



УДК 517. 977. 52

О ВОЗМОЖНОСТИ ОРГАНИЗАЦИИ ЗНАНИЙ НА ОСНОВЕ КОГНИТИВНОЙ СЕМАНТИКИ

Кузнецов О.П.

*Институт проблем управления им. В.А.Трапезникова РАН,
г. Москва, Россия*

olkuznes@ipu.ru

В работе рассмотрены основные концепции когнитивной семантики, изложенные в книге [Lakoff, 1987]. В их основе лежит тезис, заключающийся в том, что когнитивные структуры и механизмы человека существенным образом зависят от его сенсорных механизмов, а также физического и социального опыта. Обсуждаются возможности использования этих концепций для решения задач искусственного интеллекта.

Ключевые слова: когнитивная семантика, категоризация, концептуальные структуры

ВВЕДЕНИЕ

Искусственный интеллект (ИИ), возникший в 50-х годах прошлого века как самостоятельная ветвь компьютерных наук, ставил в начале своего развития весьма амбициозные цели – моделирование интеллектуальных процессов человеческого мозга. На фоне царившей тогда романтической веры в беспредельные возможности компьютеров первые же успехи в автоматизации интеллектуальных процессов (решатель задач Ньюэлла - Саймона, подходы к распознаванию образов, моделирование некоторых простых игр) привели к формированию ведущей парадигмы ИИ, которую можно назвать компьютерной парадигмой. Она исходила из предположения, что мозг с информационной точки зрения - это большой и сложный компьютер, и поэтому интеллектуальные процессы нужно реализовать на принципах работы с информацией в компьютере, т.е. на основе символьных представлений информации и их обработки с помощью алгоритмических логико-комбинаторных процедур.

Довольно быстро стало понятно, что для решения многих интеллектуальных задач вовсе не обязательно имитировать процессы мозга. Высокие (и постоянно растущие) скорости компьютерных вычислений позволили использовать сложные последовательные алгоритмы, не характерные для мозга. Структурное несходство ИИ-процессов с процессами мозга рассматривалось как залог успеха. Ведь человек начал летать, только когда перестал пытаться имитировать полет птиц; а

вместо того, чтобы автоматизировать метлу, придумал пылесос. И действительно, в основе практически всех достижений ИИ последних десятилетий, воплощенных в интеллектуальных технологиях (экспертных системах, технологиях поддержки принятия решений, системах интеллектуального анализа данных) лежат логико-комбинаторные алгоритмы, как правило, имеющие мало общего с процессами мозга. Соответственно, компьютерная парадигма приобрела несколько другой вид: «неважно, что между процессами мозга и компьютера имеются существенные различия; важно, что все интеллектуальные задачи можно решать на компьютере».

Однако со временем стало очевидно, что не все интеллектуальные процессы поддаются компьютерному моделированию. Оставляя в стороне такие очевидные явления, как творческие процессы, озарение и интуицию, которые в компьютере не моделируются вообще, рассмотрим процессы, которые в ИИ моделируются. Эти процессы у человека и компьютера тоже различаются. Человек не только быстро распознает; он еще быстро рассуждает и быстро принимает решения. Эти скорости сравнимы с компьютерными - при том, что скорость передачи сигналов в нервных сетях в миллион раз меньше электронной. Это говорит не только о том, что информационные процессы мозга устроены не так, как компьютерные (что само по себе не слишком волнует сторонников компьютерной парадигмы), но прежде всего о том, что некоторые интеллектуальные задачи человек решает, как минимум, в миллион раз эффективнее компьютера. Несколько упрощенно это выражается

формулой «То, что сложно компьютеру – просто человеку и наоборот, то, что сложно человеку, просто компьютеру». Для задачи понимания текста об этом сказано в [Пинкер, 2004]: «Часть, относящаяся к памяти, легка для компьютеров и тяжела для людей (оперативная память – основное узкое место при переработке информации человеком), а часть, относящаяся к принятию решений, легка для людей (по крайней мере если предложение правильно построено) и тяжела для компьютеров».

