KYBERNETIKA — ТОМ 11 (1975), ВЫПУСК 1

Проблема информационной синонимии

Ладислав Тондл

В своей програмной работе „К проблеме построения машинного языка для информационной машины" [9] выдвинул известный советский специалист по информатике В. А. Успенский ряд требований, которые должны были бы решимы с помощью информационного языка в автоматизированной информационной системе. Одно из этих требований — это способность формализации и машинного решения проблемы семантического отождествления, т.е. способность решать тождественны ли с точки зрения их значения разным образом высказанные или написанные данные. (Успенский: „отождествление различным образом записанных фактов".) Данная работа представляет собой попытку реализовать это требование, причем понятие „семантически тождественный" объясняется с помощью средств разработанных теорией семантической информации, прежде всего с помощью понятия „перенесенная информация" (трансинформация). По этим причинам в данной работе вводится понятие „информационной синонимии". Так как существующие до сих пор попытки решить проблемы семантической идентификации исходили из возможностей представленных в принципе тождества сформулированном Лейбницем, будут коротко показаны пределы и трудности тех методов, которые исходят из принципа Лейбница. Необходимо будет, однако, специфицировать требования, исходящие из принципа Лейбница, если мы желаем сохранить схему этих требований тоже для понятия „информационной симонимии".

1. ТРАДИЦИОННЫЕ КРИТЕРИИ ТОЖДЕСТВА (ЛЕЙБНИЦА) И ПРОБЛЕМА СЕМАНТИЧЕСКОГО ОТОЖДЕСТВЛЕНИЯ

Общепринятые формулировки критерия Лейбница имеют более общую компетенцию, чем лишь решение семантического отождествления. Этот критерий считает тождественными все сущности, независимо идет ли речь о сущностях лингвистического или нелингвистического характера, у которых все свойства общие. Если $\mathscr{X}$ представляет класс этих сущностей и $\mathscr{P}$ класс предикатов, которые можно считать именами этих свойств, то в системе $\langle \mathscr{X}, \mathscr{P} \rangle$, в которой элементы (переменные) класса $\mathscr{X}$ обозначены малыми латинскими буквами

 $x, y, u_1, u_2, \ldots$ и элементы класса $\mathscr{P}$ обозначены прописными латинскими буквами $P_1, P_2, \ldots$, критерий Лейбница можно записать двояким способом.

(a) Если мы ограничимся одноместными предикатами, то тождественными можно потом считать любые элементы класса $\mathscr{X}$ (в символической записи $I(x, y)$, которые имеют все свойства общие. Иначе говоря, x и y тождественны тогда и только тогда, если все то, что можно высказать о x, можно высказать о y, и наоборот:

$$(1) \qquad (\forall x)(\forall y)(\forall P)\left[Ix, y \equiv (Px \equiv Py)\right].$$

Очевидно, однако, что эта формулировка критерия Лейбница (которую можно, кстати, найти даже в Топиках Аристотеля), ведет к трудностям и проблемам при решении. Уже в классической системе Principia Mathematica Уайтхеда и Рассела можно найти, в связи с определением 13.01 и формулой 13.101, которые соответствуют (1), замечание, что тождество может иметь разные степени в зависимости от того, как можно понимать то, что характеризуется Уайтхедом и Расселом как предикативная функция и что можно толковать разным образом. Иначе говоря, проблема состоит в том, как можно толковать то, что в интуитивном понимании выражается словами: „Все то, что можно высказать о ...". Также по этим причинам (1) неадэкватно, так как выражение „все то, что можно высказать о ..." можно толковать шире, чем класс всех одноместных предикатов.

(б) Если мы оставим в стороне указанные ограничения, т.е. если выражение „все то, что можно высказать о ..." понимать в том смысле, что оно включает также отношения к другим элементам класса $\mathscr{X}$, потом в соответствии с этим (1) надо изменить (причем индекс вправо наверху в круглых скобках обозначает число аргументов предиката P, например $P^{(i)}$ представляет собой предикат с i аргументами):

$$\begin{aligned}(2) \qquad & (\forall x)(\forall y)(\forall P^{(1)})(\forall P^{(2)})\ldots(\forall P^{(n)})\left[Ix, y \equiv (P^{(1)}x \equiv P^{(1)}y)\,.\right.\\ & .\,(P^{(2)}x, u_1 \equiv P^{(2)}y, u_1)(P^{(3)}x, u_1, u_2 \equiv P^{(3)}y, u_1, u_2)\ldots\\ & \left.\ldots(P^{(n)}x, u_1, u_2, \ldots u_{n-1} \equiv P^{(n)}y, u_1, u_2, \ldots u_{n-1})\right].\end{aligned}$$

