

# ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ НА ОСНОВЕ ТЕОРИИ ИНВАРИАНТНОГО ПРЕДСТАВЛЕНИЯ: ЭПИСТЕМОЛОГИЧЕСКИЕ ПРОБЛЕМЫ\*

**В.А. Ладов**

В данной статье рассматривается логическая операция субсумпции, на основе которой формируется инвариантное представление об объектах действительности, как одна из определяющих для интеллектуальной деятельности. Утверждается, что для моделирования этой операции в системах искусственного интеллекта, построенных на технических носителях, необходимо прояснить, как она осуществляется в естественном интеллекте. Анализируется, как осуществление данной операции трактуется с точки зрения натуралистического подхода, принятого в нейронауке. Эксплицируются эпистемологические проблемы философского уровня, возникающие перед натуралистической теорией. Демонстрируется, каким образом натуралистическая теория интеллекта способна решить эти проблемы.

## ARTIFICIAL INTELLIGENCE ON THE THEORY OF INVARIANT PRESENTATION: EPISTEMOLOGICAL PROBLEMS

**V.A. Ladov**

The logical operation of subsumption is considered in this paper. It is asserted that the logical operation is one of the main characteristics of intelligence because it forms the invariant presentation about world objects. To model the operation in the systems of artificial intelligence it is necessary to clear realization of the operation in the systems of natural intelligence. It is analyzed in which way realization of the operation is interpreted from the naturalistic point of view. The epistemological problems of naturalistic theories are explicated in this paper. It is demonstrated that naturalistic theory of intelligence is able to solve the problems.

В настоящее время существует достаточно большое количество разработок, которые именуются «искусственными интеллектуальными системами» (это экспертные системы, алгоритмизированные системы

---

\* Исследование выполнено при поддержке Совета по грантам Президента РФ (НШ – 5887.2008.6) и в рамках государственного контракта на выполнение поисковых научно-исследовательских работ для государственных нужд в рамках федеральной целевой программы «Научные и научно-педагогические кадры инновационной России», мероприятие 1.1., проект «Онтология в современной философии языка» (2009-1.1-303-074-018).

принятия решений, разработки в области робототехники и т.д.). При этом вполне эвристичным выглядит различие, используемое Д. Серлем, между так называемым «слабым искусственным интеллектом» и «сильным искусственным интеллектом» [1]. Разработки, осуществляемые с точки зрения «слабого искусственного интеллекта», никогда не претендовали на то, чтобы в точности воспроизвести или даже превзойти интеллект человека. Они касаются решения частных задач вычисления, использования алгоритмов при выполнении того или иного рода деятельности, т.е. в лучшем случае имитируют какой-то фрагмент рациональных способностей человека и имеют своей целью переложить определенную долю чаще всего рутинной работы на машину. То, что такие системы называются «интеллектуальными», никого не должно особенно волновать. Это просто способ именования, возможно, не совсем удачный, поскольку дает жизнь еще одному омониму в языке, но все же не претендующий на какие-то глобальные выводы общемировоззренческого характера.

С «сильным искусственным интеллектом» дела обстоят совершенно иначе. С точки зрения данной теоретической позиции техническое устройство, именуемое интеллектуальным, не просто имитирует, но действительно реализует подлинную рациональную деятельность, причем делает это не фрагментарно, затрагивает не периферийные возможности интеллекта (будь то арифметическое вычисление или выполнение работы в соответствии с алгоритмом), а, напротив, воспроизводит рациональную деятельность в ее сущностных чертах. И вот только в рамках такой теоретической установки возникают эпистемологические вопросы, имеющие философскую значимость. По праву ли эти системы в принципе называются интеллектуальными? Инженеры создают различные технические системы, но где критерии того, какие из них считать интеллектуальными, а какие нет? В соответствии с какими принципами нужно осуществлять инженерную деятельность? Что нужно пытаться воспроизвести на технических носителях в первую очередь, чтобы отразить рациональную деятельность в ее сущности?

