

Задача создания Общего искусственного интеллекта и проблема сознания

Д. И. Дубровский, Институт философии РАН, Москва, Россия

В статье рассматривается проблема создания Общего искусственного интеллекта – такой системы искусственного интеллекта, которая приближается к функциональным возможностям естественного интеллекта. Важно отметить ведущую роль Сбербанка в актуализации и организации программы исследований для разработки этой проблемы, имеющей для нашей страны стратегическое значение. Успешная реализация этой программы предполагает решение принципиальных методологических вопросов, требующих участия философов – специалистов в области эпистемологии и методологии науки. Выясняется соотношение понятий «искусственного» и «естественного», «сильного» и «слабого», «общего» и «узкого» интеллекта. Рассматриваются теоретические трудности четкого определения свойств общего интеллекта и способов его практической реализации. Показано значение исследований сознания для разработки общего искусственного интеллекта. На первом плане здесь стоят вопросы использования результатов феноменологического анализа субъективной реальности, ее ценностно-смысловых и операциональных структур. Эти вопросы подробно рассматриваются, так как имеют прямое отношение для построения новых когнитивных архитектур, которые позволяют выйти за пределы «узкого» искусственного интеллекта и создать общий искусственный интеллект – способный обладать высокой степенью автономности и самостоятельного решения широкого комплекса задач в разных условиях внешней среды.

Ключевые слова: естественный интеллект, искусственный интеллект, общий интеллект, узкий интеллект, мышление, сознание, субъективная реальность, операциональные динамические структуры субъективной реальности, факторы самоорганизации

Дубровский Давид Израилевич – доктор философских наук, профессор, главный научный сотрудник Института философии РАН

ddi29@mail.ru

www.dubrovsky.dialog21.ru

Для цитирования: Дубровский Д.И. Задача создания Общего искусственного интеллекта и проблема сознания // Философские науки. 2021

DOI:

The task of creating an Artificial General Intelligence and the problem of consciousness

D. I. Dubrovsky, Institute of Philosophy of the Russian Academy of Sciences, Moscow, Russia

Abstract

The article deals with the problem of creating General Artificial Intelligence - such an artificial intelligence system that approaches the functionality of natural intelligence. It is important to note the leading role of Sberbank in updating and organizing the research program for this problem, which is of strategic importance for our country. Its successful implementation involves the solution of fundamental methodological issues that require the participation of philosophers - specialists in the field of epistemology and methodology of science. The relationship between the concepts of "artificial" and "natural", "strong" and "weak", "general" and "narrow" intelligence is clarified. The theoretical difficulties of clearly defining the properties of general intelligence and ways of its practical implementation are considered. The importance of consciousness research for the development of general artificial intelligence is shown. In the foreground are the issues of using the results of the phenomenological analysis of subjective reality, its value-semantic and operational structures. These issues are considered in detail, since they are directly related to the construction of new cognitive architectures that allow to go beyond the "narrow" artificial intelligence and create a general artificial intelligence - capable of having a high degree of autonomy and independent solution of a wide range of tasks in different environmental conditions.

D. I. Dubrovsky, doctor of philosophy, professor, chief researcher of the Institute of Philosophy of the Russian Academy of Science

ddi29@mail.ru

www.dubrovsky.dialog21.ru

For citation: Dubrovsky D.I. (2021) The task of creating an Artificial General Intelligence and the problem of consciousness. Russian Journal of Philosophical Sciences = Filozofskie nauki

Основные цели и задачи развития в нашей стране искусственного интеллекта (далее – ИИ) четко обозначены в «Национальной стратегии развития искусственного интеллекта на период до 2030 года» (утвержденной Указом Президента Российской Федерации от 10 октября 2019 года). В настоящее время развитие ИИ подошло к новому стратегическому рубежу, который обозначается как задача создания **Общего ИИ** (в принятой международной номенклатуре – **AGI**, сокращенное обозначение от: **Artificial General Intelligence**). Эта задача соответствует мировому тренду развития ИИ в условиях нарастающей конкуренции между крупнейшими научными центрами и специализированными корпорациями, – а в более широком масштабе, - между Россией и такими лидирующими в области ИИ государствами, как США, Китай, Япония, страны Западной Европы. Факт наличия острой конкуренции надо подчеркнуть, ибо Россия пока существенно отстает от своих главных конкурентов. Догонять их по основным сложившимся направлениям развития ИИ – бессмысленно.