Не исключаются, однако, еще и другие формулировки критерия Лейбница, принимающие во внимание выражения высшего типа, или, употребляя терминологию Рассела, функции высшего типа, включая экстенсиональные и интенсиональные функции. Но нельзя сомневаться в том, что таким каким бы то ни было другим приемом формулировки критерия Лейбница становятся еще менее ясными и менее применяемыми в конкретных задачах.*

* Более подробное толкование этих трудностей и изложение проблем финитизма в задачах типа отождествления дается в работе автора [8].

Чтобы любая из приведенных версий критерия Лейбница была применяемой в конкретных задачах, надо провести финитизацию, прежде всего на основе выбора существенных и решающих элементов класса $\mathscr{P}$ и на основе ограничения числа аргументов. Однако, таким образом понятая финитистская формулировка уже не соответствует интуитивным требованиям, которые выражаются словами „все общие свойства", все то, что можно высказать о ..." и т. п. В таком случае необходимо изменить приведенное требование и сформулировать следующее требование: Практически тождественны (почти тождественны и т. п.) такие два элемента класса $\mathscr{X}$ (в символической записи $I'x, y$), которые имеют все решающие свойства общие. Если $\{P_1, P_2, \dots P_n\}$ представляет собой конечный класс этих решающих свойств, модифицированная запись версии (а) имеет следующий вид:

$$(3) \qquad (\forall x)(\forall y)\left[I'x, y \equiv (P_1x \equiv P_1y)(P_2x \equiv P_2y)\dots(P_nx \equiv P_ny)\right].$$

Аналогичным способом можно также модифицировать версию (б) с тем добавлением, что при этом надо иметь в виду конечное число аргументов и тех элементов, отношения которых учитываются во внимание. Так как семантическое отождествление является особым случаем понятия I', возникает здесь необходимость найти все те решающие свойства, которые в состоянии гарантировать семантическое отождествление. Специальным случаем определения этих решающих свойств является критерий Лейбница „salva veritate", к которому мы еще вернемся в дальнейшем изложении.

От понятия I' надо отличить другое понятие, которое можно назвать неразличимостью и которое опирается на известный критерий „identificatio indiscernabilium" (отождествление неразличимых). Два элемента класса $\mathscr{X}$ неразличимы (в символической записи $I''x, y$), если они совпадают по отношению ко всем возможным критериям различения. Критерии различения можно считать компаративными предикатами, т.е. транзитивными, ассиметрическими, C-связными а поэтому и иррефлексивными предикатами [7], которые создают подкласс $\mathscr{A}$ всех двухместных предикатов (т.е. $\mathscr{A} \subset \mathscr{P}^{(2)}$). Если $\mathscr{A} = \{A_1, A_2, \dots, A_n\}$, тогда каждое совпадение надо релативизовать по отношению к известному A_i. Если обозначить совпадение по отнсшению к A_i знаком C_{A_i}, тогда свойства пары $\langle A_i, C_{A_i}\rangle$ можно задать с помощью следующих постулатов:

$$(4) \qquad \begin{gathered}
(\forall x)(\forall y)(\forall z)\left[(A_ix, y \,.\, A_iy, z) \rightarrow A_ix, z\right],\\
(\forall x)(\forall y)\left[C_{A_i}x, y \rightarrow (\sim A_ix, y \vee \sim A_iy, x)\right],\\
(\forall x)(\forall y)(A_ix, y \vee A_iy, x \vee C_{A_i}x, y),\\
(\forall x)(C_{A_i}x, x),\\
(\forall x)(\forall y)(C_{A_i}x, y \rightarrow C_{A_i}y, x),\\
(\forall x)(\forall y)(\forall z)\left[(C_{A_i}x, y \,.\, C_{A_i}y, z) \rightarrow C_{A_i}x, z\right].
\end{gathered}$$

 Затем мы имеем

(5) $$(\forall x)(\forall y)(I''x, y \equiv C_{A_1}x, y \,.\, C_{A_2}x, y \,.\,\ldots\, C_{A_n}x, y)\,.$$

Эта формулировка, которую можно считать определением понятия „неразличимости", представляет собой, собственно, формализацию критерия „identificatio indiscernabilium". О применимости (5) в конкретных задачах можно сказать то, что было сказано о применимости (1) и (2): Чтобы критерий отождествления неразличимых был пригодным для употребления в конкретных задачах, класс критериев различения $\{A_1, A_2, \ldots, A_n\}$ должен быть конечным и овладаемым с помощью тех средств памяти, емкости и средств временного характера, которыми мы располагаем в данной задаче.