Известен ряд попыток (или хотя бы проектов) построения нестандартных моделей мыслительных процессов и понятийной системы человека (например, [Sowa, 1984], [Hofstadter et al., 1994], [Кузнецов, 1995], [Голицын и др., 1996]). К сожалению, они либо локальны (т.е. способны объяснить только незначительную часть процессов мозга), либо неконструктивны, т.е. остаются на уровне нереализованных идей. Здесь мы хотим обратить внимание, на еще одно перспективное направление когнитивных наук, которое началось с исследований Э.Росха [Rosch, 1975] и наиболее полно изложено в книге Лакоффа [Lakoff, 1987]. Нас будет интересовать не вопрос адекватности теорий этого направления реальным процессам человеческого мышления, а возможности использования этих теорий в интеллектуальных технологиях.

1. Когнитивная категоризация

Концепция Лакоффа представляет собой проект решения двух проблем: проблемы категоризации и проблемы семантики. Предлагаемый подход к этим проблемам можно назвать когнитивным. В его основе лежит тезис, заключающийся в том, что когнитивные структуры и механизмы человека существенным образом зависят от его сенсорных механизмов, а также физического и социального опыта. Когнитивный подход резко противопоставляется традиционному подходу, который Лакофф называет объективизмом и который предполагает, что когнитивные структуры человека подобны структурам научного знания и в конечном счете – структурам объективного мира, описываемого этим знанием.

Проблема категоризации – это проблема формирования понятий (категорий) и структурирования понятийной системы человека. Традиционная формальная теория понятий рассматривает понятие как класс объектов, обладающих одинаковым набором признаков. Этими объектами могут быть либо подклассы, т.е. более узкие понятия, либо единичные объекты – экземпляры. С этой точки зрения все объекты класса равноправны, т.е. любой объект класса в равной мере может служить его представителем (примером). Иерархия понятий строится от элементарных объектов (экземпляров) к классам, которые в свою очередь являются объектами более общих классов и т.д. Понятия любой степени общности имеют одну и ту же древовидную

структуру – за исключением элементарных объектов, которые наиболее просты и структуры не имеют. Росх показала, что человеческая категоризация устроена не так. В категориях человека существуют прототипы – «хорошие» (центральные, репрезентативные, типичные) примеры. Хорошо известны стандартные ответы: поэт – Пушкин, птица – воробей (или ворона), фрукт – яблоко. Это говорит о том, что человеческие категории имеют внутреннюю структуру, которая для разных видов категорий может быть различной. В человеческой иерархии понятий базовые понятия, которые когнитивно наиболее просты, находятся «в середине» иерархии общего-конкретного. Обобщение происходит вверх от базового уровня, специализация – вниз. Примерами места базовых категорий в концептуальной иерархии могут служить цепочки: животное-**собака**-овчарка; мебель – **стул** – качалка, где базовыми категориями являются «собака» и «стул».

Для категорий базового уровня характерно следующее:

- они имеют единый ментальный образ (гештальт); быстро узнаются;
- в качестве их имен используются наиболее короткие и общепотребительные слова;
- большинство признаков членов категории хранится на этом уровне;
- формирование категорий у детей начинается с категорий базового уровня.

Возможные структуры категорий весьма разнообразны и не исчерпываются множествами элементов с одинаковым набором признаков. В книге Лакоффа развита довольно детальная (хотя не вполне формальная) типология категорий, многие из которых имеют нестандартную структуру. Таковы, например, радиальные категории, имеющие структуру «центр-периферия». В этой структуре центром является некоторая наиболее представительная субкатегория, а периферию образуют ее расширения. Пример – категория «мать», где центр – субкатегория с традиционными признаками матери (родившей и воспитавшей ребенка), а периферия – субкатегории «приемная мать», «суррогатная мать» и т.д. Передача признаков от центра к расширениям в общем случае нетранзитивна, и может оказаться, что какая-то периферийная субкатегория не обладает многими существенными признаками центра. Прототипом в таких категориях служит центр. Другой вид категорий с прототипическими эффектами – градуированные категории с различными степенями членства. Такие категории удобно описывать в терминах функций принадлежности, известных в нечеткой логике.

(Интересно отметить, что идея центральности – одно из важных понятий гештальт-психологии. «Центрирование – то, как мы рассматриваем части, отдельные элементы ситуации, их значение и роль

по отношению к центру, сути или корню, - является наиболее важным фактором в мышлении» [Вертгеймер, 1987]. Существенность или несущественность признаков для гештальт-психологов связана не с их внутренними свойствами, а с их ролью в структуре образа. Если какой-то несущественный в данном образе признак объявить (или «увидеть») существенным, то образ может стать «другим» - изменится его интерпретация. Этот эффект в гештальт-психологии называется перецентрированием.)

