma_\rho(\mathfrak{M})$ имеет единственную модель $\mathfrak{V}_\rho$, естественным образом определяемую как результат правдоподобного вывода, стартующего с исходных матриц ρ_r:

$$\mathfrak{V}_\rho=\langle \mathscr{B}_1, \mathscr{B}_2, \Rightarrow_{\rho_1}, \Rightarrow_{\rho_2}\rangle,$$

где $C\Rightarrow_{\rho_r}A=\begin{cases}\langle \nu, n\rangle, & \text{если } \langle C, A\rangle \in\delta_r(\nu, n),\ \nu\in\{-1, 0, 1\},\ n\in\mathrm{N}\\ \langle \tau, \omega\rangle, & \text{в противном случае,}\end{cases}$

$r=1,2$ (подробнее см. (3)).

Напомним, что КАТ $\mathrm{T}=\langle\Sigma, \Sigma', \mathfrak{R}\rangle$. Специфицируем теперь множество аксиом Σ и множество правил вывода $\mathfrak{r}$:

$$\Sigma=\Sigma_{pr}\cup\Sigma_{dc},\quad \mathfrak{R}=\mathfrak{R}'\cup\mathfrak{R}^0,$$

где Σ_{pr}, Σ_{dc}, $\mathfrak{R}'$, $\mathfrak{R}^0$ — множества процедурных аксиом (представляющих некоторые п. п. в. в декларативной форме), множество декларативных аксиом, множество п. п. в. и множество правил достоверного вывода, соответственно.

$$\Sigma_{dc}=\Sigma_{dc,\,0}\cup\Sigma_{dc,\,1},\text{ где } \Sigma_{dc,\,0} \text{ и } \Sigma_{dc,\,1},$$

соответственно, аксиомы структуры данных, например, аксиомы булевых алгебр

$$\mathscr{B}_1=\langle 2^{U^{(1)}},\ \cap,\ \cup,\ \setminus\rangle,$$

$$\mathscr{B}_2=\langle 2^{U^{(2)}},\ \cap,\ \cup,\ \setminus\rangle$$

и аксиомы связи исходных предикатов (в нашем случае это аксиомы связи предикатов $\Rightarrow_1$ и $\Rightarrow_2$).

Следующая таблица представляет формально возможные случаи Σ, Σ' и $\mathfrak{R}$:[1]

	Σ			Σ'	$\mathfrak{R}$	
	Σ_{pr}	$\Sigma_{dc,\,1}$	$\Sigma_{dc,\,0}$		$\mathfrak{R}'$	$\mathfrak{R}^0$
(a)	+	+	+	+	+	+
(b)	+	+	+	+	—	+
(c)	+	—	+	+	+	+
(d)	+	—	+	+	—	+
(e)	—	+	+	+	+	+
(f)	—	+	+	+	—	+
(g)	—	—	+	+	+	+
(h)	—	—	+	+	—	+

Случай (h) не интересен, так как не является формализацией правдоподобных рассуждений. Случаи (c), (d) и (g) фактически были нами рассмотрены выше. Сформулируем теперь некоторые декларативные аксиомы, могущие принадлежать соответствующим множествам $\Sigma_{dc,\,1}$[2]. Предварительно напомним, что выше мы рассмотрели $\mathscr{M}^{\sigma}$-аксиомы и Π^{σ}-аксиомы, принадлежащие Σ_{pr}. Смысл понятия «процедурная аксиома» и «декларативная аксиома» очевиден.

$$(a1)^{+}:\ \forall X\forall Y_1\forall Y_2(Y_1\neq\varnothing\,\&\,Y_2\neq\varnothing\rightarrow(J_{(1,n)}(X\Rightarrow_1 Y_1\cup Y_2)\leftrightarrow$$

$$\leftrightarrow(J_{(1,n)}(X\Rightarrow_1 Y_1)\,\&\,J_{(1,n)}(X\Rightarrow_1 Y_2))))$$

((+)-аксиома аддитивности справа $\Rightarrow_1$);

$$(a1^{-}):\ \forall X\forall Y_1\forall Y_2(Y_1\neq\varnothing\,\&\,Y_2\neq\varnothing\rightarrow(J_{(-1,n)}(X\Rightarrow_1 Y_1\cup Y_2)\leftrightarrow$$

[1] «+» и «—», соответственно, обозначают, что рассматриваемое множество (например, Σ_{pr}) пусто или не пусто.

[2] Некоторые из формул, приводимых ниже, могут не принадлежать выбранному $\Sigma_{dc,\,1}$, но могут быть, однако, метасредствами управления выводом (средством обоснования порождаемых гипотез, например, $(a4)^{\sigma}$, где $\sigma\in\{+,\,-\}$).

$$\leftrightarrow (J_{(-1,n)}(X \Rightarrow_1 Y_1) \& J_{(1,n)}(X \Rightarrow_1 Y_2))))$$

((—)-аксиома аддитивности справа $\Rightarrow_1$);

$$(a2)^+: \forall X_1 \forall X_2 \forall Y_1 \forall Y_2 ((J_{(1,n)}(X_1 \Rightarrow_2 Y_1) \& (J_{(1,n)}(X_2 \Rightarrow_2 Y_2)) \& \\ \& \neg \exists Z_2 (Z_2 \subset Y_2 \& J_{(-1,n)}(X_1 \Rightarrow_2 Z_2)) \& \neg \exists Z_1 (Z_1 \subset Y_1 \& \\ \& J_{(-1,n)}(X_2 \Rightarrow_2 Z_1))) \rightarrow J_{(1,\omega)}(X_1 \cup X_2 \Rightarrow_2 Y_1 \cup Y_2)$$

((+)-аксиома аддитивности $\Rightarrow_2$);

$$(a2)^-: \forall X_1 \forall X_2 \forall Y_1 \forall Y_2 ((J_{(-1,n)}(X_1 \Rightarrow_2 Y_1) \& J_{(-1,n)}(X_2 \Rightarrow_2 Y_2) \& \\ \& \neg \exists Z_2 (Z_2 \subset Y_2 \& J_{(1,n)}(X_1 \Rightarrow_2 Z_2)) \& \neg \exists Z_1 (Z_1 \subset Y_1 \& \\ \& J_{(1,n)}(X_2 \Rightarrow_2 Z_1))) \rightarrow J_{(-1,\omega)}(X_1 \cup X_2 \Rightarrow_2 Y_1 \cup Y_2))$$

((—)-аксиома аддитивности $\Rightarrow_2$);

$$(a3)^+: \forall X \forall Y (J_{\langle 1,0 \rangle}(X \Rightarrow_1 Y) \rightarrow \neg \exists Z \exists W (Z \subset X \& Z \neq \varnothing \& W \neq \\ \neq \varnothing \& W \subseteq Y \& (J_{\langle -1,n \rangle}(Z \Rightarrow_2 W) \vee J_{\langle 0,n \rangle}(Z \Rightarrow_2 W))))$$

+)-аксиома каузальной непротиворечивости);

$$(a3)^-: \forall X \forall Y (J_{\langle -1,0 \rangle}(X \Rightarrow_1 Y) \rightarrow (\neg \exists Z \exists W (Z \subset X \& Z \neq \\ \neq \varnothing \& W \subseteq Y \& W \neq \varnothing \& (J_{\langle 1,n \rangle}(Z \Rightarrow_2 W) \vee J_{\langle 0,n \rangle}(Z \Rightarrow_2 W))))$$

((—)-аксиома каузальной непротиворечивости);

$$(a4)^+: \forall X \forall Y \exists Z (J_{\langle 1,0 \rangle}(X \Rightarrow_1 Y) \rightarrow (J_{(1,\omega)}(Z \Rightarrow_2 Y) \& Z \subset X \& Z \neq \varnothing))$$

«(+)-аксиома Б. Спинозы»: «знание действия зависит от знания причины»).

При $n \geqslant 1$ $(a4)^+$ истинна в КАТ, однако при $n \geqslant 0$ может быть ложна, так как при $n = 0$ может не существовать Z, являющегося причиной Y. Однако $(a4)^+$ можно использовать как принцип достаточного основания правдоподобного вывода. А именно: если для данного $\Rightarrow_1^*$ с $M_1^{(0)}$ и для полученных $M_1^{(n)}$, $M_2^{(n)}$, где n — номер стабилизированного состояния с $(\tau, n) = (\tau, n+1)$ для всех τ-случаев $\Rightarrow_1^*$ и $\Rightarrow_2^*$ выполняется $(a4)^+$, то порожденные **гипотезы** выведены на *квазидостаточном основании.*

$$(a4)^-: \forall X \forall Y \exists Z (J_{(-1,0)}(X \Rightarrow_1 Y) \rightarrow (Z \neq \varnothing \& Z \subset X \& J_{(-1,n)}(Z \Rightarrow_2 Y)) \vee \\ \vee \forall X \forall Y \forall Z ((J_{\langle -1,0 \rangle}(X \Rightarrow_1 Y) \& Z \neq \varnothing \& Z \subset X) \rightarrow J_{(\tau,\omega)}(Z \Rightarrow_2 Y)))$$

(«(—)-аксиома Б. Спинозы»).

Аналогично (+)-случаю $(a4)^+$ будем говорить, что порождаемые гипотезы выведены на квазидостаточном основании, если выполняется $(a4)^-$.

Будем говорить, что гипотезы выведены на *достаточном основании*, если выполняются $(a4)^+ \& (a4)^-$.

Интересны приводимые ниже $(a5)^\sigma$ $(\sigma \in \{+, -\})$, связанные с п. п. в. II-го рода:

$$(a5)^+: \forall X \forall Y ((\exists Z (J_{(1,n)}(Z \Rightarrow_2 Y) \& Z \subseteq X) \& \forall U \forall W ((J_{(-1,n)}(U \Rightarrow_2 W) \& \\ \& W \subseteq Y) \rightarrow \neg (U \subset X))) \rightarrow (J_{(1,n)}(X \Rightarrow_1 Y) \vee J_{(\tau,n)}(X \Rightarrow_1 Y)));$$

(a5)$^-$: $\forall X\forall Y((\exists Z(J_{(-1,n)}(Z\Rightarrow_2 Y)\& Z\subset X)\&\forall U\forall W((J_{(1,n)}(U\Rightarrow_2 W)\&$
$\& W\subseteq Y)\to \neg(U\subset X)))\to(J_{(-1,n)}(X\Rightarrow_1 Y)\vee J_{(\tau,n)}(X\Rightarrow_1 Y)))$.

Примеры КАТ:

(a) Σ_{pr} содержит аксиомы для представления п. п. в. I-го рода

$$\Sigma_{dc,1}=\{(a3)^+,\ (a3)^-\};$$

$$\mathfrak{R}'=\{(II^{(+)}),\ (II^{(-)}),\ (II^{(0)}),\ (II^{(\tau)})\};$$

(b) Σ_{pr} содержит аксиомы для представления п. п. в. II-го рода;

$$\Sigma_{dc,1}=\{(a3)^+,\ (a3)^-\};$$

$$\mathfrak{R}'=\varnothing;$$

(c) Σ_{pr} содержит аксиомы для представления п. п. в. II-го рода

$$\Sigma_{dc,1}=\{(a3)^+,\ (a3)^-\},$$

$$\mathfrak{R}'=\{(I^{(+)}),\ (I^{(-)}),\ (I^{(0)}),\ (I^{(\tau)})\};$$

(d) Σ_{pr} содержит аксиомы для представления п. п. в. I-го и II-го рода, $\Sigma_{dc,1}=\varnothing$, $\mathfrak{R}'=\varnothing$;

(e) $\Sigma_{pr}=\varnothing$, $\Sigma_{dc,1}=\{(a3)^+,\ (a3)^-,\ (a4)^+,\ (a4)^-\}$;

$$\mathfrak{R}'=\{(I^{(+)}),\ (I^{(-)}),\ (I^{(0)}),\ (I^{(\tau)}),\ (II^{(+)}),\ (II^{(-)}),\ (II^{(0)}),\ (II^{(\tau)})\};$$

(f) $\Sigma_{pr}=\varnothing$; $\Sigma_{dc,1}=\{(a1)^+,\ (a1)^-,\ (a2)^+,\ (a2)^-,\ (a3)^+,\ (a3)^-,\ (a4^+),\ (a4)^-,\ (a5^+),\ (a5)^-\}$;

$$\mathfrak{R}'=\varnothing;$$

(g) $\Sigma_{pr}=\Sigma_{dc,1}=\varnothing$; $\mathfrak{R}'=\{(I^{(+)}),\ (I^{(-)}),\ (I^{(0)}),\ (I^{(\tau)}),\ (II^{(+)}),\ (II^{(-)}),\ (II^{(0)}),\ (II^{(\tau)})\}$;

(h) $\Sigma_{dc,1}=\Sigma_{pr}=\varnothing$; $\mathfrak{R}'=\varnothing$.

Отметим, что ДСМ-рассуждение (т. е. рассуждение, состоящее из последовательного применения п. п. в. I-го и II-го рода или содержащее их дедуктивную имитацию посредством Σ_{dc}, Σ' и $\mathfrak{R}$) может быть формализовано в случаях (a), (c), (d), (e) и (g).

Замечание. Покажем, что $\Sigma_{dc,1}=\{(aI)^+,\ (aI)^-\}$ и $\mathfrak{R}'=\{(II)^+,\ (II)^-,\ (II)^0,\ (II)^\tau\}$ образуют противоречивую КАТ.

1. Воспользуемся теоремами $\Pi_n^\sigma(V,W)$, $\sigma\in\{+,-\}$:

$(T1)$: $\forall n\forall V\forall W(\Pi_n^+(V,W)\to\exists! U\Pi_n^+(V,U))$,

$(T2)$: $\forall m\forall V\forall W(J_{\langle 1,m+1\rangle}(V\Rightarrow_1 W)\leftrightarrow J_{(\tau,m)}(V\Rightarrow_1 W)\&\Pi_m^+(V,W)$.

