Ladislav Tondl

К проблемам семантической информации

Kybernetika, Vol. 8 (1972), No. 3, (189)--212

Persistent URL: http://dml.cz/dmlcz/125773

К проблемам семантической информации

Ладислав Тондл

Статья дает обзор и оценку основных направлений современной теории семантической информации. Автор показывает некоторые возможности развитии тех понятий семантической информации, которые разработала финская школа в логике.

1. ВВЕДЕНИЕ. ИНФОРМАЦИЯ И ДАННЫЕ. СЕМАНТИЧЕСКАЯ ИНФОРМАЦИЯ

Термин „информация" стал в последнее время не только часто употребляемым словом, но также чем-то, что возбуждает определенные ожидания. Как заметил Й. Гинтикка [7], речь идет прежде всего об ожидании о проблематике, которая имеет нечто общее с математической теорией информации в смысле Шэннона и с обширной и уже сейчас математически и технически хорошо вооруженной проблематикой передачи и обработки данных. Этим иллюзиям и надеждам способствует в большой мере неточное употребление термина „информация" в не очень специальных или в популярных работах, где, обычно, происходит смешение двух категорий понятий: (1) тех понятий, которые можно выразить как „данные", „сведения", „высказывания" и т. д., (2) понятий, которые в той или иной форме связаны с понятием „информация", либо в смысле понимания математическо-статистической теории информации, либо в понимании так называемой логическо-семантической теории информации. Эта неразбериха, смешение обоих типов, к сожалению, проникло и в специальную литературу, в которой, например, неоправданно применяется термин „машины для обработки информации" (в действительности речь идет о машинах для обработки данных) и т. д.

Интуитивно разницу между обоими типами понятий можно выразить следующим образом: термин „информация" (от латинского in-formare) всегда связан с изменениями энтропических уровней, т. е. со снижением неопределенности при решении, со снижением риска, с повышением качества тех шагов, которые опираются на полученную информацию. Очевидно, что не каждое сведение способно „in-formare", т. е. в буквальном смысле слова приводить к определенной (желаемой) форме, создавать, в современном же смысле слова снижать энтропическую уровень в связи с данным комплексом задач и целей, связанных с этими задачами. Для врача, который устанавливает диагноз пациента, самыми важными сведениями являются наблюдаемые симптомы, однако, никакой роли не играет цвет потолка в комнате, где лежит пациент. Такое сведение или другие, аналогичные сведения неспособны снизить его неуверенность в сявзи с данной задачей, т. е. в связи с установлением диагноза.

Другими трудностями и проблемами, которые связаны с попытками уточнения категории (2), являются до сих пор еще неудовлетворительно разработанное и объясненное отношение так называемой математическо-статистической и логическо-семантической информации, или смешение возможностей и компетенций обеих теорий. Известно, что математическо-статистическая теория информации в своем первоначальном виде, который ее придал К. Шеннон, вообще не претендовала на постижение так называемого семантического аспекта. Определенная вина в возникновении некоторых надежд и ожиданий, что теория Шеннона будет очень скоро обобщена или расширена на семантические аспекты, лежит на В. Вивере, работа которого была напечатана как дополнение к более поздним изданиям основополагающего произведения Шеннона [12]. Эта работа, так же как и некоторые другие популярные изложения теории (например, Ц. Чери [8]), очень близко придвинули семантические, а в некоторой степени и прагматические аспекты информации к математическо-статистической теории информации. В данной работе оставлена в стороне обширная и глубоко разработанная проблематика математическо-статистической теории информации, главное внимание уделено некоторым основным шагам, которые вели к развитию так называемой семантической теории информации.

Первыми систематическими работами о понятии „семантическая информация" обычно считаются раобты Р. Карнапа и И. Бар-Хиллела [1] в начале пятидесятых лет. Эти работы основаны на результатах логической (или индуктивной) теории вероятностей, которую представил Р. Карнап [2]. И. Бар-Хиллел и Р. Карнап, что в противоположность этому теория информации Шеннона опирается на частотную концепцию вероятонсти, так что можно полагать, что отличие теории информации Шеннона от логическо-семантической теории информации заключается в различии между так называемой логической вероятностью и так называемой частотной вероятностью (по терминологии Карнапа вероятность$_1$ и вероятность$_2$). Как показал Й. Гинтикка, такая позиция является определенным упрощением и отличие обеих теорий является намного более глубоким.

Основной идеей логическо-семантической теории информации является предположение, что данное высказывание тем более богато, содержательно с информационной точки зрения, чем больше снижается неопределенность, или, говоря другими словами, чем большее количество возможностей или альтернатив исключается. Необходимо подчеркнуть, что эта идея, которую можно найти в работах Карнапа по логической семантике, впервые была высказана еще в тридцатых годах К. Поппером [11]. Поппер набросал схему понятия „содержание“ (Gehalt) предложения либо „содержание“ теории. Он предполагал, что содержание растет с ростом состояний, которые данное высказывание или данная теория исключает. Это понимание также связано с Попперовской концепцией фальсификации и с ролью фальсификации при оценке высказывания или научной теории: содержание данной теории является тем же самым, что ее невероятность, причем, содержание теории растет с классом фальсификационных возможностей этой теории, т. е. с ростом класса состояний, которые исключаются.

Эта исходная идея, которая в принципе содержится и в дальнейших, более поздних работах, может быть уточнена в еще большей степени на основании того, как мы понимаем то, что Поппер назвал „фальсификацией возможности“, или то, что Карнап назвал „логическими состояниями“, которые исключаются. В работах Й. Гинтикки и других представителей финской школы очень часто упоминаются альтернативы, что является определенной модификацией классической концепции „возможных миров“ Лейбница. С этой точки зрения семантическо-информационное содержание высказывания тем богаче, чем большее количество альтернатив или „возможных миров“ данное высказывание не допускает или исключает. Другими словами, семантическо-информационной содержание высказывания определяется не тем, что содержит даннле высказывание, но тем, что оно исключает. Именно потому, что тавтологические высказывания содержат все возможные альтернативы или все „возможные миры“, семантическо-информационное содержание тавтологических высказываний является нулевым.*

Формально точное развитие этой концепции зависит от того, каким способом выражено то, что соответствует „фальсификационным возможностям“ Поппера, либо то, что соответствует „логическим состояниям“ Карнапа или альтернативам Гинтикка. Самыми известными и наиболее употребляемыми являются два способа: первый основывается на методе описания состояния и был намечен в вышеуказанных работах Карнапа и Бар-Хиллера. Второй способ, который опирается на схемы так называемых конституентов, был намечен в работах Й. Гинтикка.**

* С этой точки зрения максимальным (по Карнапу равно 1) является семантическо-информационное содержание контрадикторных, т. е. логически ложных высказываний, которые исключают все альтернативы или „возможные миры“. Этот результат принятой концепции, однако, не находится в полном согласии с повседневным пониманием информационного содержания. Их этих соображений целесообразно ограничить компетенцию данной концепции только лишь на правдивые высказывания.