Среди механизмов, формирующих категории и их организацию в понятийные структуры, можно выделить 4 типа: *классический*, организующий категории по принципам, близким к традиционным классификациям; *образно-схематический*, в основе которого лежат схематические образы, такие как траектории, вместилища, верх-низ; *метафорический*, который отображает категории одной области в соответствующие структуры другой области; *метонимический*, в котором часть структуры (субкатегория) замещает или обозначает категорию в целом (например, «Кремль» может пониматься как «Администрация президента», а «Москва» - как обозначение федеральной власти России). Многочисленные примеры образования понятий с помощью метафор имеются в математике: пространства, числовая прямая, линейность, выпуклость, рост (функции), теория графов (пути, связность т.д.). Интересно, что теория бинарных отношений, которые формально изоморфны графам, имеет совершенно другую структуру - именно потому, что в отличие от теории графов, она не имеет общего гештальта.

2. Когнитивная семантика

Еще более радикальным является предлагаемый Лакоффом проект когнитивной семантики. Этот проект принципиально отказывается от общепринятого в настоящее время формально-логического подхода, согласно которому независимо существуют синтаксис, модельные структуры и интерпретация, т.е. принципы отображения синтаксиса на модели. Формально-логический подход, основанный на идеях Гильберта - Тарского, сыграл существенную роль в становлении компьютерной парадигмы, на которой основаны практически все достижения в области интеллектуальных систем (если не считать нейросетевых методов). Однако он неспособен объяснить многие особенности человеческого мышления, которыми мозг отличается от компьютера.

В общих чертах когнитивный подход к семантике заключается в следующем. Термины, в которых мыслит человек, значимы с самого начала. В отличие от компьютера, люди не оперируют незначимыми символами. «Когда человек доказывает утверждение, он опирается на его смысл, не строя формального доказательства. Разумный человек, не изучавший логику, способен

точно выражаться и быстро определять противоречия, но часто не может построить формальное доказательство» [Sowa, 1984]. Более того, по мнению Лакоффа, значения возникают раньше, чем формируются концептуальные структуры: они возникают из нашего доконцептуального телесного опыта. Доконцептуальные структуры - это гештальты и образно-схематические схемы, такие как вместилище, верх-низ, часть-целое, центр-периферия и т.д. Человеческая логика рождается из этих схем. Например, схема «вместилище» порождает рассуждения типа «если вместилище *A* находится во вместилище *B* и *x* находится в *A*, то он находится в *B*», обобщением которых является *Modus Ponens*. Известные в теории множеств диаграммы Венна - фактически гештальты этой схемы.

Концепты (категории), связанные с доконцептуальными структурами, являются непосредственно значимыми. Они-то и являются категориями базового уровня. Все такие категории имеют гештальтную структуру. В этих терминах можно сформулировать одну из важнейших проблем психологии, совершенно чуждую компьютерной парадигме - *что такое понимание*. Предложение понимается непосредственно, если концепты, содержащиеся в нем, непосредственно значимы. Понимание - это способность соотносить концепты со своим опытом, включая доконцептуальный. Абстрактные структуры мы понимаем в терминах образных схем. В частности, иерархическая структура понимается в терминах схем часть-целое и верх-низ, радиальная структура категорий - в терминах центр-периферия, количественная шкала - в терминах *верх-низ* и *линейный порядок*.

Различия концептуальных систем в разных культурах проистекают в основном из различий в доконцептуальном опыте. Например, концептуальная система эскимосов сильно отличается от концептуальных систем горных племен. Именно с различиями концептуальных систем связаны трудности перевода.

В заключение укажем еще на одну проблему, которая у Лакоффа только намечена - проблему когнитивной сложности. Это понятие подразумевает объем когнитивных усилий, которые прилагаются при запоминании и использовании различных концептов. С этой точки зрения наиболее когнитивно просты концепты базового уровня.

3. Когнитивный подход к задачам искусственного интеллекта

Резюмируя, выделим следующие основные идеи и проблемы когнитивной семантики:

- центральная роль гештальта,
- наличие понятий базового уровня,
- нестандартные структуры категорий:

радиальные, метонимические и т.д.,

- доконцептуальное структурирование и базовые структуры опыта;
- проблема понимания;
- проблема когнитивной сложности.