Исследователи, работающие в парадигме «сильного искусственного интеллекта», вынуждены так или иначе обращаться к рассмотрению интеллекта естественного, к исследованию человеческого разума, чтобы зафиксировать существенные признаки разумного поведения. Таким образом, разработки в области искусственного интеллекта оказываются зависимыми от исследований в области нейронауки, нейрофизиологии головного мозга, поскольку именно эта отрасль когнитивной науки способна ответить на вопрос об устройстве естественного интеллекта. При этом вполне очевидным выглядит тот факт, что в рамках когни-

тивной науки доминирующей является материалистическая трактовка сознания, поэтому инженеры искусственных интеллектуальных систем обращаются именно к нейронауке, а не к каким-либо другим философским теориям разумной деятельности.

При исследовании естественного интеллекта одним из важнейших моментов является материалистическое объяснение операции субсумпции – подведения конкретных чувственных данных под понятие, формирование абстракций. Это одна из важнейших способностей человеческого разума. Реализуя эту способность, человек структурирует данные опыта, создает иерархические концептуальные каркасы, позволяющие ему наилучшим образом ориентироваться в окружающей среде. В этом смысле воспроизведение способности к субсумпции на технических носителях оказывается важнейшей задачей для успешного развития робототехники, поскольку именно она сможет обеспечить гибкость интеллекта в оперировании с объектами и тем самым приблизить поведение искусственной интеллектуальной системы к естественной.

Тем не менее как раз в данном существенном аспекте инженерная деятельность по-прежнему не может похвастаться внушительными достижениями. По словам М. Мински [2], одного из известных специалистов в области искусственного интеллекта, машина никогда не увидит в торчащем из-за кресла хвосте кошку. Человеческий интеллект, напротив, оказывается чрезвычайно гибким и подвижным в отношении субсумпции. Все это заставляет ученых и инженеров с большим вниманием всматриваться в деятельность естественной интеллектуальной системы в надежде получить необходимые сведения о ее работе для того, чтобы затем попытаться воспроизвести ее на технических носителях.

Одной из самых последних и достаточно широко обсуждаемых работ в нейронауке является теория интеллекта Д. Хокинса [3]. Здесь автор как раз ставит вопрос о материалистическом основании процессов концептуализации, считая его одним из краеугольных для своей теории. Он называет создание абстракций «инвариантным представлением». Его задача показать, как это происходит на нейрофизиологическом уровне.

Хокинс утверждает существование иерархической структуры кортекса (коры головного мозга). Причем непосредственно с информацией, идущей от органов чувств, связаны только нижние слои кортекса. Здесь происходит возбуждение комплексов нейронов на основе раздражения нервных окончаний в органах чувств. В частности, если мы возьмем зрительное восприятие, то увидим, что глаза воспринимают объект, формируя некоторые дискретные порции информации – так называемые саккады. Если я смотрю на лицо собеседника, то мои глаза

совершают постоянные движения, концентрируясь на отдельных дискретных элементах объекта. Я вижу нос, губы, фокусируюсь на одном глазе, на другом и т.д. Вся эта информация порциями передается в нижние слои кортекса, приводя к возбуждению различные группы нейронов.

Встает вопрос, за счет чего я вижу лицо собеседника в целом? Как происходит отвлечение от конкретного чувственного материала, передаваемого саккадами? Как формируется абстракция лица? У Хокинса эта проблема получает такое специфическое наименование, как поиск «нейронов Билла Клинтона», если мы предположим, что разглядываем лицо бывшего американского президента.

Далее гипотеза разворачивается следующим образом. Если нижние слои кортекса чрезвычайно подвижны и реагируют на любые изменения информации от органов чувств, то высшие слои более инертны. Дендриты нейронов высших слоев кортекса соединяются посредством синапсов сразу со многими аксонами нейронов нижних слоев. В результате получается, что определенная конфигурация нейронов на высшем уровне оказывается постоянно активной вне зависимости от изменения нейронной активности на низших уровнях. Так возникает восприятие лица Билла Клинтона. Чтобы мы ни воспринимали в каждый конкретный саккадный момент – ухо, глаз, цвет кожи и т.д. – определенная группа нейронов на высшем уровне кортекса окажется постоянно активной. Это и будут нейроны Билла Клинтона.