Нам нужен прорыв! Прорыв на принципиально новый уровень развития ИИ, о чем говорил Президент России В.В. Путин. Всем понятно стратегическое значение такого прорыва, от которого зависит наше экономическое развитие, решение проблем безопасности, многих важных социальных вопросов.

Но возможно ли это? Теоретически такой прорыв (несмотря на нынешнее наше положение) возможен в силу того, что проблемное поле ИИ в высшей степени многомерно. В нем зреет сейчас множество мыслимых и немыслимых пока новаций разного рода и разного порядка. Вопрос в том, как своевременно диагностировать и использовать или создать такие из них, которые позволили бы добиться выхода на принципиально новый уровень разработки Общего интеллекта

(AGI), Ведь наша страна обладает высоко развитой научно-технологической базой, сильной математической школой для таких исследований, располагает рядом крупных достижений и преимуществ в решении теоретических, организационных и практических проблем развития ИИ, которые могут быть успешно использованы в нынешней конкурентной ситуации.

Фундаментальный прорыв означает формирование новой теоретической и методологической основы ИИ, разработку новых подходов и методов, которые позволили бы создать AGI, т.е. систему, приближающуюся по своим функциональным возможностям к естественному интеллекту (ЕИ).

О соотношении понятий «Общего» и «Узкого», «Сильного» и «Слабого», «Естественного» и «Искусственного» интеллекта

Вначале необходимо особо отметить, что в актуализации, организации, стимулировании исследований в области AGI первостепенную, ведущую роль сыграла активная позиция Сбербанка России, которому удалось создать на свои средства мощную научно-исследовательскую команду и вычислительную инфраструктуру в области ИИ и робототехники. В рамках этой структуры, объединяющей ведущих специалистов, были созданы экспертные советы, на которых подробно обсуждались ключевые проблемы ИИ и AGI, ежегодно проводится одна из крупнейших мировых конференций по искусственному интеллекту «Artificial Intelligence Journey» (8 – 9 ноября 2019 г., 4-5 декабря 2020 г.), на которой постоянно выступает Президент России В.В. Путин. Недавно вышла первая у нас книга, специально посвященная основательному обзору проблематики AGI, анализу состояния ее разработки, основных трудностей и возможных способов их преодоления. В предисловии к ней Президент, Председатель Правления Сбера Герман Греф следующим образом оценивает нынешнее состояние разработок ИИ: «В большинстве своем существующие решения являются примерами реализации технологий узко специализированного искусственного интеллекта, требующего

настройки и перепроверки со стороны человека. Чтобы так же хорошо решать разнообразные комплексные задачи, как это делают люди, машины должны научиться строить причинно-следственные модели окружающей среды и ориентироваться в разных контекстах, а не просто максимизировать успех при решении какой-то узкой задачи. Они должны понимать физические, психологические и другие законы нашего мира и уметь связывать новую информацию в общую картину с тем, что уже знают. Чтобы добиться этого, нам необходимо преодолеть очередной технологический рубеж — создание Общего искусственного интеллекта или AGI» [Сильный ИИ 2021,5–6]. Герман Греф подчеркивает принципиальное значение разработки в этих целях междисциплинарных проблем AGI и формулирует стратегическую установку для достижения успеха: «Крайне важным фактором создания и развития технологий Общего искусственного интеллекта является обеспечение сквозного целеполагания между прикладными (или индустриальными) задачами, фундаментальными исследованиями и системой образования — так называемая триада Practice–Education–Research» [Сильный ИИ 2021,7]. Эти три составляющие в силу их необходимой взаимообусловленности должны быть в центре внимания государственных органов, призванных содействовать реализации программы AGI.

Рассмотрим теперь теоретико-методологические вопросы философского плана. Для создания AGI прежде всего надо преодолеть ограниченности традиционной методологии разработки ИИ, основы которой заложены А. Тьюрингом. Согласно этой методологии понятие интеллекта трактуется в сугубо функционально-операциональном смысле, исключая роль сознания, вся сложная проблематика которого выносится за скобки. Исследование сознания полагается здесь излишним: есть когнитивная или практическая задача, и она формулируется и решается чисто операциональными методами с помощью компьютерных программ.

