

Когнитивная семантика и искусственный интеллект

Аннотация. В работе рассмотрены основные концепции когнитивной семантики, изложенные в [14]. 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, концептуальная система, понимание, когнитивная сложность.

Введение

Искусственный интеллект (ИИ) возник в 50-х годах прошлого века как самостоятельная ветвь компьютерных наук. В начале своего развития он ставил весьма амбициозные цели – моделирование интеллектуальных процессов человеческого мозга. Первые же успехи в автоматизации интеллектуальных процессов (решатель задач Ньюэлла - Саймона, моделирование некоторых эвристических процедур и простых игр) привели к формированию ведущей парадигмы ИИ, которую можно назвать компьютерной парадигмой. В основе этой парадигмы лежит предположение, что мозг с информационной точки зрения - это большой и сложный биокомпьютер, 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 совсем не обязательно имитировать процессы мозга. Высокие и постоянно растущие скорости компьютерных вычислений, превосходящие скорости сигналов в мозге примерно в миллион раз, позволили использовать сложные последовательные алгоритмы, не характерные для мозга. В основе практически всех достижений ИИ последних десятилетий, реализованных

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 этот тезис становился все менее убедительным. Во-первых, известны интеллектуальные процессы, которые не поддаются компьютерному моделированию. К ним относятся творческие процессы, озарение и интуиция, генерация различного рода экспертных суждений и т.д. Во-вторых, многие интеллектуальные задачи человек решает быстрее компьютера, несмотря на то, что скорости передачи сигналов в мозге в миллион раз меньше, чем в компьютере. Наиболее характерны два круга задач, решение которых занимает особое место среди когнитивных процессов. Человек быстро распознает не только отдельные объекты, но и сложные ситуации. Он умеет быстро рассуждать, получая при этом логически верные следствия, притом, что не всегда способен выстроить корректное доказательство.

Специфическая природа человеческих механизмов распознавания имеет общепризнанное объяснение (которое до сих пор остается

декларативным и не имеет конструктивного описания): у человека существует специфическая образная система восприятия, хранения и обработки информации.

«Человеческое сознание, человеческий способ познания мира отличается от компьютера наличием двух параллельных систем познания. Компьютер обладает одной системой познания, а именно, символьной, символьно-логической. Человек же обладает двумя совместно работающими системами познания. Одна из них – это рассудок, интеллект или символьно-логическое мышление, а вторая система – восприятие и образное мышление. Эти две системы существуют у человека как одно, неразрывное целое» [9].

Образную систему человека характеризуют два свойства, не имеющие аналогов в рамках компьютерной парадигмы. Первое свойство – высокая скорость обработки образов, которая особенно поразительна, если учесть отмеченные выше низкие скорости передачи сигналов в мозге. Второе свойство – целостность образных представлений («гештальтов»). Это свойство легло в основу гештальт-психологии – направления, активно развивавшегося в 30-40-х гг. [2]. По-видимому, эти два свойства неразрывно связаны друг с другом, т.е. обеспечиваются единым механизмом.

Гораздо меньше внимания уделялось способности человека быстро рассуждать. Однако некоторые важные соображения по этому поводу время от времени высказывались. В книге [18] обращается внимание на то, что в человеческом мышлении огромную роль играют заготовки-схемы стандартных ситуаций, использование которых существенно ускоряет рассуждения. Например, исследование мыслительных процессов профессиональных шахматистов показывает, что они имеют в своем распоряжении несколько десятков тысяч схем. Известная теория фреймов М. Минского фактически является одним из вариантов теории таких схем, хотя, по-видимому, скорость человеческих рассуждений его не интересовала.

Различия в эффективности компьютерных и человеческих механизмов решения интеллектуальных задач можно упрощенно выразить формулой «то, что просто человеку, сложно компьютеру и наоборот» [5]. Для задачи понимания текста об этом сказано в [7]: «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». Все это говорит о том, что исследования когнитивных структур и механизмов человека важны не только для тех когнитивных наук, для которых процессы мозга являются непосредственным объектом изучения (психологии, нейрофизиологии и лингвистики), но и для ИИ, для которого эти исследования могут привести к новым интеллектуальным технологиям. Однако до сих пор не только практически все основные результаты ИИ, но и доминирующие идеи его развития остаются в рамках компьютерной парадигмы. В то же время за последние десятилетия в других когнитивных науках появились, по крайней мере, две примечательные концепции – когнитивная семантика [14] и теория колонок [13, 15], представляющие непосредственный интерес для ИИ, но пока не привлечшие должного внимания специалистов ИИ. Задача настоящей статьи – обсудить концепцию когнитивной семантики и ее перспективы для ИИ.