** Как показал Й. Гинтикка [7], идея так называемых конституентов принадлежит Булю.

Метод описания состояния оперирует языком $\mathbf{L}_n^\pi$, в котором π есть количество предикатов и n есть количество индивидов, т. е. языком, в словаре которого имеются эти мимологические клнстанты: индивиды $a_1, a_2, \ldots, a_n$, и предикаты $P_1, P_2, \ldots, P_\pi$. (Для простоты обычно предполагаем, что речь идет об одноместных предикатах.) Каждое описание состояния представляет коньюнкцию всех возможных отдельных предложений или их отрицаний. Описание состояния таким образом явно определяет, какие индивиды относятся к отдельным подмножествам, какие – образованы коньюнктивным соединением всех предикатов или их отрицанием. Если предположить, что отдельные предложения создаются последовательностью произвольного предиката и произвольного индивида (речь идет всегда о константах индивидов), то тогда все возможные ситуации, которые можно выразить на языке $\mathbf{L}_n^\pi$, выражает дизъюнкция всех возможных описаний состояний, т. е.

$$
\begin{array}{l}
P_1a_1 \,.\; P_1a_2 \ldots\; P_1a_n \,.\; P_2a_1 \,.\; P_2a_2 \ldots\; P_2a_n \ldots\; P_\pi a_1 \,.\; P_\pi a_2 \ldots\; P_\pi a_n \vee \\
\vee P_1a_1 \,.\; P_1a_2 \ldots\; P_1a_n \,.\; P_2a_1 \,.\; P_2a_2 \ldots\; P_2a_n \ldots\; P_\pi a_1 \,.\; P_\pi a_2 \ldots\; P_\pi a_n \vee \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\vee \sim P_1a_1 \,.\sim P_1a_2 \ldots \sim P_1a_n \,.\sim P_2a_1 \,.\sim P_2a_2 \ldots \sim P_2a_n \ldots \sim P_\pi a_1 \sim P_\pi a_2 \ldots \sim P_\pi a_n \,.
\end{array}
$$

Очевидно, что в случае приведенного друзначного решения в системе, которую можно описывать языком $\mathbf{L}_n^\pi$, имеется в целом $2^{\pi n}$ возможных описаний состояний. В случае k-значного решения здесь имеется в целом $k^{\pi n}$ возможных описаний состояний.

Очевидно, что метод описаний состояний наталкивается на препятствие тогда, когда нельзя гарантировать финитность исследуемых систем, которые могут быть описаны сравнительно небольшим количеством терминов, которые можно разложить на конечное число подклассов.

Каждое из описаний представляет одно из возможных состояний, в котором может находиться система, описываемая языком $\mathbf{L}_n^\pi$. Это, однако, не означает, что все возможные состояния действительно реализуемы. Другими словами, некоторым описаниям состояний в данной исследуемой системе не обязательно должны соответствовать реальные состояния. Для того, чтобы можно было ввести основные понятия теории семантической информации, т. е. понятие „содержательная мера“ и понятие „семантическая информация“, необходимо пробабилизировать метод описания состояния. При этом мы предполагаем, что каждому высказанию на языке $\mathbf{L}_n^\pi$ соответствует определенное описание состояния или несколько определенных описаний состояний. (Тогда тавтологическим высказаниям соответствуют все возможные описания состояния, контрадикторным высказываниям не соответствует никакое описание состояния.) Если можно каждому описанию состояния поставить в соответствие определенную вероятностную меру (которая, естественно, может быть и нулевой) таким образом, чтобы сумма всех возможных описаний состояния была равна единице и чтобы эта мера была аддитивной, то тогда с помощью этой меры можно определить приведенные основные понятия семантической теории информации.

Понятие содержательной меры высказывания S_i (в символической записи cont (i)) характеризуется вероятностной мерой тех описаний состояний, которые данное высказывание исключают. Причем предполагается, что вероятностная мера произвольного высказывания (которое не является внутренне противоречивым) равна сумме всех вероятностных мер тех описаний состояний, которые соответствуют данному высказыванию. Если известна вероятностная мера высказывания S_i, которую обозначим $p(i)$, тогда

$$\operatorname{cont}(i) = 1 - p(i)\,.$$

Аналогичным образом можно ввести понятие „условная содержательная мера"' т. е. содержательная мера высказывания S_j по отношению к высказыванию S_i, которую обозначим cont (j/i). Эта условная содержательная мера определяет, собственно говоря, во сколько раз больше информации приносит одно высказывание по сравнению с другим высказыванием. Условная содержательная мера определяется следующим образом:

$$\operatorname{cont}(j/i) =_{\mathrm{df}} \operatorname{cont}(i\,.\,j) - \operatorname{cont}(i)\,.$$

(В данной записи cont $(i\,.\,j)$ является содержательной мерой конъюнкции высказываний $S_i\,.\,S_j$.) Интуитивно ясно, что условная содержательная мера равна нулю, если высказывание S_j логически вытекает из высказывания S_i. Максимум, который может достигнуть условная содержательная мера cont (j/i), выражается числом cont (j).

Кроме содержательной меры и условной содержательной меры Бар-Хиллел и Карнап [1] ввели еще понятие „информационная мера" или также „семантическая информация". Это решение опирается на те же самые исходные предположения, что и понятие „содержательная мера", однако выбирает другой способ количественных подсчетов. В связи с тем, что cont (i) может иметь значения в интервале между 0 и 1, причем 0 — это содержательная мера тавтологии, а 1 — это содержательная мера контрадикторного высказывания, предлагается другое решение, в котором соответствующая мера могла бы принимать значения от 0 до бесконечности. Эта мера обычно обозначается inf (i), что является информационной мерой высказывания S_i, которая определена следующим образом:

$$\inf(i) =_{\mathrm{df}} \log \frac{1}{1 - \operatorname{cont}(i)}\,.$$

В связи с тем, что cont $(i) = 1 - p(i)$ имеет место следующее равенство

$$\inf(i) = \log \frac{1}{p(i)} = -\log p(i)\,.$$

 Аналогично можно определить условную информационную меру $\inf(j/i)$, которая определяется как:

$$\inf(j/i) =_{\mathrm{df}} \inf(j \,.\, i) - \inf(i)$$

(где $\inf(j \,.\, i)$ есть информационная мера конъюнкции $S_j \,.\, S_i$).

При применении содержательной меры и информационной меры при оценке собранных данных необходимо учитывать условия аддитивности для $\mathrm{cont}(i)$ а $\inf(i)$. При этом желательно, чтобы операция соединения понималась как конъюнктивное соединение первичных высказываний. Это означает, что $\mathrm{cont}(i \,.\, j)$, или $\inf(i \,.\, j)$ есть содержательная мера или информационная мера соответствующая конъюнкции высказываний $S_i \,.\, S_j$.*

Содержательная мера $\mathrm{cont}(i \,.\, j)$ является аддитивной, т. е.

$$\mathrm{cont}(i \,.\, j) = \mathrm{cont}(i) + \mathrm{cont}(j)$$

тогда и только тогда, когда высказывания S_i и S_j являются логически дизъюнктивными. В противоположном случае имеет место**

$$\mathrm{cont}(i \,.\, j) = \mathrm{cont}(i) + \mathrm{cont}(j) - \mathrm{cont}(i \vee j)\,.$$

Информативная мера $\inf(i \,.\, j)$ аддитивна, т. е.

$$\inf(i \,.\, j) = \inf(i) + \inf(j)$$

тогда и только тогда, если высказывание S_i иррелевантно по отношению, к высказыванию S_j с учетом вероятностной меры p, на основании которой информационная мера сконструирована. Это означает, что

$$\inf(i \,.\, j) = \inf(i) + \inf(j)$$

тогда и только тогда, если

$$p(i \,.\, j) = p(i)\, p(i)\,.$$

Мы уже указывали на то, что содержательная мера высказывания S_i т. е. $\mathrm{cont}(i)$, характеризуется вероятностной мерой тех описаний состояний, которые данное высказывание S_i не допускает или исключает. Напротив, $\inf(i)$, как показал Гинтикка [6], является мерой величины неожиданности S_i.

* Формально корректная запись должна была бы аргументы содержательной и информационной меры записывать на метаязыке. С точки зрения упрощения записи в качестве аргументов приводятся лишь индексы первичных высказываний. Эта мысль содержится в нашей работе „Проблемы семантики“ [13].