Исследования, проводимые в рамках такой предметной области, как философия науки (к коим относится и материал настоящей статьи), как правило, заинтересованы в прояснении значений терминов, входящих в теорию, в прояснении последовательности рассуждений и в экспликации некоторых скрытых предпосылок, на которые опираются ученые. И в данном конкретном случае важно прояснить, присутствуют ли в гипотезе Хокинса какие-либо скрытые эпистемологические предпосылки. С какими существенными затруднениями эпистемологического характера эта гипотеза сталкивается?

При обсуждении этих вопросов нам бы хотелось провести параллель с феноменалистской теорией абстрагирования Д. Локка [4]. Как известно, Д. Локк предлагает естественную теорию абстракции, противостоящую классическому идеалистическому представлению платонизма, в соответствии с которым абстракция антиципирует фиксацию чувственного материала. Локк как сенсуалист утверждает, что изначально данными в опыте являются только конкретные ощущения, а абстракции возникают в разуме после обработки конкретных чувственных идей, после их сравнения и выявления общих характеристик. Од-

нако по отношению к такой феноменалистской концепции Локка можно выдвинуть возражение платонистского типа. Дело в том, что для сравнения конкретных чувственных данных и для выявления их подобия друг другу мы уже должны опираться на некоторые критерии подобия, которые сами могут выступать только как абстрактные идеи. Так, пытаясь собрать некоторую множественность предметов в класс, нам нужно выделить то свойство, в соответствии с которым этот класс будет образован. Например, я вижу эти два стола как подобные на основе того, что я уже каким-то образом фиксирую идею прямоугольности (если я обращаю внимание на крышки столов) и более фундаментальную идею формы вообще. Таким образом, Локк не может избавиться от платонистских предпосылок в своей теории абстрагирования. Он не может, исходя из принципов сенсуализма, объяснить объединение чувственных данных в классы.

В нашем конкретном примере с формированием абстракции лица Билла Клинтона на основании последовательного созерцания его частных элементов данная проблема формулируется следующим образом. Чтобы увидеть лицо целиком, необходимо все частные ощущения, формирующиеся в зрительном восприятии, интерпретировать как части единого целого. Если этого не произойдет, то мы будем иметь в восприятии только неопределенное хаотическое скопление ощущений, за которыми нельзя будет увидеть цельные вещи. Поскольку мы в своем опыте в качестве очевидных данностей все же представляем вещи как цельные единства, постольку это должно предполагать, что мы имеем интеллектуальное созерцание таких абстрактных сущностей, как часть и целое. Созерцание данных абстракций первично по отношению к конкретному чувственному восприятию. Мы оказываемся способными разглядеть за саккадой носа, губ, глаз и т.д. лицо Билла Клинтона только на основе того, что уже интерпретируем эти разрозненные ощущения в рамках системы «часть – целое», расценивая их как интегральные частные элементы цельного объекта. Естественная теория абстрагирования не сможет и в этом случае вывести абстрактное из конкретного, поскольку созерцание абстракции эпистемологически должно быть первичным по отношению к классификации конкретного чувственного материала. Прежде чем станет возможным группировка носа, губ, глаз и т.д. в лицо Билла Клинтона, должна иметь место интеллектуальная интуиция, в которой осуществляется созерцание таких абстрактных объектов, как часть и целое.

В материалистических концепциях нейронауки в эпистемологическом отношении интригует тот факт, что они, как кажется, способны обойти этот объективно-идеалистический контраргумент, сформулиро-

ванный выше. Хокинсу не нужно объяснять, на основании какого абстрактного принципа конкретные чувственные данные объединяются в общую идею и за счет каких эпистемических ресурсов происходит познание этого принципа. Ему нужно лишь указать на экспериментальные данные об исследованиях головного мозга. Так устроен мозг. Низшие слои нейронов кортекса связаны с высшими слоями именно таким образом, что, видя ухо, нос, глаз, я каждый раз вижу лицо. Здесь не нужно допущение какого-либо созерцания абстракции. Абстрактный принцип классификации конкретных чувственных данных интерпретируется как определенный «коридор» нейронных связей. Как возник этот «коридор»? В процессе эволюции.

