

Секция I. ЭПИСТЕМОЛОГИЧЕСКИЕ, МЕТОДОЛОГИЧЕСКИЕ И ЛОГИЧЕСКИЕ ПРОБЛЕМЫ МОДЕЛИРОВАНИЯ ИНТЕЛЛЕКТА

Руководители:

*академик РАН В.А. Лекторский (ИФ РАН);
д. филос. н., проф. М.В. Лебедев (ММСУ);
к. филос. н., доц. Е.А. Никитина (МИРЭА)*

ЭВОЛЮЦИЯ ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ В КОНТЕКСТЕ КОГНИТИВНЫХ НАУК

Бураков А.А.

Вологодский государственный педагогический университет

Современный мир невозможно представить без информационных технологий, которые применяются во всех сферах жизнедеятельности. С их развитием стал меняться социум и формы поведения человека. Появление интеллектуальных информационных технологий, а также развитие сетевых технологий привело к тому, что влияние информационных технологий приобрело глобальный характер.

Эволюция информационных технологий во многом основана на появлении цифровых компьютеров и развитии интеллектуальных систем. Первые разработчики интеллектуальных систем исходили из того, что интеллектуальную деятельность человека можно свести к логико-вычислительной. Данный подход на сегодняшний день реализовался в разработке различных экспертных систем, которые, основываясь на знаниях, ранее заложенных в них, могут решать узконаправленные задачи в различных областях без привлечения специалиста.

Сферами применения таких систем может быть планирование, обучение, интерпретация, управление, прогнозирование и диагностика (примером может служить система DENDRAL разработанная в Стэнфорде). Важнейшими проблемами в разработке таких систем

являются извлечение и формализация знаний, то есть необходимость перевода знаний специалиста в доступный для понимания машины язык. Непонимание машиной естественного языка — следствие символичности и контекстности информации, используемой в речи, а так же того, что сознание и психика человека представляет собой «уникальный эффект биологической самоорганизации, найденный в процессе эволюции» [1, с. 87].

Поэтому другое направление разработки интеллектуальных систем опирается на результаты биологических исследований работы мозга и процессов эволюции.

Так, модель нейронной сети способна обучаться подобно живому организму. Она формируется посредством взаимодействия с окружающим миром и не требует отражения действительности на языке символической модели. Нейронная сеть использует принцип параллельно-распределенной обработки информации. Интеллектуальная деятельность такой сети обусловлена взаимодействием простых компонентов (нейронов) и формированием новых связей между ними в процессе обучения. Организация компонентов сети происходит несколькими слоями, при этом нейроны в одном слое обрабатывают входные данные независимо друг от друга.

Развитие модели нейронных сетей привело к появлению теории сознания Дэннетта, которая была основана на представлении сознания через взаимодействие агентов в распределенной архитектуре интеллекта. Ее принцип заключался в том, что «во время перцепции управления двигательными функциями, решением задач, обучения и другой психической активности формируются объединения взаимодействующих агентов» [2, с. 791]. В этой теории сознание выступает как связующий фактор объединений агентов, совокупность которых зависит от выполняемой ими задачи. Развитие теории нейронных сетей показало, что работа мозга может рассматриваться как совокупность большого числа отдельных простых структур, независимых друг от друга.

В настоящее время приходит осознание того, что интеллект, сознание и разум человека являются продуктами его социальной эволюции. Они связаны с опытом, традициями, развитием культуры, искусства, общественных отношений. В концепции глобального эволюционизма разум, сознание и психику человека можно рассматривать как уникальные феномены социальной эволюции, которые появляются именно на этом этапе развития.

Кроме того, сознание обладает свойством «субъективной реальности» [1, с. 75], тогда как машина не обладает этим свойством. Это означает что информационный процесс, протекающий в ней, «отличается по своей организации, по своим структурным, оперативным и

целевым характеристикам от того информационного процесса, который специфичен для эго-системы головного мозга» [там же, с. 81].

Сознание и интеллект человека являются результатом всей эволюции мира, человека, поэтому их моделирование в искусственном интеллекте представляется практически невозможным. Однако на современном этапе формируются основания для перехода на следующий эволюционный уровень, и именно интеллектуальные информационные технологии могут сыграть решающую роль для дальнейшего развития этапа социальной эволюции.

Литература:

1. Искусственный интеллект: междисциплинарный подход. Под ред. Д.И. Дубровского и В.А. Лекторского — М.: ИИнтелЛ, 2006. — 448 с.
2. *Люгер Дж.* Искусственный интеллект: стратегии и методы решения сложных проблем, 4-е издание.: Пер. с англ. — М.: Издательский дом «Вильямс», 2003. — 864 с.
3. История информатики и философия информационной реальности. Под ред. Р.М. Юсупова и В.П. Котенко — М.: Академический проспект, 2007. — 429 с.

ТЕРМИНОЛОГИЧЕСКАЯ БАЗА ДАННЫХ В ОБЛАСТИ НАНОТЕХНОЛОГИЙ: ВОЗМОЖНОСТИ И СФЕРЫ ПРИМЕНЕНИЯ

Бездорожев С.В.

*Московский государственный лингвистический университет
(гуманитарный университет)*

Развитию нанотехнологий будущего уделяется разными странами все большее внимание. В России Московский государственный лингвистический университет впервые в мире поставил вопрос о взаимодействии лингвистики с нанотехнологиями. В 2008 году на одной из конференций, проходившей в стенах этого университета, прозвучал доклад Потаповой Р. К. «Нанотехнологии и лингвистика: прогнозы и перспективы взаимодействия». Вдохновленные идеями, витавшими на конференции, группа студентов, среди которых есть и сам докладчик, решила посвятить себя созданию терминологической базы данных в области нанотехнологий. Тем самым, мы пытаемся решить одну из тех проблем нанотехнологий, выделенных Потаповой Р.К. [1; 2], ко-

торые под силу решить лингвистам, а именно — разработку лингвистического обеспечения информационного поля нанотехнологий.

Как справедливо отмечают специалисты научно-исследовательского Центра на базе Курчатовского института, которым поручена реализация программы по строительству нанотехнологической сети [3], для системного развития всей инфраструктуры nanoиндустрии необходимо, прежде всего, упорядочение терминологии. С нашей точки зрения, необходимо определить в первую очередь лексические единицы семантического поля — «нанотехнологии». В данном случае целесообразно теснейшее взаимодействие специалистов в области нанотехнологии и лингвистов. Для успешного развития проекта следует уделить первостепенное внимание терминологии профессионального языка (подъязыка) в области нанотехнологии и nanoиндустрии.

По моему мнению, огромным плюсом является тот факт, что мы создаём вполне *материальный продукт* — терминологический словарь-справочник, который может реально помочь самым разным специалистам переводить тексты и получать дополнительную справочную информацию, собранную нами из многих источников. Его уникальность достигается за счет того, что он пока мы обрабатываем большие массивы информации. Во-вторых, он пока единственный в своем роде.

На базе словаря была создана память переводов (translation memory) для использования в системах автоматизированного перевода¹, таких как SDL Trados или IBM Translation Manager. Концепция Translation Memory предполагает выявление в переводимом тексте фрагментов, переводы которых уже имеются в базе данных переводов, и за счет этого происходит сокращение объема работы переводчика. Это выявление получило название выравнивания или сопоставления (alignment). Фрагменты, оставшиеся непереуведенными после выравнивания (сопоставления), передаются дальше для ручной обработки переводчику или системе машинного перевода (Machine Translation, MT). Переводчик на этом этапе может выделить вновь переуведенные фрагменты и занести новые пары параллельных текстов на двух языках в базу данных. Такая схема наилучшим образом работает на однотипных текстах, где повторяемость словосочетаний достаточно высока: например, инструкциях для пользователей, технических описаниях.

¹ Автоматизированный перевод (АП, англ. Computer-Aided Translation) — перевод текстов на компьютере с использованием компьютерных технологий. От машинного перевода (МП) он отличается тем, что весь процесс перевода осуществляется человеком, компьютер лишь помогает ему произвести готовый текст либо за меньшее время, либо с лучшим качеством [4].

Таким образом, в нашей разработке мы предлагаем создать память переводов, наполненную на основе базы данных, которая, в совокупности со специализированным ПО, будет самостоятельно находить нанотермины и переводить их (или предлагать варианты перевода). В данный момент такая translation memory уже создана на основе предварительной выборки из базы данных.

Литература:

1. *Потапова Р. К.* Нанотехнологии и лингвистика: прогнозы и перспективы взаимодействия.
2. *Потапова Р.К.* Новые информационные технологии и лингвистика. 4-е изд.. М.. КомКнига, 2005.
3. *Нарайкин О.* Узловые узлы (нанотехнологий не вписываются в вертикаль) // Поиск. №43. М., 2007.
4. Wikipedia, 2009 — http://ru.wikipedia.org/wiki/Translation_memory.

СЛАБАЯ ЭФФЕКТИВНОСТЬ «СИЛЬНОЙ» РАЦИОНАЛЬНОСТИ

Горбатов В.В.

*Государственный университет Высшая школа экономики
(Москва)*

Постклассическая модель рациональности предполагает осознание неустранимости познающего индивида из его картины мира. Выявление гносеологических оснований теории в таких условиях предполагает анализ абстракций и идеализаций, касающихся не только её предметной области, но и тех неявных ограничений, которые накладываются на действия самого познающего субъекта. Данные ограничения трудно эксплицируемы в силу того, что они носят преимущественно операциональный характер, определяя способы конструирования субъекта, а вовсе не его эмпирически обнаружимые свойства. Но вместе с тем, косвенным образом эти «конструктивные паттерны» нашего познания могут иметь эмпирические следствия в широкой области наук, связанных с моделированием рациональной деятельности субъекта — в кибернетике, информатике, исследованиях проблем искусственного интеллекта и пр.

К середине XX века казалось, что универсалистская парадигма в логике потерпела полный крах. На глазах рушились представления о едином логическом универсуме, единственном наборе истинностных значений, единственно возможном отношении логического следова-

ния, единственном и неизменном наборе логических законов. Воцарился логический плюрализм, ознаменованный появлением не только множества новых логических теорий, но и множества трактовок самой природы логического. Однако если для алетической модальной логики преодоление универсализма почти автоматически означало наступление эпохи расцвета, то эпистемической логике суждено было столкнуться с дополнительными принципиальными трудностями внутреннего характера. Оказалось, что мало разрушить классические представления о реальности, лежащие в основе старой модели логического знания — надо подвергнуть переосмыслению еще и само понятие субъекта. Дело в том, что невозможно сформулировать ни одну достаточно богатую и интересную в прикладном плане систему эпистемической логики, не приписав познающему субъекту минимальные рациональные способности, выражаемые, как правило, аксиомами:

(i) дедуктивной связности $K_a(p \rightarrow q) \rightarrow (K_a p \rightarrow K_a q)$

(ii) истинности $K_a p \rightarrow p$

(iii) позитивной интроспекции $K_a p \rightarrow K_a K_a p$

(iv) негативной интроспекции $\neg K_a p \rightarrow K_a \neg K_a p$

Но насколько идеализированный субъект познания, обладающий рациональными характеристиками в смысле приведенных выше аксиом, совпадает с реальным субъектом познания? Практика показывает, что применимость используемых стандартной эпистемической логикой конструкций в сколь-нибудь реалистичном исследовании человеческого познания весьма сомнительна. Дело в том, что лежащие в её основе идеализации базируются на понятии «сильной рациональности», главные слагаемые которого, на наш взгляд таковы:

- 1) Пассивность познающего субъекта.
- 2) Статичность знания, игнорирование его изменчивости и принципиальной незавершенности.
- 3) Полнота и непротиворечивость «миров знания»
- 4) Прозрачность собственного менталитета для рефлексии сознания.
- 5) Вытекающая из (2–4) прозрачность сознания других, следствием которой, в свою очередь, является принципиальный
- 6) Монологизм.

Необходимо отметить, что по историческому совпадению ранняя эпистемическая логика, опиравшаяся на описанную выше модель познания, находилась в стороне от тенденций, наметившихся в сфере когнитивных наук, где сложилось целое семейство неклассических концепций, куда вошли: «процедурная», «операциональная», «ситуативная», «интерактивная», «селективная», «коммуникативная» и т.п. теории рациональности. Несмотря на то, что поначалу представления об ограниченной рациональности были довольно смутными и пред-

ставляли собой скорее некие эмпирические обобщения [1] — опытом путем нащупанные нормы приемлемой аппроксимации — чем стройные теоретические модели, довольно скоро они стали популярными в кибернетике, AI-исследованиях, нелинейном программировании, лингвистике, когнитивной психологии и даже в биологии. И пока теоретики эпистемической логики соревновались в построении изящных аксиоматических систем, время отреагировать на вызов было упущено.

Вероятно, именно этот гандикап послужил одной из причин невнимательного отношения к эпистемической логике со стороны специальных дисциплин, а впоследствии привело и к тому, что наиболее острые, актуальные философские вопросы о рациональности чаще ставились не в абстрактно-логических терминах, а в терминах этих частных наук.