** Очевидно, что эти условия соответствуют отношению сложения и объединения множеств, когда элементы двух множеств можно складывать тогда и только тогда, когда оба множества дизъюнктивны.

Метод описания состояния, который является следствием вышеприведенных рассуждений, ведущих к конструированию так называемой содержательной меры и информационной меры, оперирует конечным числом индивидов и конечным числом предикатов. При конструивовании приведенных понятий обычно не предполагается, что отдельные предикаты могут зависеть друг от друга, например, в следующем смысле: если определенному индивиду соответствует определенный предикат, то ему соответствует и другой определенный предикат и т. д. Если нам известна зависимость отдельных предикатов на языке $\mathbf{L}_n^\pi$, которые мы способны выразить с помощью постулатов значения, можно соответствующим способом модифицировать конструкции cont (i) и inf (i).*

Прежде всего необходимо модифцировать вероятностную меру высказывания на языке $\mathbf{L}_n^\pi$, в котором введены постулаты значения. Если $\mathfrak{P}$ — класс всех постулатов значения, то тогда можно от вероятностной меры p перейти к модифицированной вероятностной мере p', которая должна удовлетворять следующим условиям:

(1) Вероятностная мера p' равна нулю для каждого описания состояния в $\mathbf{L}_n^\pi$, в котором $\mathfrak{P}$ не имеет места.

(2) Для каждого описания состояния в $\mathbf{L}_n^\pi$, в котором имеет место $\mathfrak{P}$, модифицированная вероятностная мера p' прямо пропорциональна p для каждого описания состояния. (Коэффициент этой пропорциональности обозначим $\boldsymbol{K}$.)

(3) Сумма модифицированных вероятнсстных мер для всех описаний состояния в $\mathbf{L}_n^\pi$ равна единице.

В связи с тем, что предполагается, что $\mathfrak{P}$, т. е. класс всех постулатов значения, является логическим истинным, то имеет место следующее равенство:

$$p'(\mathfrak{P}) = \boldsymbol{K} \,.\, p(\mathfrak{P}) = 1\,,$$

откуда вытекает формула:

$$\boldsymbol{K} = \frac{1}{p(\mathfrak{P})}\,.$$

Модифицированные содержательные меры и информационные меры для высказывания S_i, т. е. меры, учитывающие класс постулатов значения $\mathfrak{P}$, можно ввести таким образом: В связи с тем, что

$$\text{cont}\,(i) = 1 - p(i)$$

можно ввести модифицированную содержательную меру cont' (i) так:

$$\text{cont}'\,(i) = 1 - p'(i) = 1 - \frac{p(i)}{p(\mathfrak{P})}\,.$$

Аналогично можно ввести модифицированную информационную меру выска-

* Этот способ модификации cont (i) и inf (i) был намечен в работе автора [13].

 зывания S_i:

$$\mathrm{inf}'(i) = \log \frac{1}{p'(i)} = \log \frac{p(\mathfrak{P})}{p(i)} = \log p(\mathfrak{P}) - \log p(i)\,.$$

Приведенное рассуждение, которое приводит к конструированию модифицированной содержательной меры и модифицированной информационной меры, естественно применимо там, где класс постулатов значения точно фиксирован, причем, соответствующие зависимости, которые выражены в отдельных постулатах, остаются неизменными.

Кроме рассуждений, которые опираются на метод описания состояния и языки, оперирующие с индивидами и предикатами, можно применить и другие рассуждения, которые опираются на схемы так называемых конституентов. Эти рассуждения, намеченные Гинтиккой [7], позволяют в качестве исходного языка использовать более простой язык логики высказываний. Для того, чтобы результаты такой позиции были удовлетворительными, необходимо, чтобы учитывалось требование финитности. Язык, в котором введены содержательная мера и информационная мера, содержит следующие элементы:

(1) конечное количество непроанализированных отдельных высказываний $A_1, A_2, \ldots, A_k$;

(2) связки исчисления высказываний $\sim$ (отрицание), . (конъюнкция), $\vee$ (дизъюнкция) и некоторые другие связки, которые могут определяться на основании выше приведенных, например, $\rightarrow$ (материальная импликация) и т. д.

В системе, которая может быть выражена приведенными средствами, можно различать отдельные возможности, которые выражаются следующей формой высказываний, называемых Гинтиккой конституентами:

$$(\pm)\, A_1 \,.\, (\pm)\, A_2 \,.\, \ldots \,.\, (\pm)\, A_k\,,$$

где символ$(\pm)$означает, что отдельное высказывание, стоящее за этим символом, может иметь знак отрицания или не иметь этого знака. Очевидно, это количество всех возможных конституентов равно 2^k. Таким образом, конституенты соответствуют всем возможным альтернативам либо „возможным мирам“ в традиционном понимании Лейбница. Каждое истинное высказывание можно считать решением, какие из альтернатив исключаются, а какие — допускаются. Затем, это высказывание можно считать дизъюнкцией определенного количества конституентов (хотя бы одного конституента — в крайней случае тавтологического высказывания всех конституентов). Если конституенты будем означать C, тогда данное высказывание эквивалентно дизъюнкции

$$C_1 \vee C_2 \vee \ldots \vee C_{w(h)}\,,$$

где $w(h)$ — объем высказывания.*

* Ясно, что понятие „объем высказывания“ соответствует тому, что в семантике Карнапа характеризуется как логическое пространство данного высказывания.

Если же для этих рассуждений введем вероятностные меры, то можно использовать следующее исходное положение: определенное высказывание S_h тем более вероятно, чем больше альтернатив представляет, иил, другими словами, чем больше его объем $w(h)$. Это позволяет ввести следующую вероятностную меру для высказывания S_h:

$$p(h) = \frac{w(h)}{2^k} .$$

На основании рассуждений Гинтикка вводит следующее определение информационной меры высказывания S_h:

$$\inf(h) = -\log \mathbf{p}(h) = -\log \frac{w(h)}{2^k} = k - \log w(h)$$

(при предположении, что речь идет о логарифмах на базе 2).

Аналогично, на этой же базе можно ввести содержательную меру высказывания S_h таким образом:

$$\operatorname{cont}(h) = \frac{2^k - w(h)}{2^k} = 1 - p(h) .$$

Приведенные рассуждения, исходящие из конечного количества отдельных высказываний, естественно предполагают конечность вероятностного поля, где каждой альтернативе соответствует определенная численная величина. Иначе говоря, такой подход позволяет измерять содержательную меру или информационную меру всех возможных высказываний на данном языке при предположении, что мы в состоянии количественно оценить все возможные альтернативы, которые соответствуют отдельным конституентам, причем эта оценка образует финитное вероятностное поле.

Очевидно, что метод конституентов является лишь определенным упрощением метода описания состояния, причем, этот метод в конце концов приводит к аналогичным результатам. Точно также, предпосылки и границы применимости у этих методов в принципе одинаковые: строго определенная финитность, возможность вероятностной количественной оценки отдельных состояний или альтернатив, учет основных постулатов теории вероятностей.

2. АБСОЛЮТНАЯ А ПЕРЕНЕСЕННАЯ ИНФОРМАЦИЯ

Приведенные рассуждения, конечно, при условии, что выполняются соответствующие предположения — позволяют проводить количественную семантическо-информационную оценку произвольных высказываний в данной языковой системе. Однако, как показал Гинтикка [6], чаще всего нас не интересует та информация, которая выражается с помощью $\operatorname{cont}(i)$ или $\inf(i)$. Обычно требyется дальнейшая информация в зависимости от той информации, которой мы уже располагаем. Иначе говоря, нас интересует, насколько снизится первоначальная неопределенность тем, что мы получим дальнейшие сведения. Если обозначить новые полученные сведения высказыванием S_h, (что, естественно, может быть одним высказыванием или непустым конъюктивным классом высказываний), а то, что было до сих пор известно высказыванием S_g, то тогда состоит в том, как выразить, насколько уменьшается наша неопределенность по отношению к тому, к чему относится S_g, тем, что мы получаем S_h.