Всем нам хорошо известно, что именно эта методология Тьюринга открыла компьютерную эру, обеспечила выдающиеся успехи ИИ, развитие информационных технологий, качественное изменение всей земной цивилизации. Однако сейчас мы вступаем в новый этап, который настоятельно требует существенных теоретико-методологических обновлений. Парадигма функционализма, в формирование которой А. Тьюрингом внесён первостепенный вклад, естественно, остаётся в силе, но она требует более широкой интерпретации, с учетом ее роли в объяснении сознания и способов использования новейших результатов исследований сознания для развития ИИ (об этом подробнее будет сказано далее).

Тенденция "реабилитации" сознания и потребность коррективов методологии А. Тьюринга отчетливо проявлялась в разных формах за последние два десятилетия. Однако в этом направлении нам предстоит серьезная теоретико-методологическая работа. Дело еще в том, что довольно большое число специалистов в области ИИ по понятным причинам сильно укоренено в традиционных методах, позволяющих и сегодня успешно решать многие практические задачи, в силу чего они убеждены, что для реализации намечаемых больших задач, достаточно совершенствование, улучшение существующих методов. Они скептически относятся к привлечению проблемы сознания, считая, что здесь множество неопределенностей, "сплошной туман", что нет никакого смысла в разработке так называемой "Трудной проблемы сознания", ибо это только запутывает дело. Убеждения такого рода проявились у некоторых видных специалистов, выступавших на экспертных обсуждениях проекта разработки AGI, предложенного Сбербанком. Такое впечатление сложилось у меня, поскольку я участвовал в этих обсуждениях и полемизировал по этим вопросам.

Однако подобные настроения присущи, конечно, лишь некоторым специалистам. Существуют разные взгляды, разные направления разработки ИИ, среди них и такие концептуальные подходы, в которых внимание обращено как раз на изучение ЕИ, моделирование его специфических свойств, создание новых когнитивных архитектур с целью сближения функциональных возможностей ИИ с ЕИ. Одним

из тех, кто считает ЕИ главным источником и ресурсом нового этапа развития ИИ (что отвечает задачам AGI), является крупный отечественный специалист В.К. Финн. Он основательно анализирует функциональные когнитивные структуры ЕИ и, главное, придает большое значение философско-методологическому подходу к решению актуальных проблем современного этапа развития ИИ, использует его в своих исследованиях [Финн 2011]. В.К. Финном создан широко известный ДСМ-метод, на основе которого автором получены весьма значительные результаты [Финн 2016]. Развитие этого направления может быть перспективным в ряде отношений и для разработки AGI. Существуют и многие другие подходы, использующие в этом плане результаты исследований ЕИ (см., например: [Кузнецов 2016]).

Несомненно, традиционные методы с учетом их совершенствования еще долго будут успешно использоваться для решения типичных прикладных задач. Но нам ведь нужен прорыв! А для этого требуются большие творческие и волевые усилия, связанные с преодолением привычных ментальных клише, с выбором (и созданием) наиболее перспективных направлений исследований, концентрации сил и средств для разработки принципиально новых теоретических и методологических концепций развития ИИ, новых подходов к решению междисциплинарных проблем.

Перед нами на первом плане оказываются проблемы теории, методологии, междисциплинарности, такие понятия как интеллект, мышление, познание и сознание, проблема и гипотеза, объяснение и понимание, соотношение эмпирического и теоретического знания, характер связи ИИ с ЕИ. Первостепенная роль в их осмыслении принадлежит философам – специалистам в области эпистемологии (теории познания), методологии науки, феноменологии, изучения сознания в широком плане, включая те направления, которые в западной номенклатуре именуется аналитической философией сознания.

Важно подчеркнуть, однако, что здесь нужен *особый вид специализации философа*, работающего на стыке философии и науки, требующий глубоких методологических знаний и теоретических навыков, особого аналитического опыта и, конечно, достаточной компетенции в соответствующей области научного знания.

Центральным пунктом тут выступает понятие интеллекта и соотношение всех его специальных модификаций, связанных с понятиями ИИ и ЕИ: сильный и слабый, общий (AGI) и узкий ИИ, особенно *соотношение AGI с ЕИ*. К этому добавляются вопросы о прото-AGI и узкого AGI, которые рассматриваются в связи со стадиями продвижения к полноценному AGI. Все это получило широкое освещение в указанной книге [Сильный ИИ 2021], которая, подчеркнем это еще раз, вносит исключительно большой вклад в разработку AGI .