1. Когнитивная категоризация

Термин «когнитивная семантика» появился в рамках направления «когнитивная лингвистика», сложившегося в лингвистике в конце 70-х – начале 80-х гг. Это направление возникло в результате осознания недостаточности формального подхода к построению модели естественного языка, основанного на теории порождающих грамматик Хомского. Если формальный подход явно или неявно предполагает, что язык может порождаться на основе формальных, алгоритмических процедур и потому инвариантен по отношению к природе порождающего механизма – будь то человек или компьютер, – то основная идея когнитивной лингвистики заключается ровно в обратном: основные свойства естественного языка существенно связаны с телесной природой человека, а также его физическим и социальным опытом. Но если это так, то естественно пойти дальше, полагая, что этот тезис справедлив не только для языка, но для мышления человека.

Об этом довольно ясно сказано в работе [10]: «в процессе построения формальной алго-

ритмической модели языка на каждом ее этапе усложняется алгоритм и правила. И, несмотря на это, всякий раз оказывается, что какие-то фрагменты естественного языка все равно не порождаются, а то, что порождается, наоборот, не существует в языке, так что требуются новые и новые усовершенствования. Поэтому именно рамки узкой алгоритмической модели заставляют задать вопрос: а что, если человек думает и говорит совершенно иначе — т.е., не алгоритмически».

Именно осознание того, что для понимания природы языка необходимо выйти за пределы собственно лингвистических проблем, и привело к тому, что многие открытия когнитивной лингвистики представляют значительный интерес для всех когнитивных наук. Наиболее полно эти открытия нашли свое выражение в концепции Дж. Лакоффа, представленной в его книге [14].

Концепция Лакоффа представляет собой проект решения двух проблем – категоризации и 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 всех 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 структуру (древовидную, как в таксоно-

миях, или решеточную, как в теории формальных понятий [20]) – за исключением элементарных объектов, которые наиболее просты и структуры не имеют. При этом понятия не должны зависеть от особенностей существа, производящего категоризацию: его нейрофизиологии, особенностей восприятия, формирования мысленных образов и т.д.

Когнитивный подход смотрит на проблему категоризации совершенно по-другому. Его принципы категоризации можно сформулировать следующим образом.

1. В категориях человека существуют прототипы – «хорошие» (центральные, репрезентативные, типичные) примеры. Это было убедительно показано в исследованиях Э. Рош [17], развивших теорию прототипов. Принадлежность объекта к категории определяется на основе не признаков, а расстояния от прототипа. Когда людей просят называть элементы категории, частота типичных элементов значительно выше частоты других элементов: фрукт – яблоко, птица – воробей или ворона, поэт – Пушкин. (Отметим, что стандартные прототипы могут быть различными в разных культурах: понятно, что Пушкина в качестве поэта назовут только русскоязычные люди; это же касается и фрукта, и птицы). Ориентация на прототипы имеет место и при обратной процедуре: при отнесении объекта к какой-то категории очень часто рассматривают не наличие обязательных признаков, а сходство с каким-нибудь «хорошим примером». При этом время ответа существенно меньше в случае примеров, типичных для категории. Наличие «лучших» примеров говорит о неравноправии элементов категории, т.е. об *асимметрии* внутри категории. Явления асимметрии называются *прототипическими эффектами*.

2. Категория – не обязательно класс эквивалентности (по отношению «иметь одинаковый набор признаков»). Еще Л. Витгенштейн [3] обратил внимание на категории, у элементов которых общих свойств нет, а близость (*нетранзитивная*) есть. Яркий пример – понятие «игра».

3. Существует базовый уровень категоризации: категории, базовые с когнитивной точки зрения, находятся в середине иерархии общегонкретного. Обобщение происходит вверх от базового уровня, специализация – вниз. Пример: собака – базовая категория, хищник – обобщение, овчарка – специализация. Другой пример:

мебель – стул – качалка, где базовой категорией является стул. Категории базового уровня имеют не только для объектов, но для действий и признаков. Действия типа *бег* относятся к базовому уровню; *движение* – к вышестоящему, *бег трусцой* – к нижестоящему уровню.