Концепции Лакоффа дают возможность по-новому взглянуть на некоторые проблемы ИИ. В первую очередь следует выделить две важных области ИИ: организация знаний и формализация рассуждений. Для организации знаний и, в частности, построения онтологий представляют существенный интерес идеи, связанные с категоризацией (на это обращается внимание в недавнем докладе [Гаврилова и др., 2011]). Что касается формализации рассуждений, то здесь важно иметь в виду следующее. Человек не рассуждает по законам формальной логики. Такие рассуждения нереализуемы в реальном времени ввиду огромного количества последовательных элементарных шагов. Даже строгие математические доказательства на много порядков короче рассуждений, формализованных в стиле логики предикатов. Еще в книге [Sowa, 1984] отмечалось, что человеческая способность быстро рассуждать (на которую, к сожалению, обращается гораздо меньше внимания, чем на способность быстро узнавать) основана на использовании образно-схематических структур. В книге Лакоффа, по-видимому, впервые развита подробная типология таких схем.

Можно наметить следующие направления исследований в русле изложенных идей:

- формализация типологий когнитивных категорий и организация знаний на их основе;
- исследование и формализация проблемы гештальта и связи гештальтов с понятиями базового уровня; на решение этой проблемы была, в частности, ориентирована модель псевдооптической нейронной сети [Кузнецов, 1996];
- формализация быстрых рассуждений на основе образно-схематических структур;
- формализация понятия когнитивной сложности;
- реализация формализованных моделей на микроуровне, в частности, на уровне различных видов нейронных сетей.

БИБЛИОГРАФИЧЕСКИЙ СПИСОК

[Вертгеймер, 1987] Вертгеймер. М. Продуктивное мышление. / М.: Прогресс, 1987.

[Гаврилова и др., 2011] Гаврилова Т.А., Болотникова Е.С., Гулякина Н.А. 2011. Категоризация знаний для создания онтологий. / Материалы 4-й Всероссийской мультikonференции по проблемам управления МКПУ-2011, Т.1. Таганрог: изд. ТТИ ЮФУ, с. 62-66.

[Голицын и др., 1996] Голицын Г.А., Фоминых И.Б. Нейронные сети и экспертные системы: перспективы интеграции. // Новости искусственного интеллекта, 1996, №4.

[Кузнецов, 1995] Кузнецов О.П. Неклассические парадигмы искусственного интеллекта. //Теория и системы управления, 1995, N5, с. 3-23.

[Кузнецов, 1996] Кузнецов О.П. Псевдооптические нейронные сети - прямолинейные модели. //Автоматика и телемеханика, N12, с.160-172.

[Пинкер] Пинкер С. Язык как инстинкт. / М.: Едиториал УРСС, 2004.

[Hofstadter et al., 1994] Hofstadter D and Fluid Analogies Research Group. Fluid concepts and creative analogies: computer models of fundamental mechanisms of thought. - Basic Books, 1994.

[Lakoff, 1987] Lakoff J. Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. - University of Chicago Press. 1987. (Русский перевод: Лакофф Д. Женщины, огонь и опасные вещи: Что категории языка говорят нам о мышлении, - М. 2004).

[Rosch, 1975] Rosch E. Cognitive representations of semantic categories. Journal of Experimental Psychology, 1975, 104, pp.192-233.

[Sowa, 1984] Sowa J.F. Conceptual Structures - Information Processing in Mind and Machines. - Addison-Wesley Publ.Comp. 1984.

THE POSSIBILITY OF KNOWLEDGE STRUCTURING BASED ON COGNITIVE SEMANTICS

Kuznetsov O. P.

*Trapeznikov Institute of Control Sciences, Russian
Academy of Sciences, Profsoyuznaya ul., 65,
Moscow, 117997 Russia*
olkuznes@ipu.ru

The essential conceptions of cognitive semantics [Lakoff, 1987] are considered. These conceptions based on the thesis that cognitive structures and mechanisms of a man depend on his sensor mechanisms, physical and social experience. The possibilities of applications of these conceptions for artificial intelligence problems are discussed.

Keywords: cognitive semantics, categorization, conceptual structures.