Но на самом деле, сложившаяся ситуация вовсе не является фатальной — по крайней мере, со стороны самих логиков давно прикладываются усилия к тому, чтобы её преодолеть. Уже в 1960–70-х гг. были сделаны первые шаги к созданию эпистемической логики «второго поколения» (EL-2): начали интенсивно развиваться интеррогативная логика и эротетика, помимо крипкевской семантики возможных миров (PWS) появились ситуационная (SS) и теоретико-игровая семантики (GTS); в 1980-х гг. стали набирать силу идеи динамической логики и *belief revision*. Наконец, в 1996 г. в книге Я. Хинтикки «Пересмотренные принципы математики» впервые в систематическом виде была изложена идея так называемой IF-логики (Independence Friendly Logic) [подробнее см. 2]. Собственно, применение аппарата IF-логики и теоретико-игровой семантики (GTS) к традиционным проблемам логики познания Хинтикка и называет «эпистемической логикой второго поколения» [3]. Есть веские основания считать, что последняя успешно преодолевает если не все, то многие концептуальные недостатки систем «первого поколения».

Беря за образец сократическую модель познания, эпистемическая логика «второго поколения» позволяет обосновать широкий круг неклассических концепций рациональности. Она дает инструментарий, пригодный для объяснения ситуационности и телесной (отелесенной) природы познания [4], его энантизивированности и эмерджентности.

Литература:

1. Саймон Г. Рациональность как процесс и продукт мышления // THESIS, 1993, вып. 3.

2. *Hintikka J.* Hyperclassical logic (a.k.a. IF Logic) and it's implications for logical theory. // The Bulletin of Symbolic Logic, Vol. 8, №3 (2002).
3. *Hintikka J.* Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning, Cambridge University Press, 2007.
4. *Varela F., Thompson E., and Rosch E.* The Embodied Mind: Cognitive science and human experience. Cambridge, MA: MIT Press, 1991.

СОВРЕМЕННЫЕ ПРОБЛЕМЫ ПОСТРОЕНИЯ ЭКСПЕРТНЫХ СИСТЕМ

Ефремов А.С.

Тверской государственной технической университет

С развитием науки и техники все острее чувствуется потребность в интеллектуальных системах способных в автоматическом или полуавтоматическом режиме заменить рутинный труд человека. Уже сейчас трудно представить нашу жизнь без подобных технологий. Это и «умные» станки, повышающие эффективность производства и снижая издержки, и множество программных средств, анализирующих огромные объемы информации и т.д. Подобные системы называют экспертными.

Экспертная система — это система, оперирующая знаниями в определенной предметной области с целью выработки рекомендаций или решения проблем [1, 4].

Все экспертные системы имеют следующие особенности: наличие специфических знаний; система не просто выполняет некоторый алгоритм, но и создает собственные паттерны поведения; знания сконцентрированы на определенную область; знания предполагают определенную организацию и интеграцию — то есть отдельные сведения соотносятся друг с другом и образуют логические цепочки; знания «заточены» на решение проблемы; система не просто демонстрирует свои знания, но и умеет ими распорядиться.

Существует ряд типовых задач, над которыми трудятся экспертные системы: извлечение информации из первичных данных (обработка сигналов); диагностика неисправностей (как в технических системах, так и в человеческом организме); структурный анализ сложных объектов (например, химических соединений); выбор конфигурации сложных многокомпонентных систем (например, распределенных компьютерных систем); планирование последовательности выполнения операций, приводящих к заданной цели (промышленные роботы).

Основными характеристиками экспертных систем являются: 1) моделирование механизма моделирования мышления человека в конкретной области. Это существенно отличает экспертные системы от систем математического моделирования или компьютерной анимации. Программа, конечно, полностью не воспроизводит психологическую модель эксперта в этой предметной области, но, тем не менее, основное внимание уделяется воспроизведению компьютерными средствами методики решения поставленных задач; 2) получение выводов на основе заложенных знаний. Знания в системе представлены, как правило, на специальном языке и хранятся отдельно от собственно программного кода, который и формирует выводы и соображения; 3) применение эвристических и приближенных методов. Несмотря на то, что методы используемые системами являются приближительными, они не требуют исчерпывающей исходной информации, и с определенной степенью уверенности выдают верный результат.

Рассмотрим базовые функции экспертных систем: приобретение знаний; представление знаний; недостаток ресурсов; прочие проблемы [2, 3].

Существует ряд проблем, которые необходимо решить для корректной работы системы.

Первая проблема — приобретение знаний.

- Трудность интерпретирования специфических знаний в математические термины.
- Трудность передачи «личного опыта эксперта» экспертной системе.
- Сложность ограничения области знаний для решения поставленной задачи.

Вторая проблема — представление знаний.

- Логическая адекватность. (Представление должно содержать все отличия, которые закладываются в исходную сущность).
- Эвристическая мощьность. (Это означает, что наряду с наличием выразительного языка представления должно существовать некоторое средство использования представлений, сконструированных и интерпретируемых таким образом, чтобы с их помощью можно было решить проблему.)
- Естественность нотации. (Большинству приложений, на базе экспертных систем, необходимо накопление большого объема знаний. Выражения, которыми описываются знания, должны быть простыми для написания и их смысл должен быть понятен без компьютерной интерпретации.)

Третья проблема — недостаток ресурсов.

- Недостаток квалифицированного персонала.
- Средства построения экспертных систем.
- Длительное время разработки.

Прочие проблемы включают:

- Нереалистические ожидания. (ЭС могут давать не самые лучшие решения на границе их применимости, при работе с противоречивыми знаниями и в рассуждениях на основе здравого смысла.)
- Неадекватность инструментальных средств решаемой задаче.

Таким образом, существует немало проблем в области создания экспертных систем, решение которых является достаточно нетривиальной задачей, но есть надежда, что дальнейшее развитие науки и техники поможет повысить эффективность данной технологии.

Литература:

1. *Джексон П.* Введение в экспертные системы. — М.: Изд. дом «Вильямс», 2001. — 624 с.
2. Основы теории нейронных сетей // <http://www.intuit.ru>
3. *Медведев В.С.* Нейронные сети. MATLAB 6 / Под общ. ред. к.т.н. В.Г. Потемкина. — М.: Диалог-МИФИ, 2002. — 496 с.
4. Экспертная система // <http://ru.wikipedia.org/wiki/>

ИНФОРМАЦИОННО-ПОИСКОВАЯ СИСТЕМА ДЛЯ ИССЛЕДОВАНИЯ МЕТОДОВ РАЗРЕШЕНИЯ ЛЕКСИКО-СЕМАНТИЧЕСКОЙ ОМОНИМИИ

Казенников А.О., Панков А.В., Трифонов Н.И.

*Московский государственный институт радиотехники,
электроники и автоматики (технический университет)*

В докладе приведены исследования алгоритмов вычислительной лингвистики, которые могут использоваться для повышения эффективности информационного поиска. Описана разработанная информационно-поисковая система (ИПС), которая кроме традиционных методов информационного поиска может реализовывать лингвистический разбор документов и использовать информацию, извлекаемую в ходе семантического анализа.

За последние 15 лет были исследованы различные способы интеграции алгоритмов разрешения лексико-семантической омонимии в алгоритмы информационного поиска для повышения его эффективности [1–8]. Информационный поиск на основе булевой алгебры реализует простой алгоритм обработки запросов. Он возвращает все документы, которые удовлетворяют заданному выражению булевой ал-

гебры. Кроме булевой алгебры возможно указание элементарных позиционных параметров. Например, по запросу система может найти все документы, где «ключ» и «источник» разделены не более чем тремя словами.

Задача описываемого модуля — выделить подмножество потенциально релевантных документов. Ранжирование и усечение этого множества осуществляют модули традиционного и семантического ранжирования документов.

В области информационного поиска сформировалось два основных метода ранжирования документов: с использованием векторной модели представления документа и с использованием вероятностной модели представления документа.

При получении набора документов при булевом поиске все слова в запросе обладают одинаковыми весами. Недостатки использования булевой алгебры для обработки запросов разрешаются в этом модуле. Все полученные документы ранжируются в порядке убывания релевантности. Для этого документ и запрос представляются с помощью векторов в пространстве термов. Каждый терм представляет собой размерность. Ранг документа определяется подобием документа запросу.

В векторной модели предполагается, что термины независимы между собой, тогда пространство ортогонально. Это упрощает вычисление расстояния между векторами.

Разработанная информационно-поисковая система включает следующие основные модули:

- модуль индексирования документов;
- модуль обработки запросов;
- булев информационный поиск;
- модуль традиционных методов ранжирования;
- методы лингвистической обработки;
- ранжирование с применением лингвистических методов;
- комбинация результатов;
- расширение запросов.

Выделение стоп-слов производится по словарю, в котором находятся часто употребляемые слова, не несущие содержательной нагрузки — различные предлоги и союзы и т.п.

Процедура лемматизации определяет основу слова по словарю. Процедура может выделить несколько основ, поскольку лемматизация производится на основе только синтаксических параметров. Получившиеся основы слов называются термами.

После этого составляется инвертированный список термов, встречающихся во всех документах.

Для каждого термина собираются следующие данные:

- число документов, в которые входит этот терм;
- общее число вхождений текущего токена во всех документах;
- частота термина в документе;
- список позиций токена в документе.

В ИПС используется два источника данных. Имеется внутренняя база данных, которая содержит проиндексированные документы в некотором представлении. На текущий момент планируется простой сбор статистики по терминам. Будет также использоваться и внешний источник данных, представляющий собой набор полнотекстовых версий используемых документов (корпус).

Литература:

1. *Robert Krovetz and W. Bruce Croft.* Lexical ambiguity and information retrieval. *Information Systems*, 10(2):115–141, 1992.
2. *Mark Sanderson.* Word sense disambiguation and information retrieval. In *Proceedings of SIGIR-94, 17th ACM International Conference on Research and Development in Information Retrieval*, pages 49–57, Dublin, IE, 1994.
3. *Sanderson M.* Word Sense Disambiguation and Information Retrieval. PhD thesis, University of Glasgow, 1997.
4. *Julio Gonzalo, Felisa Verdejo, Irina Chugur, and Juan Cigarran.* Indexing with wordnet synsets can improve text retrieval. In *Proceedings of the COLING/ACL '98 Workshop on Usage of WordNet for NLP*, pages 38–44, Montreal, Canada, 1998.
5. *Mark Sanderson.* Retrieving with good sense. *Information Retrieval*, 2:47–67, 2000.
6. *Salton G., Fox E., and Wu H.* Extended boolean information retrieval. *Communications of ACM* 26, 1983.
7. *G. Salton, C. Buckley.* Term-weighting approaches in automatic text retrieval. *IPM*, 1988.
8. *R. Losee, A. Bookstein, and C. T. Yu.* Probabilistic models for document retrieval: A comparison of performance on experimental and synthetic databases. In *Proceedings of 9th ACM SIGIR Conference on Research & Development in Information Retrieval*, 1986.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ГНОСЕОЛОГИЧЕСКИЙ И ОНТОЛОГИЧЕСКИЙ АСПЕКТЫ

Козлов А.Е.

Тихоокеанский государственный экономический университет

Гносеологический анализ проблемы искусственного интеллекта вскрывает роль таких познавательных орудий, как категории, специфическая семиотическая система, логические структуры, ранее накопленное знание. Они обнаруживаются не посредством исследования физиологических или психологических механизмов познавательного процесса, а выявляются в знании, в его языковом выражении. Орудия познания, формирующиеся, в конечном счете, на основе практической деятельности, необходимы для любой системы, выполняющей функции абстрактного мышления, независимо от ее конкретного материального субстрата и структуры. Поэтому, чтобы создать систему, выполняющую функции абстрактного мышления, т. е. в конечном счете, формирующую адекватные схемы внешних действий в существенно меняющихся средах, необходимо наделить такую систему этими орудиями [1]. Развитие систем искусственного интеллекта за последние десятилетия идет по этому пути. Однако степень продвижения в данном направлении в отношении каждого из указанных познавательных орудий неодинакова и в целом пока незначительна [2].