Для категорий базового уровня характерно следующее:

- они имеют единый, целостно воспринимаемый ментальный образ (гештальт); они быстро узнаются;
- с ними связана общая двигательная программа, т.е. сходные физические действия для взаимодействия с различными членами категории;
- в качестве их имен используются наиболее короткие и общеупотребительные слова, первичные с точки зрения вхождения в словарный запас языка: в лингвистике известен «список Свадеша» [8], содержащий сто наиболее устойчивых слов (за тысячу лет меняются около 15 из ста): отец, мать, вода, камень, сердце, кровь, белый, два, три, ...; этот список одинаков почти для всех языков;
- 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 и слабо формализованная, типология категорий (которые он называет идеализированными когнитивными моделями – ИКМ). В ней выделяются 4 типа моделей:

Пропозициональные модели не используют механизмы воображения (метафоры и ментальные образы). Они специфицируют элементы, определяют их признаки и отношения между ними.

Образно-схематические модели специфицируют схематические образы, такие как траектории, вместилища, верх-низ (об образных схемах подробнее будет сказано в следующем разделе).

Метафорические модели – отображения пропозициональных или образно-схематических моделей одной области в соответствующие структуры другой области.

Метонимические модели – модели описанных выше типов с дополнительным отношением между частями структуры, при котором часть структуры (субкатегория) замещает или

обозначает категорию в целом (например, «Кремль» может пониматься как «Администрация президента», а «Москва» – как обозначение всей федеральной власти России).

Пропозициональные модели включают в себя классические категории, определяемые множеством общих признаков, однако, кроме того, содержат категории, имеющие нестандартную структуру. Таковы, например, радикальные категории, имеющие структуру «центр-периферия». В этой структуре центром является 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. Принципы расширения: нетранзитивная связь по цепочке; области опыта, важные для категоризации и определяющие характер связей в цепочке; мотивация (нетранзитивность цепочки означает следующее: если элементы a и b категории имеют множество общих свойств P_1 , то в следующем звене цепочки элементы b и c имеют другое множество общих свойств P_2 (быть может, сильно пересекающееся с P_1); соответственно, хотя соседние элементы цепочки имеют много общих свойств, элементы на разных концах цепочки могут вообще не иметь общих свойств). Эти принципы позволяют понять, почему категория такова, но не порождают ее, т.е. не позволяют однозначно предсказать ее устройство. В процессе распада категории в первую очередь выпадают наиболее удаленные звенья цепи. Поскольку передача признаков от центра к расширениям в общем случае нетранзитивна, то может оказаться, что какая-то периферийная субкатегория не обладает многими существенными признаками центра. Прототипом (лучшим примером) в таких категориях служит центр.

Еще один вид пропозициональных 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 цен-

тру, сути или корню, - является наиболее важным фактором в мышлении» [2]. Существенность или несущественность признаков для гештальт-психологов связана не с их внутренними свойствами, а с ролью в структуре образа. Если какой-то несущественный в данном образе признак объявить (или “увидеть”) существенным, то образ может стать “другим” - изменится его интерпретация. Хорошо известны оптические примеры, иллюстрирующие этот эффект. В гештальт-психологии он называется пересортировкой.

Многочисленные примеры образования понятий с помощью *метафор* имеются в математике: пространства, числовая прямая, линейность, выпуклость, рост (функции), пути и связность в теории графов и т.д. Интересно, что теория бинарных отношений, которые формально изоморфны графам, имеет совершенно другую систему понятий – именно потому, что в отличие от теории графов, она не имеет общего гештальта.

Среди метонимических моделей также выделяются несколько видов. Один из них – это *социальные стереотипы*, когда подкатегория наделена статусом представителя категории в целом. Пример - стереотип матери-домохозяйки. Во многих культурах матери-домохозяйки считаются более хорошими примерами матерей. Соответственно, в этих культурах данная подкатегория определяет ожидания, какой должна быть мать, и порождает прототипические эффекты. В совокупности с признаками центральной субкатегории понятия «мать» (см. выше) образуется структура с составным стереотипом. Лучшим примером матери является биологическая мать-домохозяйка, которая растит ребенка, не работает и замужем за отцом ребенка. Чем ближе женщина к прототипу, тем более она репрезентативна как пример матери.