2. КРИТЕРИЙ „SALVA VERITATE"

Если же каждая из приведенных версий должна быть употребляема как критерий семантического отождествления, надо уточнить класс $\mathscr{X}$ и, более того, выбрать те свойства или критерии различения, которые являются решающими для семантического отождествления, т.е. выбрать элементы класса $\mathscr{P}$, или же класса $\mathscr{A}$. Такое решение этой спецификации решающих критериев можно найти тоже у Лейбница, где, кстати, тождество скорее является отношением между знаками, чем отношением между обозначаемыми объектами, или же оба аспекта не совсем ясно различимы. Эта спецификация гласит: Eadem sunt quorum unum potest substitui alteri salva veritate. (Тождественны те, которые взаимно заменимы при сохранении истинности.)

Если эта спецификация полностью применима, предполагается, что элементы класса $\mathscr{X}$ – слова или словесные знаки. В этом случае тождественными можно считать те словесные знаки, которые:

(1) взаимно заменимы,

(2) при чем сохраняется истинностное значение предложения, в котором замен осуществляется.

Возможно легко показать, что эта формулировка известная в литературе под названием salva veritate test для решения семантического отождествления слов (синонимии), ведет к антиномиям, которые характеризуются как антиномии синонимичных имен. Эти трудности, которые были одним из импульсов построения логической семантики и семантических основ современной логики и математики, пытались решать Фреге, Рассел, Карнап, Квайн а ряд других пионеров логической семантики.* В целом, главные проблемы, которые здесь выделяются, можно подытожить в следующих вопросах:

* Критический обзор этих различных решений дает работа автора [5].

(а) Какие выражения можно принимать во внимание как выражения, которые семантически отождествляемы (т.е. можно их считать синонимами)?

(б) В каком контексте или в каких условиях возможно принимать во внимание заменимость синонимичных выражений?

(в) Что должно сохраниться при заменимости?

Первоначальная оригинальная версия ответов на эти вопросы, которую можно ожидать у Лейбница, гласит: (а) слова, (б) в контексте предложения, (в) денотация слов и истинностное значение предложения. Фреге открыл трудности этого решения и заменил в ответе (в) денотацию (Bedeutung) смыслом (Sinn) слов. Квайн показал, что даже это решение не удовлетворяет и заменил „salva veritate" Лейбница критерием „salva analycitate", т.е. требованием логической эквивалентности синонимичных слов. Карнап попытался расширить ответ (а) и принять во внимание все выражения, которые могут обозначать (он их назвал десигнатами), т.е. предложения и комплексы предложений, но он столкнулся с другими трудностями. Его решение, которое опирается на понятие „интенсионального изоморфизма", не в состоянии решить проблемы синонимии тех выражений, которые оперируют с т. наз. неэкстенсиональным контекстом, например выражений о взглядах кого-нибудь, о том, что этот человек убежден, что ... и т. п., т.е. выражений, связанных с неэкстенсиональной контекстовой цепью. Сюда включаются высказывания об убеждениях, верах, знаниях, предпочтении и другие выражения неклассических логик.

Чтобы любая удовлетворительная спецификация критерия „salva veritate" была применима для целей, которые имел в виду В. А. Успенский, надо ввести некоторые условности, которые представляют дальнейшую модификацию ответов на вопросы (а), (б), (в):

(а) В качестве выражений, которые мы обоснованно можем принимать во внимание как выражения семантически отождествляемые, мы будем брать все основные формы данных: высказывания (написанные, высказанные или другим способом сообщенные) или классы высказываний, или известным образом упорядоченные последовательности высказываний, т.е. тексты. Это значит, что мы исключаем отдельные слова, не имеющие вид высказываний (например сокращенный ответ на вопрос). Эти основные формы высказываний мы будем называть *данные-объекты* (или кратко объекты), обозначающие класс $\mathscr{X}$ с элементами $x, y, z, \ldots$ (переменные), i, j (константы).