Тогда информационная мера того, что приносит S_h, по отношению к тому, чего касается S_g, определяется так:

$$\inf(g) - \inf(g/h) = \log \frac{p(g/h)}{p(g)} = \log \frac{p(g \,.\, h)}{p(g)\, p(h)} .$$

Это выражение представляет сдвиг в неопределенности, который был достигнут тем, что к имеющимся сведениям, которые выражаются S_g, присоединяются другие сведения, выражаемые S_h.

Эти рассуждения, позволяющие выражать содержательную меру или информационную меру отдельных высказываний, Гинтикка характеризовал как рассуждения, определяющие абсолютную информацию. Все те рассуждения, которые ведут к характеристике сдвига в первоначальной неопределенности, предполагают определенную релятивизацию информации. Поэтому приведенное выражение, определяющее разницу между первичной абсолютной информационной мерой и вновь полученной условной информационной мерой, представляет один из элементарных видов релятивизации.

Выражение $\inf(g) - \inf(g/h)$ Гинтикка характеризует как *перенесенная информация* и записывает $\operatorname{transinf}(h/g)$. Это выражение с интуитивной точки зрения соответствует тому, что дает высказывание S_h по отношению к тому, чего касается S_g. Это можно проиллюстрировать на следующих примерах. Предположим, что на следующую неделю, в течение которой мы хотим поехать в отпуск, имеется долгосрочное предсказание, что будет солнечная погода (S_g). В конце предшествующей недели стоит солнечная погода (S_h). Речь идет о том, чтобы обсудить, что дает нам высказывание S_h по отношению к тому, чего касается S_g. Разница между абсолютной информацией и перенесенной информацией особенно выступает в том случае, если S_h является сингулярным эмпирическим высказыванием о данном объекте x, а S_g — некоторым генерализированным высказыванием о классе определенных объектов X, причем имеет место следующее соотношение $x \in X$. Обычно нас интересует не только наблюдаемый объект x, но также и другие элементы класса X, хотя в данных условиях всех их наблюдать мы не можем. Нам также интересно знать, что может принести высказывание об объекте x по отношению к тем высказываниям, которые касаются всего класса X.

Понятие „перенесенная информация" transinf (i/j), которое позволяет определить, что дает высказывание S_i по отношению к тому, к чему относится S_j,* особенно важно для тех ситуаций, когда высказывание S_i относится к отдельным объектам или к определенной малой группе объектов, являющиеся элементами класса, к которому относится S_j. Например, в исследовательской деятельности речь идет о наблюдениях, измерениях, об отдельных экспериментах, причем бывает обычно желательным, чтобы на основании результатов этих или подобных процедур придти к выводу о большем количестве объектов или, точнее говоря, о том, что об этом большем количестве объектов до сих пор предполагаем. Очевидно, что понятие „перенесенная информация" связано с ядром проблематики научной индукции и индуктивных логик.

В приведенных процедурах обычно является желательным решить, дадут ли дальнейшие наблюдения, измерения, эксперименты не только дальнейшие сведения (в том, что повторение наблюдений, измерений, экспериментов на основании имеющихся программ этих и дальнейших процедур дает нам новые сведения, обычно никто не сомневается. Однако, спорным является то, дает ли эта деятельность дальнейшую информацию, т. е. снижает ли она неопределенность в нашем решении, в наших знаниях об определенных классах явлений и т. д.) но и дальнейшую информацию. Именно из этих соображений целесообразно ближе заняться рассмотрением понятия „перенесенная информация", в особенности тогда, когда имеется возможность оперировать количественной формой этого понятия.

Развитием этого понятия можно считать все концепции, которые пытаются — с семантическо-информационной точки зрения — взвешивать то, что дают результаты наблюдений и экспериментирования по отношению к предметной области, которой касается определенная гипотеза. Уточнить понятие „информация, полученная от данного наблюдения по отношению к данной гипотезе" попытался Р. Гилпинэн [5]. Работа Гилпинэна в сущности состоит в дальнейшей спецификации понятия „перенесенная информациа" и в определении его свойств.

Предположим, что результаты наблюдений резюмированы в высказывании S_e, в то время как гипотеза о всей области, которой касается наблюдение, выражена в высказывании S_h. Тогда Гилпинэн вводит понятие $Q(e/h)$ как объяснение информации, которую дает S_e о том, к чему относится гипотеза S_h.** $Q(e/h)$ считается количественным понятием, причем, естественно, что исключаются те случаи, когда S_h является L-истинным. Свойства $Q(e/h)$ определяются с помощью постулатов и выводимых из них теорем, которые соответствуют

* Очевидно, эти и дальнейшие аналогичные формулировки аддитивных стандартов для меры семантической информации, предполагают детонационные концепции семантики высказываний S_i, S_j, и т. д.

** В отличие от языка Гилпинэна выберем такой способ выражения, который соответствует детонационно-семантической концепции.

 тому, что мы интуитивно ожидаем от этой меры перенесенной информации:

(R 1) Если $S_e \underset{L}{\equiv} S_i$ и $S_h \underset{L}{\equiv} S_k$,

тогда $Q(e/h) = Q(i/h)$, $Q(e/h) = Q(e/k)$.

Этот постулат соответствует тому, что два логических эквивалентных (синонимных) высказывания дают одинаковую информацию по отношению к S_h, а то же самое высказывание дает одинаковую информацию по отношению к двум синонимным гипотезам.

Следующий постулат касается границы значений, которые может принимать $Q(e/h)$. При этом желательно, чтобы это значение было минимальным или нулевым, когда S_e является тавтологическим высказыванием, недающим никакой информации. Это значение должно быть максимальным, если $S_e \underset{L}{\rightarrow} S_h$, т. е. высказывание S_e логически имплицирует высказывание S_h. Выбор экстремных значений $Q(e/h)$ зависит от договоренности. Наиболее естественным является следующий выбор: в качестве минимума $Q(e/h)$ принять значение 0, а в качестве максимума — выбрать 1, этому соответствует второй постулат:

(R 2) $$0 \leqq Q(e/h) \leqq 1\,.$$

Затем, необходимо потребовать, чтобы

(R 3) если $S_e \underset{L}{\rightarrow} S_h$, тогда $Q(e/h) = 1$.

Необходимо также потребовать, чтобы $Q(e/h)$ было аддитивным. С интуитивной точки зрения ясно, что речь идет о том, чтобы в случае, когда имеются два различных наблюдения, которые выражаются с помощью высказываний S_e и S_i и которые не дают никакой общей информации по отношению к S_h, имело место следующее равенство:

$$Q(e\,.\,i/h) = Q(e/h) + Q(i/h)\,.$$

Этому соответствует ситуация, когда дизъюнкция $S_e \vee S_i$ логически истинна. Если дизъюнкция $S_e \vee S_i$ логически истинна, то тогда приведенные высказывания не дают никакой общей информации. Если, кроме того, соотношение $S_e \vee S_i \vee S_h$ логически истинно, то тогда первые два высказывания не дают никакой информации по отношению к тому, к чему относится S_h. Это выражает следующий постулат:

(R 4) если $S_e \vee S_i \vee S_h$ логически истинно,

то тогда $Q(e\,.\,i/h) = Q(e/h) + Q(i/h)$.