Другой вид метонимических моделей - это *типичные примеры*, которые, в отличие от социальных стереотипов, обычно используются бессознательно и автоматически; с ними не связаны ожидания. Рассуждение на основе типичных случаев - один из важнейших аспектов мышления. Имеется асимметрия между типичными и нетипичными случаями: знание о типичных случаях распространяется на нетипичные случаи, но не наоборот, что ведет к прототипическим эффектам. Например, значения некоторых признаков, характерные для ти-

пичных примеров, по умолчанию распространяются на остальные экземпляры, хотя для них они могут и не иметь места. Так, для тигров предполагается, что они полосатые, для золота – то, что оно желтое.

Еще один вид метонимических моделей - *идеалы*. Многие категории осмысляются в терминах идеальных случаев, которые могут не быть типичными или стереотипными. Пример: разница между идеальным и стереотипным мужем. Большое количество культурных знаний организовано в терминах идеалов. Люди делают оценки и ставят цели на будущее в терминах идеалов; на идеалах могут быть основаны рассуждения.

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, порождающие семантику. 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 не способен 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» [18].

«Мозг работает не с *информацией* в компьютерном понимании этого слова, а со *смыслом*, или *значением*» [12].

«В математической логике имеется независимо существующий синтаксис, независимо существующие модельные структуры и прин-

ципы отображения синтаксиса на модельные структуры. Естественный язык устроен по-другому. Он возник вместе со значением, и когда мы мыслим при помощи естественного языка, мы мыслим о вещах в терминах, которые имеют значение, но не так, что ты сначала мыслишь, а затем открываешь, о чем мы мыслили и что значили наши понятия» [14].

Это видно даже в такой науке, как математика, формальный подход к которой склонен считать ее наукой о манипулировании символами. Казалось бы, в математике что угодно можно обозначать как угодно, лишь бы ввести нужные определения. Однако существуют традиции и культура обозначений: x, y, z – как правило, переменные; a, b, c – как правило, константы; f, g, h – функции и т.д. Это и значит, что существуют смыслы, изначально связанные с некоторыми обозначениями. Более того, по мнению Лакоффа, значения возникают раньше, чем формируются концептуальные структуры: они возникают из нашего доконцептуального телесного опыта. Наш опыт доконцептуально структурирован на базовом уровне. Доконцептуальные структуры – это гештальты и образно-схематические структуры, такие как вместилище, верх-низ, часть-целое, центр-периферия и т.д. Хотя в гештальтах можно выявить внутреннюю структуру, целое психологически является более базовым, чем части. «Базовое» не значит «элементарное», поскольку концепты базового уровня не являются атомарными блоками, лишены внутренней структуры. Категории базового уровня соотносятся с этой доконцептуальной структурой и непосредственно понимаются в ее терминах.

Образные схемы – это источник разнообразных структур категорий и схем быстрых рассуждений (см. ссылку на книгу [18] в п.1). Анализ некоторых основных схем в книге Лакоффа уделено значительное внимание.

Схема вместилища. Структурные элементы этой схемы по Лакоффу: *внутреннее, граница, внешнее*. Они порождаются нашим непосредственным телесным опытом: наши тела – это и вместилища (внутренних органов, пищи), и одновременно объекты, находящиеся во вместилищах (например, в комнатах). Эта схема порождает базовую логику повседневных рассуждений: любой объект находится либо внутри, либо снаружи. Если вместилище A находится во вместилище B и X находится в A , то он

находится в B . Обобщением такого рода рассуждений является Modus Ponens. Известные в теории множеств диаграммы Венна – фактически гештальты этой схемы.

Эта схема порождает многочисленные метафоры, помогающие осмыслить разнообразные типовые ситуации и сценарии. Например, поле зрения понимается как вместилище: вещи появляются и исчезают из поля зрения. Когда человек говорит, что он попал в трудную ситуацию (внутри) и ищет выход из нее (наружу), он использует схему вместилища. В информатике эта схема используется для хранения информации, и добавляются структурные элементы *доступ, вход и выход*.

Отметим, что в терминах базовой логики вместилища находит свое объяснение известный эксперимент Уэйсона [19]. Комментарий к нему можно найти в книге [11]. Он заключается в следующем. Испытуемым показывают 4 карты, на лицевой стороне которых написаны $A, D, 4, 7$. Сообщается правило, которое надо проверить (гипотеза): «если на одной стороне – гласная, то на другой – четное число». Требуется повернуть только те карты, которые, по мнению испытуемого, необходимы и достаточны для верификации гипотезы. Правильное решение – повернуть A и 7 . Однако большинство поворачивают либо только A , либо A и 4 . Поворот A – это поиск подтверждения гипотезы; это делают все. Поворот 7 – это поиск опровержения гипотезы. Поворот карты 4 ничего не дает для верификации гипотезы: как известно в логике, при истинном q («число четно») импликация «если p то q » истинна для любого p , т.е. любая буква на обороте не опровергнет и не подтвердит гипотезу.