(б) Эти объекты *сообщаются* потребителю и должны служить ему в виде информации для решения данной задачи, для достижения известных целей. При этом информационная ценность этих объектов обсуждается с той точки зрения, насколько эти объекты способны предоставить информацию по отношению к решению данной задачи, т.е. снизить неопределенность потребителя по отношению к решению данной задачи, повысить качество решения данной задачи, понизить потери, связанные с решением, уменьшить средний риск, связанный с решением, или другим образом снизить энтропический уровень данной задачи.

(в) Данные-объекты взаимно заменимы в условиях описанных в (б), если при этом сохраняется и информационная ценность по отношению к тем задачам, решению которых они должны служить.

3. КРИТЕРИЙ „SALVA RELATIONE" И ПОНЯТИЕ „ИНФОРМАЦИОННОЙ СИНОНИМИИ"

В связи с приведенной спецификацией можно сформулировать следующий критерий семантическго отождествления, который исходит из общей схемы понимания Лейбница, который однако „salva veritate" Лейбница и „salva analyticate" Квайна или другие аналогические формулировки заменяет сохранением перенесенной информации или трансинформации (relatio): Eadem sunt quae sibi mutuo substitui possunt salva relatione. Точнее говоря, семантически можно отождествлять те данные-объекты, которые по отношению к известной задаче или к комплексу задач взаимно заменимы, при чем сохраняется перенесенная информация, которую эти объекты предоставляют по отношению к известной задаче или комплексу задач. На основе так формулированного критерия можно ввести понятие „*информационной синонимии*". Прежде чем перейти от интуитивных рассуждений к попытке более точного определения, мы покажем некоторые трудности и возможные возражения, которые могут возникнуть против понятия „информационной синонимии".

Прежде всего надо подчеркнуть, что не имеется в виду синонимия в абсолютном смысле слова, как это принято в лингвистически понимаемых концепциях синонимии, на основе которых синонимичны два выражения, если они имеют общие семантические характеристики, прежде всего денотацию, смысл, заменимость в любом контексте и т.п. Понятие „информационной синонимии" всегда *относительно*, так как оно опирается на понятие „перенесенной информации", которое всегда относительно по отношению к данной априорной информации.

Другая трудность более серъезна: Если мы говорим, что два объекта информационно синонимичные, если они по отношению к данной задаче или комплексу задач взаимно заменимы, причем сохраняется мера перенесенной информации по отношению к данной задаче, тогда информационно синонимичны любые два лингвистически и семантически (в каком угодно понимании) различные, отличающиеся объекта, которые по отношению к данной задаче не переносят никакой информации. Чтобы возможно было преодолеть эту трудность, надо ввести условно определенную меру *достаточной перенесенной информации*, которую можно считать минимальной мерой перенесенной информации по отношению к данной задаче; мера перенесенной информации ни в коем случае не может быть ниже. Без меры достаточной перенесенной информации теряет понятие ,,информационной синонимии" свое обоснование. Достаточной перенесенной информацией можно считать такую информацию, которая в состоянии обеспечить решение данной задачи в желаемом качестве, устранить удовлетворительным способом первоначальную неопределенность или незнание, ответить удовлетворительным способом на заданные вопросы и т.п.

После этого изложения некоторых интуитивных исходных пунктов понятия ,,информационной синонимии" можно перейти к собственному определению этого понятия: Пусть x и y два разных объекта, которые потребителю должны перенести информацию по отношению к задаче или комплексу задач, которые можно описать или выразить с помощью даты или класса данных i. Объект i мы будем называть априорной базой данных для данной задачи или класса задач. Априорную базу данных представляют собой например известные инструкции решения данной задачи, известные гипотезы, алгорифмы или ранее приобретенные эмпирические данные и т.п. Объект i может представлять собой тоже базу данных при постановке вопросов, причем x является возможным ответом на вопрос. Задачей x по отношению к априорной базе i, которые подаются в какой угодно интерпретации, снизить или устранить первоначальную неопределенность, связанную с i. Потом то, что переносит x по отношению к тому, чего касается i, можно выразить с помощью понятия ,,перенесенной информации": Перенесенная информация выражает то, насколько новый объект x (или y) в состоянии снизить первоначальную неопределенность, связанную с i. Если мы обозначим первоначальную неопределенность, связанную с i, $\boldsymbol{U}(i)$ и условную неопределенность, связанную с i, располагая сверх того объектом x, $\boldsymbol{U}(i/x)$, потом перенесенная информация является функцией первоначальной и условной неопределенности и растет вместе с величиной разницы первоначальной и условной неопределенности, т.е. разницы