В наибольшей мере системы искусственного интеллекта используют формально-логические структуры, что обусловлено их неспецифичностью для мышления и, в сущности, алгоритмическим характером. Это дает возможность относительно легкой их технической реализации. Однако даже здесь кибернетике предстоит пройти большой путь. В системах искусственного интеллекта еще слабо используются модальная, императивная, вопросная и иные логики, которые функционируют в человеческом интеллекте и не менее необходимы для успешных познавательных процессов, чем давно освоенные логикой, а затем и кибернетикой формы вывода. Повышение «интеллектуального» уровня технических систем, безусловно, связано не только с расширением применяемых логических средств, но и с более интенсивным их использованием [3]. Намного сложнее обстоит дело с семиотическими системами, без которых интеллект невозможен. Языки, используемые в компьютерах, еще далеки от семиотических структур, которыми оперирует мышление. Прежде всего, для решения ряда задач необходимо последовательное приближение семи-

отических систем, которыми наделяется компьютер, к естественному языку, точнее, к использованию его ограниченных фрагментов. В этом плане предпринимаются попытки наделить входные языки компьютера универсалиями языка, например полисемией (которая элиминируется при обработке в лингвистическом процессоре). Разработаны проблемно-ориентированные фрагменты естественных языков, достаточные для решения системой ряда практических задач. Наиболее важным итогом этой работы является создание семантических языков (и их формализация), в которых слова-символы имеют интерпретацию. Однако многие универсалии естественных языков, необходимые для выполнения ими познавательных функций, в языках искусственного интеллекта пока реализованы слабо (открытость) или используются ограниченно (полисемия). Современные системы искусственного интеллекта способны осуществлять перевод с одномерных языков на многомерные. В частности, они могут строить диаграммы, схемы, чертежи, графы, высвечивать на экранах кривые. Компьютеры производят и обратный перевод (описывают графики с помощью символов). Такого рода перевод является существенным элементом интеллектуальной деятельности [2]. Но современные системы искусственного интеллекта пока не способны к непосредственному (без перевода на символический язык) использованию изображений или воспринимаемых сцен для «интеллектуальных» действий. Поиск путей глобального (а не локального) оперирования информацией составляет одну из важнейших и перспективных задач теории искусственного интеллекта.

Воплощение в информационные массивы и программы систем искусственного интеллекта аналогов категорий находится пока в начальной стадии. Аналоги некоторых категорий (например, «целое», «часть», «общее», «единичное») используются в ряде систем представления знаний, в частности в качестве «базовых отношений», в той мере, в какой это необходимо для тех или иных конкретных предметных или проблемных областей, с которыми взаимодействуют системы. В формализованном понятийном аппарате некоторых систем представления знаний предприняты отдельные попытки выражения некоторых моментов содержания и других категорий (например, «причина», «следствие»). Однако ряд категорий (например, «сущность», «явление») в языках систем представления знаний отсутствует. Проблема в целом разработчиками систем искусственного интеллекта в полной мере еще не осмыслена, и предстоит большая работа философов, логиков и кибернетиков по внедрению аналогов категорий в системы представления знаний и другие компоненты интеллектуальных систем. Это одно из перспективных направлений в развитии теории и практики кибернетики.

Современные системы искусственного интеллекта почти не имитируют сложную иерархическую структуру образа, что не позволяет им перестраивать проблемные ситуации, комбинировать локальные части сетей знаний в блоки, перестраивать эти блоки. Не является совершенным и взаимодействие вновь поступающей информации с совокупным знанием, фиксированным в системах. В семантических сетях и фреймах, использующихся при представлении знаний, пока недостаточно используются методы, благодаря которым интеллект человека легко пополняется новой информацией, находит нужные данные, перестраивает свою систему знаний [3]. Еще в меньшей мере современные системы искусственного интеллекта способны активно воздействовать на внешнюю среду, без чего не может; осуществляться самообучение и вообще совершенствование «интеллектуальной» деятельности.

Таким образом, хотя определенные шаги к воплощению гносеологических характеристик мышления в современных системах искусственного интеллекта сделаны, но в целом эти системы еще далеко не владеют комплексом гносеологических орудий, которыми располагает человек и которые необходимы для выполнения совокупности функций абстрактного мышления. Чем больше характеристики систем искусственного интеллекта будут приближены к гносеологическим характеристикам мышления человека, тем ближе будет их «интеллект» к интеллекту человека, точнее, тем выше будет их способность к комбинированию знаковых конструкций, воспринимаемых и интерпретируемых человеком в качестве решения задач и вообще воплощения мыслей.

Литература:

1. *Алексеева И.Ю.* Знание как объект компьютерного моделирования. // Вопросы философии, 2007, №3, с. 42–49.
2. *Алексеева И.Ю.* Искусственный интеллект и рефлексия над знаниями. // Философия науки и техники, 2008, №9, с. 44–53.
3. *Гаврилов А.В., Канглер В.М.* Использование искусственных нейронных сетей для анализа данных. 2005. — №3 (56).

ИССЛЕДОВАНИЕ ОТНОШЕНИЙ САМООРГАНИЗАЦИИ В КОЛЛЕКТИВАХ ИНТЕЛЛЕКТУАЛЬНЫХ АГЕНТОВ

Колесников А.В., Листопад С.В.

*Калининградский филиал Института проблем информатики
Российской академии наук*

Системы поддержки принятия решений (СППР), в отличие от их несовершенных компьютерных вариантов — лучший на сегодняшний день способ решения сложных, практических задач. Поиск ответов в подобных задачах станет возможным, если имитировать в памяти компьютера работу СППР по самоорганизации, синтезируя из имеющегося в распоряжении науки инструментария и используя новый метод всякий раз, когда подобные проблемы возникают. Коллективное принятие решений в СППР — поиск компромисса под управлением лица, принимающего решения (ЛПР), поэтому результат работы коллектива во многом определяется опытом, знаниями и личностными особенностями ЛПР. От того, насколько хорошо ЛПР сможет в ходе обсуждения устанавливать одни и разрывать другие отношения в среде участников, зависит возникновение синергетического эффекта объясняющего высокое качество коллективных решений в сравнении с решениями отдельных экспертов.

Доклад посвящен созданию интеллектуальной компьютерной системы поддержки принятия решений (КСППР) с самоорганизацией и мягкими вычислениями оценок сплоченности группы экспертов на основе анализа их целей. Чтобы установить в каких же группах: единомышленников или конкурентов — вероятность возникновения этого эффекта выше, разработан метод расчета оценок сплоченности группы экспертов.

Метод рассмотрим на примере многоагентной системы (МАС), как вычислительной модели коллективных решений и основанной на взаимодействии агентов — автономных сущностей с собственными целями, воспринимающая внешние воздействия и влияющая на окружающую среду [1]. Обобщая [2, 4], цель g агента, как субъекта управления — состояние st объекта управления, имеющее для агента субъективную ценность, определяющее его деятельность за определенный промежуток времени. Состояние st объекта управления описывается набором его свойств $P = \{p_1, \dots, p_n\}$, где n — количество свойств описывающих состояние объекта управления, с использованием аппарата теории нечетких множеств. Нечеткая цель fg агента — нечеткое множество с функцией принадлежности $\mu^{fg}(st)$, $st \in ST$. Функция

принадлежности $\mu^{fg}(st) = \mu^{fg}(p_1, \dots, p_n)$ принимает значения на множестве $[0; 1]$, при этом, чем выше ее значение, тем ближе состояние объекта управления st к цели агента fg .

Рассмотрим случай, когда состояние st описывается одним свойством p , то есть $\mu^{fg}(st) = \mu^{fg}(p)$. Исходная информация — цели интеллектуальных агентов (экспертов). Результат — тип архитектуры МАС по степени взаимодействия агентов в виде строки текста.

В первую очередь, цели агентов представляются нечеткими целями, исключаются из рассмотрения агенты, чья нечеткая цель — константа. Тогда для каждой пары агентов A и B вычисляется мера сходства целей:

$$s(A, B) = 0,5 \cdot \left(\left(\int_{p^{\min}}^{p^{\max}} \mu_A^{fg}(p) d(p) \right)^{-1} + \left(\int_{p^{\min}}^{p^{\max}} \mu_B^{fg}(p) d(p) \right)^{-1} \right) \cdot \int_{p^{\min}}^{p^{\max}} \min(\mu_A^{fg}(p); \mu_B^{fg}(p)) d(p), \quad (1)$$

где p^{\min} и p^{\max} — минимальное и максимальное значение свойства p , соответственно, и $\mu_A^{fg}(p)$ и $\mu_B^{fg}(p)$ — функции принадлежности нечетких целей агентов A и B , соответственно.

На основе значения меры сходства для каждой пары агентов рассчитывается значение лингвистической переменной «отношение» [3]:

$$cl = \langle \beta, T, U, Pr_1, Pr_2 \rangle, \quad (2)$$

где $\beta =$ «отношение» — наименование лингвистической переменной; $T = \{\text{«конкуренция»}, \text{«нейтралитет»}, \text{«сотрудничество»}\}$ — множество ее значений, каждое из которых представляет собой наименование отдельной нечеткой переменной; $U = [0; 1]$ — область определения нечетких переменных, входящих в определение лингвистической переменной; $Pr_1 = \emptyset$ — синтаксическая процедура, описывающая процесс образования новых термов из элементов множества T ; $Pr_2 = \{\mu_{\text{«конкуренция»}}(s) = (1+(3 \cdot s)^8)^{-1}, \mu_{\text{«нейтралитет»}}(s) = (1+(6 \cdot (s-0,5))^8)^{-1}, \mu_{\text{«сотрудничество»}}(s) = (1+(3 \cdot (s-1))^8)^{-1}\}$ — семантическая процедура, ставящая в соответствие каждому терму множества T , а также каждому новому терму, образуемому процедурой Pr_1 , осмысленное содержание, посредством формирования соответствующего нечеткого множества. В результате, должна быть получена матрица **CL** значений переменной cl «отношение». Затем для каждой пары конкурирующих (нейтральных) агентов ищется агент, сотрудничающий с обоими агентами.

В матрице **CL** отношения между этим агентом и каждым из конкурирующих (нейтральных) агентов заменяются на «нейтралитет». По модифицированной матрице **CL** выбирается один из типов архитектуры МАС: МАС с сотрудничающими агентами; состоит только из сотрудничающих и нейтральных агентов, в которой полностью отсутствуют отношения конкуренции; МАС с нейтральными агентами, где присутствуют только нейтральные отношения; МАС с конкурирующими агентами, где есть хотя бы одна пара агентов, между которыми установлено отношение конкуренции (в таких МАС могут быть также нейтральные и сотрудничающие агенты, при наличии сотрудничающих агентов они могут рассматриваться как «суперагент», тогда все агенты такой МАС будут конкурирующими или нейтральными). В результате в КСППР происходит идентификация архитектуры МАС, что важно для принятия решений о необходимости (или отсутствии таковой) ее модификации с целью повышения качества принимаемых коллективом решений и самоорганизации КСППР.

Реализация предлагаемого метода в КСППР позволяет рассчитать данные для последующего сравнения архитектур МАС различных типов. Результаты сопоставления различных архитектур МАС позволят определить, в каких организациях выше вероятность возникновения синергетического эффекта, когда принятое коллективом решение будет лучше, чем любое из решений, предлагаемых отдельными интеллектуальными агентами. На основе этой информации, интеллектуальная КСППР после распознавания класса коллектива агентов, сможет определить необходимые действия по совершенствованию его архитектуры, что повысит качество принимаемых с использованием КСППР решений.

Литература:

1. Multiagent systems: A modern approach to distributed artificial intelligence / edited by Gerhard Weiss. — London: The MIT Press Cambridge, 1999.
2. Козелецкий Ю. Психологическая теория решений. — М.: Прогресс, 1979. — 503 с.
3. Колесников А. В. Гибридные интеллектуальные системы. Теория и технология разработки / Под ред. А. М. Яшина. — СПб.: Изд-во СПбГТУ, 2001.
4. Сурмин Ю. П. Теория систем и системный анализ: Учеб. пособие. — Киев: МАУП, 2003. — 368 с.

СКЕПТИЧЕСКИЙ ПАРАДОКС: ОБ ОДНОМ ИЗ РЕШЕНИЙ И ПРОТИВОРЕЧИВОСТИ ЗАДАЧИ

Ламберов Л.Д.

Уральский государственный университет им. А.М. Горького

В процессе коммуникации мы производим материю (колебания воздуха, пятна чернил), но полагаем, что за ней стоит определенное содержание. Значения, ментальное содержание и «работа» языка должны, как это представляется, быть «открыты» нашему взгляду [1]. Проблематичным остается вопрос о том, являются ли значения репрезентациями чего-то экстралингвистического и отношениями между языковыми выражениями о нашем окружении [2]. Бесспорным, как представляется, остается тот момент, что язык нормативен по своему характеру (должно употреблять термин «собака» только по отношению к определенным объектам). Из нормативного характера языка и невозможности установления универалистских оснований значения некоторые философы заключали о несуществовании фактов о значении [3]. Коротко основной момент скептического парадокса состоит в следующем.

Предыдущий опыт употребления любых знаков (например, значка «+») не является достаточным для того, чтобы закрепить за данный знаком определенное значение. Хрестоматийный пример, рассматриваемый, С. Крипке — сложение чисел 68 и 57. Мы могли научиться употреблять знак «+» только на конечном числе примеров, всегда существуют такие числа, которые мы никогда не складывали (и даже никогда не будем складывать!). Если предположить, что мы не складывали раньше числа более 57, то мы не можем однозначно ответить на вопрос о том, чему будет равно выражение «68 + 57». В действительности, существует бесконечное количество вариантов ответа. Под значком «+» мы можем подразумевать обычную функцию сложения, тогда выражение «68 + 57» будет равно 125. С другой стороны, мы можем подразумевать под «+», например, функцию, которая принимает значение сложения чисел в том случае если пример написан на зеленой классной доске и 5 если это условие не выполняется. Либо мы можем подразумевать функцию, которая принимает значение 1 для всех аргументов больше 57, если же аргументы меньше либо равны 57, то эта функция принимает значение сложения аргументов. Следует заметить, что этот скептицизм не является скептицизмом относительно математики, это скептицизм именно о языке.