Почему испытуемые поворачивают A и не поворачивают 7 ? Верификация гипотезы производится с помощью двух силлогизмов: Modus ponens «если p то q ; p истинно; следовательно, q истинно» и Modus tollens: «если p то q ; q ложно, следовательно, p ложно». Modus ponens служит для подтверждения гипотезы, Modus tollens – для ее опровержения. «Если Modus ponens правильно применяют почти все, то Modus tollens – только две трети» [11]. С точки зрения когнитивной семантики можно дать такое объяснение этому: Modus ponens имеет соответствующий гештальт и дает положительное знание; Modus tollens гештальта не имеет и дает отрицательное знание. Гештальты почти всегда позитивны – это образы чего-то существующе-

го. Образ несуществования представить довольно трудно.

Схема часть-целое. Структурные элементы схемы: *целое, части, конфигурация*. Телесный опыт: наши тела представляют собой целое с частями, которыми мы можем манипулировать. Чтобы действовать в мире, мы должны иметь знание о структуре *часть-целое* других объектов. Базовая логика: схема асимметрична и нерелексивна: если A – часть B , то ни B , ни A не являются частью A . Невозможно, чтобы целое существовало, а части – нет. Однако части образуют целое, только если существует объединяющая их конфигурация. Если целое расположено в P , то и части расположены в P . Эта логика лежит в основе различных таксономий и, в частности, онтологий. Пример метафоры: семья. Брак понимается как конфигурация, создающая целое (семью) из супругов как частей; развод – как разделение, т.е. ликвидация конфигурации.

Схема центр-периферия. Телесный опыт: мы представляем наши тела как имеющие центр (голова, туловище) и периферию (пальцы, волосы). Центр важнее периферии: повреждения центра более серьезные; центр, а не периферия идентифицирует объект. Базовая логика: периферия зависит от центра, а не наоборот. Метафоры основаны на том, что важное понимается как центральное: «в центре нашей деятельности находится...», «центральной проблемой является...»

Заметим, что в теории графов существуют два понятия центра с различной семантикой. Центр неориентированного дерева (точка, для которой максимальное расстояние до других точек минимально) определяется симметричной функцией расстояния. Это понятие близко к обыденному понятию географического центра. Центр ориентированного дерева (корень) определяется антисимметричным отношением порядка. Это понятие характерно для иерархических структур (в частности, для организационных структур и таксономий), где порядок интерпретируется как подчинение или включение. Именно оно формализует схему «центр-периферия». Важность центра выражается, в частности, в том, что его удаление разрушает всю структуру.

Другие схемы: **связь** (примеры метафор – межличностные и деловые отношения: установить и разорвать связь); **источник-путь-цель**

(эта схема может быть временной или пространственной; ее образ – траектория), **верх-низ, спереди-сзади, линейный порядок** (шкала).

Указанные схемы (а также ряд других) можно назвать универсальными, поскольку они в той или иной мере присутствуют у всех людей и во всех культурах. Наряду с этим имеется большое количество схем, возникающих в результате индивидуального или коллективного опыта, который может быть связан с профессиональной деятельностью (пример – уже упоминавшиеся схемы шахматистов), с длительным проживанием в определенной среде и т.д.

Абстрактные концептуальные структуры возникают, во-первых, в результате метафорической проекции образно-схематических структур на абстрактные области (выше уже приводились примеры метафор в математике) и, во-вторых, в результате проекции категорий базового уровня на вышестоящие и нижестоящие категории. В частности, именно образные схемы структурируют описанные в п.2 типы категорий и систем категорий: радиальные категории структурированы схемой «центр-периферия», классифицирующие системы категорий (таксономии) – схемами «часть-целое» и «верх-низ» (ориентация «вниз» определяет направление наследования свойств) и т.д.