$$\boldsymbol{U}(i) - \boldsymbol{U}(i/x)\,. \tag{6}$$

В области вероятностных и индуктивных логик принято калькулировать меру первоначальной и условной неопределенности с помощью логарифмической меры inf с базой 2 или с т. наз. содержательной мерой cont определнных в известной работе Карнапа и Бар-Хиллела о теории семантической информации [1],* Однако, этим не исключаются другие возможности калькуляции мер неопределенности. Найболее элементарными способами калькуляции перенесенной информации, которые исходят из приведенных мер inf и cont, можно считать следующие способы:

(I) Если исходить из логарифмической меры inf, потом информация перенесенная объектом x по отношению к тому, к чему относится база данных i (в сим-

* Обзор некоторых альтернативных возможностей калькуляции мер неопределенности дает работа автора [6].

 волической записи $TI(x/i)$) может быть определена так:

$$(7) \qquad TI(x/i) = \text{transinf}(x/i) = \inf(i) - \inf(i/x) = \\ = \log p(i/x) - \log p(i).$$

В этом случае имеет место:

$$(8) \qquad U(i) = -\log p(i),$$

$$(9) \qquad U(i/x) = -\log p(i/x).$$

(II) Если исходить из содержательной меры cont. потом информация, перенесенная x по отношению к тому, к чему относится i, может быть определенна следующим образом:

$$(10) \qquad TI(x/i) = \text{transcont}(x/i) = \text{cont}(i) - \text{cont}(i \to x) = \\ = 1 - p(x \vee i).$$

В этом случае имеет место:

$$(11) \qquad U(i) = 1 - p(i),$$

$$(12) \qquad U(i/x) = p(x \vee i) - p(i) = p(x) - p(x \,.\, i).$$

Обе приведенные меры перенесенной информации* симметричны по отношению к x и i. Эти меры возможно нормализировать, чем устраняется их симметричность. Если обозначить нормализированную меру перенесенной информации $TI_i(x/i)$, тогда на основе концепций (I) и (II) получается:**

$$(13) \quad \text{(I)} \quad TI_i(x/i) = \text{transinf}_i(x/i) = \frac{\inf(i) - \inf(i/x)}{\inf(i)} = \frac{\log p(i) - \log p(i/x)}{\log p(i)},$$

$$(14) \quad \text{(II)} \quad TI_i(x/i) = \text{transcont}_i(x/i) = \frac{\text{cont}(i) - \text{cont}(i \to x)}{\text{cont}(i)} = \\ = \frac{1 - p(x \vee i)}{1 - p(i)} = \frac{[1 - p(x)][1 - p(i/{\sim}x)]}{1 - p(i)} = p({\sim}x/{\sim}i).$$

Возможно доказать, что

$$(15) \qquad -\infty \leqq \text{transinf}_i(x/i) \leqq 1$$

* Эти меры соответствуют формулам 6.7 и 6.8 в [4].

** Эти нормализированные меры перенесенной информации соответствуют формулам 6.10 и 6.11 в [4]. Сверх того (14) соответсвует мере систематизационной мощности Гемпеля [3].

причем $\text{transinf}_i(x/i) = \max \text{transinf}_i = 1$ тогда и только тогда, если i логически вытекает из x, т.е. если имеет место, что $x \underset{L}{\rightarrow} i$, и $\text{transinf}_i(x/i) = \min \text{transinf}_i = -\infty$ тогда и только тогда, если отрицание i логически вытекает из x, т.е. если имеет место, что $x \underset{L}{\rightarrow} \sim i$.

Что касается предельных значений $\text{transcont}_i(x/i)$, имеет место, что

$$(16) \qquad 0 \leqq \text{transcont}_i(x/i) \leqq 1$$

причем $\text{transcont}_i(x/i) = \max \text{transcont}_i = 1$ тогда и только тогда, если i логически вытекает из x, т.е. если $x \underset{L}{\rightarrow} i$, и $\text{transcont}_i(x/i) = \min \text{transcont}_i = 0$ тогда и только тогда, если i вытекает из отрицания x, т.е. если $\sim x \underset{L}{\rightarrow} i$.