Тем не менее, факты о значениях могут существовать. Согласно теории значения как употребления, нормативность значения заключается в нашем обязательстве (нашем должновании) следовать правилам употребления, но это последнее не является ни самим значением, ни даже какой-либо его частью. Это последнее — скорее наша прагматическая приверженность определенным принципам, которые гарантируют взаимопонимание носителей языка, если мы желаем принадлежать определенному сообществу — мы употребляем слова строго определенным образом. Следование правилу подразумевает желание того, чтобы другие говорящие могли понять сказанное. Поэтому, чтобы удовлетворить свое желание говорящий обязан употреблять выражения определенным образом. А это вовсе не предполагает, что само значение обладает какой-либо внутренней нормативной природой.

Так, факты о значении могут существовать [2], например, в качестве корреляции субъектных диспозиций к употреблению слов определенным образом в определенных обстоятельствах. И данный подход выдерживает аргументацию С. Крипке против понимания фактов о значении как диспозиций. В действительности, отождествление фактов о значении с фактами корреляции субъектных диспозиций не подпадает под критику С. Крипке, так как факты о значении в данной интерпретации понимаются не как сами диспозиции, а как их гарантированное прагматическими соображениями некоторого данного сообщества сходство. Данная интерпретация скептического аргумента, по мнению В. А. Ладова [4], не может, как то полагает П. Хорвич, считаться прямым решением, но оно вполне подходит в качестве скептического. Сам же В. А. Ладов предлагает использовать для *прямого* решения скептической проблемы понятие математической индукции. По его мнению, из множества фактов об употреблении некоторого выражения (например, математического) можно по принципу математической индукции вывести обоснованность приписывания *одного определенного* значения всем выражениям некоторой группы. Здесь, однако, следует заметить, что данное решение не может считаться убедительным без дополнительного доказательства того, что система, в которой формулируется обобщение, не является ?-противоречивой. Так как если она является таковой, то формулирование указанного обобщения приведет к тому, что наша система станет противоречивой в полном смысле [6], что приведет, в конце концов, в тому, что в ней будет доказуемо любое утверждение. В отношении этого можно возразить, что рассуждения о формальных системах не применимы при рассуждении о естественном языке. Но тогда как такое формальное понятие как математическая индукция применима к естественному языку?

Однако, дело в том, что сам С. Крипке приводит подобное решение этой проблемы, данное Л. Витгенштейном (в действительности, как кажется, данная проблема использовалась Л. Витгенштейном для критики традиционной точки зрения на язык и обучение языку). Коротко говоря, так как не может существовать индивидуальный язык, так как все мы являемся членами сообщества, то и слова мы употребляем сообразно общим правилам. Однако, если возможно такое скептическое решение, то должны существовать и факты о значении, а именно — схожесть субъектных диспозиций к употреблению. Последнее же не замечается С. Крипке, который из данной проблемы пытается вывести, что фактов о значениях не существует вовсе. Но так как то, что «скептическое» решение проблемы возможно, влечет за собой то, что факты о значении существуют, получается, что, во-первых, С. Крипке ошибается в своей интерпретации проблемы, поставленной Л. Витгенштейном, как скептической, а, во-вторых, фактически, формулирует свою «скептическую» проблему противоречивым образом. Следует отметить, что, конечно, факты о схожести субъектных диспозиций употребления могут быть недоступны нашему познанию, но последнее, однако, не означает, что фактов о значении не существует. В данном ключе аргументация несколько похожа на контраргументы Д. Дэвидсона против неопределенности перевода. Коротко говоря, контраргументы Д. Дэвидсона сводятся к тому, что тот, кто утверждает, что перевод невозможен, должен иметь критерий правильности перевода, но последнее предполагает, что ему есть с чем сравнивать, и поэтому перевод возможен (он уже, фактически, есть у того, кто утверждает, что предъявляемый ему перевод неверен). С. Крипке последовательно рассматривает кандидатов на факты о значении, но также последовательно отбрасывает их. Откуда, спрашивается, он может знать, какими должны быть факты о значениях, если он приходит в конце концов к выводу, что фактов о значениях не существует?

Литература:

1. *Dummett M. What Is a Theory of Meaning? (I)* // Dummett M. *The Seas of Language*. — N. Y.: Oxford University Press, 1996. — pp. 1–33.
2. *Horwich P. Meaning*. — N. Y.: Oxford University Press, 1998. — 256 pp.
3. *Kripke S. Wittgenstein on Rules and Private Language*. — Cambridge, Mass.: Harvard University Press, 1984. — 160 pp.
4. *Ладов В.А. Иллюзия значения: проблема следования правилу в аналитической философии*. — Томск: Изд-во Том. ун-та, 2008. — сс. 326.
5. *Tarski A. Some Observations on the Concepts of ω -Consistency and ω -Completeness* // Tarski A. *Logic, Semantics, Metamathematics: Papers from 1923 to 1938*. — Oxford: Clarendon Press, 1956. — pp. 279–295.

ПОДХОДЫ К ИЗУЧЕНИЮ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Мазуров П.В.

Тверской государственной технической университет

В области искусственного интеллекта (ИИ) решается ответственная задача: специалисты в этой области пытаются не только понять природу интеллекта, но и смоделировать интеллектуальные сущности [1, 2]. В создании ИИ выделяют два направления исследований. Предметом первого направления являются системы, которые «думают» и «действуют», подобно людям. Наглядным примером может служить разработанная А. Ньюэллом и Г. Саймоном программа GPS («General Problem Solver» — универсальный решатель задач). В междисциплинарной области когнитологии совместно используются компьютерные модели, взятые из искусственного интеллекта, а также экспериментальные методы, взятые из психологии, для разработки точных и обоснованных теорий работы человеческого мозга.

Второе направление исследований связано с системами, которые «думают» и «действуют» рационально. Своими корнями это направление уходит в философское наследие Аристотеля. В основе этих систем лежит предположение Аристотеля, о том, что законы мышления управляют работой ума, на основе чего получила свое развитие логика. В XIX столетии ученые, работавшие в области логики, создали точную систему логических обозначений для утверждений о предметах любого рода, которые встречаются в мире, и об отношениях между ними. К середине 60-х гг. XX века были уже разработаны программы, которые могли в принципе решить любую разрешимую проблему, описанную в системе логических обозначений. Исследователи в области искусственного интеллекта, придерживающиеся так называемых традиций логицизма, надеются, что им удастся создать интеллектуальные системы на основе подобных программ. Но при осуществлении указанного подхода возникают два серьезных препятствия. Во-первых, довольно сложно взять любые неформальные знания и выразить их в формальных терминах, требуемых для системы логических обозначений, особенно если эти знания не являются полностью достоверными. Во-вторых, возможность сравнительно легко решить проблему «в принципе» отнюдь не означает, что это действительно удастся сделать на практике. Даже такие задачи, в основе которых лежит несколько десятков фактов, могут исчерпать вычислительные ресурсы любого компьютера, если не используются опреде-

ленные методы управления тем, какие этапы проведения рассуждений должны быть опробованы в первую очередь [3].

Между учеными, развивающими названные направления, конечно существуют разногласия. Подход, опирающийся на изучение человека, представляет собой область науки, развитие которой происходит по принципу выдвижения гипотез и их экспериментального подтверждения. Тогда как второй подход, основанный на понятии рациональности, представляет собой сочетание математики и техники. Вместе с тем, исследователи работающие в рамках разных подходов, взаимодействуют друг с другом.

Литература:

1. Искусственный интеллект: междисциплинарный подход / Под. ред. Д.И. Дубровского и В.А. Лекторского. — М.: ИИнтелЛ, 2006. — 448 с.
2. *Рассел С., Норвиг П.* Искусственный интеллект: современный подход. — М.: Изд. дом «Вильямс», 2006. — 1408 с.
3. *Ясницкий Л.Н.* Введение в искусственный интеллект. — М.: Академия, 2005. — 176 с.

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ДВА ВЗГЛЯДА НА ПРОБЛЕМУ

Михайлов Д.В.

Таганрогский филиал ОАО НИИ Системотехники

В двадцатом веке осуществилась давняя мечта человека: человечество вышло в космос. Казалось, что и поставленная совсем недавно задача создания искусственного интеллекта (ИИ) близка к решению. Однако это впечатление оказалось ошибочным. Вот уже срок восемь лет, как человек побывал на орбите, но искусственный интеллект все еще остается делом неблизкого будущего. Причем нельзя даже приблизительно сказать, сколько времени пройдет до момента его появления.

В чем же причина этого? Почему человечество всего через пятнадцать лет после страшной войны сумело выйти в космос, а сейчас, в несравненно более спокойных условиях, не может решить задачу создания ИИ? Причин множество, но в данной статье хотелось бы остановиться на одной из них — отсутствии четкой постановки задачи.

Прежде, чем понять, что же такое ИИ, необходимо уточнить понятие интеллекта естественного. Брокгауз и Ефрон отождествляют интеллект и ум и связывают его с познанием. Словарь Ушакова дает

схожее определение, дополняя его противопоставлением интеллекта чувствам и воле. Статья в большой советской энциклопедии гласит: «Интеллект, способность мышления, рационального познания, в отличие от таких, например, душевных способностей, как чувство, воля, интуиция, воображение и т.п.». Наконец, словарь общей психологии связывает интеллект с решением задач и приспособлением к текущим условиям обитания. Таким образом, можно выделить следующие основные признаки интеллекта: возможность рационального познания и решения задач, связанных с поддержанием своего существования. Можно пойти чуть дальше и дополнить эти признаки следующими: возможность на основе полученной информации создавать нечто принципиально новое и постановка задач.

Действительно, человек всю свою историю занимается тем, что изменяет окружающую среду (к лучшему или к худшему — это другой вопрос), привносит в нее те элементы, которых в ней не было, ставит перед собой все более сложные задачи. Следует заметить, что практически вся эта деятельность, так или иначе связана с решением проблемы выживания индивида или, в некоторых случаях, популяции в целом. Получение огня, изобретение колеса, письменности — все это прикладные задачи. Даже такая, казалось бы, далекая в древности от практики задача, как астрономия, несла также и практическую ценность: календарь, составленный с помощью астрономических знаний, помогал египетским жрецам предсказывать разливы Нила, а это уже прямая связь с сельским хозяйством.

Какой можно сделать вывод? Интеллект человека всегда был прикладным и лишь изредка использовался в целях, никак не связанных с решением практических задач.

Теперь посмотрим, что понимают под ИИ в наше время. Здесь можно выделить две большие «школы мысли». Первая из них называет ИИ инструментом для решения узкого круга задач: распознавание изображений, игра в шахматы, вождение автомобиля и т.д. Вторая же считает ИИ неким подобием человеческого интеллекта. Соответственно и цели, к достижению которых стремятся представители этих двух «школ», работая над ИИ, различаются. Первые создают еще один инструмент, который принципиально не отличается, скажем, от автомобиля, самолета или даже молотка. В чем принципиальное отличие того же молотка от программы, распознающей текст? Ни в чем — и ту, и другую задачу человек может решать и без инструмента, хотя это и сопряжено с определенными неудобствами. Оба инструмента, выполняя свой процесс, не ставят перед собой задач, их определяет человек. Можно возразить: программа, управляющая автомобилем, скажем, принимает решения, прокладывая маршрут машины. Но речь идет о том, что до тех пор, пока она не выбирает начальную и конеч-

ную точки маршрута, в плане разумности она принципиально не отличается от самого автомобиля, который может, допустим, блокировать зажигание, определив повышенное содержание алкоголя в крови владельца. Если принять такое определение ИИ, то кризиса в работах по его созданию не наблюдается: принципиальный подход к каждой из конкретных задач определен, работа ведется, уже есть практические результаты. Например, в 1997 году компьютер Deep Blue обыграл в шахматы Г. Каспарова.

Сложнее с критериями второй «школы». Фактически, предлагается создать систему, которая будет так же, как и человек, не только решать поставленную задачу, но и ставить задачи перед собой самостоятельно. Не только повторять то, что было до нее (как сейчас действуют, скажем, шахматные программы), но и создавать нечто принципиально новое, как это делал и делает человек. И, наконец, заниматься всем этим не для развлечения, не потому, что человек ввел такую программу, а для поддержания собственного существования. То есть ИИ в том смысле, который в это понятие вкладывают сторонники второй «школы» должен иметь нечто вроде инстинктов самосохранения и размножения. Именно это и будет служить мотивом его действий: сбора информации, постановки задач и их решения. В этом кроется одна из причин того, что ИИ в таком понимании не только не создан, но и не обещает появиться в ближайшем будущем. Получается, что нужно не просто научить робота думать, но и мотивировать его. Без этого ИИ останется на уровне программы для игры в шахматы — будет хорошо решать однажды поставленную человеком задачу. Рискну предположить, что решение этой задачи в ближайшее время не произойдет и, как следствие, ИИ «второго рода» не появится.