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 когнитивной психологии, совершенно чуждую компьютерной парадигме – *что такое понимание*. Предложение понимается непосредственно, если понятия, содержащиеся в нем, непосредственно значимы. Предложение «Книга лежит в сумке» непосредственно понимается, потому что *книга* и *сумка* – понятия базового уровня, а *лежит в* – базовое отношение, образованное схемой *вместилще*. Если мы непосредственно воспринимаем ситуацию «Книга лежит в сумке», то у нас будет гештальтное восприятие книги, сумки и отношения между ними.

Понимание – это способность соотносить понятия со своим опытом, реализованном в накопленных схемах. Абстрактные структуры мы понимаем либо непосредственно – в терминах образных схем, которые их структурируют (см. выше), либо опосредованно – через уже

понятые структуры. Хорошо известно различие между «полным непониманием» (отсутствием понимания) и неправильным (ложным) пониманием. Непонимание – это отсутствие какой-либо образной схемы. Процесс понимания проблемы или объекта заключается в том, что ищется схема, а затем задаются уточняющие вопросы в терминах элементов этой схемы. Если нет схемы, непонятно, какие вопросы надо задавать. Неправильное понимание связано либо с выбором неадекватной схемы (например, выбор линейного порядка вместо центра-периферии), либо неправильным соотношением с элементами адекватной схемы (например, неудачный выбор центра – скажем, при попытке понять в терминах «центр-периферия», как устроены отношения в некоторой малой группе, неправильно выбран лидер группы).

Человеческое знание организовано в виде концептуальной системы – системы категорий и отношений между ними. Эта система структурирована радиально: существуют центральные (базовые) и нецентральные категории и истины. Центральные истины характеризуются в терминах непосредственно воспринимаемых понятий, соответствующих доконцептуальной структуре опыта. К ним относятся категории базового уровня материальной области и схемы, возникающие из обыденного и профессионального опыта. Базовое знание приобретается в ходе взаимодействия с окружением через восприятие и манипулирование материальными объектами. Другие виды знания получаются либо непосредственно (эмоциональное и социальное знание), либо опосредованно (знание, полученное другими и переданное через книги и т.д.).

Все люди обладают способностью концептуализации, т.е. построения концептуальных систем. Наша способность концептуализации включает:

- способность формировать базовые символические и образно-схематические структуры, коррелирующие с доконцептуальными структурами нашего опыта;

- способность к метафорической проекции структур материальных областей на структуры абстрактных областей, обусловленная структурными корреляциями между ними;

- способность формировать сложные категории, используя в качестве структурных образцов образные схемы.

Однако одинаковые способности концептуализации могут порождать различные концептуальные системы. Различия концептуальных систем в разных культурах проистекают в основном из различий в доконцептуальном опыте – как материальном, так и социальном. Например, концептуальная система эскимосов сильно отличается от концептуальных систем горных племен, причем эти различия близки к базовому уровню. Индивидуальные концептуальные системы людей одной культуры также могут отличаться достаточно сильно, однако эти различия, как правило, ближе к периферии.

Различия в концептуальных системах, сформированных на разных языках *A* и *B*, приводят к трудностям перевода с чужого языка *B* в родной язык *A*. Однако следует различать два вида трудностей: трудность *понимания* предложения языка *B* и трудность построения предложения в языке *A*, семантически эквивалентного исходному предложению в *B*. Понимание – это установление соответствия между чужой и своей концептуальными системами. Перевод – это реализация установленного соответствия в корректном предложении на своем языке. Понимание возможно без перевода; адекватный перевод без понимания невозможен. Поэтому при изучении другого языка начинают с базовых категорий материального опыта, которые универсальны. В этом случае перевод почти всегда возможен, и понимание достигается легко. Трудности начинаются при выходе за пределы этих категорий.

Проблема установления соответствия между чужой и своей концептуальными системами возникает не только при переводе в лингвистическом смысле слова. Хорошо известно выражение «мы говорим на разных языках», которое в терминах когнитивной семантики означает «у нас разные концептуальные системы». Сложность установления этого соответствия зависит от различий в концептуальных системах. Сравнивать концептуальные системы можно по разным критериям. Например, понятия, существующие в одной системе, могут отсутствовать в другой; категория, базовая в одной системе, не является базовой в другой, и т.д.