Пусть $J_{(\tau,n)}(C\Rightarrow_1 A)$, $\Pi_n^+(C,A)$ истинны, а $A=A_1\cup A_2$, $A_i\neq\varnothing$ $(i=1,2)$, $A_1\neq A_2$. Тогда в силу $(II)^+$ имеем $J_{\langle 1,n+1\rangle}(C\Rightarrow_1 A)$, согласно $(a1)^+$ имеем

$$J_{\langle 1,n+1\rangle}(C\Rightarrow_1 A_1\cup A_2)\leftrightarrow(J_{\langle 1,n+1\rangle}(C\Rightarrow_1 A_1)\& J_{\langle 1,n+1\rangle}(C\Rightarrow_1 A_2),$$

применяя ($T2$), получим

$$J_{\langle 1, n+1 \rangle}(C \Rightarrow_1 A_i) \leftrightarrow (J_{(\tau, n)}(C \Rightarrow_1 A_i) \,\&\, \Pi_n^+(C, A_i)), \quad i = 1, 2,$$

используя двузначную логику высказываний (и, следовательно, $\mathfrak{R}^0$), получим $\Pi_n^+(C, A_1)$, $\Pi_n^+(C, A_2)$, что противоречит ($T1$). Легкая модификация п. п. в. I-го рода устраняет это противоречие (28):

$$(\widetilde{II})^{(\sigma)}: \ \frac{J_{(\tau, n)}(V \Rightarrow_1 U),\ \Pi_n^\sigma(V, U) \,\&\, W \subseteq U \,\&\, W \neq \varnothing}{J_{\langle \nu, n+1 \rangle}(V \Rightarrow_1 W)},$$

где $\nu = \begin{cases} 1, & \text{если } \sigma = + \\ -1, & \text{если } \sigma = -; \end{cases}$

$$(\widetilde{II})^{(0)}: \ \frac{J_{(\tau, n)}(V \Rightarrow_1 U),\ \Pi_n^0(V, U) \,\&\, W \subseteq U \,\&\, W \neq \varnothing}{J_{\langle 0, n+1 \rangle}(V \Rightarrow_1 W)}.$$

Таким образом, для устранения указанного противоречия достаточно определить $\overline{\Pi}^\sigma$-предикат следующим образом:

$$\overline{\Pi}_n^\sigma(V, W) \rightleftharpoons \Pi_n^\sigma(V, A) \,\&\, W \subseteq A \,\&\, W \neq \varnothing,$$

где A — константа.

Дадим теперь краткую теоретико-модельную характеризацию КАТ, содержащую п. п. в.

Декларативная теория Σ_{dc} имеет класс моделей. Σ' выражает информацию об экспериментальных данных (факты из КАТ). Возникает задача: согласовать Σ' с данной декларативной теорией Σ_{dc}. Модель этой теории и Σ' определяется следующим образом.

Константы фактоподобных формул выберем в качестве порождающего множество термов и породим обычным способом модель (т. е. породим эрбрановскую модель), где функциональными символами являются символы из языка теории (в рассматриваемых выше примерах: $\cup$, $\cap$, $\setminus$). В порожденном указанном выше универсуме модели зададим отношения согласно экспериментальным данным Σ'. Процедуры переработки данных в построенной таким образом модели (в рассмотренном выше случае это п. п. в. I-го и II-го рода) порождают (расширяют) отношения над универсумом (в рассматриваемом выше случае $\Rightarrow_1^*$ и $\Rightarrow_2^*$) так, чтобы эти отношения соответствовали аксиомам теории Σ_{dc}.

Выше рассмотрели простой случай модели

$$\langle \mathscr{B}_1, \mathscr{B}_2, \Rightarrow_1^*, \Rightarrow_2^* \rangle,$$

где $\mathscr{B}$ — булевы алгебры множеств, так что объекты задаются как множества. Обобщим теперь этот случай. Имеются два типа моделей: одна модель структур (структур объектов), универсумом которой являются структуры (конечные множества, кортежи, слова, графы, образующие конструктивные объекты), а вторая модель — модель объектов. Переработка данных осуществляется с целью установления связи между структурами и объектами.

При этом предполагается наличие двух подъязыков:

(1) языка объектов, в котором выражаются сами объекты (т. е. их имена), а также свойства и отношения между ними;

(2) языка описания структур, в котором содержатся имена элементов и отношения между ними.

Отметим в связи с этим, что язык $\mathscr{L}_t$ имеет квантификацию по двум универсумам моделей, упомянутых выше. Эти две модели связаны посредством частично определенных отношений $\Rightarrow^*_j$ ($j=1, 2$), представляющих информацию об экспериментальных данных Σ'. Назначение п. п. в. из $\mathfrak{R}'$ состоит в том, чтобы по возможности уменьшить неопределенность в задании частично определенных отношений $\Rightarrow^*_j$.

Если теория задает полностью отношения, связывающие две указанные подмодели, то тогда отношения, порожденные средствами п. п. в., используются для *коррекции интерпретации* отношений, теоретически заданных в данной модели. В силу юмовского принципа мы не можем корректировать саму теорию, так как результаты применения п. п. в. будут формулами с внутренними оценками из V_t, т. е. будут формулами другого характера, чем формулы теории. Таким образом, результаты применения п. п. в. не будут иметь статуса[1] аксиом, но могут быть использованы для коррекции конкретных отношений в конкретной модели. Если ввести некоторый порог действия юмовского принципа, то тогда этот случай соответствует концепции индукции Поппера [69] и ее компьютерной реализации, предложенной Е. Шапиро [73].

Модель КАТ, как мы уже отметили, строится следующим образом: из имен объектов, описанных в Σ', генерируем универсум, а отношения над этим универсумом задаются либо в силу теории (как интерпретация), либо порождаются посредством п. п. в. Имеются две возможности в случае, если отношения заданы теорией (т. е. Σ_{dc}).

I. На универсуме интерпретируются отношения, заданные теорией, а дальнейшее применение п. п. в. базируется на них. При этом возможные несоответствия появляются при порождении гипотез.

II. Порождается отношение как гипотезы на базе переработки фактов из Σ' посредством п. п. в. и в то же время задается интерпретация отношения, заданного в теории. Аналогично случаю I может возникать несоответствие отношения, заданного в теории (точнее, в ее интерпретации), и отношения, порождаемого посредством п. п. в.

---

[1] Имеется в виду зависимость оценок формул от Σ'.

§ 6. Некоторые примеры индуктивных правдоподобных выводов

Как мы уже упоминали выше, п. п. в. I-го и II-го рода являются примерами абдукции, т. е. правилами правдоподобного вывода «от частного — к частному». Во введении мы привели неформализованный пример индуктивного правдоподобного вывода — метод сопутствующих изменений Д. С. Милля [26] (см. также [32]).

Для формализации метода сопутствующих изменений (метода с. и.) требуется расширить язык $\mathscr{L}'_e$. Пусть $\mathfrak{b} = \langle \mathbf{B}, \prec_1 \rangle$ и $\mathscr{W} = \langle \mathbf{A}, \prec_2 \rangle$ частично упорядоченные множества такие, что они содержат соответственно параметры, характеризующие объекты C $(C \subset \mathrm{U}^{(1)})$ и свойства A $(A \subseteq \mathrm{U}^{(2)})$. Пусть $\mathbf{B} = \{b_1, b_2, \ldots, b_n, \ldots\}$, а $\mathbf{A} = \{\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots\}$. Положим, что $\mathbf{B} \cap \mathrm{U}^{(1)} = \varnothing$ и $\mathbf{A} \cap \mathrm{U}^{(2)} = \varnothing$. Введем два новых сорта (сорт 4 и 5) переменных $x, z, x_1, z_1, \ldots, x_n, z_n, \ldots$ и $y, w, y_1, w_1, \ldots, y_n, w_n, \ldots$, соответственно имеющие областями определения $\mathbf{B}$ и $\mathbf{A}$. Соответственно введем двуместные предикаты $x \prec_1 z$ и $y \prec_2 w$ и кванторы по указанным переменным сортов 4 и 5. Объекты и свойства будем теперь представлять в следующем виде: $\{t_i\} \cup T$, где t_i — терм сорта 4 (т. е. переменная или константа сорта 4), $\{t_i\}$ — одноэлементное множество, а T — терм сорта 1 (т. е. переменная сорта 1 или константа $C \subset \mathrm{U}^{(1)}$). Множество свойств будем представлять посредством $\{t'_i\} \cup T'$, где t'_i — терм сорта 5 (т. е. переменная или константа сорта 5), $\{t'_i\}$ — одноэлементное множество, а T' — терм сорта 2 (т. е. переменная сорта 2 или константа $A \subseteq \mathrm{U}^{(2)}$). Тогда в рассматриваемом расширении $\mathscr{L}'_e$ введем атомарные формулы вида

$$(\{t_i\} \cup T) \Rightarrow_1 (\{t'_j\} \cup T'),$$
$$(\{t_i\} \cup T) \Rightarrow_2 (\{t'_j\} \cup T').$$

Определим теперь семейство предикатов, называемых предикатами метода с. и., зависящих от параметров k и n:

$$\tilde{\mathscr{M}}_{c,n,k}(x_1, \ldots, x_k, y_1, \ldots, y_k, V, W, k) \rightleftharpoons$$
$$\rightleftharpoons \exists Y_1 \ldots \exists Y_k \left(\left(\mathop{\&}_{l=1}^{k} J_{(1,n)} ((\{x_l\} \cup V) \Rightarrow_1 (\{y_l\} \cup Y_l)) \right) \& \right.$$
$$\& (Y_1 \cap \ldots \cap Y_k = W) \& \forall x \forall V \forall w_1 \forall w_2 \forall Y ((x \prec_1 z \& W \subseteq Y \&$$
$$\& J_{(1,n)} ((\{x\} \cup V) \Rightarrow_1 (\{w_1\} \cup Y)) \& J_{(1,n)} ((\{z\} \cup Z) \Rightarrow_1 (\{w_2\} \cup Y))) \to$$
$$\to \left(\left(\bigvee_{l=1}^{k} (x = x_l) \right) \& \left(\bigvee_{l=1}^{k} (z = x_l) \right) \& \left(\bigvee_{l=1}^{k} (w_1 = y_l) \right) \& \right.$$
$$\left. \left. \& \left(\bigvee_{l=1}^{k} (w_2 = y_l) \right) \& w_1 \prec_2 w_2 \right) \& k \geqslant 2 \right).$$

Определим, далее, непараметрический предикат $\mathscr{M}_{c,n,\bar{k}}(V, W)$, где $\bar{k}$ — некоторая константа из N:

$$\mathscr{M}^{+}_{c,n,\bar{k}}(V, W) \leftrightharpoons \exists k\, \exists x_1 \ldots \exists x_k \exists y_1 \ldots \exists y_k \exists Y_1 \ldots$$

$$\ldots \exists Y_k \tilde{\mathscr{M}}^{+}_{c,n,k}(x_1, \ldots, x_k, y_1, \ldots, y_k, V, W, k).$$

Аналогично формулируется предикат $\mathscr{M}^{-}_{c,n,\bar{k}}(V, W)$ заменой соответствующих вхождений $J_{(1,n)}$ на $J_{(-1,n)}$.

Сформулируем уточнение метода с. и. Д. С. Милля — п. п. в. следующего вида:

$$\frac{J_{(\tau,n)}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W)),\ \mathscr{M}^{+}_{c,n,\bar{k}_1}(V, W) \,\&\, \neg \mathscr{M}^{-}_{c,n,\bar{k}_2}(V, W)}{\forall x \exists y (\{x\} \subset \mathbf{B} \to J_{\langle 1, n+1 \rangle}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W)))},$$

$$\frac{J_{(\tau,n)}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W)),\ \mathscr{M}^{-}_{c,n,\bar{k}_1}(V, W) \,\&\, \neg \mathscr{M}^{+}_{c,n,\bar{k}_2}(V, W)}{\forall x \exists y (\{x\} \subset \mathbf{B} \to J_{\langle -1, n+1 \rangle}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W)))}.$$

Аналогичны п. п. в. для случаев с заключениями

$$\forall x \exists y (\{x\} \subset \mathbf{B} \to J_{\langle 0, n+1 \rangle}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W))),$$

$$\forall x \exists y (\{x\} \subset \mathbf{B} \to J_{(\tau, n+1)}((\{x\}\cup V) \Rightarrow_2 (\{y\}\cup W))).$$

Формализация метода с. и. Д. С. Милля посредством п. п. в. типа ДСМ может быть осуществлена и в модифицированном языке $\mathscr{L}'_e$ отличном от указанного выше. А именно, в качестве термов можно выбрать пары $\langle t_i, T \rangle$; $\langle t'_j, T' \rangle$, где t_i, t'_j — термы сорта 4 и сорта 5 соответственно, а T, T' — термы сорта 1 и сорта 2 соответственно. Модификация языка логики предикатов 1-ого порядка с кортежными термами была предложена в [33]. Используя кортежные термины вида $\langle t_i, T \rangle$, $\langle t'_j, T' \rangle$, определим атомарные формулы $\langle t_i, T \rangle \Rightarrow_1 \langle t'_j, T' \rangle$ и $\langle t_i, T \rangle \Rightarrow_2 \langle t'_j, T' \rangle$. Очевидно, что отношения $\Rightarrow_1^*$ и $\Rightarrow_2^*$, соответствующие введенным выше предикатам $\Rightarrow_1$ и $\Rightarrow_2$, являются частично определенными отношениями на

$$(\mathbf{B} \times 2^{\mathrm{U}^{(1)}}) \times (\mathbf{A} \times 2^{\mathrm{U}^{(2)}}).$$

Формула в модифицированном языке $\mathscr{L}'_e$ определяется в соответствии с [33] аналогично определению формулы в $\mathscr{L}'_e$.