Плохо ли это? Отнюдь. Практическая польза от ИИ «первого рода» очевидна — это еще один инструмент, облегчающий жизнь человека, помогающий ему в работе. А какова практическая польза от гипотетического пока изобретения ИИ «второго рода»? Избавление человека от рутинной работы? С этим справится и ИИ «первого рода». Работа в условиях, опасных для человеческой жизни? Для этого даже не обязателен ИИ, достаточно надежной системы дистанционного управления. Колонизация дальнего космоса? В любом случае, для создания колоний необходимы люди, а с задачей сбора информации может справиться ИИ «первого рода». А вот определенный вред ИИ «второго рода» в случае своего создания нанести сможет: одной из доминант его поведения станет выживание и размножение, а для этого ему потребуются ресурсы, которых не всегда хватает и самому человечеству.

Единственное преимущество ИИ «второго рода» над ИИ «первого рода» — универсальность. Однако практика показывает, что про-

фессионалы востребованы более, чем универсалы. Что делает перспективу создания ИИ «второго рода» весьма туманной.

Литература:

1. Джефф Хокинс, Сандра Блейкли. Об интеллекте. Пер. с англ. — М.: ООО, «И.Д. Вильямс», 2007.
2. Чернухин Ю.В. Искусственный интеллект и нейрокompьютеры. Таганрог: Изд-во ТРТУ, 1997.
3. Плахов А.А. О проблеме создания искусственного интеллекта. <http://www.membrana.ru/articles/readers/2002/12/16/210400.html>
4. Общая психология. Словарь. Под ред. А.В. Петровского. М.: ПЕР СЭ, 2005.

ИНТЕРСУБЪЕКТИВНОСТЬ ИНТЕЛЛЕКТА КАК ОБЪЕКТ КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ

Оботурова М.А.

Вологодский государственный педагогический университет

Одним из актуальных направлений когнитивных наук на современном этапе развития является решение проблемы моделирования форм представления знания естественного интеллекта в искусственном, а также «получение нового знания методом компьютерного моделирования» [5, с. 5].

Решение этой задачи носит междисциплинарный характер, предполагает необходимость не только частно-научного, но и философского осмысления проблемы, так как понимание сущности, природы сознания, субъекта познания является одним из центральных эпистемологических философских вопросов.

В связи с этим, философия, по справедливому утверждению Л.А. Микешиной, «становится дисциплиной, сопричастной когнитивной науке, экспериментальной деятельности, осуществляемой, в частности, при разработке программ искусственного интеллекта» [6, с. 49]. Именно на стыке философии и общенаучных когнитивных проблем оказалась возможной «проверка самых тонких и абстрактных гипотез о природе человеческого разума» [там же], ибо для того, чтобы создать модель естественного интеллекта в искусственном, необходимо, прежде всего, адекватное понимание сущности самого субъекта познания.

В отличие от парадигмы изоляционизма классической эпистемологии, рассматривающей субъекта познания автономно, вне взаимосвязи с окружающей средой, на современном этапе развития когнитивных наук человек познающий рассматривается как «расширенный субъект», содержащий в себе как «естественные объекты, так и культурные артефакты (в том числе язык, миф, науку и т.д.); соответственно познание и сознание должны быть поняты в рамках именно этого «расширенного» субъекта, а не индивидуального организма» [4, с. 19].

В современной теории познания субъект рассматривается не изолированно, а контекстуально, так как он «включен в сам объект жизни общества, формы культуры, гуманитарные и социальные науки, виды искусства, религии» [7, с. 58]. Поэтому в современной эпистемологии и когнитивных исследованиях речь идет не просто о субъективности процесса познания, а о его интерсубъективном характере и специфике моделирования.

Интерсубъективность понимается в современной теории познания как независимость сознания, опыта субъекта о мире только от личностных особенностей и ситуаций. В связи с этим, проблема моделирования естественного интеллекта в искусственном связана, прежде всего, с пониманием того, «каким образом индивидуальное сознание выходит к опыту другого «Я», а через это — к универсальному горизонту опыта» [2, с. 170].

Сознание познающего субъекта, с точки зрения Д.И. Дубровского, является четырехмерным, оно «есть реальность, которая включает параметры знания, ценности и активность» [1, с. 59]. Кроме того, процесс познания человеком мира включает в себя не только явные, рациональные, но и неявные, иррациональные формы, возможность «умолчания», то есть «невербализации информации, лишь предполагаемой, подразумеваемой или даже неосознаваемой субъектом речевой деятельности...» [6, с. 51]. Понимание этого препятствует созданию упрощенных трактовок сознания, которые абсолютизируют одну из его сторон.

В результате этого, проблема моделирования естественного интеллекта в искусственном не сводится только к воспроизведению формально-логических способностей, но и контекстуально-игрового, коммуникативного, ситуационного характера сознания, языка. Поэтому представление знания с помощью таких когнитивных структур как схемы, карты, фреймы, в отличие от формально-логических понятий, позволяют «осуществить обоснование нового видения знания, отличное от традиционных гносеологических представлений о познавательной деятельности...» [6, с. 50]

Так, в теории фреймов М. Минского ставится проблема моделирования ситуационно-смысловой интерсубъективной структуры представления знаний [8]. При таком подходе модель внешнего мира представляет собой не набор абстрактных, независимых друг от друга понятий, а совокупность стереотипных, повторяющихся, различных, но близких ситуаций. Часто и устойчиво проявляющие себя ситуации представлены в верхней части графовой структуры фрейма, а неявные, случайные, неустойчивые — в его нижнем уровне. Поэтому «понимание» в искусственном интеллекте представляет собой процесс конкретизации неявно заданных узлов (терминалов) и согласование возможных понятий с определенной ситуацией.

Таким образом, проблема моделирование интерсубъективного естественного интеллекта в искусственном предполагает решение весьма сложных задач, так как сознание человека определяется не только его индивидуальностью, но и укорененностью в бесконечное многообразие внешних связей и отношений.

Литература:

1. *Дубровский Д.И.* Основные категориальные планы проблемы сознания // Вопросы философии. — 2008. — № 12. — С. 59–75.
2. *Калиниченко В.В.* Интерсубъективность // Современная западная философия: Словарь. — М.: ТОН — Остожье, 2000. — С. 170–171.
3. *Касавин И.Т.* Текст. Дискурс. Контекст. — М.: Канон+, 2008. — 543 с.
4. *Лекторский В.А.* Философия, искусственный интеллект и когнитивная наука // Искусственный интеллект: междисциплинарный подход. Под ред. Д.И.Дубровского и В.А.Лекторского. — М.: ИИнтелЛЛ, 2006. — С. 12–21.
5. *Макаров В.Л.* Получение нового знания методом компьютерного моделирования // Там же. — С. 5–11.
6. *Микешина Л.А.* // Новые информационные технологии и судьбы рациональности в современной культуре // Вопросы философии. — 2003. — №12. — С. 47–52.
7. *Микешина Л.А.* Трансцендентальные измерения гуманитарного знания // Вопросы философии. — 2006. — №1. — С. 49–66.
8. *Минский М.* Фреймы для представления знаний: Пер. с англ. — М.: Энергия, 1979. — 151 с.

НЕКЛАССИЧЕСКИЙ ПОДХОД К ИССЛЕДОВАНИЮ СОЗНАНИЯ В ЭПИСТЕМОЛОГИИ РАДИКАЛЬНОГО КОНСТРУКТИВИЗМА

Плужникова Н.Н.

Волгоградский государственный педагогический университет

Традиционная эпистемология заложила механистический взгляд на человека и его сознание. Согласно традиционной редукционистской модели в эпистемологии вся целостность (полнота) знания определяется количеством обработанной информации. Классический наблюдатель как бы «расщепляет» реальность на отдельные элементы, которые существуют в его сознании как «качества» отдельных вещей. Такой подход в постнеклассической (эволюционной) эпистемологии, согласно отечественному исследователю М.В. Бурцеву, можно обозначить как «обсервационизм» [1]. В данной методологии субъект признается единственным гарантом объективной реальности, организовывающий и структурирующий действительность в контурах своего мышления. Истоки данной методологической позиции можно проследить в психофизическом дуализме Р. Декарта и механистических концепциях XVIII века.

Так, Декарт, разделив тело и душу, собственно разбил реальность, как указывает американский исследователь сознания Д. Деннет, на два мира: на мир «*res cogitans*» и мир «*res extensa*», то есть на сферу физического и сферу интенционального [4, 102]. При этом сознание выступило как физическая структура, отражающая в себе «объективно» существующий мир. Сознание превратилось в «площадку» для сбора «качеств» внешнего мира. Человеческое сознание не изобретает, а комбинирует идеи, поступающие через органы чувств. Все наше знание о мире исходит из органов чувств и помещается в резервуаре, именуемым сознанием.

При этом механистический взгляд на мир не исчезает вместе с механистическими концепциями. Можно сказать, что XX век, как никакой другой, пережил воодушевление в попытках построить человека-робота по модели «воспринимающей статуи» Э.Б. Кондильяка. Так, отечественный исследователь А.А. Грудинкин описывает одну из новейших разработок чувствующего робота «*Kismet*», который способен не только распознавать предметы, но и адекватно/неадекватно реагировать на возможных собеседников: «однако изъясняет он эти чувства не потому, что ощущает их. Нет, так велит программа, вложенная в него. В отличие от таких сложных, умных существ как кошка или собака, этот робот — с его лукавыми глазками и подмигивающими уша-

ми — все равно иллюзорное, механическое создание. В его чертах по-прежнему нет жизни» [3].

Переосмысление редукционистской модели познания и неклассический подход к природе сознания появляется с развитием теории систем и квантовой механики в середине XX века. На базе данных наук появляется новое междисциплинарное научное течение — радикальный конструктивизм (П. Вацлавик, Э. фон Глазерсфельд, Г. Бейтсон). Основным постулатом радикального конструктивизма становится тезис о том, что любое знание активно конструируется субъектом. При этом под субъектом может подразумеваться наблюдатель (человек), организм или живая система. Радикальность же конструктивизма, по мнению Э. фон Глазерсфельда, состоит в том, что он отграничивает себя от классической эпистемологии в вопросе о соотношении знания и объективной действительности. Его позицию можно назвать шокирующей классическую эпистемологию, поскольку она подрывает основы «метафизического реализма» [2, 67].

Радикальный конструктивизм имеет много общего с квантовой механикой. В первую очередь это касается вопроса интерпретации реальности. В 70–80-е гг. XX века ко времени возникновения радикального конструктивизма общая теория систем получает свое окончательное оформление благодаря развитию экономико-математического моделирования, неравновесной термодинамики, синергетики, теории операций, теории катастроф и концепции сложности.

Привнесение качественно нового системного взгляда на окружающий мир изменило научные представления о природе знания, о способах функционирования сложных биологических и социальных систем. Существенным методологическим сдвигом явилось и новое сложное понимание реальности как целостного множества иерархически связанных между собой систем. Эта точка зрения объединяет научно-теоретические положения радикального конструктивизма и некоторые выводы из теории квантовой механики (Н. Бор, Луи де Бройль, В. Гейзенберг, Э. Шредингер, П. Дирак). Согласно радикальному конструктивизму, реальность творится в результате измерения и осознания наблюдателем результата измерения. При этом такие измерения действительности могут создавать совершенно разные, равноправные реальности, существующие параллельно друг другу. Альтернативных миров столько, сколько результатов измерения. При этом сознание наблюдателя как бы рассеяно между созданными им самим измерениями. Это положение объединяет квантовую механику с тезисом радикального конструктивиста Г. Бейтсона о существовании системного разума. При этом и в квантовой механике и в ради-

кальном конструктивизме утверждается, что наблюдатель ограничен в познании, поскольку индивидуальное сознание воспринимает происходящее так, будто существует лишь одна альтернатива, в которой он живет. Человек видит лишь ограниченную часть мира или, согласно квантовой механике, живет лишь в одной из альтернативных реальностей.

Так, немецкий философ Курт Хюбнер пишет, что, несмотря на значительные сдвиги в познании мира, универсум все еще по большей части мыслится в соответствии с картезианской моделью: «реальность понимается как данность, состоящая из частей и их свойств, но не как из взаимоотношений между этими частями. Эта точка зрения отменяет системный взгляд на мир, что наше существование есть только возможность стать реальностью, но не сама реальность» [5, 80].

Реальность представляет собой неразрывное единство составляющих ее частей: сам наблюдатель является неразрывной частью своих наблюдений. В роли наблюдателя могут выступать как социальные субъекты, так и сама система. Именно в рамках подобного неклассического подхода радикальный конструктивизм делает попытку переосмысления таких понятий как жизнь, обучение и культура, расширяя, тем самым, предметную область квантового представления о реальности и постулата о системном характере мироздания и о природе сознания.