Концепция Лакоффа заново ставит проблему сложности. С традиционной (формально-алгоритмической) точки зрения структура тем сложнее, чем больше в ней элементов и отношений между ними; самые простые концепту-

альные сущности – это элементарные сущности (примитивы), не имеющие структуры. С позиции когнитивной семантики сложность должна определяться величиной интеллектуальных усилий, затрачиваемых при распознавании, запоминании, понимании и использовании концептуальных сущностей. С этой точки зрения наиболее простыми оказываются базовые сущности - гештальты, которые когнитивно элементарны (непосредственно воспринимаются и понимаются), но в процессе концептуализации оказываются в середине таксономической иерархии и имеют сложную внутреннюю структуру. Поэтому то, что алгоритмически сложно, может оказаться когнитивно простым.

3. Когнитивный подход к задачам искусственного интеллекта

Резюмируя, можно выделить следующие основные идеи концепции Лакоффа:

- центральное место понятия гештальта;
- наличие категорий базового уровня;
- нестандартные структуры категорий: радиальные, метонимические и т.д.;
- доконцептуальное структурирование и базовые образно-схематические структуры;
- подход к проблеме понимания на основе базовых категорий и схем;
- постановка проблемы когнитивной сложности.

Эти идеи открывают путь к объяснению реальных (т.е. наблюдаемых) когнитивных процессов мозга, которые не имеют удовлетворительного объяснения в рамках компьютерной парадигмы. Что такое типичный пример? Что такое понимание? Почему для понимания нужны наглядные примеры? С чем связаны трудности перевода? На все эти вопросы концепция Лакоффа дает ответ. Поэтому есть все основания говорить о высокой объяснительной силе этой концепции. Более того, ее объяснительная сила выходит за рамки проблем когнитивных наук; в ее терминах можно формулировать многие серьезные социальные проблемы. Упомянем здесь ряд таких проблем, связанных с механизмами усвоения знаний.

Методические проблемы образования. В свое время выдающийся российский математик В. Арнольд ярко и аргументированно критиковал «бурбакистский» подход к преподаванию математики, внедренный во французской шко-

ле [1]. Этот подход предлагает преподавать математику как формальную систему; наглядность, как минимум, не приветствуется или просто изгоняется. Очевидно, что с позиций когнитивной семантики наглядность совершенно необходима, поскольку она формирует гештальты, которые, в свою очередь способствуют пониманию и уменьшают когнитивную сложность работы с абстрактными понятиями.

Манипулирование сознанием: от коммерческой рекламы до политической демагогии. Эти разнообразные виды деятельности объединяет стремление внушить человеку определенные устойчивые представления, т.е. сделать их частью его концептуальной системы. Успех приемов манипулирования часто достигается за счет создания ярких гештальтов (логотипов, слоганов, броских выражений типа «предательство», «беспредел», «распродажа Родины» и т.д.), которые воспринимаются непосредственно и не стимулируют рациональные размышления над воспринятым.

Искусство и познание: содержится ли новое знание в произведениях искусства? Если ограничиться литературой, то этот вопрос звучит так: можно ли считать знанием о реальном мире неформальные описания вымышленных людей, мест и ситуаций, которые в принципе не верифицируемы? (При этом речь не идет о беллетризованных биографиях реальных личностей или об исторических романах, где верификация возможна.) С позиции когнитивной семантики ответ, безусловно, - да. Яркие образы литературных героев часто используются в нашем общении как гештальты-маркеры типичных человеческих черт (Отелло – ревность, Плюшкин – бессмысленная скупость и т.д.), а переходящие из века в век сюжеты мировой литературы (например, сюжет «Короля Лира») представляют собой образные схемы типовых бытовых ситуаций. Именно благодаря наглядности (т.е. близости к гештальтам) литературные образы и сюжеты запоминаются и используются в реальной жизни часто гораздо эффективнее, чем бесстрастно-объективные описания реальных человеческих характеров и ситуаций.

Концепции Лакоффа позволяют 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 ча-

стности, построения онтологий представляют существенный интерес идеи, связанные с категоризацией и различными видами категорий (на это, в частности, обращается внимание в недавнем докладе Гавриловой и др. [4]). Что касается формализации рассуждений, то здесь важно иметь в виду, что 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. Человеческая способность быстро рассуждать (на которую, к сожалению, обращается гораздо меньше внимания, чем на способность быстро узнавать) основана на использовании образных схем, роль которых отмечалась еще в книге [18] и которые подробно рассмотрены в концепции Лакоффа. Хорошо организованная концептуальная система также способствует быстрым рассуждениям.