Обобщим формализованный выше метод сопутствующих изменений Д. С. Милля, комбинируя средства распознавания эмпирических зависимостей между параметрами x сорта 4 и y сорта 5 и между структурированными подобъектами (термы сорта 1) и множествами свойств структурированных объектов (термы сорта 2). Приведем ниже пример такого обобщения для прямого положительного метода сходства Д. С. Милля, для формализации которого определим предикат $\tilde{\mathscr{M}}^{+}_{c,a,n}(V, W, k)$.

$$\tilde{\mathcal{M}}^{+}_{ca,n}(V, W, k) \leftrightharpoons \exists x_1 \ldots \exists x_k \exists y_1 \ldots \exists y_k \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots$$

$$\ldots \exists U_k \Big(\Big(\mathop{\&}_{l=1}^{k} J_{(1,n)} ((\{x_l\} \cup Z_l) \Rightarrow_1 (U_l \cup \{y_l\})) \& \forall U \, (J_{(1,n)} ((\{x_l\} \cup Z_l) \Rightarrow_1$$

$$\Rightarrow_1 (\{y_l\} \cup U) \to U \subseteq U_l) \& (Z_1 \cap \ldots \cap Z_k = V)) \& \forall i \forall j \, ((i \neq j) \& 1 \leqslant i,$$

$$j \leqslant k) \to ((x_i \neq x_j) \vee (Z_i \neq Z_j)) \& V \neq \varnothing \& \Big(\mathop{\&}_{l=1}^{k-1} (x_l \dashv_1 x_{l+1}) \Big) \&$$

$$\& \Big(\mathop{\&}_{l=1}^{k-1} (y_l \dashv_2 y_{l+1}) \Big) \& \forall X \forall Y \forall x \forall z \forall w_1 \forall w_2 \, (((x \dashv_1 z \, \& \, J_{(1,n)} ((\{x\} \cup X) \Rightarrow_1$$

$$\Rightarrow_1 (\{w_1\} \cup Y)) \& \forall U \, (J_{(1,n)} ((\{x\} \cup X) \Rightarrow_1 (\{w_1\} \cup U)) \to U \subseteq Y) \&$$

$$\& J_{(1,n)} ((\{z\} \cup X) \Rightarrow_1 (\{w_2\} \cup Y) \& \forall U \, (J_{(1,n)} ((\{z\} \cup X) \Rightarrow_1 (\{w_2\} \cup U) \to$$

$$\to U \subseteq Y) \& V \subset X) \to \Big(W \neq \varnothing \& W \subseteq Y \& \Big(\bigvee_{l=1}^{k} (X = Z_l) \Big) \Big) \&$$

$$\& \Big(\bigvee_{l=1}^{k} (x = x_l) \Big) \& \Big(\bigvee_{l=1}^{k} (z = x_l) \Big) \& \Big(\bigvee_{l=1}^{k} (w_1 = y_l) \Big) \&$$

$$\& \Big(\bigvee_{l=1}^{k} (w_2 = y_l) \Big) \& w_1 \dashv_2 w_2 \Big) \Big) \& k \geqslant 2 \Big).$$

Далее, $\mathcal{M}^{+}_{ca,n}(V, W) \leftrightharpoons \exists k \tilde{\mathcal{M}}^{+}_{ca,n}(V, W, k)$. Сформулируем соответствующее п. п. в., являющееся правилом индуктивного вывода типа «сопутствующие изменения + сходство»:

$$\frac{J_{(\tau,n)} ((\{x\} \cup V) \Rightarrow_2 (\{y\} \cup W)), \ \mathcal{M}^{+}_{ca,n}(V, W) \& \neg \mathcal{M}^{-}_{ca,n}(V, W)}{\forall x \exists y \, (\{x\} \subset B \to J_{\langle 1, n+1 \rangle} ((\{x\} \cup V) \Rightarrow_2 (\{y\} \cup W)))}.$$

Аналогично формулируются правила индуктивного вывода для случаев

$$\mathcal{M}^{-}_{ca,n}(V, W) \& \neg \mathcal{M}^{+}_{ca,n}(V, W), \quad \mathcal{M}^{+}_{ca,n}(V, W) \& \mathcal{M}^{-}_{ca,n}(V, W),$$

$$\neg \mathcal{M}^{+}_{ca,n}(V, W) \& \neg \mathcal{M}^{-}_{ca,n}(V, W).$$

Очевидно, что эти правила в свою очередь могут быть обобщены для предикатов $\mathcal{M}^{\sigma}_{x,n}(V, W)$, где $\sigma \in \{+, -\}$, а имя предиката $x \in \mathbf{I}$, где $\mathbf{I}$ — множество имен (+)- и (−)-предикатов $\mathcal{M}^{\sigma}_{x,n}$, на котором, как было показано выше, определено отношение предпорядка. П. п. в. и соответствующие стратегии могут быть сформулированы аналогично рассмотренному выше случаю с отношением предпорядка на $\mathbf{I}$.

Весьма интересен тот факт, что п. п. в. типа «сопутствующие изменения + сходство», хотя и имеют в качестве заключения индуктивное обобщение, но содержат в посылках встроенные средства абдукции, используемые при формализации $\mathcal{M}^{+}_{a,n}(V, W) \& \neg \mathcal{M}^{-}_{n,n}(V, W)$. Это обстоятельство говорит о том, что фигуры правдоподобных выводов весьма нетривиальны и могут быть использованы при формализации «усиления умственной деятельности» в смысле Эшби [47].

Отметим теперь одну интересную возможность развития автоматизированных ДСМ-рассуждений, связанную с применением программы, аналогичной программе «Бэкон» [60]. Дело в том, что в заключении индуктивного вывода типа «сопутствующее изменение параметров и + структурное сходство» (например, в случае $\mathcal{M}_{ca,}{}^{a}{}_{n}$-предикатов) x и y могут принадлежать числовому множеству $\mathbf{B}=\mathbf{A}$. Тогда возникает интересная проблема нахождения явного вида функциональной зависимости $y=f(x)$ посредством программы, подобной программе «Бэкон». Таким образом, поиск эмпирических зависимостей в БД с неполной информацией, содержащей как структурированные (нечисловые) объекты, так и их числовые характеристики, может быть осуществлен комбинированной программой типа «Милль+ +Бэкон». Создание соответствующей программной системы обогатит инструментальные средства интеллектуальных систем.

Очевидно, что возможно усложнение сформулированных методов типа «сопутствующие изменения параметров $\bar{x}_i$ и y_i + + структурное сходство», где $y_i=f_i(x_{i1}, \ldots, x_{is_i})$, $i \geqslant 2$, $\bar{x}_i=(x_{i1}, \ldots, x_{is_i})$.

§ 7. Замечания о фальсификаторах для правдоподобных выводов типа ДСМ

Вернемся теперь к классификации недостоверных выводов, предложенной во введении к настоящему разделу, и отметим, что п. п. в. I и II рода являются абдукциями типа (в), т. е. соответствующие правдоподобные выводы имеют обучаемое (порождаемое) множество процедурных фальсификаторов в смысле Поппера [31], что видно из рассмотренных выше фигур правдоподобных выводов типа ДСМ. Например, для п. п. в. I рода

$$\frac{J_{(\tau,n)}(V \Rightarrow_2 W),\ \mathcal{M}^{+}_{x,n}(V, W)\ \&\ \neg \mathcal{M}^{-}_{x,n}(V, W)}{J_{\langle 1,n+1\rangle}(V \Rightarrow_2 W)},$$

$$\frac{J_{(\tau,n)}(V \Rightarrow_2 W),\ \mathcal{M}^{-}_{x,n}(V, W)\ \&\ \neg \mathcal{M}^{+}_{x,n}(V, W)}{J_{\langle -1,n+1\rangle}(V \Rightarrow_2 W)}$$

фальсификаторами будут, соответственно, пары $\langle C_1', A_1 \rangle$ и $\langle C_2', A_2 \rangle$, если $\mathcal{M}^{-}_{x,n}(C_1', A_1)$ и $\mathcal{M}^{+}_{x,n}(C_2', A_2)$ истинны. Т. е. $\langle C_1', A_1 \rangle$ и $\langle C_2', A_2 \rangle$ фальсифицируют, соответственно, $J_{\langle 1,n+1\rangle}(C_1' \Rightarrow_2 A_1)$ и $J_{\langle -1,n+1\rangle}(C_2' \Rightarrow_2 A_2)$.

Отметим важное обстоятельство, что фальсификация правдоподобного рассуждения типа ДСМ может быть осуществлена посредством декларативных аксиом каузальной непротиворечивости $(a3)^+$, $(a3)^-$ и условий каузальной полноты $(a4)^+$, $(a4)^-$, которые можно назвать декларативными средствами фальсификации гипотез в КАТ.

Процедурная и декларативная нефальсифицируемость результатов правдоподобных рассуждений типа ДСМ образуют достаточное основание правдоподобного рассуждения (или, как говорят, критерий рациональности вывода), о котором мы говорили как во введении, так и при рассмотрении строения и примеров КАТ.

Интересно отметить, что и индуктивные выводы, представляющие собой комбинирование п. п. в. I рода и метода сопутствующих изменений Д. С. Милля [26], [32], имеют процедурно порождаемое множество фальсификаторов, т. е. являются правдоподобными выводами типа (в), фальсификаторами которых являются необходимые части структурных (нечисловых параметров) причин [32] и следствий, являющихся значениями переменных пары $\langle V, W\rangle$ из заключения, соответствующего п. п. в.:

$$\forall x \exists y (\{x\} \subset \mathrm{B} \rightarrow J_{\langle 1, n+1\rangle}((\{x\} \cup V) \Rightarrow_2 (\{y\} \cup W))).$$

Приведем пример декларативной фальсифицируемости порождаемых гипотез в экспертных системах типа ДСМ [28].

Пусть применение п. п. в. I-го рода ($\mathrm{I}^{(-)}$) породило гипотезу $J_{\langle -1, \overline{m}\rangle}(C' \Rightarrow_2 A')$, где константа $\overline{m} \in \mathrm{N}$, т. е.:

$$\frac{J_{(\tau, m-1)}(C' \Rightarrow_2 A'),\ \mathscr{M}^{-}_{a, m-1}(C', A') \,\&\, \neg \mathscr{M}^{+}_{a, m-1}(C', A')}{J_{\langle -1, \overline{m}\rangle}(C' \Rightarrow_2 A')}.$$

Рассмотрим в некоторой КАТ (+)-аксиому каузальной непротиворечивости

$$(a3)^+\colon \forall X \forall Y (J_{\langle 1,0\rangle}(X \Rightarrow_1 Y) \rightarrow \neg \exists Z \exists W (Z \subset X \,\&\, Z \neq \varnothing \,\&\, W \neq$$
$$\neq \varnothing \,\&\, W \subseteq Y \,\&\, (J_{\langle -1, n\rangle}(Z \Rightarrow_2 W) \vee J_{\langle 0, n\rangle}(Z \Rightarrow_2 W)))).$$

Пусть $C \in 2^{\mathrm{U}^{(1)}}$, $A \in 2^{\mathrm{U}^{(2)}}$, тогда получим

$$J_{\langle 1,0\rangle}(C \Rightarrow_1 A) \rightarrow \neg \exists Z \exists W (Z \subset C \,\&\, Z \neq \varnothing \,\&\, W \neq$$
$$\neq \varnothing \,\&\, W \subseteq A \,\&\, (J_{\langle -1, \overline{m}\rangle}(Z \Rightarrow_2 W) \vee J_{\langle 0, n\rangle}(Z \Rightarrow_2 W)))).$$

Пусть Σ' из КАТ содержит $J_{\langle 1,0\rangle}(C \Rightarrow_1 A)$, которое в экспертной системе принадлежит БД с неполной информацией. Применяя к предыдущим формулам правило modus ponens и очевидные преобразования логики предикатов 1-го порядка [25] и эквивалентность

$$\forall Z \forall W (J_{\langle 1, \overline{m}\rangle}(Z \Rightarrow_2 W) \vee J_{(\tau, m)}(Z \Rightarrow_2 W) \vee$$
$$\vee J_{\langle -1, \overline{m}\rangle}(Z \Rightarrow_2 W) \vee J_{\langle 0, \overline{m}\rangle}(Z \Rightarrow_2 W)) \leftrightarrow t$$

получим

$$\forall Z \forall W (\neg (Z \subset C) \vee Z = \varnothing \vee W \neq \varnothing \vee \neg (W \subseteq A) \vee$$
$$\vee (J_{\langle 1, \overline{m} \rangle} (Z \Rightarrow_2 W) \vee J_{(\tau, \overline{m})} (Z \Rightarrow_2 W))).$$

Откуда следует, в силу допущений из Σ' (т. е. БД), что так как $C' \subset C$, $A' \subseteq A$, $A' \neq \varnothing$ и $C' \neq \varnothing$, то имеет место

$$J_{\langle 1, \overline{m} \rangle} (C' \Rightarrow_2 A') \vee J_{(\tau, \overline{m})} (C' \Rightarrow_2 A'),$$

что противоречит порожденной гипотезе $J_{\langle -1, \overline{m} \rangle} (C' \Rightarrow_2 A')$.