Литература:

1. Бурцев М.В. Эволюционно-кибернетический подход к проблеме познания [Электронный ресурс] / М.В. Бурцев. — [2004]. — <http://mbur.narod.ru/philosophy/epistref.html>.
2. Глазерсфельд Э. фон. Введение в радикальный конструктивизм // Вестник МГУ. — Сер. 7, Философия. — 2001. № 4. — С. 59–81.
3. Грудинкин А.А. Исчезнет ли человек к XXI веку? [Электронный ресурс] / А.А. Грудинкин. — [2005]. — http://www.znanie-sila.ru/online/_ovr0/issue_988.html.
4. Деннет Д. Виды психики: на пути к пониманию сознания. — М.: Идея-Пресс, 2004. — 184 с.
5. Цит. по: Rasch W. Nicklas Luhmann's Modernity. The Paradoxes of Differentiation. — Stanford, Calif.: Stanford Univ. Press, 2000. — 246 s.

ОТ ФИЛОСОФИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА К ФИЛОСОФИИ МОДЕЛИРОВАНИЯ МОЗГА

Савельев А. В.

*Уфимский государственный авиационный
технический университет*

В России и за рубежом в последнее десятилетие можно наблюдать интенсивное возрождение философствования вокруг самых современных проблем познания и воспроизведения свойств высшей нервной деятельности человека и смежных с этим вопросов философского характера [1]. Неудачи и дефицит парадигмальной новизны в попытках исследования и искусственного воплощения свойств интеллекта приводят к новым исканиям на уровне философских оснований. В последние годы во всем мире прошел целый ряд профессиональных конференций, а Комитет по философии и компьютерам (КФК) Американской философской ассоциации (АФА) поддержал создание Международной ассоциации по компьютерной науке и философии (МАКНФ). МАКНФ активно действует в Европе, Азиатско-Тихоокеанском регионе и в Северной Америке, поддерживая международные конференции, такие как ежегодные Е-КНФ для Европы, АТ-КНФ для Азиатско-Тихоокеанского региона и СА-КНФ для Северной Америки. В нашей работе аргументировано доказывается, что философские проблемы искусственного интеллекта (ИИ), фактически состоящие в попытке информационно-компьютерного воспроизведения психологизма человека как его внешних функциональных проявлений, трансформируются в новые проблемы философии моделирования мозга (ФММ). Исследованы причинно-следственные отношения появления этих соответствующих философских направлений в плане, их онтологического оправдания и содержания в них трансформированных традиционных философских проблем.

Основные из них таковы: 1. Несмотря на известную новизну, внесенную работами по ФИИ в старые представления, в целом указанная методологическая установка и соответствующая ей модель познания отбрасывают исследования по ИИ на уровень XVIII века. 2. Поскольку объектом познания является сам орган познания, чтобы познать, необходимо стандартизировать самого себя [2], т.е. познающий субъект должен элиминировать собственную индивидуальность, что является слишком очевидно невыполнимым (самосознание мешает решению проблемы «сознание — тело» [3]). Таким образом, неявно осуществляется мотив *экзистенциального моделирования* [4]. Тем бо-

лее, что трансформируя, таким образом, свою собственную субъективность во внешнюю объектно-предметную область, осуществляется как бы объективирование познающего субъекта путём такой экстернизации. 3. С развитием компьютерной техники открываются небывалые ранее возможности решения трудоёмких умственных задач и моделирования сознания и поведения человека, то есть познания собственной сущности. 4. Накопленное огромное количество нейрофизиологических, психологических и социологических знаний ещё ожидает как своего применения, так и философского осмысления и обобщения.

Полагается, что неклассический этап характеризуется активным применением идей феноменологии и герменевтики. В этом случае — постнеклассическим этапом можно считать переход к изучению внутренней структуры мыслящей материи с целью информационного воспроизведения, возможности отделения от субстрата, либо модифицирования, т.е. к ФММ. В докладе проанализированы основные направления, тенденции и парадигмы поднимаемых фундаментальных философских вопросов вокруг последних изменений в понимании моделирования мозга (ММ), мозговой деятельности, высших психических функций человека: **1.** Разработка и применение идей конструктивизма, дальнейшее их развитие и критико-полемические дополнения в ФММ. Общность взглядов этого направления выражается в утверждении неотделимости наших действий от представлений, причём, представления играют детерминирующую роль. В то же время, уходя от внешнего во внутреннее, упомянутое направление в попытке уйти от неизбежной субъектности преобразованием её в объектность, следует (или определяет?) новейшим тенденциям ИИ структуралистского характера с переводом внимания на внутреннюю структуру мыслящего субъекта. Утверждая, что структурные изменения обусловлены их собственной динамикой, выстраивается попытка решения хайдеггеровской проблемы фундирования. По словам Е. Глазерфельда: “Конструктивизм... ничего не говорит и говорить не должен о том, что может или не может существовать. С конструктивистской точки зрения знание не воссоздаёт “картину” мира и никак этот мир не репрезентирует” [6]. **2.** Попытки разрешения дихотомических проблем сознание — мозг, субъект — объект, материальное — идеальное, структура — функция, синтаксис — семантика, социальное — природное и т. д. в ФММ. Данный круг проблем традиционно для философии (несмотря на тысячелетия своей давности) продолжает, так или иначе, присутствовать практически во всех ведущих направлениях исследований по философии познания, ИИ, НИТ, моделирования мозга, высших функций человека. Довольно популярный спектр направлений попыток разрешения проблем

конструктивизма в эпистемологических основаниях философии ИИ и НИТ сводится к усилиям модификации либо устранения дихотомических проблем, либо их философии как, например, деконструктивное опровержение феноменологии Д. Деннетом, как необходимого условия преодоления дуализма [6]. **3.** Продолжение применения классических идей герменевтики и феноменологии в моделировании мозга, сознания и нейроинформационных технологиях. Проанализировав ограничения, свойственные когнитивным моделям существующих НИТ, можно сделать мета-методологические выводы о необходимости и возможной плодотворности большего присутствия бытия объекта в Хайдеггеровском смысле в когнитивной схеме представления нейрона. Повышение онтологичности, существенности в существовании и, наоборот, существования в существенности, реализуемых в виде усложнения модели нейронного элемента и увеличения таким образом его весовой функциональности, его функциональной плотности [123], может привести к новому витку усложнения возможно-решаемых задач и повышению эффективности решения известно-решаемых задач. **4.** Субстанциональные проблемы сознания. Всё более широкое обсуждение в последнее десятилетие обнаруживает поднимаемый вновь на новом уровне трансформированный комплекс вопросов, обозначенный нами как вопросы о субстанциональности сознания. Отчасти это стимулировано огромным возрастом информационной мощности обрабатывающих средств и тотальной информатизацией в связи с этим. Возникает достаточно интенсивное обсуждение принципиально ранее не звучавших вопросов отделимости сознания от его материального субстрата. Несомненным следствием этого является появление наиболее радикальных направлений, постулирующих возможность отделения сознания (души) от материального носителя и перемещения на другой материальный носитель, причём, число публикаций на эту тему лавинообразно нарастает. Большую популярность, особенно, за рубежом приобретают последнее время рассуждения на темы искусственной личности, искусственной жизни, искусственного общества и «философских зомби»: Д. Сёрл (отсутствие интенциональности), Д. Деннет (неполная интенциональность), Р. Кирк (не обладают самосознанием), Т. Моуди, О. Флананган, Т. Полджер, Т. Нагель, Д. Чалмерс, Дж. Поллок (артифакты, не обучаются), Н. Блок (не владеют феноменальным опытом), Р. Сталнейкер, С. Брингсйорд и др.

Сделаны обобщающие выводы о наступлении нового этапа в ФИИ, который можно обозначить как **философия моделирования мозга (ФММ)**.

Работа выполнена при поддержке РФФИ (грант № 07-06-11003).

Литература:

1. *Савельев А.В.* О конференциях по философии искусственного интеллекта // <http://www.sciteclibrary.ru/texts/rus/stat/st2042.htm>, 2008.
2. *Савельев А.В.* Эпистемология самопознания в нейрокомпьютерной парадигме // *Философия науки*, 2007, № 3(34), с. 42–59; http://www.philosophy.nsc.ru/journals/philsience/3_07/02.pdf.
3. *Макгинн К.* Что конституирует проблему сознание — тело? // В материалах IV Российского философского конгресса, М.: МГУ, 2005, т. 1, с. 263–264.
4. *Алексеева И.Ю., Петрунин Ю.Ю., Савельев А.В.* Философия моделирования как мета-методология в нейроинформационных технологиях и искусственном интеллекте // *Вестник МГУ, серия 7 Философия*, 2007, № 2, с. 47–61.
5. *Glaserfeld E. von Radical Constructivism. A Way of Knowing and Learning.* L., 1996, P. 114–115.
6. *Dennett D.C.*, 1995. The Practical Requirements for Making a Conscious Robot, http://www.ecs.soton.ac.uk/~harnad/Papers/Py_104/dennettrob.html

**ЛОГИКА СУЩНОСТЕЙ Э. ГУССЕРЛЯ В ГЕНЕЗИСЕ ТЕОРИИ
ИСКУССТВЕННОГО ИНТЕЛЛЕКТА****Скутин А.С., Сокол В.Б.***Тюменская государственная академия
мировой экономики, управления и права*

В последние 25–30 лет в связи с бурным ростом исследований в области искусственного интеллекта получили мощный импульс развития традиционные проблемы языка, мышления и сознания. Произошло становление когнитивной парадигмы, тематически оформились проблематика и основные направления исследований, которые сосредоточились вокруг трех основных идей:

- 1) фундаментальная идея о том, что мышление — это манипулирование ментальными репрезентациями типа планов, сценариев, моделей и пр.; это — сугубо рациональное понимание природы мышления, которое вскоре показало свою очевидную ограниченность;
- 2) единство знаний во взаимосвязи с эмоциями, чувствами и мотивациями; неаналитические способы обработки информации, человеческие навыки;
- 3) формы представлений и механизмы так называемой образной информации.

Эти три направления, несмотря на их значительные отличия, существуют в рамках единой концепции когнитивной организации, основные параметры которой следующие: открытость когнитивных систем; приобретение нового знания путем усвоения внешних схем и правил и оперирование ментальными репрезентациями, имеющими аналоги в реальности.

Вместе с тем, существуют и принципиально иные стратегии когнитивной организации. Прежде всего, это концепция организации живых систем, развитая чилийским ученым У. Матураной. Живые системы — это когнитивные системы, которые замкнуты и самореферентны: с одной стороны, организация живой системы определяет область всех ее возможных взаимодействий, в которые она может вступать без утраты собственной идентичности, с другой стороны — идентичность сохраняется до тех пор, пока единство взаимодействий системы не нарушено. С этой точки зрения появление естественного языка не является принципиально новым моментом в становлении человека, а представляет собой просто дальнейшее расширение его когнитивной области. В такой системе основная функция языка заключается не в передаче информации и осуществлении референции к независимым от него сущностям, а в *ориентации ориентируемого в его собственно когнитивной области*. Для слушателя важно не само содержание сообщения, а те операции внутри когнитивной области, которое оно вызывает. На самом деле, слушатель *сам создает информацию* и формирует те значения, которые призваны обеспечить его оптимальное «сопряжение» с окружающей средой.

В этом свете по-другому видится и сам процесс познания в целом. С точки зрения традиционного когнитивного подхода, познание — формирование новых ментальных репрезентаций, отображающих реальность. Согласно же Матуране, знать — значит *вести себя адекватным образом*, сохраняя идентичность живой системы. Матурана указывает, что основной функцией языка является коннотация, а не денотация, как привыкли полагать: «...никого и никогда нельзя убедить рациональными доводами в истинности того, что, в конечном счете, не присутствует неявно в комплексе верований этого человека...».

Таким образом, происходит корректировка исходной гипотетической установки, идущей еще от античности, о том, что возможно однозначное разграничение на объективный мир физических явлений и субъективный ментальный мир мыслей и чувств. В соответствии с таким пониманием, роль познания сводится к «сбору информации» и построению ментальных моделей. В последнее время этому традиционному представлению противопоставляется ряд концепций принципиально иного подхода к сущности познания. Кроме идей упомянуто-

го У. Матураны, уже в начале XX века к похожим выводам пришел Мартин Хайдеггер в своей экзистенциальной герменевтике. По Хайдеггеру, наиболее значимым для понимания мира и человека является каждодневная практика, до-рефлексивный опыт. Окружающий мир не может быть жестко ограничен от человека, поскольку он структурирован посредством его целей, обстоятельств и намерений. Мы всегда находимся внутри, а не вне той или иной ситуации — отмечали Хайдеггер и позднее Г.-Х. Гадамер — и поэтому не можем иметь о ней чисто объективного знания.