Иначе имеет место равенство:

$$Q(e\,.\,i/h) = Q(e/h) + Q(i/h) - Q(e \vee i/h)\,.$$

Затем, желательно, чтобы информация, которую дает дизъюнкция $S_e \vee S_i$ по отношению к тому, к чему относится S_h, была меньше или равна информации, которую дает одно из высказываний по отношению к тому, к чему относится S_h, т. е.

$$Q(e \vee i/h) \leqq Q(e/h) \quad \text{и}$$

$$Q(e \vee i/h) \leqq Q(i/h)\,.$$

Этому соответствует и следующий постулат: $Q(e/h)$

(R 5) $$Q(e \vee i/h) = Q(e/h)\,Q(i/h \vee e) \quad \text{или}$$

$$Q(e \vee i/h) = Q(i/h)\,Q(e/h \vee i)\,.$$

Мы уже упоминали о том, что необходимо исключить те ситуации, когда S_h является тавтологическим высказыванием. В конце концов, такое высказывание немыслимо как гипотеза о данной предметной области, и поэтому не имеет смысла размышлять над информацией, которую дает S_e по отношению к тому, к чему относится S_h, в том случае, когда S_h является тавтологическим высказыванием.

Понятие $Q(e/h)$ является, однако, лишь одним из возможных видов понятия „перенесенная информация". $Q(e/h)$ соответствует разнице, которую с помощью информационной меры можно выразить как:

$$\inf(h) - \inf(h/e)\,.$$

Необходимо строго отличать это понятие от понятия условной информационной меры $\inf(h/e)$, которые мы определили следующим образом:

$$\inf(h/e) = \inf((h\,.\,e) - \inf(e)\,.$$

Если $Q(e/h)$ выражает информацию, которую дает S_e по отношению к тому, к чему относится гипотеза S_h, то тогда можно поставить вопрос о том, в какой степени информация, которая дается гипотезой S_h, передается также наблюдением, измерением, экспериментированием, результаты которого выражает высказывание S_e. Этому соответствуют различные концепции степеней подтверждений, которые были развиты в системах индуктивных и вероятностных логик. С семантическо-информационной точки зрения этому соответствует понятие „силы очевидности" (evidential strenght), свойства которого обсуждает Г. Торнебом [14].

3. ИНФОРМАЦИОННЫЕ МЕРЫ НОМОЛОГИЧЕСКИХ ВЫСКАЗЫВАНИЙ

Уже имеевшее место рассуждение о том, что дает с информационной точки зрения определенная гипотеза по отношению к тому, что мы узнали в течении наблюдений или экспериментов, обращает внимание к общим высказываниям. С информационной точки зрения нас, естественно, интересуют те общие высказывания, какими являются, например, научные законы, научные гипотезы или определенные эмпирические обобщения, имеющие эмпирическую интерпретацию. (Речь идет об области таких высказываний, которые характеризуются как номологические высказывания, несмотря на некоторую неопределенность этого термина, несмотря на то, что до сих пор не удалось удовлетворительно решить проблему разграничения номологических высказываний с одной стороны и других общих высказываний, в особенности так называемых случайных обобщений с другой стороны.)

Исходным положением для определения информационной ценности номологических высказываний может быть роль, которую эти высказывания играют в различных научных процедурах. Вообще говоря, функцией номологического высказывания является снижение неопределенности в решении об объектах, которые попадают в предметную область данного номологического высказывания. Речь может идти, например, об объяснении того, что данный предмет имеет определенные свойства, об объяснении того, почему произошли определенные события, о предсказании достоверного или в высокой степени вероятного события и т. д. Таким образом целесообразно выбрать для создания масштабов информационной ценности номологических высказываний определенные общие схемы тех процедур, в которых имеются номологические высказывания. Обычно, речь идет о схемах научного объяснения, научных предсказаний, схемах диагностических рассуждений и т. д. Согласно этим схемам (например, дедуктивно — номологическая схема научного объяснения и научного предсказания Гемпеля, соответствующие статистические аналогии этой схемы и т. д.) ценность номологических высказываний, т. е. научных законов, гипотез, теорий и т. д., состоит в том, что эти высказывания способны в большей или в меньшей мере объяснять определенные факты, позволяют осуществить предсказание этих фактов, их классификацию. Если эти процедуры назовем согласно Гемпелю научной систематизацией, то тогда можно говорить о *систематизационной* мощности номологического высказывания (теории, закона, гипотезы и т. д.). Особым случаем систематизационной мощности являются объяснительная мощность, предсказательная моъность, диагностическая мощность и т. д.*

* В связи с тем, что процедуры научного объяснения и научного предсказания не являются полностью симметричными, нельзя отождествлять объяснительную мощность и предсказательную мощность.

То, что с интуитивной точки зрения ожидается от понятия „систематизационная мощность", совершенно аналогично вышеприведенным требованиям к стандартам семантической информации: речь идет о том, чтобы выразить, в какой степени определенная научная гипотеза S_h способна снизить нашу неопределенность в решении о факте, выраженном в высказывании S_e, например, о факте, который мы должны объяснить, предсказать и т. д. (В случае научного объяснения S_e представляет так называемый экспланандум, т. е. высказывание о том, что должно быть объяснено. В случае предсказания S_e — высказывание о будущем факте, который мы должны определить, предсказать и т. д.) Очевидно, что при конструировании этого понятия предполагается наличие релятивизованного подхода, т. е. подхода, который пытается оценить семантическую ценность номологического высказывания S_h по отношению к тому, к чему относится высказывание S_e. Нри этом целесообразно предполагать, что S_e — сингулярное истинное синтетическое высказывание и что исключаются случаи, когда S_e является логически истинным, т. е. когда является тавтологическим высказыванием, либо когда является логически ложным (контрадикторным). Систематизационная мощность данного номологического высказывания определяет, насколько данное номологическое высказывание снижает неопределенность по отношению к тому, что объяснено, предсказано и т. д. Предпосылкой такого подхода к конструированию понятия систематизационной мощности является возможность определения первичной неопределенности, связанной с высказыванием S_e, которую обозначим $\boldsymbol{U}(e)$, и условной непределенности $\boldsymbol{U}(e/h)$, т. е. неопределенности, связанной с высказыванием S_e при предположении, что можно полностью положиться на истинность номологического высказывания S_h.* Неопределенность $\boldsymbol{U}(e)$ может также считаться условной неопределенностью, если мы обуславливаем ее знанием тавтологического высказывания S_t. Тогда имеется $\boldsymbol{U}(e) = \boldsymbol{U}(e/t)$.

Пока предположим, что мы в состоянии количественно определить $\boldsymbol{U}(e)$ и $\boldsymbol{U}(e/h)$. Одновременно предположим, что S_h является номологическим высказыванием и S_e не является ни логически истинным, ни логически ложным. Таким образом обоснованную конструкцию меры систематизационной мощности разработали Пиетаринэн и Туомела [9] и Пиетаринэн [10].

Если обозначить меру систематизационной мощности номологического высказывания S_h по отношению к тому, к чему относится S_e, знаком $\boldsymbol{E}(h/e)$, то тогда желательно, чтобы свойства $\boldsymbol{E}(h/e)$ удовлетворяли следующим постулатам:

$$\text{(г 1)} \qquad \boldsymbol{E}(h/e) = f[\boldsymbol{U}(e), \boldsymbol{U}(e/h)] .$$

* Естественно, возможен и такой подход к конструированию понятия „систематизационная мощность", когда надежность номологического высказывания определена вероятностным способом. Тогда вышеприведенная ситуация, когда мы можем полностью полагаться на истинность S_h, является частным случаем, когда надежность S_h максимальна и обычно равна 1, если имеет место утверждение, что значения надежности S_h лежат в интервале нежду 0 и 1.