Таким образом, пара $\langle C', A' \rangle$ — фальсификатор (в смысле К. Поппера) КАТ, содержащей соответствующее множество полуфактов Σ', аксиому $(а3)^+$ и процедурную аксиому

$$\forall V \forall W ((J_{(\tau, n)} (V \Rightarrow_2 W) \,\&\, \mathcal{M}^-_{a, n} (V, W) \,\&\, \neg \mathcal{M}^+_{a, n} (V, W)) \rightarrow$$
$$\rightarrow J_{\langle -1, n+1 \rangle} (V \Rightarrow_2 W)$$

(или соответствующее п. п. в. I-го рода из $\mathfrak{R}'$).

§ 8. Примеры правдоподобных рассуждений и их строение

Цель настоящего параграфа описать строение индуктивного обобщения, являющегося результатом правдоподобного рассуждения, содержащего абдукцию, дедукцию и модельные вычисления (т. е. проверки истинности или ложности формул в соответствующей семантике, например, в БД для экспертных систем).

Предварительно обратим внимание на следующее утверждение:

$$\vdash (\mathcal{M}^+_{a, n} (C, A) \rightarrow \forall X \forall Y ((J_{(1, n)} (X \Rightarrow_1 Y) \,\&\, C \subset X) \rightarrow A \subseteq Y). \qquad (1)$$

Если выполняются посылки п. п. в. I-го рода, то абдукция дает гипотезу $J_{\langle 1, n+1 \rangle} (C \Rightarrow_2 A)$, т. е.

$$\frac{J_{(\tau, n)} (C \Rightarrow_2 A),\ \mathcal{M}^+_{a, n} (C, A) \,\&\, \neg \mathcal{M}^-_{a, n} (C, A)}{J_{\langle 1, n+1 \rangle} (C \Rightarrow_2 A)}.$$

Используя modus ponens и приведенное выше утверждение дедуктивно выводим

$$\vdash (\forall X \forall Y ((J_{(1, n)} (X \Rightarrow_1 Y) \,\&\, C \subset X) \rightarrow A \subseteq Y). \qquad (2)$$

Отметим при этом, что этот дедуктивный вывод мы совершим при условии выполнения абдукции (т. е. п. п. в. I-го рода), приведенной выше, которая термам C и A придает ранг термов гипотезы, порожденной правдоподобным выводом.

Пусть $\mathrm{U}^{(2)}=\{a_0, a_1, \ldots, a_{s_2}\}$, причем положим, что $A=\{a_0, \ldots, a_m\}$, $m<s_2$. Сопоставим каждой формуле вида $\{a_i\}\subseteq Y$ $(i=1, \ldots, s_2)$ булевскую переменную ξ_i, а формуле $A\subseteq Y$ сопоставим соответствующую булевскую конъюнкцию $\xi_1\& \ldots \&\xi_m$.

Рассмотрим теперь булевские функции G (называемые «скелетными»), имеющие вид $G=\rho\&\chi_1\&\urcorner\chi_2$, где $\rho=\xi_1\&\ldots\&\xi_m$, а χ_1 и χ_2 — монотонные булевские функции [48], содержащие переменные из списка $\xi_{m+1}, \ldots, \xi_{s_2}$. Интерпретируем вид булевской функции, исходя из содержательных соображений.

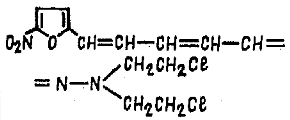
$\rho=\xi_1\&\ldots\&\xi_m$ соответствует наличию свойств, коррелирующих с необходимой частью причины (ср. [32]), χ_1 — монотонная булевская функция свойств из $\mathrm{U}^{(2)}\setminus A$, сопутствующих необходимой части причины; а χ_2 — монотонная булевская функция (переменные которой не пересекаются с переменными из ρ), отвечающая свойствам, несовместимым со свойствами из A при наличии необходимой части причины в рассматриваемых объектах [20]. Специальный алгоритм [19], применяемый к Σ' (т. е. к БД), устанавливает истинность формулы

$$\forall X\forall Y\,((J_{(1,n)}(X\Rightarrow_1 Y)\&C\subset X)\to G(A\subseteq Y,\ A_1\subseteq Y,\ \ldots,\ A_q\subseteq Y), \quad (3)$$

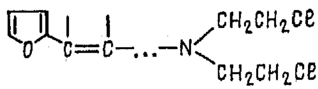
где $A_j\subseteq\mathrm{U}^{(2)}\setminus A$. Установление истинности формулы [3] мы и называем *модельным вычислением* (т. е. вычислением в БД). Легко видеть, что для эмпирических зависимостей (ЭЗв) [2] и [3] имеет место следующее: $[2]\not\vdash[3]$, но $[3]\not\vdash[2]$. Таким образом, ЭЗв [3] дедуктивно не выводится из ЭЗв [2], а является индуктивным обобщением ЭЗв [2], обнаруженным вычислениями в модели (БД). Сама же ЭЗв [2] дедуктивно выводима из [1]: $[1]\vdash[2]$.

Приведем некоторые данные, представленные здесь в неполном и схематическом виде, которые были реализованы в экспертной системе типа ДСМ [28] при решении задачи прогнозирования причин (фармакофоров) биологических активностей. В приводимых экспериментах объекты и подобъекты суть специально закодированные химические соединения и их фрагменты, а свойства — соответствующие биологические активности [27]. Ниже приведем один из экспериментов, выполненный на соединениях ряда фуранов (массив подготовлен к. б. н. Д. А. Бодягиным в ВОНЦ АМН СССР). Этот эксперимент был реализован программами Е. С. Панкратовой и В. Г. Ивашко. Свойства [29]: противоопухолевые биологические активности, a_0 — токсичность, a_1, a_2, a_3, a_4, соответственно, противоопухолевые активности относительно опухолей Йенсена, саркомы —45, опухоли Уокера, рака РС—1.

Общий вид БД:

	объекты	наличие свойств	отсутствие свойств	факты
C_1		(+) $\{a_1, a_4\}$	(−) $\{a_0, a_3\}$	$J_{(1,n)}(C_1 \Rightarrow_1 A_1)$
C_2		$\{a_1, a_3, a_4\}$	$\{a_0, a_2\}$	$J_{(1,n)}(C_2 \Rightarrow_1 A_2)$
⋮	....	......	.....	....
C_m		$\{a_1, a_3\}$	$\{a_0, a_2, a_4\}$	$J_{(1,n)}(C_m \Rightarrow_1 A_m)$

Где: $A_1 = \{a_1, a_4\}$, $A_2 = \{a_1, a_3, a_4\}, \ldots, A_m = \{a_1, a_3\}$; пусть $A_1' = \{a_1\}$, $A_0' = \{a_0\}$, $A_3' = \{a_3\}$, $A_4' = \{a_4\}$,

$C' =$ 

Индуктивное обобщение, полученное ЭВМ:

$$\forall X \forall Y \, ((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, C' \subset X) \rightarrow$$

$$\rightarrow ((A' \subseteq Y) \,\&\, ((A_3' \subseteq Y) \vee (A_4' \subseteq Y)) \,\&\, \neg (A_0' \subseteq Y))).$$

Ниже мы изобразим схематически порождение вышеприведенного индуктивного обобщения:

$$
\begin{tikzcd}[column sep=small]
\neg \mathcal{M}^{-}_{a,n}(\mathbb{C}',A') \arrow[dr, "\text{абдукция}"] & , & \mathcal{M}^{+}_{a,n}(\mathbb{C}',A') \arrow[dl] \arrow[dr, "\text{дедукция}"] & \\
& J_{(1,n)}(\mathbb{C}' \Rightarrow_2 A') \arrow[dr, "\text{индуктивное обобщение}"'] & & \forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, \mathbb{C}' \subset X) \to A' \subseteq Y) \arrow[dl] \\
& & \forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, \mathbb{C}' \subset X) \to ((A' \subseteq Y) \,\&\, ((A'_3 \subseteq Y) \vee (A'_4 \subseteq Y)) \,\&\, \neg (A'_0 \subseteq Y))) . &
\end{tikzcd}
$$

В этом же эксперименте [29] были получены и другие индуктивные обобщения, заинтересовавшие экспертов, и, в частности:

$$\forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \,\&\, C'_1 \subset X) \to$$
$$\to ((A'_4 \subseteq Y) \,\&\, ((A'_1 \subseteq Y) \vee (A'_0 \subseteq Y)) \,\&\, \neg ((A'_1 \subseteq Y) \,\&\, (A'_0 \subseteq Y)))).$$

Результатом охарактеризованного выше рассуждения является индуктивное обобщение, следовательно, одним из использованных правдоподобных выводов является индукция. Однако понятие «индукции» в свете сказанного оказывается весьма нетривиальным и многосторонним. В связи с этим индукцию в нашем смысле [54] мы будем называть эмпирической структурной индукцией (ЭСИ). Охарактеризуем специфические черты ЭСИ:

1. Эмпирическая структурная индукция (в отличие от математической индукции) есть результат сравнения фактов — примеров и контрпримеров;

2. ЭСИ есть процедура выделения существенного сходства и различия структурированных конструктивных объектов (конечные множества, кортежи, слова, графы, пространственные графы); что требует от формализации ЭСИ выбора соответствующих структур данных [14];

3. Определение существенного сходства объектов нетривиально и зависит от определения эмпирических операций локального сходства [14], детерминированных «наивной семантикой» предметной области, формализуемой в КАТ;

4. ЭСИ, рассмотренная нами, отличается от популярной индукции (которую мы назовем прямой), имеющей вид

$$\frac{\varphi(a_1), \ldots, \varphi(a_n)}{\forall x \varphi(x)}$$

так как она имеет иную структуру, использующую примеры вида $\varphi(a_1), \ldots, \varphi(a_n)$, но также и специальные условия, характеризующие эти примеры (ср. абдукции п. п. в. I-го рода и используемые для получения индуктивных обобщений в методе «сопутствующие изменения+сходство» или в рассуждении с индуктивным обобщением, полученным выше); кроме того, ЭСИ, изучаемая нами, в отличие от популярной индукции, является не прямой, а косвенной, ибо используя примеры вида $\varphi(a_i)$, $i=1, \ldots, n$, косвенная индукция имеет заключением формулу $\forall x \psi(x)$, где $\psi(x)$ не эквивалентна $\varphi(x)$;

5. Рассмотренные нами примеры ЭСИ имеют как встроенные абдукции, так и множество фальсификаторов (как устанавливаемых посредством декларативных аксиом в КАТ, так и порождаемых процедурно в результате обучения на примерах и контрпримерах);

6. Индуктивное обобщение, получаемое применением правдоподобного рассуждения к начальному состоянию, заданному матрицами $M_1^{(0)}$ и $M_2^{(0)}$, отвечающим частично определенным отношениям $\Rightarrow_1^*$ и $\Rightarrow_2^*$, будем называть *внешней индукцией*, результатом которой и является само индуктивное обобщение, выражающее некоторую эмпирическую зависимость (ЭЗв); однако если рассмотреть последовательность вложенных исходных множеств $\mathrm{U}_1^{(1)} \subset \mathrm{U}_2^{(1)} \subset \ldots \subset \mathrm{U}_n^{(1)} \subset \ldots$, то сохраняющиеся ЭЗв (при рассмотренных расширениях) будем называть эмпирическими закономерностями, найденными (устойчивыми) в соответствующих расширениях состояний КАТ $\Sigma_1', \Sigma_2', \ldots, \Sigma_n', \ldots$, отвечающих состояниям БД (очевидно, что внешняя индукция является примером немонотонных рассуждений).

Резюмируя сказанное, подчеркнем, что ЭСИ=абдукции+ +дедукции+модельным вычислениям, характеризующим строение данного типа индуктивного рассуждения.

§ 9. Анализ строения правил правдоподобного вывода II-го рода

Целью настоящего подраздела является установление связи между отношениями локального сходства [14], являющегося отношением толерантности [45], и п. п. в. II-го рода.

Пусть R_1, R_2 — отношения толерантности такие, что $R_1 \subseteq \mathrm{U}^{(1)} \times \mathrm{U}^{(1)}$ и $R_2 \subseteq \mathrm{U}^{(2)} \times \mathrm{U}^{(2)}$, где R_1 и R_2 — отношения, характеризующие сходства термов языка $\mathscr{L}_e$. Обозначим, соответственно, посредством $T^{(1)}, T_1^{(1)}, T_2^{(1)}, \ldots$ и $T^{(2)}, T_1^{(2)}, T_2^{(2)}, \ldots$ — термы сортов 1 и 2. Тогда $T_1^{(i)} R_j T_2^{(i)}$ означает, что $T_1^{(i)}$ и $T_2^{(i)}$ сходны ($i=1, 2$; $j=1, 2$).

Пусть φ, ψ — формулы $\mathscr{L}_e$, а $\tilde{\varphi}$, $\tilde{\psi}$ — их предваренные нормальные формы (25), соответственно: $\varphi \leftrightarrow \tilde{\varphi}$, $\psi \leftrightarrow \tilde{\psi}$, где $\tilde{\varphi} \doteq \mathfrak{K}_1 \varphi_1$, $\tilde{\psi} \doteq \mathfrak{K}_2 \psi_1$, где $\mathfrak{K}_1$, $\mathfrak{K}_2$ — кванторные приставки, а «$\doteq$» — отношение графического равенства.