Истоки же идей Хайдеггера и Гадамера, как известно, связаны с феноменологией Э. Гуссерля, обращение к которой может задать новое продуктивное измерение теории искусственного интеллекта. Гуссерль, пытаясь обозначить принципы «строгой науки», видел причину «кризиса европейских наук» в ущербности «логики фактов», предзаданной европейской научной рациональности еще Аристотелем. Тезис Гуссерля в том, что: «... истины относительно чистых сущностей не содержат ни малейших утверждений касательно фактов, а следовательно, из них *одних* невозможно извлечь даже и самой незначительной истины, относящейся к фактам» [1, с. 7]. Другими словами, Гуссерль строит свою феноменологию на принципиальном различении факта и сущности: сущность имманентно содержится в фактах, ее эксплицирующих, но отнюдь не «открывает» себя в этих фактах. Знать факт не значит знать сущность явления, стоящего за этим фактом. Логика всех традиционных опытных наук строится на фактах, поэтому уводит исследование от сущности познаваемых вещей и явлений. Гуссерлевский призыв «назад к вещам» означает редукцию «мира фактов» ради обращения взора ученого на «мир сущностей», являющихся как феномен сознания. Таким образом, задается основа принципиально иной логики, «логики сущностей», на которой, по Гуссерлю, возможна истинно «строгая наука». Феноменологическая наука принципиально отвергает монополию оторвавшихся от сущностей понятий, в том смысле, что строится на высвобождении сущностных смыслов от искажающих ее слов-выражений. «Слово является нам как нечто в себе безразличное, смысл же — как то, на что «нацелены» с помощью этого слова, то, что имеется в виду посредством этого знака» [2, с. 27].

Утверждая новые принципы научной логики (логики сущностей), Гуссерль начинает с феноменологического анализа смысловых построений коммуникативного акта слушания-говорения. Уже в «Логических исследованиях», анализируя природу конституирования сознанием смысла, Гуссерль придает *первостепенное значение слушанию и говорению* как базовым факторам феноменологического бытия: «мысль должна быть выражена не просто как некоторое значение, но также со-

общена посредством извещения, что, конечно, возможно только при действительном говорении и слушании»[2, 28]. Интересно, что сначала у Гуссерля «прийти к данности» означало возможность быть увиденным («я вижу в чистой рефлексии»). Поэтому, чтобы изменить позицию сознания от естественной установки к *феноменологической*, необходимо научиться видеть незримые для *естественного* взгляда сущности вещи. Однако, когда возникает необходимость проанализировать феноменологию осознания одним видящим «другого» такого же видящего, больше кажется уместной метафора *слушания*. Интерсубъективность Гуссерля подразумевает не столько способность увидеть «другого», сколько умение *услышать* его. Естественная установка в отношении к «другому» заставляет человека слышать в словах «другого» только свои собственные мысли, феноменологическая же позиция сознания возможна только в искусстве *вслушиваться* в голос и мысли «другого». Открытость «другому», готовность *расслышать* его является внутренним истоком «жизненного мира», забвением которого Гуссерль объяснял причину кризиса современной культуры, науки в частности.

Конечно же, общая парадигма традиционной европейской научной рациональности наложила свои ограничения и на сферу проблематики теории искусственного интеллекта. Поэтому обращение к феноменологическому методу и логике сущностей, несомненно, способствовало бы преодолению множества указанных выше противоречий и «тупиков», «накопившихся» внутри современной когнитивной науки.

Литература:

1. Гуссерль Э. Идеи к чистой феноменологии и феноменологической философии. Глава 1. <http://elenakosilova.narod.ru/>
2. Гуссерль Э. Логические исследования. Том 2. Параграф 8//Философско-литературный журнал «Логос» №9. — М., 1997, с. 25–41.

ЭПИСТЕМОЛОГИЧЕСКИЕ И СОЦИО-КУЛЬТУРНЫЕ ИСТОЧНИКИ ПРЕДСТАВЛЕНИЯ ОБ ИНТЕЛЛЕКТЕ

Тарасов И.П.

Саратовский государственный университет им. Н.Г. Чернышевского

Принято считать, что отличительным признаком, который конституирует привилегированное положение человека в окружающем мире, является его «разумность». Даже биологическое определение вида человеческого содержит в себе это прилагательное на латин-

ском языке — Homo sapiens — человек «разумный». Предполагается, что эта «субстанция» делает нас отличными от других существ, не количественно, коннотируя те или иные особенности физиологического строения человека, но прежде всего, качественно: животные не просто обладают меньшей степенью разумности по сравнению с нашей, они в принципе ей не обладают¹.

Таким образом, мы можем утверждать, что наличие у какого-либо существа сознания, мышления или разума конституирует его особую близость или родственность с человеком², т.е. если мы докажем наличие «сознания» или «разума» у некоторого существа — пусть даже искусственно созданного, то оно автоматически приобретает статус человекоподобного существа. В соответствии с этим замечанием, можно сделать вывод, что возможным продуктом, который получится, при удачной попытке создания искусственного интеллекта, будет существо (или артефакт), которое смело может быть названо гуманоидом (или включено в сообщество людей на равных с ними правах). На это указывает тот факт, что основным методологическим принципом при моделировании искусственного интеллекта является создание такого артефакта, который по интересующим нас параметрам мог бы имитировать человека³ (конечно, имитация может иметь превосходную степень). Одновременно интересующие нас параметры выступают критериями разумности или источниками, из которых исследователи формируют свои представления относительно интеллекта, который им предстоит моделировать.

Так что же нас заставляет верить в то, что человеку присуща эта загадочная субстанция, которая столь резко выделяет его из всего остального окружающего мира? Существуют несколько стандартных типов ответа на этот вопрос, каждый из которых мы кратко проанализируем:

- Сфера познания (в более узком смысле, это научное познание)

¹ Однако не все согласны, с этой достаточно распространенной точкой зрения см.: Дарвин Ч. Происхождение человека и половой отбор. М.: Издание Академии Наук СССР, 1953. с. 186-187, с. 195.

² Интересное замечание по этому поводу делает Д. Деннет, который пишет, что «корректировка границ класса носителей психики имеет большое этическое значение» // Деннет Д. Виды психики: на пути к пониманию сознания. М.: Идея-Пресс. 2004. с. 13.

³ Классическим примером может служить тест Тьюринга, который выделяет вербальные навыки, в качестве интересующих нас параметров, и связывает их наличие с существованием интеллекта см.: Turing A. Computing Machinery and Intelligence // Mind, Vol. 59, 1950. p. 434. Некоторые философы предлагают ввести т.н. полный тест Тьюринга, который состоял бы в имитации всех человеческих навыков см.: Harnad S. Other bodies, Other minds: A machine incarnation of an old philosophical problem. // Minds and Machines, Vol. 1, 1991. pp.43-54.

- Рациональные способы поведения (в последнее время говорят об адаптивных типах поведения)
- Культурная принадлежность
- Вербальные навыки.

Обратимся к первому источнику веры в существование разума — к знанию, которое человек получает с помощью него — это т.н. «истины разума» или, по-другому, научные истины. Самый простой пример — это арифметическое выражение « $2 + 2 = 4$ ». На каком бы языке мы не говорили, к какой бы культуре мы не принадлежали, для нас — это выражение является несомненной и необходимой истиной, как и многие другие положения физики, математики, логики и т.д. Необходимость, вневременность, принципиальная непогрешимость — вот основные черты научного знания, которые делают его отличным от обычных и случайных мнений людей, иногда и могущих быть истинными. Таким образом, разум с помощью применения ряда методологических процедур и правил способен получать научное знание о мире. В этом отношении разумность приравнивается к этим методологическим правилам и понимается в широком смысле как причина познания.

Следующий источник представления об интеллекте — это особый стиль поведения, его еще часто называют «рациональным», якобы мы с помощью разума, можем контролировать наши чувства, обуздывать инстинкты и поступать в соответствии с его вердиктами и предписаниями, добиваясь поставленных целей. В этом отношении разумность (или интеллектуальность) приравнивается к эффективному способу достижения целей.

Другой источник — это культурная принадлежность или включенность в определенное сообщество. Здесь важным критерием идентификации интеллекта является принятие или наличие определенных культурных ценностей, таких как религиозная принадлежность, образование, воспитание, политические взгляды и т.д.

И последняя распространенная точка зрения, интуитивно связывает наличие разума со способностью разговаривать, общаться, вступать в контакт и коммуницировать.

Первое, что мы хотели бы заметить в этой связи: все перечисленные критерии, в большей или меньшей степени релятивизированы относительно тех сообществ, в которых они существуют. Этим мы хотим сказать, что не бывает абсолютных критериев разумности, а все они социо-культурно зависимы и релятивизированы относительно определенного исторического уровня развития общества и естественно они постоянно изменяются. Методологические трудности, которые при этом возникают, указывают на невозможность строить на таком рыхлом фундаменте серьезную научную теорию. Второе заме-

чание говорит о привязанности этих критериев к человеку, и таким образом создание искусственного интеллекта всегда будет происходить с оглядкой «назад», т.е. оно будет теоретически пассивно. Таким образом, другие возможные (нечеловеческие) виды интеллекта остаются ему, недоступны.

Возможный выход, который может быть представлен, состоит в том, чтобы оценивать интеллектуальность алгоритма по достижению целей, которые перед ним ставятся⁴. Данный выход ведет к тривиализации понятия интеллекта и к его имплицитному отрицанию или к замене субстанциональной трактовки этого концепта на конструктивистскую.

Литература:

1. *Деннет Д.* Виды психики: на пути к пониманию сознания. — М.: Идея-Пресс. 2004. — 184 с.
2. *Дарвин Ч.* Происхождение человека и половой отбор. — М.: Издание Академии Наук СССР, 1953. — 1040 с.
3. *Turing A.* Computing Machinery and Intelligence // *Mind*, Vol. 59, 1950. pp. 433–460
4. *Harnad S.* Other bodies, Other minds: A machine incarnation of an old philosophical problem // *Minds and Machines*, Vol. 1, 1991. pp. 43–54.
5. *Schmidt M., Lipson H.* Distilling Free-Form Natural Laws from Experimental Data // *Science*, Vol. 324, 2009. pp. 81–85.

СТАТУС «ГЁДЕЛЕВОГО АРГУМЕНТА»

Филипповский В.А.

*Самарский государственный аэрокосмический
университет им. ак. С.П. Королёва*

История «гёделевого аргумента» против создания искусственно-го интеллекта берёт своё начало в книге Э. Нагеля и Дж. Ньюмена «Торема Гёделя» (1958), где утверждается, что основной вывод из огра-

⁴ Во многом это отражает некоторые современные тенденции в создании ИИ, которые больше не направлены на глобальное создание искусственного интеллекта, а занимаются решением локальных проблем, зачастую носящих прикладной характер. Как, например, создание «физических калькуляторов», которые могут по снятым данным с физических систем, выдавать функциональные выражения, которые эквивалентны физическим законам, которыми описываются эти системы см.: Schmidt M. Lipson H. Distilling Free-Form Natural Laws from Experimental Data // *Science*, Vol. 324, 2009. pp. 81-85.

ничительных теорем состоит в том, что «природа и возможности человеческого разума неизмеримо тоньше и богаче любой из известных пока машин» [3]. Через три года после издания этой книги, в статье Дж. Лукаса «Рассудок, машины и Гёдель» (1961) гёделевский аргумент формулируется уже более точно. Лукас открывает свою работу тезисом: «Мне кажется, что гёделева теорема доказывает, что механизм является фальшивым, т.е. что рассудок не может быть объяснён как машина» [2]. Далее эту мысль подхватывает Р. Пенроуз в книге «Новый ум короля. О компьютерах, мышлении и законах физики» (1989). Следуя Р. Пенроузу, под «гёделевым аргументом» понимается тезис об алгоритмической невычислимости функции сознания, основанием которой служат теоремы о неполноте: «... в формировании нашего сознания с необходимостью есть элементы, которые не могут быть получены из какого бы то ни было набора вычислительных инструкций; что, естественно, даёт нам веские основания считать, что сознательное восприятие — процесс существенно «невычислимый»» [4]. Среди отечественных исследователей, которые поднимают эту проблему, следует назвать, по крайней мере, Е.М. Иванова из Саратова, который ставит под сомнение «гёделевый аргумент» на некотором содержательном уровне.

Цель настоящей заметки — обратить внимание на формальный статус «гёделевского аргумента». Последний утверждается как следствие теорем К. Гёделя о неполноте. Это означает, что корректность «гёделевского аргумента» находится в непосредственной зависимости от корректности указанных теорем. Недавние результаты автора показали, что достоверность теорем К. Гёделя о неполноте должна быть поставлена под серьёзное сомнение.

Более подробное рассмотрение результатов можно обнаружить в [5, 6]. В настоящей заметке лишь кратко представлены результаты, о которых идёт речь в указанных источниках.

Перед автором стояла задача возобновить корректный разговор о теоремах, вследствие чего автор обратился к первоисточнику — статье К. Гёделя «О формально неразрешимых предложениях *Principia Mathematica* и родственных систем I» [1] (английский перевод). В результате проведённого исследования, автор пришёл к выводам, которые могут быть вкратце изложены.

1) Подлинное доказательство Первой теоремы К. Гёделя о неполноте, содержащееся в статье «О формально неразрешимых предложениях *Principia Mathematica* и родственных систем I», является некорректным в силу наличия в нём некорректно введённого определения отношения $Q(x, y)$, которое не удовлетворяет требованиям некреативности и элиминируемости, предъявляемым всякому определению.