Мера систематизационной мощности номологического высказывания S_h, релятивизированная по отношению к тому, к чему относится S_e, здесь функцией двух аргументов, т. е. неопределенности, связанной с S_e, и условной неопределенности S_e по отношению к S_h.

(г 2) $\quad \boldsymbol{E}(h/e) \gtreqless 0$ тогда и только тогда, когда $\boldsymbol{U}(e/h) \lesseqgtr \boldsymbol{U}(e)$,

(г 3) $\quad \boldsymbol{E}(h/e) = \max \boldsymbol{E}$ тогда и только тогда, когда $\boldsymbol{U}(e/h) = \min \boldsymbol{U} = 0$,

(г 4) $$\max \boldsymbol{E} = 1\,,$$

(г 5) $\quad \boldsymbol{E}(h/e) \geqq \boldsymbol{E}(k/e)$ тогда и только тогда, когда
$$\boldsymbol{U}(e/h) \leqq \boldsymbol{U}(e/k)\,.$$

На основании (г 5) функция f монотонно убывает, если растут значения ее второго аргумента. Нри этом целесообразно предполагать, что f является линейной функцией второго аргумента.

На основании приведенных постулатов можно определить $\boldsymbol{E}(h/e)$ таким образом:

(D1) $$\boldsymbol{E}(h/e) =_{\mathrm{df}} \frac{\boldsymbol{U}(e) - \boldsymbol{U}(e/h)}{\boldsymbol{U}(e)}\,.$$

Согласно такому определению меры систематизационной мощности все номологические высказания, которые снижают неопределенность по отношению к тому, к чему относится S_e, на одну и ту же степень, имеют одинаковую систематизационную мощность.

Естественным является требование, чтобы эта мера систематизационной моъности, аналогично всем приведенным выше мерам семантической имформации, была аддитивной. Если имеются два различных номологических высказывания S_h и S_k, то тогда имеет место следующее утверждение:

$$\boldsymbol{E}(h\,.\,k/e) = \boldsymbol{E}(h/e) + \boldsymbol{E}(k/e)\,,$$

если

$$\boldsymbol{U}(e/h) + \boldsymbol{U}(e/k) - \boldsymbol{U}(e/h\,.\,k) = \boldsymbol{U}(e)\,.$$

Определение (D 1), очевидно, не накладывает никаких ограничений на минимальное значение $\boldsymbol{E}(h/e)$, в то время как максимум $\boldsymbol{E}$ согласно постулату (г 4), всегда равен 1. Возникают две возможности которые обозначим (I) и (II), причем либо (I) либо (II) есть следущюий постулат, который необходимо присоединить к вышеприведенным постулатам для удовлетворительной теории систематизационной мощности:

(I) $\quad \boldsymbol{E}(h/e) = \min \boldsymbol{E}$ тогда и только тогда, когда $\boldsymbol{U}(e/h) = \max \boldsymbol{U}$.

Альтернативной возможностью для $\min \boldsymbol{E}$ является то, чтобы максимальная величина условной неопределенности, связанной с высказыванием S_e, в том случае, если можно надежно полагать на S_h, равнялась первичной неопределенности $\boldsymbol{U}(e)$. Это означает, что

$$\text{(II)} \qquad \boldsymbol{E}(h/e) = \min \boldsymbol{E} \quad \text{тогда и только тогда, когда} \quad \boldsymbol{U}(e/h) = \boldsymbol{U}(e)\,.$$

Эта другая возможность соответствует интуитивному предположению, когда научный закон или научная гипотеза S_h совершенно не имеет никакого значения по сравнению с S_e. (Представим себе, например, что S_e есть прогнозное высказывание о будущем затмении Луны, а S_h — научный закон, который относится к совершенно другим явлениям. Тем самым, однако, не утверждается, что если $\boldsymbol{E}(h/e) = \min \boldsymbol{E}$, то $\boldsymbol{E}(h/i)$ также минимально, если S_i — отличное истинное синтетическое высказывание.)

Обе альтернативы, т. е. (I) и (II) совершенно различны и взаимно несовместны, т. к. они предполагают различную интерпретацию понятия „непрпеделенность". Осевидно, что понятие „неопределенности", связанной с определенным высказыванием, является многозначным и для дальнейшего уточнения меры систематизационной мощности необходимы дальнейшие объяснения.

Теперь мы обратим внимание на некоторые свойства $\boldsymbol{E}(h/e)$: В связи с тем, что задачей номологического высказывания является снижение неопределенности по отношению к S_e, имеет место утверждение, что если S_e логически вытекает из S_h, то тогда S_h не оставляет никакой неопределенности по отношению к S_e. (Это касается, например, ситуации, когда в смысле дедуктивно-номологической модели научного объяснения Гемпеля экспланадум можно однозначно вывести из эксплананс, или ситуации, когда сведения, включая соответствующие номологические высказывания, позволяют нам осуществлять однозначно точное и определенное предсказание.) Поэтому:

$$\text{если} \quad S_h \underset{L}{\rightarrow} S_e\,, \quad \text{тогда} \quad \boldsymbol{E}(h/e) = 1\,.$$

Если же S_h — тавтологическое высказывание, то оно не может дать никакой информации по отношению к S_e. Тогда:

$$\text{если } S_h \text{ логически истинно, то} \quad \boldsymbol{E}(h/e) = 0\,.$$

Эти свойства $\boldsymbol{E}$ являются полностью тривиальными и соответствуют интуитивному подходу. Более трудным является определение условий для $\min \boldsymbol{E}$. И здесь возникают два альтернативных положения, которые соответствуют (I) и (II). Одной из возможных альтернатив условий для $\min \boldsymbol{E}$ является тот случай, когда из S_h логически вытекает отрицание S_e. В связи с тем, что S_h логически имплицирует $\sim S_e$, определенность, с которой мы ожидаем, что S_e, является минимальной, и соответствующая условная неопределенность максимальной, т. е.

(I) если $S_h \underset{L}{\rightarrow} \sim S_e$, то тогда $E(h/e) = \min E$.

Возможно, однако, и другое решение, которое оперирует с отрицанием S_h. Мера систематизационной мощности S_h по отношению к S_e минимальна тогда, когда S_e вытекает из отрицания S_h, т. е.:

(II) если $\sim S_h \underset{L}{\rightarrow} S_e$, то тогда $E(h/e) = \min E$.

Очевидно, что эти приведенные условия для $\min E$ соответствуют вышеприведенным альтернативам. Обе альтернативы, как уже об этом говорилось выше, предполагают различную интерпретацию понятия „неопределенность" данного высказывания:

Альтернатива (I), которая считает $E(h/e)$ минимальным тогда, когда неопределенность $U(e/h)$ максимальна, и при предположении, когда из S_h логически вытекает отрицание S_e, интерпретирует меру неопределенности, связанную с данным высказыванием, как меру неожиданности того, к чему относится S_e.

Альтернатива (II), которая считает $E(h/e)$ минимальным тогда, когда условная неопределенность S_h по отношению к S_e равна первичной неопределенности S_e и когда S_e логически вытекает из отрицания S_h, интерпретирует неопределенность данного высказывания S_e как недостаточность значия того, к чему относится S_e.

Теперь обратим свое внимание на условия, которым должна подчиняться мера неопределенности, связанная с высказыванием S_e, при предположении, что дано S_h. Речь идет о тех условиях, из которых последнее опять имеет два альтернативных вида, соответствующих (I) и (II).