Под сильным интеллектом имеют в виду человеческий интеллект, наделенный сознанием, т.е. ЕИ. Все существующие формы ИИ обычно относят к слабому интеллекту. Но все, что касается общего интеллекта (AGI) и его отношения к ЕИ и к обычному ИИ, встречает ряд вопросов, ответы на которые истолковываются неоднозначно.

Термин «общий интеллект» появился в 1997 году, как раз когда компьютер Deep Blue обыграл чемпиона мира в шахматы. На основе понятия общего интеллекта было предложено делить ИИ не на сильный и слабый, а на *общий* (способный адаптироваться к решению разных проблем) и *узкий* (способный решать лишь определенную задачу). В 2007 г. состоялась первая конференция, специально посвященная проблемам развития общего ИИ, после чего эти проблемы стали предметом возрастающего внимания среди специалистов в области ИИ. Ведь подавляющее большинство систем ИИ являются узкоспециализированными, способными качественно решать лишь определенную узкую задачу. Поэтому со стороны критиков мы издавна слышали, что компьютер, запрограммированный на решение тысячи задач, не способен самостоятельно научиться решать тысячу первую.

В книге проводится подробное рассмотрение истории развития ИИ, выделяются ее значимые этапы и показано, что это развитие шло в направлении снижения зависимости технологий ИИ от человека, их способности к решению более широкого круга задач. Однако эти достижения все же остаются далекими от решения проблемы AGI. Авторы дают следующую общую оценку нынешнего положения дел.

«Во-первых, с точки зрения адаптивности все существующие системы ИИ на основе любых известных подходов предполагают функционирование в ограниченном наборе задач и условий и не способны самообучаться функционированию в условиях существенно новых. То есть они являются программируемыми, хотя сложность их программирования заметно снизилась, а способность к обучению возросла. Во-вторых, с точки зрения автономности систем ИИ все существующие системы не являются автономными и не могут полноценно функционировать без живого оператора, отвечающего за запуск и остановку, техобслуживание, целеполагание и определение режимов работы в зависимости от тех или иных условий или задач. То есть они остаются управляемыми» [Сильный ИИ 2021, 52]. «В-третьих, с точки зрения интегративности современные системы ИИ являются не системами, обладающими интеллектом как таковым (даже ограниченным), а системами компьютерного зрения, обработки естественного языка, анализа данных (машинного обучения), обработки символьной информации (рассуждений на основе знаний) и т. д., то есть интегративными не являются» [там же]. Тем самым четко обозначены те главные, пока отсутствующие свойства, которые необходимы для построения действительного AGI.

Но посмотрим, как решается в книге вопрос об определении AGI, его степени общности и его отношения к ЕИ. Авторы предпочитают такие определения: «Общий интеллект — это способность достигать сложных целей в сложных средах» (Бен Герцель); «Интеллект — это способность системы адаптироваться к своей среде, работая при недостаточных знаниях и ресурсах» (Пей Ванг) [там же, 28]. «Делая

акцент на широком диапазоне сред, область AGI позволяет нам избавиться от антропоцентричных предпочтений и предлагает сфокусироваться на общих решениях, пригодных для разных агентов (человека, животных, роботов, ботов и т. д.), действующих в разных условиях...»[там же, 29] (в том числе «на неизведанных планетах»). Такая широкая трактовка интеллекта теоретически правомерна, соответствует принципу изофункционализма систем и принципу инвариантности информации по отношению к физическим свойствам ее носителя (т.е. возможности одной и той же информации кодироваться разными способами, иметь различное физическое воплощение). Не исключено, что такая трактовка интеллекта может быть полезной за пределами не только антропологического, но и биологического контекста, оставляя простор для творчества технологических суперноваций, для осмысления гибридных и иных новообразований в ходе идущей нарастающими темпами антропотехнологической эволюции.

Однако возникает вопрос: почему сразу отодвигается на задний план «сильный» интеллект, наш ЕИ, как, несомненно, высшая из всех известных нам форма интеллекта, наделенная сознанием, благодаря которому мы создаем все искусственные интеллекты. Ведь именно в ЕИ безусловно имеются столь драгоценные и желательные для AGI свойства.