Пусть, далее, Q — отношение сходства формул ($\tilde{\varphi} Q \tilde{\psi}$), определяемое посредством 1° — 3°, где P — символ n-арного предиката ($n \in \mathrm{N}$); t_{ij} — термы сортов 1 или 2, R либо R_1, либо R_2, тогда:

1°. $P(\bar{t}_1) Q P(\bar{t}_2)$, если $t_{1i} R t_{2i}$, $i=1,\ldots,n$, где $\bar{t}_1 = \langle t_{11},\ldots,t_{1n}\rangle$, $\bar{t}_2 = \langle t_{21},\ldots,t_{2n}\rangle$;

2°. если $\chi_1, \chi_2, \chi_3, \chi_4$ — формулы $\mathscr{L}_e$, и $\chi_1 Q \chi_3$, $\chi_2 Q \chi_4$, то $(\chi_1 \Delta \chi_2) Q (\chi_3 \Delta \chi_4)$, где $\Delta \doteq \&$, $\vee$, $\rightarrow$;

3°. если χ_1, χ_2 — формулы $\mathscr{L}_e$ и $\chi_1 Q \chi_2$, то $(J_{\bar{v}} \chi_1) Q (J_{\bar{v}} \chi_2)$;

4°. если $\chi_1 \doteq \mathfrak{K} \varphi$, $\chi_2 \doteq \mathfrak{K} \Psi$ и $\varphi Q \psi$, то $\chi_1 Q \chi_2$.

Рассмотрим теперь, следуя Пойа [30] 2 схемы вывода по аналогии:

$$(\mathrm{I})\ \frac{t_1 R t_2,\ \varphi(t_1)}{\varphi(t_2)}, \qquad (\mathrm{II})\ \frac{\varphi Q \psi,\ \psi}{\varphi}.$$

Именно с помощью (I) и (II) мы проанализируем ниже п. п. в II-го рода. Но для этой цели нам нужно будет рассмотреть п. п. в. I-го рода с несколько иной точки зрения, учитывающей то обстоятельство, что $\mathscr{M}^{\sigma}_{x,n}$-предикаты выразимы в языке логики предикатов с кванторами по кортежам переменной длины (34).

Рассмотрим частный случай $\mathscr{M}^{\sigma}_{x,n}$-предиката $(x=a)$ и представим $\mathscr{M}^{+}_{a,n}(V, W, k)$ следующим образом:

$$\tilde{\mathscr{M}}^{+}_{a,n}(V, W, k) \rightleftharpoons \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots \exists U_k P_{1,n,k}(V, W, Z_1,\ldots,Z_k, U_1,\ldots,U_k),$$

где k — переменная сорта 3, т. е. $k \in \mathrm{N}$ (множество натуральных чисел есть область изменения k), где

$$P_{1,n,k}(V, W, Z_1,\ldots,Z_k, U_1,\ldots,U_k) \rightleftharpoons$$

$$\rightleftharpoons \Bigg(\Bigg(\mathop{\&}_{i=1}^{k} (J_{(1,n)}(Z_i \Rightarrow_1 U_i) \& \forall U (J_{(1,n)}(Z_i \Rightarrow_1 U) \rightarrow U \subseteq U_i)\Bigg)\Bigg) \&$$

$$\& \left(\bigcap_{i=1}^{k} Z_i = V\right) \& \forall i \forall j (i \neq j \& 1 \leqslant i,\ j \leqslant k) \rightarrow Z_i \neq Z_j) \& V \neq \varnothing \&$$

$$\forall X \forall Y ((J_{(1,n)}(X \Rightarrow_1 Y) \& \forall U (J_{(1,n)}(X \Rightarrow_1 U) \rightarrow U \subseteq Y) \& V \subset X) \rightarrow$$

$$\rightarrow \left(W \neq \varnothing \& W \subseteq Y \& \left(\bigvee_{i=1}^{k}(X = Z_i)\right)\right)\Bigg) \& k \geqslant 2\Bigg),$$

$K_{1,k}^{(n)}(Z_1, \ldots, Z_k) \Leftrightarrow \exists V \exists W \exists U_1 \ldots \exists U_k P_{1,n,k}(V, W, Z_1, \ldots, Z_k, U_1, \ldots, U_k)$, где $k \in \mathbb{N}$.

Рассмотрим $\mathrm{K}_1 = \{K_{1,k}^{(n)}(Z_1, \ldots, Z_k)\}_{\substack{n \in \mathbb{N} \\ k \in \mathbb{N}}}$ — семейство $K_{1,k}^{(n)}$-формул, где $K_{1,k}^{(n)}(Z_1, \ldots, Z_k)$ есть «обращение» предикатов $\tilde{\mathscr{M}}_{a,n}^{+}(V, W, k)$ (заметим, что длина кортежей $\langle Z_1, \ldots, Z_k \rangle$ зависит от $k \in \mathbb{N}$). Множество же всех кортежей $\langle Z_1, \ldots, Z_k \rangle$, удовлетворяющих $K_{1,k}^{(n)}(Z_1, \ldots, Z_k)$, *не образует отношения*, но образует конструкцию, которая была названа *глобальным сходством* [13, 14, 15] (критерий представимости глобального сходства локальным, т. е. толерантностью, найден в указанных выше работах).

$$K_{1,\bar{k}}^{(n)}(Z_1, \ldots, Z_{\bar{k}}) \leftrightarrow$$

$$\leftrightarrow \exists U_1 \ldots \exists U_{\bar{k}} P_{1;n,\bar{k}}(C', A, Z_1, \ldots, Z_{\bar{k}}, U_1, \ldots, U_{\bar{k}}),$$

где индивидная константа $\bar{k} \in \mathbb{N}$, V/C', W/A, Z_i/C_i образуют соответствующую подстановку ($i = 1, \ldots, \bar{k}$), а $\langle C_1, \ldots, C_{\bar{k}} \rangle$ удовлетворяют $K_{1,\bar{k}}^{(n)}(Z_1, \ldots, Z_{\bar{k}})$. Существование случаев непредставимости глобального сходства локальным (т. е. отношениями) связано с тем фактом, что выразительные возможности языка $\mathscr{L}'_e$ превышают выразительные возможности языка логики предикатов 1-го порядка, ибо $\mathscr{L}'_e$ содержит кванторы по кортежам переменной длины, как это было показано выше.

Наблюдения над $\tilde{\mathscr{M}}_{a,n}^{+}(V, W, k)$ (и вообще над предикатами и $\tilde{\mathscr{M}}_{x,n}^{\sigma}(V, W, k)$, $\sigma \in \{+, -\}$) показывают, что

(1) $Z_1 \cap \ldots \cap Z_k = V \& V \neq \varnothing$ выражает некоторое сходство, которое мы называем *локальным*;

(2) условие, выражающее эмпирическую зависимость в $\mathscr{M}_{a,n}^{+}(V, W)$ между V и W (прогнозируемую как причинное отношение), представляет уникальность сходства, удовлетворяющего $K_{1,k}^{(n)}(Z_1, \ldots, Z_k)$.

(3) а упоминавшееся выше условие исчерпываемости $\bigvee_{i=1}^{k}(Z_i = X)$ показывает, что находятся все объекты, порождающие гипотезу, соответствующую паре $\langle V, W \rangle$.

Неформальная суть сказанного выше состоит в следующем. Для порождения гипотезы о том, что подобъект V является причиной множества свойств W, необходимо найти множество глобально сходных объектов $Z_1, \ldots, Z_k$, результатом применения к которому некоторой операции «локального сходства» является

V, а результатом применения этой же операции к соответствующему множеству свойств является W [14].

Приступим теперь к непосредственному анализу строения п. п. в. II-го рода (например, $(\text{ПИО})^+$ — «принцип индуктивного обобщения» (+) — правила правдоподобного вывода II-го рода, целью которого является доопределение случаев $J_{(\tau,n)}(C\Rightarrow_1 A)$ посредством найденных гипотетических причин посредством п. п. в. I-го рода). Итак, $(\text{ПИО})^+$ или

$$(\text{II}^{(+)}):\ \frac{J_{(\tau,n)}(C\Rightarrow_1 A),\ \Pi_n^{(+)}(C,\ A)}{J_{\langle 1,n+1\rangle}(C\Rightarrow_1 A)}$$

Рассмотрим информацию, содержащуюся в $\Pi_n^{(+)}(C,\ A)$ (см. соответствующее определение $\Pi_n^{(+)}$-предикатов):

$$J_{(1,n)}(C_1\Rightarrow_2 A_1),\ \ldots,\ J_{(1,n)}(C_k\Rightarrow_2 A_k);$$

$$C_1\subset C,\ \ldots,\ C_k\subset C, \qquad (*)$$

$$A_1\cup\ldots\cup A_k=A, \qquad (**)$$

$$\vdash (J_{(1,n)}(C_i\Rightarrow_2 A_i)\leftrightarrow J_{(\tau,n-1)}(C_i\Rightarrow_2 A_i)\,\&\,\mathcal{M}^+_{a,n-1}(C_i,\ A_i)$$

(теорема ДСМ-теории).

Обращение $\mathcal{M}^+_{a,n-1}(C_i,\ A_i)$ дает

$$J_{(1,n-1)}(C_{i1}\Rightarrow_1 A_{i1}),\ \ldots,\ J_{(1,n-1)}(C_{ir_i}\Rightarrow_1 A_{ir_i}),$$

$$C_{i1}\cap\ldots\cap C_{ir_i}=C_i,\quad i=1,\ \ldots,\ k,$$

см. определение $\mathcal{M}^+_{a,n}$-предиката).

Рассмотрим ЭЗ (для метода сходства с предикатом $\mathcal{M}^+_{a,n}(V,\ W)$):

$$\forall X\forall Y\,((J_{(1,n-1)}(X\Rightarrow_1 Y)\,\&\,C_i\subset X)\rightarrow A_i\subseteq Y),$$

$$A_i\subseteq A_{ij},\quad A_i=\bigcap_{j=1}^{r_i} A_{ij},$$

следовательно, $A_{ij}\cap A=A_{ij}$ (в силу (**)).

Таким образом, для $W_1R_2W_2\leftrightharpoons W_1\cap W_2\neq\varnothing$ имеем $A_{ij}R_2A$.

Аналогично: $J_{(1,n-1)}(C_{ij}\Rightarrow_1 A_{ij})$, $i=1,\ \ldots,\ k$, $\bigcap_{j=1}^{r_i} C_{ij}=C_i$, $C_{ij}\subset\subset C_i\subset C$ (в силу (*)).

Таким образом, $C_{ij}\cap C=C_{ij}$ для $V_1R_1V_2\leftrightharpoons V_1\cap V_2\neq\varnothing$. Следовательно, имеем $C_{ij}R_1C$. Таким образом:

(1) $\dfrac{C_{ij}R_1C,\ A_{ij}R_2A}{(C_{ij}\Rightarrow_1 A_{ij})\,Q\,(C\Rightarrow_1 A)}$, где $i=1,\ \ldots,\ k$; $j=1,\ \ldots,\ r_i$; $r=r_1+\ldots+r_k$,

$$\overbrace{(C \Rightarrow_1 A)}^{\varphi} Q \overbrace{(C_{ij} \Rightarrow_1 A_{ij})}^{\psi}$$

$$(2) \qquad \frac{J_{(1,n-1)}(C_{ij} \Rightarrow_1 A_{ij})}{J_{(1,n-1)}(C \Rightarrow_1 A)}$$

схема Д. Пойа:
φ аналогична ψ

ψ более правдоподобно

$\left\{ \begin{array}{l} \varphi \text{ несколько более} \\ \text{правдоподобно} \end{array} \right.$

Следовательно, (ПИО)$^{(+)}$ содержит r аналогий, а аналогия есть принцип правдоподобного вывода, используемый для поиска сходства.

Мы заканчиваем анализ п. п. в. II-го рода сформулированным принципом Пойа [30], приведенным ранее в подразделе «Краткая история вопроса»: предположение становится несколько более правдоподобным, когда становится более правдоподобным аналогичное предположение. Для формализации этого принципа нам понадобилось развить специальный логико-математический аппарат, содержащий бесконечнозначную многосортную логику предикатов с кванторами по кортежам, теорию сходства объектов и конструкции квазиаксиоматических теорий с процедурами правдоподобных выводов, имитирующих фигуры рассуждения естествоиспытателей, обнаруженные еще Ф. Бэконом и Д. С. Миллем.

§ 10. Некоторые итоги: современное представление об исследуемой проблематике

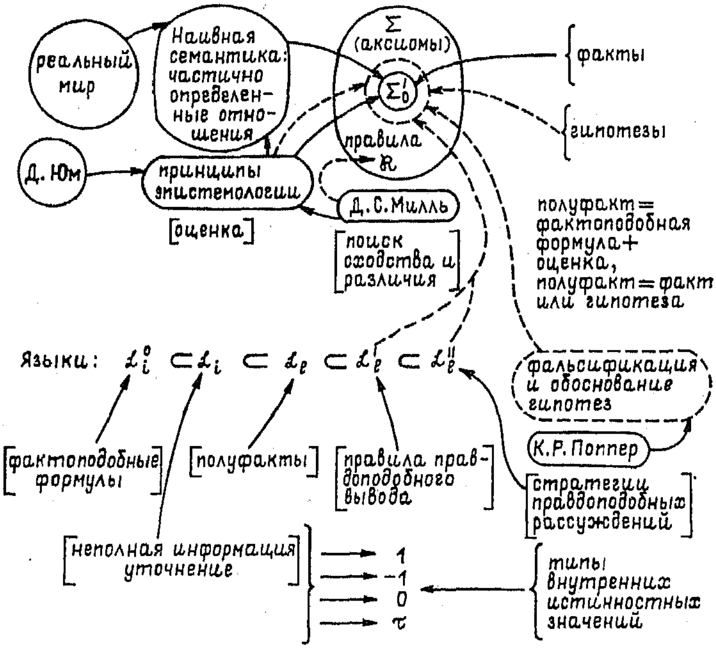
В заключение раздела «Правдоподобные выводы и правдоподобные рассуждения» отметим плодотворную связь логико-философских и логико-математических идей, способствовавших развитию теории правдоподобных рассуждений:

1. Теория индукции Ф. Бэкона (идея таблиц присутствия и таблиц отсутствия);
2. Принцип индукции Д. Юма — невозможность верификации индуктивных обобщений логическими средствами, принятие гипотез на основе «ассоциации идей»;
3. Так называемые индуктивные методы Д. С. Милля — методы сходства и методы различия, с одной стороны, и метод сопутствующих изменений, с другой стороны. «Мы должны умозаключать на основании законов природы, т. е. тех единообразий сходства или несходства, какие можно в действительности наблюдать» [26, с. 487];
4. Плодотворная идея фальсификации гипотез как критерий демаркации научного рассуждения от ненаучного (Поппер);
5. Идея поиска фигур правдоподобных рассуждений Пойа, упомянутых выше;

6. И, наконец, формализации немонотонных рассуждений, стимулированные идеями Маккарти, применение которых необходимо для использования пополняемых баз знаний интеллектуальных компьютерных систем.