Мы привели пока только наиболее важные свойства двух мер перенесенной информации, исходящих из логарифмической меры inf и содержательной меры cont. Из других свойств этих мер необходимо обратить внимание на условия аддитивности и условия для информационной иррелевантности. Что касается условий аддитивности перенесенной информации двух объектов x и x' (связанных с их конъюнктивным соединением), является действительным, употребляем ли концепцию (I):

(17) $\text{transinf}_i(x \,.\, x'/i) = \text{transinf}_i(x/i) + \text{transinf}_i(x'/i)$,
если выполнены следующие условия:

(а) $p(x \,.\, x') = p(x)\, p(x')$, т.е. независимость x и x' по отношению к вероятностной мере p_i;

(б) $p(x \,.\, x'/i) = p(x/i)\, p(x'/i)$, т.е. условная независимость x и x' по отношению к i, имея в виду вероятностную меру p.

Если применять концепцию (II), является действительным следующее условие аддитивности:

(18) $\text{transcont}_i(x \,.\, x'/i) = \text{transcont}_i(x/i) + \text{transcont}_i(x'/i)$,
если выполнено следующее условие: дизъюнкция $x \vee x'$ является логически истинной (всегда истинной).

Мы приведем еще условия для информационной иррелевантности (связанные опять с конъюнктивным соединением):

(19) $\text{transinf}_i(x \,.\, x'/i) = \text{transinf}_i(x/i)$, т.е. x' является информационно иррелевантным по отношению к $\text{transinf}_i(x/i)$, если выполнены следующие условия:

(а) $p(x \,.\, x') = p(x)\, p(x')$

(б) $p(x \,.\, x'/i) = p(x/i)\, p(x'/i)$

(в) $p(x' \,.\, i) = p(x')\, p(i)$ т.е. сверх того еще независимость x' и i по отношению к вероятностной мере p.

 Для $\mathrm{transcont}_i$ имеет место следующее условие информационной иррелевантности:

(20) $\mathrm{transcont}_i(x \,.\, x'/i) = \mathrm{transcont}_i(x/i)$, если выплонено условие, что дизъюнкция $x' \vee i$ является логически истинной.

Этот обзор важнейших свойств приведенных мер перенесенной информации ясно показывает некоторые преимущества (II). По этим причинам мы будем в избранных примерах калькуляции пользоваться концепцией (II).

К определению понятия „информационной синонимии" нам нужно ввести еще понятие „достаточной перенесенной информации". Объект x в состоянии предоставлять достаточную информацию по отношению к i, если он в состоянии гарантировать решение данной задачи в желаемом качестве, отстранить удовлетворительным образом первоначальную неопределенность, гарантировать решение с риском, который не будет выше допустимого уровня, гарантировать удовлетворительным способом ответы на заданные вопросы и т.п. Перенесенная информация, предоставляемая объектом x по отношению к тому, к чему относится i, является достаточной, если имеет место:

(21) $$\boldsymbol{TI}_i(x/i) \geqq \varepsilon$$

причем ε ($\varepsilon \leqq 1$) условно определенная величина меры перенесенной информации по отношению к i. Тогда действительно то, что два (отличные) объекта информационно синонимичные по отношению к i, тогда и только тогда, если они переносят по отношению к i эту же меру информации и если эта мера больше или равна ε. Если мы понятие „информационной синонимии" двух объектов по отношению к i обозначим знаком $S_i x, y$, мы можем определить:

(22) $$S_i x, y =_{\mathrm{df}} [\boldsymbol{TI}_i(x/i) = \boldsymbol{TI}_i(x/i)] \,.\, [\boldsymbol{TI}_i(x/i) \geqq \varepsilon] \,.$$

Это определение выражает ранее приведенные условия: Два объекта x и y информационно синонимичны по отношению к базе данных i тогда и только тогда, если они взаимно заменимы при сохранении меры перенесенной информации и эта мера больше или равна достаточной перенесенной информации.