Авторы разъясняют свою позицию следующим образом: «Определение интеллекта в области AGI может казаться слишком абстрактным и далеким от наших представлений о естественном интеллекте. А еще широко распространено мнение, что человеческий интеллект—единственный пример интеллекта. Так почему бы не опираться на него? Даже если мы предположим, что для характеристики AGI необходимы или достаточны критерии человеческого интеллекта, описание последнего тоже основано на довольно зыбких понятиях» [там же, 30]. И в подтверждении о «зыбкости» приводится определение интеллекта из Википедии (как известно, далеко не надежного в научном отношении источника). В нем, наряду с указанием, что он является «качеством психики», перечисляется в

несистематизированном виде длинный ряд его «способностей», начиная от приспособления к новым ситуациям, до ощущений, мышления, воображения, внимания, воли и рефлексии.

Действительно, с таким определением трудно иметь дело, непонятно, каким образом можно упорядочить, соотнести между собой эти способности и использовать их для целей AGI. В итоге приводится такое, не вполне ясное заключение: «Определение интеллекта в области AGI не включает эти способности, но и не отвергает их» [там же, 31]. Однако заслуживает внимания следующий весомый аргумент: «Как повышение уровня решения узких задач вплоть до сверхчеловеческого не потребовало «сильных» качеств, так и расширение общности методов решения задач вовсе не обязательно подразумевает преднамеренное движение в сторону сильного ИИ» [там же, 32]. Под «сверхчеловеческим» имеются в виду такие действительные достижения «узкого» ИИ, как, например, победа компьютера над чемпионом мира по игре в Го, многие другие результаты, связанные с задачами анализа больших данных, быстрого действия, многих результатов использования ИИ в производственных и коммуникативных процессах.

Но далее мы читаем: «Можно предположить, что *некоторые аналоги некоторых «сильных» качеств у действительно общего ИИ должны быть.* Например, наверняка общий ИИ должен иметь ***способность к интроспекции*** — анализу ***собственных мыслительных процессов*** или даже оптимизации лежащих в их основе алгоритмов. При этом у такого ИИ будет некий ***образ себя как часть картины мира.*** Но это не обязательно означает, что у него будет самосознание в философском смысле и уж тем более личность сродни человеческой, хотя по глубине рефлексии он вполне может и превосходить человека. Наверняка он проявит «понимание» тех областей, в которых действует успешнее человека. Но это не значит, что такое понимание будет сопровождаться у него субъективными переживаниями, схожими с человеческими. Наверняка у него будет многомерная система мотивации, включающая аналоги, например, любопытства и удивления. Но

его вовсе не обязательно пытаться наделить всеми человеческими эмоциями. Хотя, скажем, для социальных роботов это может быть полезно, но даже они способны лишь симулировать чувства и эмоции, а не испытывать их » [там же, 32 – 33. *Курсив мой. Д.Д.*].

Эти суждения весьма показательны. Они свидетельствуют, что специфические свойства и функции сознания, присущие всякому нашему мыслительному процессу, должны привлекаться для построения AGI. Но возможно ли такое за вычетом состояния **субъективной реальности** – этого **неотъемлемого качества сознания**?

Теоретически, это возможно и, скорее, даже необходимо, однако лишь при условии адекватного вычленения и описания соответствующих оперативных аналогов определенных состояний субъективной реальности. Но, что имеется в виду под «интроспекцией» и «собственными мыслительными процессами» или «образом себя как части картины мира»? Эти понятия выражают специфические особенности реальной сознательной деятельности (нашего ЕИ), они являются весьма сложными, многомерными, взаимодополняющими друг друга по содержанию. В когнитивном плане каждое из них включает возможность множества аналитических и синтетических оперативных действий. Поэтому недостаточно ясно, о каком аналоге идет речь. Приведенные понятия берутся слишком абстрактно, в силу чего указания на соответствующие аналогии вряд ли могут иметь предполагаемое значение для моделирования когнитивных структур AGI.

Однако, несмотря на это, сама установка на поиск подобных аналогий заслуживает поддержки, так как обращена к задаче использования специфических свойств сознательной деятельности для разработки AGI. Здесь требуются результаты профессионального философского и специального феноменологического анализа операциональных структур субъективной реальности, допускающих корректное определение их аналогов, пригодных для построения когнитивных архитектур AGI.