Связь этих идей мы попытались выразить эскизно в приводимой ниже схематической таблице, пользуясь развитой нами системой понятий.

Логическая конструкция процесса познания, имитируемого в теории правдоподобных рассуждений (наглядное резюме)



Внешние истинностные значения: t, f (истина, ложь) 1 — «фактическая истина», —1 — «фактическая ложь», 0 — «фактическое противоречие», τ — «недоопределенность». Пример формализации: объекты=множества; типы полуфактов: 1) «объект имеет множество свойств» ($X \Rightarrow_1 Y$); «подобъект есть причина наличия (отсутствия) множества свойств» ($X \Rightarrow_2 Y$); структура данных: $\mathscr{B}_i = \langle 2^{(i)}, \cap, \cup, \setminus \rangle$, $i=1, 2$.

§ 11. О немонотонном варианте ДСМ-метода автоматического порождения гипотез

Еще один вариант понимания немонотонности как возможности изменения истинностных значений в модели возникает при рассмотрении различных стратегий автоматического порождения гипотез в базах данных с неполной информацией [39, 1]. В исходном состоянии база данных представлена матрицами частично определенных предикатов «$X^{(1)} \Rightarrow_1 Y^{(2)}$» и «$X^{(1)} \Rightarrow_2 Y^{(2)}$», интерпретируемых, соответственно, как объект $X^{(1)}$ обладает (не обладает) множеством свойств $Y^{(2)}$ и подобъект $X^{(1)}$ есть причина наличия (отсутствия) множества свойств $Y^{(2)}$. Истинностными значениями, отвечающими исходному состоянию являются $-1, 1$ и τ, интерпретируемые, соответственно, как «ложь», «истина», «недоопределенность». Значение τ играет особую роль: в дальнейшем при преобразовании матриц на некоторых местах τ будет заменено одной из «оценок неопределенности», появляющихся в процессе вычислений. Исходные матрицы преобразуются с помощью специальных процедур, которые фактически являются формализацией индуктивного вывода и порождают «гипотезы» — истинностные значения вида $\langle \varepsilon, n \rangle$ (где n — номер шага вычисления, $\varepsilon \in \{-1, 1, 0\}$ — «тип» истинностного значения: «ложь», «истина» и «эмпирическая противоречивость»). Эти процедуры представлены в языке логики предикатов с кванторами по кортежам переменной длины [34], содержащем описание булевых алгебр $2^{\mathrm{U}^{(1)}}$ и $2^{\mathrm{U}^{(2)}}$ (где $\mathrm{U}^{(1)}$ и $\mathrm{U}^{(2)}$ — исходные множества элементов, из которых строятся, соответственно, объекты $X^{(1)} \subseteq \mathrm{U}^{(1)}$ и множества свойств $Y^{(2)} \subseteq \mathrm{U}^{(2)}$). Процедуры порождения гипотез сформулированы в некотором смысле симметрично, т. е. каждому «положительному» правилу $\mathscr{M}^+(X^{(1)}, Y^{(2)})$, порождающему значения типа «истина», соответствует «отрицательное» правило $\mathscr{M}^-(X^{(1)}, Y^{(2)})$, порождающее истинностные значения типа «ложь». Точнее, «положительную» процедуру можно представить в виде

$$\frac{\mathscr{M}_n^+(X^{(1)}, Y^{(2)}),\ \neg\mathscr{M}_n^-(X^{(1)}, Y^{(2)})}{(X^{(1)} \Rightarrow_r Y^{(2)}) \text{ принимает значение } \langle 1, n+1 \rangle}, \tag{1}$$

где $\mathscr{M}_n^+$ — представление правила $\mathscr{M}^{\pm}$ на n-м шаге [1], а $r \in \{1, 2\}$.

Отрицательная процедура представляется симметрично, с взаимной заменой $+$ на $-$. Если же одновременно выполняются $\mathscr{M}_n^+(X^{(1)}, Y^{(2)})$, $\mathscr{M}_n^-(X^{(1)}, Y^{(2)})$, то порождается «эмпирическое противоречие»:

$$\frac{\mathscr{M}_n^+(X^{(1)}, Y^{(2)}),\ \mathscr{M}_n^-(X^{(1)}, Y^{(2)})}{(X^{(1)} \Rightarrow_r Y^{(2)}) \text{ принимает значение } \langle 0, n+1 \rangle}. \tag{2}$$

При монотонных стратегиях порождения гипотез [39, 1] значения, полученные на n-м шаге, в дальнейшем не подвергаются сомнению и пересмотру на последующих шагах. В то же время могут оказаться полезными иные стратегии, обеспечивающие, так сказать, «проверку гипотез на устойчивость». А именно: все гипотезы (или некоторые из них — например, наиболее интересующие или наиболее недостоверные) пересматриваются, т. е. соответствующие значения вычисляются повторно, исходя на каждом шаге из нового (сформировавшегося к этому моменту) распределения предположений об истинности, ложности или противоречивости. При этом некоторые гипотезы могут подтвердиться, а для некоторых возможно изменение ранее полученных истинностных значений (пример такого изменения продемонстрирован в [18]). Ясно, что гипотезы, подтверждаемые всегда (или в подавляющем большинстве случаев), представляются более достоверными, чем те, для которых на разных шагах примерно с одинаковой частотой получаются разные истинностные значения. Таким образом, предлагаемый подход позволяет получить более дифференцированную оценку правдоподобности выдвигаемых гипотез. Фактически, результатом индуктивного вывода при таком подходе оказывается не одна гипотеза (об истинности, ложности или противоречивости предположения о корреляции), полученная на некотором шаге, а целая последовательность значений (т. е. цепочка гипотез, последовательно выдвигаемых на каждом шаге процесса «проверки на устойчивость»). Иначе говоря, вместо истинностных значений вида $\langle \varepsilon, k \rangle$ (где k — натуральное число, а ε принадлежит множеству $V_\tau^{(0)}=\{1, -1, 0, \tau\}$) теперь значениями оказываются последовательности вида $\bar{\varepsilon}=(\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots)$, где $\varepsilon_i \in V_\tau^{(0)}$ выражает гипотезу, полученную на i-м шаге проверки. Здесь при теоретическом моделировании, не связанном с конкретными ограничениями размера моделей и конкретными способами организации вычислительной процедуры, целесообразно допускать в качестве $\bar{\varepsilon}$ произвольные, возможно, и бесконечные последовательности. В то же время в реальных, прикладных ситуациях, требующих алгоритмической обработки, необходимо, чтобы все последовательности $\bar{\varepsilon}=(\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$ были конечны, причем достаточно обозримой длины. Это выдвигает задачу нахождения приемлемых способов завершения (или, возможно, прерывания) процесса индуктивного вывода. Нетривиальность этой задачи связана исключительно со сложностными соображениями. При чисто теоретическом взгляде на вычислимость завершение рассматриваемого процесса «переопределения» значений (разумеется, в случае конечных моделей) легко определить алгоритмически. Действительно, достаточно, чтобы на двух различных шагах n_1 и n_2 ($n_1<n_2$, (n_2-n_2) произвольно) повторилось все распределение значений (гипотез, выдвинутых на этих шагах) целиком во всей доопределяемой матрице. После этого последовательность вычислений будет повторяться периодически, т. е. $\varepsilon_i=\varepsilon_{i-k(n_2-n_1)}$ (для всех $i\geqslant n_2$) (где k таково, что $n_1\leqslant i-k(n_2-n_1)<n_2$). Однако при этом наилучшая верхняя оценка «длины периода» (точнее: верхняя оценка для n_2) при «тотальной» перепроверке на каждом шаге всех недоопределенных элементов исходной базы данных: 4^{n_0}, где n_0 — число эле-

ментов, недоопределенных на начальном (нулевом) шаге. А поскольку каждый шаг индуктивного вывода сам достаточно сложен (он состоит из большого числа «элементарных операций» по проверке всех правил индуктивного вывода), получаемый алгоритм может оказаться практически не способным работать в обозримое время. Поэтому интересно рассматривать альтернативные варианты «досрочного» прекращения индуктивного вывода при достижении некоторых более или менее естественных условий. Рассмотрим одну из таких возможностей. Для этого вспомним, что при «монотонной» стратегии число шагов индуктивного вывода было ограничено числом n_0, так как на каждом шаге количество недоопределенных элементов матрицы уменьшалось (хотя бы на единицу), а все «доопределившиеся» элементы как бы «исключались из рассмотрения» и дальнейшему «пересмотру» не подвергались. При «немонотонной» стратегии аналогичное свойство не выполняется. Элемент, принявший значение $\neq\tau$ на некотором шаге $n>0$, в дальнейшем снова может оказаться недоопределенным. Тем не менее, можно ввести некоторое множество «существенно недоопределенных» элементов, которое в процессе индуктивного вывода не расширяется. А именно, назовем элемент матрицы (базы данных) существенно недоопределенным на шаге n, если он ни разу (до этого шага) не был доопределен (т. е. ни разу не принимал значения $\neq\tau$). Ясно, что на каждом шаге индуктивного вывода число существенно недоопределенных элементов не увеличивается. Поэтому, если условиться обрывать процесс индуктивного вывода, как только на некотором шаге множество существенно недоопределенных элементов окажется неизменным (т. е. не произойдет доопределения ни одного существенно недоопределенного элемента), то можно будет гарантировать, что это случится на шаге, не большем n_0 (как и в случае «монотонной» стратегии). Если при этом процесс индуктивного вывода окажется «фактически монотонным», т. е. каждая однажды выдвинутая гипотеза на всех дальнейших шагах «подтверждается» ($\varepsilon_i\neq\tau\rightarrow\varepsilon_{i+1}=\varepsilon_i$ для всех элементов матрицы), то результат по существу совпадает с результатом индуктивного вывода при «монотонной» стратегии: достаточно сопоставить истинностное значение $\langle\varepsilon, k\rangle$ последовательности — $(\underbrace{\tau, \tau, \ldots, \tau}_{k}, \varepsilon, \varepsilon, \ldots)$ (и значение $\langle\tau, \omega\rangle$ — последовательности $(\tau, \tau \ldots \tau, \ldots)$).

Для последовательностей $\overline{\varepsilon}=(\varepsilon_0, \varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n)$, получаемых в результате индуктивного вывода, определим различные функции, описывающие «оттенки правдоподобия». Например, $f_c(\overline{\varepsilon})=\left(1-\frac{\delta_c}{n}\right)$, где δ_c — число изменений истинностного значения — оценка «стабильности» гипотезы (например, $f_c(\overline{\varepsilon})=1$ для «абсолютно стабильных» гипотез $\bar{\varepsilon}=(\varepsilon, \varepsilon, \ldots, \varepsilon)$, $\varepsilon \in V_{\tau}^{(0)}$ или $f_{\tau}(\bar{\varepsilon})=\left(1-\frac{\delta_{\tau}}{n}\right)$, где δ_{τ} — количество вхождений τ в $\bar{\varepsilon}$, — оценка «доопределенности» гипотезы (например, $f_{\tau}(\bar{\varepsilon})=1$, если $\bar{\varepsilon}=(\varepsilon_0, \varepsilon_0, \ldots, \varepsilon_0)$, где $\varepsilon_0 \neq \tau$ — значение, доопределенное в исходной базе данных).

Для дедуктивной формализации этой стратегии можно рассмотреть континуумзначную логику со множеством истинностных значений $(V_{\tau}^{(0)})^{\mathrm{N}}$ (множество функций из N в $V_{\tau}^{(0)}$, т. е. последовательностей вида $\bar{\varepsilon}=(\bar{\varepsilon}(0), \bar{\varepsilon}(1), \bar{\varepsilon}(2), \ldots)$). В ней определяются логические операторы (связки и кванторы), позволяющие представить процедуры, подобные (1) и (2), в форме аксиоматического исчисления (подробнее см. в [2]).

В заключение отметим, что хотя техническая реализация «немонотонности» в описанном здесь смысле несколько отличается от других работ по немонотонным логикам, фактически эти рассмотрения связаны с общим подходом к немонотонности, состоящем в возможном изменении значений истинности (переходе от истинных значений к неистинным) при включении в рассмотрение дополнительных сведений (в условиях неполной информации).

§ 12. Перспективы и прогноз (нерешенные задачи)

Перспективы развития теории правдоподобных рассуждений связаны с развитием автоматизированных интеллектуальных информационных систем (АИИС) и ролью последних не только как имитаторов в смысле Тьюринга [35], но и как усилителей умственной деятельности экспертов в смысле Эшби [47].

Следующая схема характеризует строение АИИС:

АИИС=РИС+ПИС+САП, где РИС, ПИС и САП, соответственно, рассуждающая информационная система, поисковая информационная система и система автоматического пополнения БД (базы фактов).