Мы могли бы сказать, что натуралистическая концепция Д. Хокинса, как и иные теории нейронауки, представляет собой специфический вид трансценденталистского рассуждения. Это – натуралистический трансцендентализм. Здесь основные свойства разума, отвечающие за формирование опыта, интерпретируются как укорененные в созданных в процессе эволюции материалистических образованиях. При этом сама эпистемологическая схема описания процесса формирования опыта оказывается вполне соответствующей классическому идеалистическому трансцендентализму И. Канта [5]. Как у Канта разрозненные данные ощущений нуждаются в объединяющем их принципе, за счет которого производится структурирование и классификация ощущений, так и у Хокинса объяснение цельной картины чувственного созерцания основано на экспликации принципа объединения частного в целое. Разница только в том, что у Канта фундирующие чувственный опыт понятия и категории относятся к идеальному трансцендентальному субъекту, тогда как в натуралистических концепциях сфера трансцендентального, т.е. того, что обеспечивает возможность опыта, коренится в нейрофизиологическом устройстве головного мозга человека. Трансцендентальным субъектом здесь выступает определенная форма материи как продукт эволюции.

Подобные взгляды развивает эволюционная эпистемология, заговаривая о натуралистических трансцендентальных основаниях опыта в противовес классическому кантовскому трансцендентализму. Такое сравнение позиций И. Канта и К. Лоренса – одного из видных представителей эволюционной эпистемологии – можно найти, например, в исследованиях К.-О. Апеля [6].

Для философов, разделяющих позицию радикального или математического реализма (платонизма), должен выглядеть небезынтересным тот факт, что если феноменалистская теория абстракции может быть подвергнута критике со стороны платонизма, то для натуралистической

трактовки процесса абстрагирования явный контраргумент с ходу сформулировать сложно. Что значит созерцать такие абстракции, как часть и целое? С натуралистической точки зрения это не какие-то эфемерные невидимые объекты, а структурные образования материи, тот самый «коридор», возникший в процессе эволюции нейронных связей головного мозга, который обеспечивает надлежащий синтез разрозненных чувственных данных. Сфера метафизического в рамках такой интерпретации стремительно теряет весь свой покров таинственной значимости, который всегда тщательно старалась сохранить идеалистическая философия.

Оставляя здесь открытым вопрос о возможном платонистском ответе натурализму, хочется отметить лишь то, что наиболее продуктивными, как это видно на примере настоящего рассмотрения, эпистемологические исследования оказываются именно в точках столкновения диаметрально противоположных парадигм, ибо именно здесь могут быть сформулированы самые принципиальные критические аргументы, на которые любая эпистемологическая теория, претендующая на адекватное описание опыта познания, обязана отвечать.

Что же касается технологий искусственного интеллекта, то в рамках данной области инженерной деятельности натуралистическая интерпретация продуцирования инвариантного представления, позволяющая обойти те эпистемологические трудности, с которыми сталкивалась феноменалистская теория абстракции, конечно же, может только приветствоваться. После того как найдено объяснение процесса концептуализации опыта на основании исследования биологических носителей интеллекта, дело остается за малым – воспроизвести этот процесс на технических носителях, что, по мысли Д. Хокинса, может быть реализовано в недалеком будущем.

#### ЛИТЕРАТУРА

1. Серл Д. Мозг, сознание и программы // Аналитическая философия: становление и развитие. – М., 1998. – С. 376–400.
2. Minsky M. Will Robots Inherit the Earth? // Scientific American. – 1994. – № 4. – October.
3. Хокинс Д., Блейкли С. Об интеллекте. – М., 2007.
4. Локк Д. Опыт о человеческом разумении // Соч.: В 3 т. – М., 1985. – Т. 1.
5. Кант И. Критика чистого разума // Соч.: В 8 т. – М., 1994. – Т. 3.
6. Апель К.-О. Лингвистическое значение и интенциональность: Соотношение априорности языка и априорности сознания в свете трансцендентальной семиотики или лингвистической прагматики // Язык, истина, существование. – Томск, 2002. – С. 204–224.