Такая установка отличается от жестких высказываний о том, что общий искусственный интеллект «не лежит между слабым и сильным ИИ, а просто

находится в стороне от них и определяется без отсылки к человеческому интеллекту как ИИ, способный решать широкий круг задач» [там же, 27]. И она противостоит отмечавшимся выше случаям прямого отрицания роли сознания в разработке проблем ИИ и AGI. Интересно, что специалист по ИИ, занимающий такую позицию, демонстрирует парадоксальную ситуацию: он *как бы забывает*, что сам является носителем ЕИ и обладает искомыми свойствами AGI, постоянно использует их в своей деятельности, несмотря на то, что они скрыты для него в хитросплетениях его сознания.

Понятно, что большинство специалистов в области ИИ занимаются прикладными задачами, и они не могут быть достаточно компетентными в философско-методологических вопросах, касающихся проблемы сознания, анализа ценностно-смысловых и деятельно-волевых структур субъективной реальности; у них «хватает» своих дел, трудностей и проблем. Но сейчас важно понимание, что профессионалы, занимающиеся *теорией познания, методологией науки, феноменологией субъективной реальности* могут быть активными помощниками в общем деле, особенно в вопросах теоретического осмысления роли сознания на нынешнем этапе развития ИИ. Разумеется, многие специалисты в области ИИ отдают себе в этом ясный отчет и готовы к сотрудничеству.

Сознание и AGI. Субъективная реальность: ее общие свойства, ценностно-смысловые и интенционально-волевые оперативные структуры

Исследования сознания, взятого в его специфическом и необходимом качестве субъективной реальности, позволяют конкретно выяснить основные системные, структурные, функциональные, операциональные характеристики естественной мыслительной деятельности человека, а постольку представляют важный ресурс для моделирования когнитивных структур, отвечающих задачам построения AGI. С этой целью необходимо рассмотрение основных философских и научных положений о

сущности самого качества субъективной реальности (эти вопросы подробно освещались в ряде моих работ: [Дубровский 2002], [Дубровский 2019a] и др.). Дело в том, что это качество часто выносится за скобки теми, кто занимается задачами сугубо функционалистского, бихевиорального типа (как «сделать» что-то; в том числе, создать программу для компьютера или создать робота). В подавляющем числе таких случаев действительно нет необходимости специально выделять и анализировать качество субъективной реальности, так как внимание сосредоточено на поставленной задаче и наличном процессе деятельности; мы как бы не замечаем это качество даже когда напрягаем, корректируем собственное мышление; оно подобно воздуху, который мы тоже «не замечаем» во многих интервалах своей жизни. Качество субъективной реальности становится слишком «заметным», когда предметом исследования оказывается само мышление и сознание. Здесь возникает целый ряд принципиальных вопросов. Чтобы их четко выделить и рассмотреть, желательно использовать их нумерацию.

1. Именно качество субъективной реальности создает основные теоретические трудности проблемы сознания. Они, связаны, в первую очередь, с тем, что явлениям субъективной реальности нельзя приписывать физические свойства (массу, энергию, пространственные характеристики), и тогда возникает вопрос: как, не прибегая к идеалистическим постулатам о наличии особой духовной субстанции, построить теоретически корректное объяснение сознания с материалистических позиций, учитывая его очевидную связь с физическими процессами. Вопросы, касающиеся такого объяснения стало принято объединять под общим названием «Трудная проблема сознания». Она уже более шестидесяти лет находится в центре внимания западной аналитической философии, служит предметом острых дискуссий. Ей посвящена поистине необъятная литература (многие тысячи публикаций!). Однако нет общепринятого решения. В большинстве концепций предлагается тот или иной редукционистский подход, чаще всего в форме отождествления явлений субъективной реальности с физическими процессами (см. подробный критический

анализ этих концепций [Дубровский, 2019, 140 – 183]). В отличие от таких концепций мною был предложен нередукционистский способ решения основных теоретических вопросов «Трудной проблемы сознания» (см. [Дубровский 1971], [Дубровский 2015], [Dubrovsky 2019]), который может представлять определенный интерес и для разработки некоторых вопросов ИИ.