Базу данных, по отношению к которой информационная синонимия принимается во внимание, нет необходимости понимать как единственный объект, но можно ее понимать как конечный класс данных — объектов $\mathscr{I}$ ($\mathscr{I} = \{i_1, i_2, \ldots, i_n\}$). Речь идет например о проблемной обстановке, решение которой предполагает отдельные шаги, о тексте, который представляет собой упорядоченную последовательность отдельных высказываний и т.п. Если невозможно калькулировать меру перенесенной информации по отношению к $\mathscr{I}$ в целом, а лишь по отношению к отдельным элементам $\mathscr{I}$, тогда информационную синонимию объектов x и y по отношению к $\mathscr{I}$ можно определить таким образом:

$$(23)\quad S_{\mathscr{I}}x, y =_{\text{df}} \{[TI_{i_1}(x/i_1) = TI_{i_1}(y/i_1)] \,.\, [TI_{i_2}(x/i_2) = TI_{i_2}(y/i_2)] \,.\, \ldots$$
$$.\, [TI_{i_n}(x/i_n) = TI_{i_n}(y/i_n)]\} \text{ и для всех } i_i\,(i_i \in \mathscr{I}) \text{ имеет место, что } TI_{i_i}(x/i_i) \geqq \varepsilon.$$

Следовательно, приведенные определения информационной синонимии релативизуют понятие „информационной синонимии", а то с одной стороны по отношению к базе данных i, с другой стороны по отношению к пределам достаточной перенесенной информации. Это значит, что два объекта x и y, которые синонимичные по отношению к i, могут не быть синонимичны по отношению к другой базе данных j. К нарушению существующей синонимии может дойти тогда, если изменятся требования на пределы достаточной перенесенной информации.

Так как обе концепции опираются на применение пробабилистических параметров на данные объекты (причем мы в общем изложении не решаем вопрос о конкретной интерпретации этих параметров, например в виде частотной интерпретации вероятности, в виде субъективной вероятности, меры ожидания, меры достоверности и т.п.), надо показать, какие есть предположения для $S_i x, y$, если мы их выразим с помощью этих параметров. Если мы исходим из концепции (I), т.е. из логарифмической информационной меры, то получим, что два объекта x и y информационно синонимичны по отношению к i тогда и только тогда, если имеет место

(24) (a) $$p(i/x) = p(i/y) \quad \text{и}$$
(б) $$\log p(i/x) \leqq \log p(i)\,(1 - \varepsilon)\,.$$

Если мы исходим из концепции (II), т.е. из меры основанной на понятии cont, мы получим, что два объекта x и y информационно синонимичны по отношению к i тогда и только тогда, если имеет место

(25) (a) $$p(i \vee x) = p(i \vee y) \quad \text{и}$$
(б) $$p(i \vee x) \leqq 1 - \varepsilon[1 - p(i)]\,.$$

4. ИНФОРМАЦИОННАЯ РЕЛЕВАНТНОСТЬ И ПОНЯТИЕ „СТРОГОЙ ИНФОРМАЦИОННОЙ СИНОНИМИИ"

Понятие „информационной синонимии", которое основано на принципе той же меры перенесенной информации, которая больше или равна достаточной перенесенной информации, является в известном смысле слабее чем критерий Квайна „salva analycitate". Это значит, что из $S_i x, y$ никоим образом не вытекает логическая эквивалентность x и y. Единственно, что нам гарантирует $S_i x, y$, представляет собой в случае применения концепции (I) равенство условных вероятностей $p(i/x) = p(i/y)$ и в случае применения концепции (II) равенство $p(i \vee x) = p(i \vee y)$.

Условия информационной синонимии можно сделать более строгими, если мы принимаем во внимание информационную релевантность объекта x по отношению к y и к базе данных i: Строго информационно синонимичными можно считать два объекта x и y по отношению к i тогда и только тогда, если информация, передаваемая с помощью x по отношению к i равна информации, передаваемой с помощью y по отношению к i, и если сверх того аггрегация y к x или наоборот не повысит или не понизит уровень этой перенесенной информации. Иначе говоря, строго информационно синонимичными (в символической записи $Syn_i\, x, y$) по отношению к i мы считаем два объекта x и y, если они информационно синонимичны и если аггрегация одного к другому или наоборот не повысит или не понизит меру перенесенной информации по отношению к i, т.е.