(U 1) $U(e/h) \geqq 0$,

(U 2) если $S_h \underset{L}{\rightarrow} S_e$, тогда $U(e/h) = 0$,

(U 3) если S_h — тавтологическое высказывание, тогда $U(e/h) = U(e)$.

Интуитивный смысл этих условий совершенно ясен. В отлисие от первых трех условий четвертое — записывается альтернативно:

(U 4)$^{\mathrm{I}}$ если $S_h \underset{L}{\rightarrow} \sim S_e$, то $U(e/h)$ максимально,

(U 4)$^{\mathrm{II}}$ если $\sim S_h \underset{L}{\rightarrow} S_e$, то $U(e/h) = U(e)$.

Необходимо подчеркнуть, что конкретный вид применяемой меры неопределенности зависит от ряда обстоятельств прагматического характера. Можно рассматривать зависимости характеризованной выше меры неопределенности от тех стандартов семантической информации, о которых мы говорили раньше, или от связанных с ними вероятностных мер.

Й. Пиетаринэн [10] предложил в целом четыре вида определения меры неопределенности, которые он обозначил следующим образом: unc_1, unc_2, unc_3, unc_4.

Из них unc_1 и unc_2 удовлетворяют условиям (U 1), (U 2), (U 3) а $(\mathrm{U}\ 4)^{\mathrm{I}}$,
а unc_3 и unc_4 удовлетворяют условиям (U 1), (U 2), (U 3) а $(\mathrm{U}\ 4)^{\mathrm{II}}$:

(1) $$\mathrm{unc}_1\,(e/h) = -\log\,(e/h)$$

в случае условной неопределенности; в случае простой неопределенности предполагается, что S_h — тавтологическое высказывание, т. е. S_t, также имеет место равенство:

$$\mathrm{unc}_1\,(e) = \mathrm{unc}\,(e/t) = -\log p(e)\,.$$

Очевидно, что unc_1 соответствует условной информационной мере, или абсолютной информационной мере.

(2) $$\mathrm{unc}_2\,(e/h) = 1 - p(e/h)\,, \quad \text{или}$$
$$\mathrm{unc}_2\,(e) \quad = 1 - p(e)$$

Очевидно, что unc_2 соответствует условной содержательной мере, или абсолютной содержательной мере.

В случае, когда вместо $(\mathrm{U}\ 4)^{\mathrm{I}}$ исходим из $(\mathrm{U}\ 4)^{\mathrm{II}}$, имеются следующие возможности:

(3) $$\begin{aligned}\mathrm{unc}_3\,(e/h) &= p(h \vee e) - p(e) = \\ &= p(h) - p(h\,.\,e) = \\ &= p(h)\,[1 - p(e/h)]\,,\end{aligned}$$

(4) $$\mathrm{unc}_4\,(e/h) = \log \frac{p(h \vee e)}{p(e)}\,.$$

Оставив в стороне варианты (3) и (4), в которых исходим из $(\boldsymbol{U}\ 4)^{\mathrm{II}}$, в связи с тем, что для этих вариантов трудно подыскать ясную интуитивную интерпретацию, можно ввести меру неопределенности или меру условной неопределенности на основе содержательной меры cont и информационной меры inf следующим образом:

$$\boldsymbol{U}(e) \quad =_{\mathrm{df}} \mathrm{cont}\,(e) \quad = 1 - p(e)\,,$$
$$\boldsymbol{U}(e) \quad =_{\mathrm{df}} \inf\,(e) \quad = -\log p(e)\,,$$

$$U(e/h) =_{\mathrm{df}} \mathrm{cont}\,(e/h) = \mathrm{cont}\,(e\,.\,h) - \mathrm{cont}\,(h) =$$
$$= \mathrm{cont}\,(h \rightarrow e) = p(h)\,p(\sim e/h)\,,$$
$$U(e/h) =_{\mathrm{df}} \inf\,(e/h) \quad = \inf\,(e\,.\,h) - \inf\,(h) = -\log p(e/h)\,.$$

В связи с тем, что мера систематизационной мощности высказывания S_h по отношению к тому, к чему относится высказывание S_e, определена так

$$E(h/e) =_{\mathrm{df}} \frac{U(e) - U(e/h)}{U(e)}$$

можно в определении компенсировать меру неопределенности U содержательной мерой cont или информативной мерой inf, что приведет нас к соответствующему определению систематизационной мощности. Если исходить из другой альтернативы, т. е. вместо $(U\,4)^{\mathrm{I}}$ использовать $(U\,4)^{\mathrm{II}}$, то получим определение меры систематизационной мощности, которое соответствует unc_3 и unc_4.

Любое практическое приложение этих или аналогично сконструированных мер систематизационной мощности (объяснительной мощности, предсказательной мощности, диагностической мощности и т. д.) номологических высказываний зависит таким образом от возможностей количественного определения соответствующей вероятностной меры или меры неопределенности. Нам кажется, что для разных задач необходимо выбирать различные способы определения этой меры. В случаях, где можно предполагать высокую меру структурной подобности объяснительных и предсказательных рассуждений (как это, например, предполагает классическая дедуктивно — номологическая модель научного объяснения Геммеля и Оппенгайма [4]), в качестве меры можно, например, взять неопределенность степени точности предсказания, взять неопределенность степени точности предсказания, которое возможно благодаря данному номологическому высказыванию. Если речь идет о большом количестве предсказаний, то эта мера может определяться аналогично известной Шенноновской мере множества информации, т. е.

$$U(e) = -\sum_{i=1}^{n} p_i \log p_i(e_i)\,.$$

Если данное номологическое высказывание является исхондим для других типов решения, т. е. для выбора различных альтернатив из множества всех возможных альтернатив, то определение неопределенности можно опереть на установление качества этих решений.

Приложение может касаться самых различных задач. Приведем некоторые абстрактные типы возможных задач, в которых появится значение приведенных мер, прежде всего меры объяснительной мощности.

Если дан определенный класс номологических высказываний $\{S_{h_1}, S_{h_2}, \ldots$ $\ldots, S_{h_n}\}$, которые представляются как возможные систематизационные средства по отношению к S_e, т. е. как средства, которые позволяют объяснить, предсказать, диагностировать и т. д. то, к чему относится S_e, то необходимо выбрать те элементы класса $\{S_{h_1}, S_{h_2}, \ldots, S_{h_n}\}$, которые могут создавать части потенциального экспланандс для S_e, которые могут быть, например, базой для предсказаний или для других типов решения (обычно речь идет о решении объяснительного типа, которое опирается на номологические высказывания), и тогда целесообразно сравнивать систематизационные (объяснительные, предсказательные) мощности отдельных элементов приведенного класса или отдельных подклассов.

Если, например, имеет место соотношение:

$$\boldsymbol{E}(h_i/e) > \boldsymbol{E}(h_j/e)\,,$$

то тогда, естественно, предпочитать $S_{h_i} > S_{h_j}$.

Если соблюдены предпосылки для того, чтобы мера систематизационной мощности была аддитивной, тогда более целесообразно оперировать парой номологических высказываний, чем одним высказыванием, в случае, если имеет место

$$\boldsymbol{E}(h_i\,.\,h_j/e) > \boldsymbol{E}(h_i/e) \quad \text{или}$$

$$\boldsymbol{E}(h_i\,.\,h_j/e) > \boldsymbol{E}(h_j/e)\,.$$

Необходимость оперировать обеими номологическими высказываниями отпадает, если одно из них иррелевантно отношению к S_e, таким образом оно не в силах повысить меру систематизационной мощности другого высказывания. Определение иррелевантности номологического высказывания зависит от вышеприведенных альтернатив (I) и (II). Если в качестве альтернативы выбрать (II), то S_h иррелевантно по отношению к S_e, если

$$\boldsymbol{U}(e/h) = \boldsymbol{U}(e)\,.$$

В этом случае имеет место соотношение:

$$\boldsymbol{E}(h_i\,.\,h_j/e) = \boldsymbol{E}(h_i/e) + \boldsymbol{E}(h_j/e) =$$

$$= \frac{\boldsymbol{U}(e) - \boldsymbol{U}(e/h_i)}{\boldsymbol{U}(e)} + \frac{\boldsymbol{U}(e) - \boldsymbol{U}(e)}{\boldsymbol{U}(e)} = \boldsymbol{E}(h_i/e)\,.$$

Таким образом, мы имеем право заменить пару $\{S_{h_i}, S_{h_j}\}$ лишь первым из обоих элементов.