Реализация рассуждения на ЭВМ (в смысле описанного выше понимания «рассуждения») предполагает представление знаний в виде КАТ, которые характерны для предметных областей, данные которых хорошо структурированы, а знания — слабо формализованы. Слабая формализованность знаний, характерная для экспериментального естествознания, компенсируется в практике рассуждений естествоиспытателей правдоподобными умозаключениями (аналогиями, индукцией), которые в неформализованном виде являются «техническим арсеналом» продуктивного (творческого) мышления [9]. Формальная аппроксимация продуктивного мышления требует, однако, модификации известных средств математической логики, центром исследований которой являются аксиоматические теории и теория доказательств. В русле потребностей «компьютерной логики» возникают проблемы формальной имитации правдоподобных рассуждений модифицированными средствами математической логики [24]. По-видимому, одним из основных принципов формализации правдоподобных рассуждений является принцип: «семантика задачи определяет средства ее формализации, если таковые отсутствуют — их следует выдумать» [39].

РИС есть реализация в компьютерной системе комбинации «правдоподобный вывод+достоверный вывод», создание таких «автоматических рассуждателей» означает развитие архитектуры и функций решателей задач в интеллектуальных системах. Но реализация рассуждения, включающего правдоподобный вывод, требует поиска релеватных посылок в БД и базах знаний АИИС, т. е. использования ПИС. Однако АИИС существует в «информационной среде» (в смысле Ю. А. Шрейдера и М. В. Арапова), что требует автоматического пополнения БД из множества текстов на естественном языке, образующего расширяющуюся «информационную среду». Таким образом, открытость множества посылок правдоподобного рассуждения, характеризующая его немонотонность, специфична для АИИС и адекватна применениям подходящих средств формализации машинно-ориентированных правдоподобных выводов. Правдоподобное рассуждение является специфическим средством человеко-машинных систем, осуществляющих симбиоз «экспертная система (ЭС)+пользователь». ЭС II-го поколения есть РИС, использующая базы данных и знаний с неполной информацией и обладающая подсистемой обоснования результата, которая применяет некоторый критерий достаточного основания правдоподобного рассуждения (например, рассмотренные выше) (в связи с понятием ЭС II-го поколения см. [20]).

Упомянем в заключение некоторые проблемы развития теории правдоподобных рассуждений типа ДСМ.

Проблема 1. Требуется создать программные системы, реализующие так называемые обобщенные методы автоматического порождения ДСМ-гипотез [39]. Это означает, что элементарные ЭЗ типа «причина (подобъект)$\Rightarrow_2$следствие (множество свойств)», которые порождаются применением п. п. в. I-го рода, должны быть обобщены так, что они будут представимы *тернарным* отношением $T^+(V, \mathscr{X}, W)$, которое интерпретируется следующим образом: «V есть (+)-причина множества свойств W при *отсутствии «тормозов»* из множества $\mathscr{X}$». Естественно, что порождение гипотез вида $T^+(V, \mathscr{X}, W)$ потребует обобщения п. п. в. I-го рода, что повлечет за собой обобщение п. п. в. II-го рода.

Проблема 2. Включение в структуры данных для БД частично упорядоченных параметров (в том числе — линейно упорядоченных) потребует создания программ, реализующих п. п. в. I-го рода типа «сопутствующие изменения+структурное сходство». Создание таких программ может привести к возникновению РИС типа «Милль+Бэкон», где «Милль»—ДСМ-подсистема, описанная выше, а «Бэкон» — подсистема типа [60].

Проблема 3. Необходимость обобщения п. п. в. I и II рода до немонотонных процедур [2] потребовала создания соответствующих программных систем, их реализующих.

Проблема 4. Требуется создать программные средства трансляции $\mathscr{L}_e'$ и $\mathscr{L}_e''$ в соответствующие алгоритмические языки.

Проблема 5. Распространить ДСМ-метод автоматического порождения гипотез на различные структуры данных (например, на объекты, являющиеся кортежами, словами, графами и пространственными графами), определив соответственно локальное сходство для объектов указанной структуры.

Проблема 6 (связана с Проблемой 4). Требуется построить язык $\mathscr{L}_e''$ такой, чтобы он обладал металогическими средствами для логического проектирования упорядоченного множества возможных стратегий, реализующих функционирование решателя типа «правдоподобный вывод+достоверный вывод» [54]. $\mathscr{L}_e''$ должен содержать декларации (высказывания и высказывательные формы [44]), вопросы и императивы как специфические термы языка, в котором формализуются вопросы, ответы на которые посредством соответствующих процедур, выполняющих императивы, дадут возможность решать задачи соответствующего класса. Этот класс задач определит соответствующий тип РИС.

$\mathscr{L}_e''$ должен быть «шире» наличных в РИС процедур, что даст возможность пополнять средствами $\mathscr{L}_e''$ АИИС новыми процедурами в соответствии с принципом открытости интеллектуальных систем [42; 3] (в последней работе принцип открытости АИИС содержится в понимании рассуждения как процесса, в котором посредством металогических средств формируются новые посылки, релевантные цели рассуждения).

Таким образом, в РИС с неполной информацией, содержащей правила правдоподобного вывода, открытость РИС и понимание правдоподобия связаны в следующем смысле.

Если в посылках п. п. в. имеются утверждения типа «$(C \Rightarrow_i A)$ имеет истинностное значение $\langle\tau, n\rangle$» (где $i=1,2$), а заключением является утверждение «$(C \Rightarrow_i A)$ имеет истинностное значение $\langle\nu, n+1\rangle$» (где $\nu \in \{-1, 0, 1\}$), то в КАТ, содержащей базу знаний Σ_{pr} (т. е. только процедурную БЗ), утверждение «$(C \Rightarrow_i A)$ имеет истинностное значение $\langle\nu, n+1\rangle$», выводимо и истинно. Однако при интерпретации формулы $(C \Rightarrow_i A)$ в реальной действительности эта формула может иметь истинностное значение, отличное от ν. Например, выводимо «$(C \Rightarrow_i A)$ имеет истинностную оценку $\langle 1, n+1\rangle$» (т. е. C обладает множеством свойств A), а в действительности установлено, что объект C не обладает множеством свойств A и т. п.

Таким образом, в КАТ выводимо истинное внешнее утверждение о формулах $(C \Rightarrow_1 A)$, $(C' \Rightarrow_2 A)$, полученное посредством п. п. в. из $\mathfrak{R}'$, а в действительности $(C \Rightarrow_1 A)$ (или $(C' \Rightarrow_2 A)$) имеет другое истинностное значение, установленное неформально посредством дополнительной информации о действительности (например, эта информация может быть получена в новом эксперименте). Указанный факт согласуется с условием открытости множества Σ', соответствующего БД с неполной информацией. Итак, оценки внутренней формулы $(C \Rightarrow_1 A)$ посредством п. п. в и посредством сопоставления с информацией о действительности могут быть различны. Следовательно, правдоподобие указанных формул (гипотез) не может быть охарактеризовано одними только метасредствами КАТ. Это обстоятельство вызывает следующую проблему.

Проблема 7. Требуется формулировать архитектуру АИИС как *человеко-машинной системы*, допускающей посредством ПИС+САП с использования критерия достаточного основания правдоподобного рассуждения в РИС полуавтоматическое пополнение БД и БЗ.

Проблема 8. Требуется развить аппарат автоматического расширения фрагментов БЗ посредством решателя задач типа «правдоподобный вывод+достоверный вывод», порождающего индуктивные обобщения эмпирических зависимостей (ЭЗ) при расширениях БД, тем самым выдвигая гипотезы об *эмпирических закономерностях*, обладающих определенной *устойчивостью* [28].

Проблема 9. Применить ДСМ-метод автоматического порождения гипотез для новых задач установления ЭЗ (генотип—фенотип» (структура данных—«тексты генетического кода»); распространить указанный метод для порождения продукций, посредством которых можно породить сети реакций метаболизмов.

Данная проблема связана с банальным тезисом: «индукция должна обосновываться индуктивно».

Еще раз следует обратить внимание на различие между приближенными и правдоподобными выводами и рассуждениями, отмеченное в начале данного раздела. И несмотря на то, что это различие формулируется здесь не по «канонам строгости», принятой в точных науках, мы считаем полезным подчеркнуть следующие обстоятельства.

(1) Формализация приближенных выводов требует уточнения и введения средств оценки *ошибки* заключения (например, наличие множеств «малой меры» в смысле [10]). Таким образом, целью «хорошей» формализации приближенных выводов является *минимизация ошибки заключений* для соответствую-

щих выборок обучающих примеров и контрпримеров. Точная формулировка минимизации ошибки заключения образует достаточное основание (или критерии рациональности) приближенных выводов и рассуждений. Целью «хорошей» формализации правдоподобных выводов и рассуждений является порождение множества возможных фальсификаторов заключений и установление факта нефальсифицируемости устойчивого множества заключений относительно «наибольших» (т. е. самых «опасных») фальсификаторов на множествах расширяющихся обучающих выборок примеров и контрпримеров (полуфактов). Точная формулировка указанного обстоятельства и является достаточным основанием (или критерием рациональности) в смысле Поппера [31]. Некоторые попытки формализации достаточного основания правдоподобных рассуждений были предприняты в основной части статьи (см. также в связи с этим [41] и [21], где подробно охарактеризованы экспертные системы типа ДСМ, содержащие правдоподобные рассуждения, управляемые некоторым критерием достаточного основания принятия заключения).

В связи со сказанным возникает еще одна естественная проблема:

Проблема 10. Создать решатель задач в РИС типа «недостоверный вывод+достоверный вывод», где под недостоверным выводом понимается «правдоподобный вывод+приближенный вывод». Практически в ЭС это означает, в частности, комбинирование правдоподобных рассуждений со статистическими методами, что может оказаться весьма полезным средством человеко-машинных интеллектуальных систем.

Говоря фигурально, создатели интеллектуальных систем не чуждаются выдвижения «лозунгов на злобу дня». Пойдя на поводу у этой «моды», скажем, что специфическим качеством интеллектуальной системы является реализация механизма «имитатор+усилитель умственных способностей» (ср. [35], [47]).

В связи с этим создание интеллектуальных систем типа «Милль+Бэкон» (подробнее о программе «Бэкон» см. в [52]) расширяет возможности автоматизированной обработки данных и знаний, являющейся сферой деятельности информатики. Отметим при этом, что формализация правдоподобных рассуждений, являющаяся по существу некоторой «вычислительной эвристикой», является «логическим стержнем» управления процессом автоматизированной обработки данных и знаний посредством ЭВМ.

Правдподобное рассуждение, использующее средства достоверного вывода, образует сферу применений методов искусственого интеллекта, имеющую условное название «компьютерная логика», которая предназначена быть научным аппаратом человеко-машинного познания природы. Именно, «игра с природой» в человеко-машинных системах, открывающая новые эмпирические зависимости в БД и БЗ, недоступные возможностям человеческого рассуждения, имеет обратную связь, стимулирующую развитие средств теории правдоподобного рассуждения, столь необходимых для проблематики искусственного интеллекта как дисциплины, изучающей имитацию рассуждения и восприятия информации посредством ЭВМ.

В заключение автор считает своим приятным долгом отметить вклад в развитие теории ДСМ-рассуждений своих коллег О. М. Аншакова, Т. Гергея, С. М. Гусакову, Д. П. Скворцова, а также отметить вклад своих коллег, реализовавших ДСМ-рассуждения в программных системах, которые были созданы М. И. Забежайло, М. А. Михеенковой, К. П. Хазановским, В. Г. Ивашко, Е. С. Панкратовой и С. О. Кузнецовым. Важная роль в апробации результатов машинных экспериментов принадлежит экспертам В. В. Авидону, В. Г. Блиновой, Д. А. Бодягину, С. П. Козловой, В. М. Красновой.

В настоящей работе были использованы материалы препринта: Т. Гергей, В. К. Финн «К теории правдоподобных рассуждений», Будапешт, 1987 год.