В доказательстве Теоремы играют важную роль два импликативных отношения:

$$\overline{x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]} \rightarrow Bew_k \left[Sb \left(q_{z(x)}^{17} \quad z(y)^{19} \right) \right] \text{ и}$$

$$x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right] \rightarrow Bew_k \left[Neg \left(Sb \left(q_{z(x)}^{17} \quad z(y)^{19} \right) \right) \right].$$

Несколько модифицированные, эти отношения непосредственно влияют на вывод теоремы о существовании неразрешимых высказываний. Для того, чтобы получить эти отношения, Гёдель вводит определение произвольного отношения $Q(x, y)$, отождествляя его с $\overline{x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]}$. Применяя к дефиниендуму определения $Q(x, y)$ ранее доказанную Теорему V и ранее введённое утверждение (8), Гёдель получает некоторые импликативные отношения. При обратной подстановке дефиниенса определения $\overline{x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]}$ вместо дефиниендума $Q(x, y)$, Гёдель получает импликативные отношения, которые были указаны выше.

Определениям предъявляют два требования: элиминированности и некреативности [5]. Требование некреативности отождествляется с условием консервативности расширения, т.е. условием, что новое определение не добавляет к уже имеющимся доказуемым суждениям новые, ранее не доказуемые. Условие элиминированности требует, чтобы все суждения, в которые входит дефиниендум определения, были эффективно переводимы в другие суждения, в которые дефиниендум не входит. Последнее условие, в частности, требует, чтобы дефиниендум и дефиниенс были взаимозаменяемы. Для проверки некреативности и элиминированности определения отношения $Q(x, y)$ необходимо применить Теорему V и утверждение (8) к его дефиниенсу, т.е. к $\overline{x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]}$. Поскольку Теорема V устанавливает импликации из утвердительной и отрицательной форм одного высказывания, за основное отношение нужно взять его утвердительную форму. В нашем случае — это отношение $x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]$, которое является двухаргументным, первый аргумент которого — x , второй — $\left[Sb \left(y_{z(y)}^{19} \right) \right]$. В силу Теоремы V и утверждения (8) этому отношению соответствует знак отношения q со свободными переменными 17 и 19, такой, что имеют место отношения:

$$x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right] \rightarrow Bew \left(Sb \left[q_{z(x)}^{17} \quad z \left(Sb \left(y_{z(y)}^{19} \right) \right) \right] \right) \text{ и}$$

$$\overline{x B_k \left[Sb \left(y_{z(y)}^{19} \right) \right]} \rightarrow Bew \left(Neg \left(Sb \left[q_{z(x)}^{17} z_{z(Sb(y_{z(y)}^{19}))}^{19} \right] \right) \right).$$

Эти отношения — это не те же самые отношения, которые получил Гёдель. Введённое Гёделем определение отношения $Q(x, y)$ не удовлетворяет требованиям некреативности и элиминированности.

Для корректного проведения доказательства необходимо отказаться от этого определения. Однако если в доказательстве элиминировать определение отношения $Q(x, y)$ и использовать непосредственно дефиниенс этого определения, тогда в выводе не возникают противоречия, которые приводят к заключению о существовании неразрешимых высказываний. Таким образом, основной тезис Первой теоремы К.Гёделя о неполноте, остаётся необоснованным, т.е. может быть признан в качестве гипотезы, но не в качестве теоремы.

2) Доказательство Второй теоремы К. Гёделя о неполноте, содержащееся в статье «О формально неразрешимых предложениях *Principia Mathematica* и родственных систем I», является несамостоятельным: зависит от доказательства Первой теоремы. Конкретно это выражается в том, что обоснование для введения формулы $Wid\{k\} \supset Bew_k(17 Gen r)$ доказательства Второй теоремы, представляет собой ссылку на доказательство Первой теоремы. Вторая теорема является следствием Первой теоремы, а её доказательство — продолжением доказательства Первой теоремы. Это означает, в частности, что некорректность доказательства Первой теоремы влечёт за собой также некорректность доказательства Второй теоремы.

Требуется тщательное исследование теорем и уточнение их значения для науки. Это же исследование окончательно уточнит статус «гёделевого аргумента».

Литература:

1. *Gödel K. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter System I.* In: Monatshefte für Mathematik und Physik, 38. PP. 173–198; English translation: Gödel K. On formally undecidable propositions of Principia Mathematica and related systems I. In: Frege and Gödel: Two Fundamental Texts in Mathematical Logic. Cambridge, Massachusetts, 1970. PP. 87–108.
2. *Lucas J.R. Minds, machines and Gödel.* In: Minds and Machines, Prentice-Hall, 1954. P. 43.
3. *Нагель Э., Ньюмен Дж.Р.* Теорема Гёделя. М., 1970. С. 60.
4. *Пенроуз Р.* Новый ум короля. О компьютерах, мышлении и законах физики. М., 2008. С. 14.
5. *Смирнов В.А.* Логические методы анализа научного знания. М., 2002, С. 264.

6. Филипповский В.А. Доказательство Второй теоремы К.Гёделя о неполноте. // В мире научных открытий, №4, 2009. С. 62–65.
7. Филипповский В.А. Некорректность аутентичного доказательства Первой теоремы К.Гёделя о неполноте. // Наука. Философия. Общество. Материалы V Российского философского конгресса. Том 1. Новосибирск, 2009. С. 136–137.

ОГРАНИЧЕНИЯ ИНФОРМАЦИОННЫХ МЕТОДОВ

Чечулин В.Л.

Пермский государственный университет

Структура отражения действительности в сознании человека, формирующаяся поуровнево при переходе от одного психологического возраста к другому, требует удвоения образа действительности [1].

Вместе с тем, в теории алгоритмов имеется следующий сильный результат [2]: «Для всякого алфавита A может быть указан такой нормальный алгоритм U над A , что невозможен нормальный алгоритм в A , эквивалентный U относительно \langle этого же алфавита $\rangle A$.

В качестве такого алгоритма можно, например взять удваивающий алгоритм над алфавитом A , т.е. такой что нормальный алгоритм U над A , что $U(P)=PP$, где P — слово в A .»

То есть, алгоритм удвоения слова в алфавите A обязательно содержит буквы вне этого алфавита (по крайней мере одну), что означает, что полное удвоение образа действительности имеющееся в сознании, неосуществимо путём некоторого алгоритма, — алгоритмически нереализуемо, что показывает ограниченность информационно-алгоритмических методов в представлении процесса познания, и тем самым их ограниченность в процессе приложения к системам содержащим неотъемлемо самого человека, для управления такими системами.

Эти ограничения аналогичны ограничениям приложения предикативных формальных систем к описанию сложных процессов действительности (теореме Гёделя о неполноте предикативной формальной системы [3]).

Таким образом, и с математической стороны показана созерцательно очевидная ограниченность информационно алгоритмических методов для описания высшего уровня деятельности человека — реализации свобод, предполагающей наличие самоприменимых ценно-

стных суждений, находящихся на самом глубоком уровне отражения окружающего мира в сознании.

Литература:

1. *Чечулин В.Л.* О гносеологических основаниях 6-ти стадийного научно-инновационного цикла // Инновации РАН — 2008, материалы научно-практической конференции РАН, Н. Новгород, 2008, сс. 51–52.
2. *Нагорный Н.М.* К усилению теоремы приведения теории алгоритмов // Доклады Академии Наук СССР, 1953 г., т. 90, №3, сс. 341–342.
3. *Чечулин В.Л.* О приложениях семантики самопринадлежности // Вестник Пермского университета, сер. Математика. Механика. Информатика, вып. 3 (29), 2009 г., сс. 10–17.

ПРОБЛЕМА СУБЪЕКТА В ФИЛОСОФИИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Ястреб Н.А.

Вологодский государственный педагогический университет

Философия искусственного интеллекта представляет собой направление в современной философии, связанное с осмыслением наиболее общих проблем в моделировании познавательных способностей человека в искусственных системах. Одной из основных проблем этой области является проблема субъективности искусственного интеллекта.

В истории философии понятие субъекта рассматривалось со стороны различных подходов. Для античности была характерна онтологическая трактовка субъекта. Он понимался как носитель субстанциальных свойств и характеристик, способный определять качественные особенности объектов. Начиная с Нового времени, становится распространенным гносеологический подход, в рамках которого субъект понимался как разум, осуществляющий процесс познания. В философии сознания субъект связывается с «чистым сознанием», интенциональность которого вносит индивидуальный смысл в бытие объектов. В постмодернизме субъект рассматривается не как творец, а как комбинатор объективных элементов, что нашло отражение в идее «смерти субъекта».

С нашей точки зрения, субъект может быть определен через понятие двойного отражения, а ключевой категорией может выступать как познание, так и сознание. В зависимости от выбора категории,

под субъектом будем понимать *познающую систему, способную познавать не только внешние объекты, но и свой процесс познания, либо систему, осознающую себя и осознающую сам факт осознания*. Данные формулировки нетождественные, и их введение ставит актуальную проблему — соотношение когнитивных компетенций и субъективной реальности.

Применительно к искусственному интеллекту понятие субъективности можно сформулировать, исходя из сильной и слабой его трактовки. С позиции *слабого ИИ* он представляет собой искусственную систему, способную выполнять операции, результаты которых аналогичны отдельным интеллектуальным действиям человека. Тогда «*слабой субъективностью*» будем считать актуальную или потенциальную способность искусственной системы выполнять когнитивные операции.

Вторая трактовка, *сильный ИИ*, подразумевает некоторую искусственную систему, способную не только выполнять отдельные операции, но и понимающую и осознающую то, что она делает. Следовательно, «*сильная субъективность*», применительно к искусственным системам, есть способность выделять объект познания, выбирать методы и средства, самостоятельно ставить цели, *понимать и осознавать сам факт познания*. Далее имеется в виду сильная трактовка

С учетом данного определения, можно поставить ряд вопросов, связанных с проблемой субъективности в философии искусственного интеллекта. Прежде всего, может ли в принципе искусственная система обладать субъективностью (в сильной трактовке)? На данный вопрос есть несколько, иногда противоположных, ответов. Противники приписывания искусственным системам субъективности обращают внимание, что машины программируются и способны выполнять только те алгоритмы, которые заложил человек; имеют не такую физическую организацию, как мозг человека, следовательно, не могут обладать человеческими когнитивными способностями, хотя и могут достаточно успешно имитировать субъективность (например, тест А. Тьюринга, китайская комната Д. Серля). В то же время, если рассматривать субъективность как одну из функций интеллекта, в основе которой лежат информационные процессы, а их можно воспроизводить на любой физической основе, то ее моделирование становится возможным. Кроме того, существует естественный субъект (человек), следовательно, нет запрета на возникновение «искусственного субъекта» в принципе.

Другой важный вопрос связан с тем, каким образом можно определить, обладает ли система субъективностью, понимает ли она то, что делает. Здесь возникает проблема критериев оценки субъективности вообще. Почему мы считаем, что другие люди обладают субъек-

ективностью, т.е. понимают или осознают свои действия? Мы оцениваем их внешнее поведение и приписываем субъективность по аналогии с собой (антропоморфизм). По-видимому, мы можем это предполагать, но не знать. Данная проблема схожа с проблемой возможности наличия субъективности у животных. Традиционно считалось, что ни сознания, ни субъективности у животных нет, однако в современной этологии имеют место попытки опытной проверки, например, эксперименты с узнаванием животными себя в зеркале. В результате, наличие субъективности у животных и возможность ее определения в принципе не отрицается.

Ключевой вопрос состоит в том, обладают ли современные системы искусственного интеллекта субъективностью (в какой-либо трактовке)? С достаточной степенью уверенности можно приписать многим искусственным системам субъективность в слабом смысле, так как они способны осуществлять отдельные когнитивные операции. О субъективности в сильном смысле в отношении искусственного интеллекта, без его связи с человеком, речи пока не идет, однако отрицать принципиальную возможность этого нельзя.

Проблема возможности субъективности искусственных систем неизбежно приводит к вопросу об их взаимодействии с человеком и об отношении человека к ним. По-видимому, потребуются выстраивание новых видов социальных отношений правовых, этических норм, регламентирующих поведение «естественных» и искусственных интеллектов.

Литература:

1. *Дубровский Д.И.* Сознание, мозг, искусственный интеллект: сб. статей. М.:, 2007. 272 с.
2. *Кочергин А.Н.* Искусственный интеллект и мышление // *Философия искусственного интеллекта. Материалы всероссийской междисциплинарной конференции*, г.М., МИЭМ, 17-19 января 2005г. — М.: ИФ РАН, 2005. — С. 37–39.
3. *Новейший философский словарь: 2-е изд., переработ. и дополн.* — Мн.: Интерпрессервис; Книжный дом. 2001. — 1280с.
4. *Пенроуз Р.* Новый ум короля: О компьютерах, мышлении и законах физики: Пер. с англ. / *Общ. ред. В.О. Малышко.* — М.: Эдиториал УРСС, 2003. — 384 с.
5. *Серль Дж. Р.* Сознание, мозг и программы // *Аналитическая философия: становление и развитие.* М., 1998
6. *Тьюринг А.М.* Может ли машина мыслить?: С приложением статьи Дж. Фон Неймана «Общая и логическая теория автоматов» / *Перевод с английского: Ю.А.Данилов. Редактор и автор предисловия С.А. Яновская.* — М.: Физматгиз, 1960. — 112 с.