Тем самым мы коснулись одной из самых элементарных форм редукции элементов класса $\{S_{h_1}, S_{h_2}, \ldots, S_{h_n}\}$. Можно, однако, представить себе и другие формы редукции. Например, часто необходимо определить такой минимальный подкласс класса $\{S_{h_1}, S_{h_2}, \ldots, S_{h_n}\}$, который обеспечивает max $\boldsymbol{E}$ по отношению к S_e. (Естественно, что не исключено, что все элементы приведенного класса иррелевантны по отношению к S_e, так что искомый класс пуст.)

Очень важной практической задачей является установление того, что можно было бы характеризовать как *достаточная систематизационная мощность по отношению* к S_e. Предположим, что наши возможности реализации систематизационных процедур (например, научного объяснения, предсказания, диагноза и т. д.) ограничены. Точно также, требования к качеству решения этих процедур, например, к желаемой точности предсказания, могут быть ограниченными. На основании следующих интуитивных рассуждений можно ввести понятие „достаточная систематизационная мощность":

Систематизационная мощность S_h по отношению к S_e тем больше, чем более снижена условная неопределенность по отношению к тому, к чему относится S_e если известно S_h. (Если $\boldsymbol{U}(e/h) = \boldsymbol{U}(e)$, то $\boldsymbol{E}(h/e)$ минимально, т. е. равно нулю.) Тогда можно исходить из разницы значений $\boldsymbol{U}(e) - \boldsymbol{U}(e/h)$. Если максимально допустимая неопределенность, которую можно, например, интерпретировать как максимальный уровень допустимого среднего риска, как допустимый уровень неточности при предсказании и т. д., имеет величину ε, то это означает, что справедливо отношение:

$$\boldsymbol{U}(e) - \boldsymbol{U}(e/h) \geqq \varepsilon\,.$$

Меру достаточной систематизационной мощности, которую обозначим $\boldsymbol{SE}$, можно характеризовать таким образом:

$$\boldsymbol{E}(h/e) \text{ есть } \boldsymbol{SE}(h/e) \text{ тогда и только тогда, когда}$$

$$\boldsymbol{U}(e) - \boldsymbol{U}(e/h) \geqq \varepsilon(\varepsilon > 0)\,.$$

Примером полезности понятия „достаточной систематизационной мощности" может быть следующее рассуждение: предположим, что необходимо предсказать координаты движущегося тела. Предсказание можно осуществить на основании классической динамики или, более точным образом, на основании релятивистской динамики. Все зависит от того, как велика разница между этими двумя подходами, либо, точнее говоря, как велика разница между результатами обоих подходов, не превышает ли она допустимый уровень неточности. В связи с тем – образно говоря –, что мы не будем использовать логарифмическую линейку там, где нам хватит таблица умножения, и не будем использовать вачислительную машину там, где достаточна логарифмическая линейка, мы можем удовлетвориться достаточной систематизационнойм ощностью. Не будем учитывать, например, преобразования Лоренца, где скорость движения иррелевантна по отношению к скорости c и где совершенно достаточны классические преобразования. Это также означает, что из класса $\{S_{h_1}, S_{h_2}, \ldots, S_{h_n}\}$, который может создавать общий компонент потенциального экспланaнс для S_e, можно опустить не только те элементы, которые иррелевантны по отношению к S_e, но также и другие элементы, неучет которых не снижает общей неопределенности ниже уровня ε.

Это дает нам право характеризовать SE следующим образом: $E(h_i \,.\, h_j, \ldots$ $\ldots, h_n/e)$ является достаточной систематизационной мощностью приведенного класса $\{S_i, S_j, \ldots, S_n\}$ по отношению к S_e тогда и только тогда, когда

$$U(e) - U(e/h_i \,.\, h_j, \ldots, h_n) \geqq \varepsilon\,.$$

Дальнейшее развитие теории семантической информации опирается на уточнение роли обюих высказываний, т. е. прежде всего, нсмологических высказываний в различных процедурах. В то время, как первые попытки создания удовлетворительной меры семантической информации исходили из ближе неспецифицированных высказываний и не обращали внимание на их функции в различных задачах, те подходы, которые были намечены в последней части данной работы, пытаются учесть специфику этих задач, требования, предъявляемые к качеству их решения или некоторые дальнейшие возможные релятивизации.

(Поступила в редакцию 1 октября 1971 г.)

ЛИТЕРАТУРА

[1] Y. Bar-Hillel, R. Carnap: Semantic Information. In: W. Jackson, ed.: Communication Theory, London 1953. (Первоначальная, более широкая версия этой работы датирована 1952 г.)
[2] R. Carnap: Logical Foundations of Probability. Univ. of Chicago Press, 1950.
[3] C. G. Hempel: Aspects of Scientific Explanation. New York 1965.
[4] C. G. Hempel, P. Oppenheim: The Lcgic of Explanation. Philosophy of Science *15* (1948).
[5] R. Hilpinen: On the Information Provided by Observations. In: J. Hintikka, P. Suppes: Information and Inference. D. Reidel Publ. Comp., Dordrecht 1970.
[6] J. Hintikka: Varieties of Information and Scientific Explanation. In: Rootselaar, Staal, eds.: Logic, Methodology and Philosophy of Science. North-Holland Publ. Comp., Amsterdam 1968.
[7] J. Hintikka: On Semantic Information. In: Physics, Logic, and History. Plenum Press, 1970.
[8] C. Cherry: On Human Communication. MIT Press and John Wiley, 1957.
[9] J. Pietarinen, R. Tuomela: On measures of the Explanatory Power of Scientific Theories. Akten des XIV. Kongresses für Phil. Wien. Herder, 1968.
[10] J. Pietarinen: Quantitative Tools for Evaluating Scientific Systemization. In: J. Hintikka, P. Suppes, eds.: Information and Interference. D. Reidel Publ. Comp., Dordrecht 1970.
[11] K. Popper: Logik der Forschung. J. Springer, Wien 1934. III. Aufl. J. C. B. Mohr, Tübingen 1969.
[12] C. E. Shannon, W.Weaver: The Mathematical Theory of Communication. The Univ. of Illinois Press, Urbana 1949.
[13] L. Tondl: Problémy sémantiky. Academia, Praha 1966.
[14] H. Törnebohm: Two Measures of Evidential Strength. In: J. Hintikka, P. Suppes, eds.: Aspects of Inductive Logic. North Holland Publ. Comp., Amsterdam 1966.

K problémům sémantické informace

LADISLAV TONDL

Stať podává přehled a zhodnocení hlavních směrů současné teorie sémantické informace. Autor naznačuje některé možnosti rozvinutí pojmů odvozených z Hintikkova pojmu „přenesené informace".

Prof. Dr. Ladislav Tondl, DrSc.; Ústav pro filosofii a sociologii ČSAV (Институт философии и социологии ЧСАН), Jilská 1, Praha 1.