§ 13. Приложение

Формулировки некоторых $\mathcal{M}^{\sigma}_{x,n}$-предикатов:

$$\tilde{\mathcal{M}}^{-}_{a,n}(V, W, k) \rightleftharpoons \exists k\, \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots \exists U_k (J_{(-1,n)}(Z_1 \Rightarrow_1 U_1) \mathbin{\&}$$
$$\mathbin{\&} \forall U J_{(-1,n)}(Z_1 \Rightarrow_1 U) \to U \subseteq U_1)) \mathbin{\&} \ldots \mathbin{\&} (J_{(-1,n)}(Z_k \Rightarrow_1 U_k) \mathbin{\&}$$
$$\mathbin{\&} \forall U J_{(-1,n)}(Z_k \Rightarrow_1 U) \to U \subseteq U_k)) \mathbin{\&} (Z_1 \cap \ldots \cap Z_k = V) \mathbin{\&}$$
$$\mathbin{\&} V \neq \varnothing \mathbin{\&} \forall i\, \forall j\, ((i \neq j) \mathbin{\&} 1 \leqslant i,\ j \leqslant k \to Z_i \neq Z_j) \mathbin{\&}$$
$$\mathbin{\&} \forall X\, \forall Y\, ((J_{(-1,n)}(X \Rightarrow_1 Y) \mathbin{\&} \forall U\, (J_{(-1,n)}(X \Rightarrow_1 U) \to U \subseteq Y) \mathbin{\&}$$
$$\mathbin{\&} V \subset X) \to \left(W \neq \varnothing \mathbin{\&} W \subseteq Y \mathbin{\&} \left(\bigvee_{i=1}^{k} (X = Z_i)\right)\right)\right) \mathbin{\&} k \geqslant 2\Big)$$

(предикат отрицательного сходства);

$$\tilde{\mathcal{M}}^{-}_{a(6),n}(V, W, k) \rightleftharpoons \exists k\, \exists Z_1 \ldots \exists Z_k \exists U_1 \ldots$$
$$\ldots \exists U_k \left(\mathop{\&}_{i=1}^{k} (J_{(-1,n)}(Z_i \Rightarrow_1 U_i) \mathbin{\&} \forall U\, (J_{(-1,n)}(Z_i \Rightarrow_1 U) \to\right.$$
$$\left.\to U \subseteq U_i))\right) \mathbin{\&} (Z_1 \cap \ldots \cap Z_k = V) \mathbin{\&} V \neq \varnothing \mathbin{\&} \forall i \forall j\, ((i \neq j) \mathbin{\&}$$
$$\mathbin{\&} 1 \leqslant i,\ j \leqslant k \to Z_i \neq Z_j) \mathbin{\&} \forall X\, \forall Y\, ((J_{(-1,n)}(X \Rightarrow_1 Y) \mathbin{\&}$$
$$\mathbin{\&} \forall U\, (J_{(-1,n)}(X \Rightarrow_1 U) \to U \subseteq Y) \mathbin{\&} V \subset X) \to$$
$$\to \left(W \neq \varnothing \mathbin{\&} W \subseteq Y \mathbin{\&} \left(\bigvee_{i=1}^{k} (X = Z_i)\right)\right)\Big) \mathbin{\&} (k \geqslant 2) \mathbin{\&}$$

$$\& \forall X \forall Y \exists Z \exists V_0 \exists W_0 ((J_{(-1,n)} (X \Rightarrow_1 Y) \& W \subseteq Y \& V \subset X) \to$$
$$\to (Z = ((X \setminus V) \cup V_0) \& \neg (V \subset V_0) \& V_0 \neq \varnothing \&$$
$$\& (J_{(1,n)} (Z \Rightarrow_1 W \cup W_0) \vee J_{(\tau,n)} (Z \Rightarrow_1 W \cup W_0)))))$$

(предикат отрицательного различия);

$$\tilde{\Pi}_n^- (V, W) \rightleftharpoons \exists Y_1 \ldots \exists Y_k \Big(\Big(\mathop{\&}_{i=1}^{k} \exists X_i (J_{(-1,n)} (X_i \Rightarrow_2 Y_i) \& X_i \subset V) \&$$
$$\& \Big(\bigcup_{i=1}^{k} Y_i = W \Big) \& \forall Y \Big(\exists X (J_{(-1,n)} (X \Rightarrow_2 Y) \& X \subset V) \to$$
$$\to \Big(\bigvee_{i=1}^{k} (Y = Y_i) \Big) \Big) \& \forall U ((U \subseteq W \& U \neq \varnothing) \to$$
$$\to \neg \exists Z (J_{(1,n)} (Z \Rightarrow_2 U) \& Z \subseteq V)) \Big)$$

(предикат $\text{ПИО}_n^{(-)}$: Принцип индуктивного обобщения для отрицательного случая).

ЛИТЕРАТУРА

1. *Аншаков О. М., Скворцов Д. П., Финн В. К.*, Некоторые семантические и синтаксические проблемы ДСМ-метода автоматического порождения гипотез. НТИ, 1984, сер. 2, № 11, 5—13
2. —, —, —, О монотонных и немонотонных формализациях ДСМ-метода автоматического порождения гипотез. В кн.: Школа-семинар «Семиотические аспекты формализации интеллектуальной деятельности». Кутаиси, 1985. Тезисы докладов. М.: ВИНИТИ, 1985, 108—114
3. —, —, —, Логические средства экспертных систем типа ДСМ. Семиотика и информатика, вып. 28, 1986, 65—101.
4. —, —, —, Логические средства ДСМ-метода автоматического порождения гипотез. НТИ, 1987, сер. 2, № 9, 9—17
5. *Барздинь Я. М.*, Некоторые правила индуктивного вывода и их применение. Семиотика и информатика, вып. 19, 1982, 59—89
6. *Бочвар Д. А.*, Об одном трехзначном исчислении и его применении к анализу парадоксов классического расширенного функционального исчисления. Мат. сб., 1938, *4*, № 2, 287—308
7. *Брудно В. А., Скворцов Д. П., Финн В. К., Цаленко М. Ш.*, Базы данных с неполной информацией. Семиотика и информатика, вып. 25, 1985, 5—45
8. *Бэкон Ф.*, Новый Органон. ОГИЗ, 1935
9. *Вертгеймер М.*, Продуктивное мышление. М.: Прогресс, 1987
10. *Гаек П., Гавранек Т.*, Автоматическое образование гипотез. М.: Наука, 1984
11. *Геловани В. А., Ковригин О. В.*, Экспертные системы в медицине. М.: Знание, 1987
12. *Гренандер У.*, Лекции по теории распознавания образов. Т. 2. М.: Мир, 1981
13. *Гусакова С. М., Финн В. К.*, О формализации локального и глобального сходств. НТИ, 1986, сер. 2, № 6, 16—19
14. —, —, Сходство и правдоподобный вывод. Изв. АН СССР. Тех. кибернет., 1987, № 5, 42—63
15. —, —, О новых средствах формализации понятия сходства. НТИ, 1987, сер. 2, № 10, 14—22
16. *Забежайло М. И., Финн В. К., Козлова С. П., Катамадзе Т. Г., Авидон В. В., Рабинков А. А.*, Об одном методе автоматического формирования гипотез и его программной реализации. НТИ, 1982, сер. 2, № 4, 20—26
17. —, —, *Авидон В. В., Катамадзе Т. Г., Блинова В. Г., Бодягин Д. А., Рабинков А. А.*, Об экспериментах с базой данных с неполной информацией

посредством ДСМ-метода автоматического порождения гипотез. НТИ, 1983, сер. 2, № 2, 28—32
18. —, Ненаследуемость эмпирического противоречия в ДСМ-методе и немонотонные рассуждения. НТИ, 1984, сер. 2, № 11, 14—17
19. *Ивашко В. Г.*, Об алгоритме индуктивного обобщения II уровня. В кн.: Школа-семинар «Семиотические аспекты формализации интеллектуальной деятельности». Кутаиси, 1985. Тезисы докладов. М.: ВИНИТИ, 1985, 141—145
20. —, *Финн В. К.*, Экспертные системы и некоторые проблемы их интеллектуализации. Семиотика и информатика, вып. 27, 1986, 25—61
21. —, *Фабрикантова Е. Ф., Финн В. К., Хазановский К. П.*, Об экспертных системах типа ДСМ. В кн.: Справочник по интеллектуальным системам, М.: Наука, 1989
22. *Кайберг Г.*, Вероятность и индуктивная логика. М.: Прогресс, 1978
23. *Мальцев А. И.*, Алгебраические системы. М.: Наука, 1970
24. *Маслов С. Ю.*, Асимметрия познавательных механизмов и ее следствия. Семиотика и информатика, вып. 20, 1983, 3—31
25. *Мендельсон Э.* Введение в математическую логику. М.: Наука, 1976
26. *Милль Д. С.*, Система логики силлогистической и индуктивной. М.: Книжное дело, 1900
27. *Михеенкова М. А., Авидон В. В., Суханова С. А.*, О программной реализации ДСМ-метода автоматического порождения гипотез с неодноэлементным множеством признаков. НТИ, 1984, сер. 2, № 11, 1984, 18—22
28. —, *Финн В. К.*, Об одном классе экспертных систем с неполной информацией. Изв. АН СССР Тех. кибернет., 1986, № 5, 82—103
29. *Панкратова Е. С., Ивашко В. Г., Авидон В. В., Блинова В. Г., Бодягин Д. А.*, Экспериментальная проверка новой версии ДСМ-метода. НТИ, 1988, сер. 2, № 2
30. *Пойа Д.*, Математика и правдоподобные рассуждения. М.: Наука, 1975
31. *Поппер К.*, Логика и рост научного знания. М.: Прогресс, 1983
32. *Поспелов Д. А.*, Ситуационное управление. Теория и практика. М.: Наука, 1986
33. *Скворцов Д. П., Финн В. К.*, Замечание об одном расширении языка многосортной логики предикатов. НТИ, 1981, сер. 2, № 8, 25—26
34. —, О некоторых способах построения логических языков с кванторами по кортежам. Семиотика и информатика, вып. 20, 1983, 102—126
35. *Тьюринг А.*, Может ли машина мыслить? М.: Гос. изд. физ.-мат. лит., 1960
36. *Уинстон П. Г.*, Построение структурных описаний по примерам. В кн.: Психология машинного зрения. М.: Мир, 1978, 185—248
37. *Финн В. К.*, О возможностях формализации правдоподобных рассуждений средствами многозначных логик. В кн.: VII Всес. симп. по логике и методологии науки. Киев: Наукова думка, 1976, 82—83
38. —, Базы данных с неполной информацией и новый метод автоматического порождения гипотез. В кн.: Диалоговые и фактографические системы информационного обеспечения. М., 1981, 153—156
39. —, О машинно-ориентированной формализации правдоподобных рассуждений в стиле Ф. Бэкона—Д. С. Милля. Семиотика и информатика, вып. 20, 1983, 35—101
40. —, Правдоподобные выводы и проблемы автоматического порождения теорий из фактов. В кн.: Интенсиональные логики и логическая структура теорий. Тез. докл. советско-финского коллоквиума по логике. Тбилиси: Мецниереба, 1985, 156—158
41. *Хазановский К. П., Финн В. К.*, Некоторые вопросы программного обеспечения для экспертных систем на основе ДСМ-метода автоматического порождения гипотез. НТИ, 1988, сер. 2, № 1
42. *Хьюитт К.*, Открытые системы. В кн.: Реальность и прогнозы искусственного интеллекта. М.: Мир, 1987
43. *Чень Ч., Ли Р.*, Математическая логика и автоматическое доказательство теорем. М.: Наука, 1983

44. *Черч А.*, Введение в математическую логику. М.: ИЛ, 1956
45. *Шрейдер Ю. А.*, Равенство, сходство, порядок. М.: Наука, 1971
46. *Этмане И. Э.*, Об одной формализации понятия многоточия. Семиотика и информатика, вып. 27, 1986, 121—141
47. *Эшби Р.*, Введение в кибернетику. М.: ИЛ, 1959
48. *Яблонский С. В.*, Введение в дискретную математику. М.: Наука, 1979
49. *Anshakov O. M., Skvortsov D. P., Finn V. K.*, On logical means of formalization of plausible inferences. Proc. VII Intern. Congress of Logic, Methodology and Philos. of Sci., 1987
50. *Barzdin J. M.*, Some Rules of Inductive Inference and their use for Program Synthesis. In: Information Processing 83, IFIP, 1983, 333—338
51. *Bohr N.*, Nature, 143, 1939
52. *Bradshaw G. L., Langley P., Simon H. A.*, The discovery of internistic properties. Proce. of the third national conference of the Canadian society for computional studies of intelligence, 1980, 19—25
53. —, —, —, Studying scientific discovery by computer simulation, Sci., *222*, 1983, 971—975
54. *Gergely T., Finn V. K.*, On solver «plausible inference+deduction» type in intellectual informational-computing systems, In: Artificial Intellegence, IFAC, Proc. Ser., № 9, 1984
55. —, —, Constructivity of plausible inferences and its role in a theory of reasoning. In: VIII Intern. Congr. of Logic, Meth. and Philos. of Sci., Moscow, 1987, 250—254
56. *Gabbay D. M.*, Intuitionistic Basis for Nonmonotonic Logic. Lect. Notes Comp. Sci., *138*, 1982, 260—273
57. *Goodman I. R., Nguen H. T.*, Knowledge-based systems. North-Holland, Amsterdam, 1985
58. *Gupta M. M., Kandel A., Bandler W., Kiszka J. V.* (eds.), Approximate Reasoning in Expert systems, North-Holland, Amsterdam, 1985
59. *Hume D.*, Treatise on human nature. Selby-Bigge ed., 1888, Book I, Part III, Section IV
60. *Langley P.*, Rediscovering Phisics Whith BACON. 3, Proc. Sixth Intern. Joint Conf. on Artif. Intell., 1979, 505—507
61. *Los J.*, The Foundations of Methodological Analysis of Mill's Methods. In: Twenty Five Years of Logic. Meth. in Poland, W., 1977
62. *McCarthy J. M.*, Circumscription — a form of non-monotonic Reasoning. Artif. Intell., 1980, *13*, 27—39
63. *McDermott D., Doyle J.*, Non-monotonic Logic I. Artif. Intell., 1980, *13*, № 1—2, 41—72
64. —, Non-monotonic Logic II. Non-monotonic Modal Theories. J. Assoc. Comput. Mach., 1982, 29, № 1, 33—57
65. *Michalski R. A.*, A Theory and Methodology of Inductive Learning. Artif. Intell., 1983, *20*, № 2
66. *Morgan J. G.*, Hypotheses generation by machine. Artif. Intell., 1971, *2*, 179—187
67. *Peirce C. S.*, Philosophical Writings. Buchler J. ed., Dover Publ. Co., N. Y., 1955
68. *Plotkin G. D.*, A Further Note on Inductive Generalization. Mach. Intell. *6*, 1971, 101—124
69. *Popper K.*, Objective Knowledge. Oxford, Clarendon Press, 1974
70. *Ramsey F. P.*, The Foundations of Mathematics. Routledge and Kegan Paul, London, 1950
71. *Reichenbach H.*, The Theory of Probability. Berkley, 1949
72. *Reiter R.*, A logic for Default Reasonings. Artif. Intell., 1980, *13*, № 1—2, 81—132
73. *Shapiro E. Y.*, Algoritmic Programm Debugging. May, 1982, Research Report 237
74. von *Wright G. H.*, A Treatise on Induction and Probability. N. Y., 1960
75. *Winograd T.*, Extended Inference Modes in Reasoning by Computer Systems. Artif. Intell., 1980, *13*, № 1—2, 5—26