(26) $$Syn_i x, y =_{\mathrm{df}} [TI_i(x/i) = TI_i(y/i) = TI_i(x \,.\, y/i)] \,.\, [TI_i(x/i) \geqq \varepsilon]\,.$$

Понятие „строгой информационной синонимии" по отношению к базе данных i предполагает, что объект y информационно синонимичен с x и одновременно информационно иррелевантен по отношению к $TI_i(x/i)$. Ранее приведенные условия в понятии „строгой информационной синонимии" изменяются так что в (24) (a) заменяется

(27) (a′) $$p(i/x) = p(i/y) = p(i/x \,.\, y)$$

и (a) в (25) заменяется

(28) (a′) $$p(i \vee x) = p(i \vee y) = p[i \vee (x \,.\, y)]\,.$$

Понятие „строгой информационной синонимии" двух объектов x и y по отношению к i также не гарантирует выполнение критерия Квайна „*salva analycitate*", т.е. логическую эквивалентность x и y. Одновременно надо подчеркнуть, что также логическая эквивалентность x и y не гарантирует, что x и y информационно синонимичны или строго информационно синонимичны по отношению к i в том случае, если имеет место

(29) $$TI_i(x/i) = TI_i(y/i) \quad \text{и одновременно}$$
$$TI_i(x/i) < \varepsilon\,.$$

Приведенные свойства можно показать на схематическом примере, в котором калькуляция TI_i основывается на концепции (II). Логическое пространство (в смысле Карнапа [2]) изображено с помощью квадрата, разделенного на 16 одинаковых частей. Экстензия отдельных объектов изображена с помощью номеров отдельных частей:

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	16

i ... 4, 8, 12, 16
x_1 ... 4, 7, 8, 11, 12
x_2 ... 7, 8, 11, 12, 16
$x_3(\equiv x_1 . x_2)$ ... 7, 8, 11, 12
$x_4(\equiv x_1 \vee x_2)$ ... 4, 7, 8, 11, 12, 16
x_5 ... 3, 4, 7, 8
x_6 ... 1, 2, 3, 4
x_7 ... 1, 2, 3, 4

$$\begin{array}{ll}
p(i) = \frac{1}{4} & \\
p(x_1) = \frac{5}{16} & p(i/x_1) = \frac{3}{5} \\
p(x_2) = \frac{5}{16} & p(i/x_2) = \frac{3}{5} \\
p(x_3) = \frac{1}{4} & p(i/x_3) = \frac{1}{2} \\
p(x_4) = \frac{3}{8} & p(i/x_4) = \frac{2}{3} \\
p(x_5) = \frac{1}{4} & p(i/x_5) = \frac{1}{2} \\
p(x_6) = \frac{1}{4} & p(i/x_6) = \frac{1}{4} \\
p(x_7) = \frac{1}{4} & p(i/x_7) = \frac{1}{4}
\end{array}$$

Если $\varepsilon = 0{,}8$, потом следует, что

$$TI_i(x_1/i) = TI_i(x_2/i) = TI_i(x_3/i) = TI_i(x_4/i) = TI_i(x_5/i) = \frac{5}{6}.$$

Это значит, что объекты x_1, x_2, x_3, x_4, x_5 информационно синонимичны по отношению к i. Если мы употребляем понятие „строгой информационной синонимии", тогда строго информационно синонимичны объекты x_1, x_2, x_3, x_4. В отличии от этого объекты x_6 и x_7, хотя они эквивалентны, не являются информационно синонимичными по отношению к i, так как $TI_i(x_6/i) < \varepsilon$. По той же причине они не являются строго инфомационно синонимичными.

Автор выражает свою благодарность профессору О. Зиху за ценные замечания.

(Поступило в редакцию 12 июня 1974 г.)

[1] Y. Bar-Hillel, R. Carnap: Semantic Information. In: W. Jackson (ed.): Communication Theory. London 1953.

[2] R. Carnap: Logical Foundations of Probability. The Univ. of Chicago Press, Chicago 1950.

[3] C. G. Hempel: Aspects of Scientific Explanation. Free Press, New York 1965.

[4] I. Niiniluoto and R. Tuomela: Theoretical Concepts and Hypothetical-Inductive Inference. D. Reidel, Dordrecht—Boston 1973.

[5] L. Tondl: Problémy sémantiky. Academia, Praha 1966.

[6] Л. Тондл: К проблемам семантической информации. Kybernetika *8* (1972), 3, 189—212.

[7] L. Tondl: Prerequisities for Quantification in the Empirical Sciences. Theory and Decision *2* (1972), 238—261.

[8] L. Tondl: Scientific Procedures. D. Reidel, Dordrecht—Boston 1973.

[9] В. А. Успенский: К проблеме построения машинного языка для информационной машины. Проблемы кибернетики, вып. 2. Под. ред. А. А. Ляпунова. Москва 1959.

Prof. Dr. Ladislav Tondl DrSc., Heřmanova 8, 170 00 Praha 7. ČSSR.