«Разум должен что-то получить от формирования категорий, и это что-то – *вывод*. Ясно, что мы не можем знать все о каждом объекте. Но мы можем объединить некоторые их свойства в категорию и из категории предсказать свойства, которые мы не наблюдали» [16].

Еще одна важная проблема – это проблема когнитивной сложности. Когда говорят о быстром рассуждении или узнавании, фактически оценивают эти процессы с точки зрения сложности. Существование базовых категорий также оправдано с точки зрения когнитивной сложности:

«Чем меньше категория, тем точнее предсказание. Однако для идентификации малой категории требуется слишком много свойств. Поэтому наши повседневные категории находятся посередине: «автомобиль», а не «Форд», «стул», а не «мебель». Э. Рош назвала эти категории базовыми» [16].

Когнитивная сложность практически не исследована; в книге Лакоффа она всего лишь поставлена. Соображения об измерении ее объемом интеллектуальных усилий или временем получения ответа, по-видимому, верны, но слишком неточны.

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, в частности, ориентирована модель псевдооптической нейронной сети [6], основанная на голографических механизмах;

- 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.

Литература

1. Арнольд В. http://scepis.ru/library/id_649.html, 2000.
2. Вертгеймер М. Продуктивное мышление. / М.: Прогресс, 1987.
3. Виттгенштейн Л. Философские исследования. // Виттгенштейн Л. Философские работы. Часть I. М: Гнозис, 1994.
4. Гаврилова Т.А., Болотникова Е.С., Гулякина Н.А. Категоризация знаний для создания онтологий. / Материалы 4-й Всероссийской мультиконференции по проблемам управления МКПУ-2011. Т.1. Таганрог: Издательство ТТИ ЮФУ, 2011. С. 62-66.
5. Кузнецов О.П. Неклассические парадигмы искусственного интеллекта. //Теория и системы управления, N5, 1995. С. 3-23.
6. Кузнецов О.П. Псевдооптические нейронные сети - прямолинейные модели. //Автоматика и телемеханика, N12. 1996. С.160-172.
7. Пинкер С. Язык как инстинкт. М.: Едиториал УРСС, 2004.
8. Плуныян В.А. Почему языки такие разные? М.: Русские словари, 1996.
9. Поспелов Д.А. Метафора, образ и символ в познании мира. Новости искусственного интеллекта, № 1, 1998. С. 94-14.
10. Рахилина Е.В. Когнитивная семантика: История. Персоналии. Идеи. Результаты. // Семиотика и информатика. Вып. 36. 1998. С. 274-323.
11. Ришар Ж.Ф. Ментальная активность. Понимание, рассуждение, нахождение решений. М.: Институт психологии РАН, 1998.
12. Роуз С. Устройство памяти - от молекул к сознанию. М.: Мир, 1995.
13. Hawkins Jeff. On Intelligence. – New York, Henry Holt and Company, 2005. Русское издание: Хокинс Джефф. Об интеллекте. М.: Вильямс, 2007.
14. Lakoff J. Women, Fire, and Dangerous Things: What Categories Reveal About the Mind. – University of Chicago Press, 1987. Русское издание: Лакофф Д. Женщины, огонь и опасные вещи: Что категории языка говорят нам о мышлении, М.: 2004.
15. Mikhailov A., Pok Yang Ming. Artificial Neural Cortex. Proceedings of the Artificial Neural Networks in

- Engineering Conference (ANNIE 2001), November, 4-7, 2001, St. Louis, Missouri, USA.
16. Pinker S. How the Mind Works. W.W.Norton&Comp., 1997.
 17. Rosch E. Cognitive representations of semantic categories. *Journal of Experimental Psychology*, 1975. 104, pp.192-233.
 18. Sowa J.F. Conceptual Structures - Information Processing in Mind and Machines. Addison-Wesley Publ.Comp. 1984.
 19. Wason P.C. Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 1968, 20, pp. 273-281.
 20. Wille R., Ganter D. Formal Concept Analysis. Berlin: Springer-Verlag. 1999.

Кузнецов Олег Петрович. Заведующий лабораторией Института проблем управления РАН. Окончил МГУ им. Ломоносова (философский ф-т в 1958 г., механико-математический ф-т в 1961 г.). Доктор технических наук, профессор. Автор 130 печатных работ, в том числе трех монографий. Область научных интересов: искусственный интеллект (моделирование когнитивных процессов, представление знаний, поддержка принятия решений). E-mail: olkuznes@ipu.ru