«Трудная проблема сознания» рассматривается в двух главных категориальных планах: *онтологическом и эпистемологическом*:

1) в *онтологическом плане* ставятся вопросы объяснения связи явлений субъективной реальности с деятельностью мозга, их способности управлять телесными процессами и, более широко, происхождения субъективной реальности, форм ее существования и роли в физическом мире (при условии, что ей нельзя приписывать физические свойства);

2) в *эпистемологическом плане* возникает проблема перехода от индивидуально-субъективного опыта, присущего только данному индивиду (выражаемого в отчетах от первого лица) к *интерсубъективным*, общезначимым утверждениям (от третьего лица) и к обоснованию истинного знания. Тем не менее, все специфические для человека мыслительные процессы, связанные с решением практических и когнитивных задач, характерных в том числе и для ИИ, осуществляются в форме субъективной реальности. Поэтому достаточно апробированные результаты ее феноменологических исследований могут служить для осмысления способов создания AGI, прежде всего для построения его когнитивных архитектур, способных преодолеть ограниченность когнитивных архитектур узкого ИИ.

С другой стороны, для проблематики AGI важное значение имеют вопросы онтологического плана, касающийся объяснения *активности сознания* (каким образом явления субъективной реальности могут служить причиной телесных, физических изменений, как объяснить способность целеполагания, свободы воли, способность моего Я управлять моими физическими действиями). Ведь для систем

узкого ИИ характерно то, что они плохо работают с причинно-следственными связями; для агента же, способного действовать в широком диапазоне сред, необходимо «понимание» причинно-следственных связей, как объективно присущих любой среде, в которой он должен учиться ориентироваться и действовать; при этом для агента важно умение различать физические причины и информационные, способность «понимать» специфику информационной причинности, видом которой являются психическая причинность. А для этого как раз и нужно иметь решение «Трудной проблемы сознания».

2. Понятие субъективной реальности обозначает всякое осознаваемое состояние в отличие от бессознательных процессов, которые существуют одновременно с ним и служат его необходимой основой. Оно охватывает как отдельные осознаваемые явления субъективной реальности и их виды (ощущения, восприятия, чувства, мысли, желания, целеполагания, волевые усилия и т.д.), так и их целостное персональное образование, объединяемое нашим Я.

Субъективная реальность *персональна*, переживается индивидом как «текущее настоящее», т.е. *сейчас*, хотя может отображать прошлое и будущее; она удостоверяет для меня реальность моего существования. Это непрерывный процесс, который временно прерывается в условия глубокого сна без сновидений, общего наркоза, обморока, в других случаях временной потери сознания (и в этих интервалах я для себя не существую). Именно состояние субъективной реальности выражает актуально протекающий процесс мышления, когнитивную и деятельную активность личности. Не говоря уже о повседневной, обыденной сознательной деятельности, все новые фундаментальные идеи в науке и культуре первично возникают, оформляются и развиваются в форме явлений субъективной реальности отдельных личностей.

3. Субъективная реальность в своем специфическом качестве, присуща не только сознанию человека, но и психике животных, о чем свидетельствует опыт общения с ними и данные зоопсихологии об их ощущениях, восприятиях и

свойственных высшим животным формам мышления (наличие качества субъективной реальности отчетливо подтверждается, например, известными опытами с воздействием галлюциногенов на собак).

Субъективная реальность – находка биологической эволюции, У первых одноклеточных организмов носителем информации и регулятором поведения служили химические процессы. Это был *допсихический уровень информационной реальности*. Следующим этапом ее развития стало зарождение психики с ее качеством субъективной реальности, возникшим у многоклеточных животных, способных активно передвигаться в постоянно изменяющейся внешней среде.

Это ознаменовало возникновение *нового типа информационного управления*, регуляции поведения, позволило решить проблему эффективной самоорганизации многоклеточного организма. Ведь его элементами являются отдельные клетки, которые также представляют собой самоорганизующиеся системы со своими весьма жесткими программами, «отработанными» эволюцией в течение многих сотен миллионов лет. Но теперь последние должны были согласовываться с общеорганизменной программой, как и наоборот.

Это – весьма сложная задача, решение которой предполагало нахождение оптимальной *меры централизации и автономизации* контуров управления, способной обеспечить сохранение и укрепление целостности сложной живой системы, т.е. такой меры централизации управления, которая не нарушает фундаментальные программы отдельных клеток, и такой меры автономности их функционирования, которая, наряду с кооперативными и конкурентными способами их взаимодействия между собой, не препятствует их содружественному участию в реализации программ управления целостным организмом.

Аналогичная проблема эффективной самоорганизации социальных и технических систем весьма актуальна, в том числе для ИИ, когда возникают подобные задачи взаимоотношения глобальных и локальных уровней управления в сложных системах, взаимодействия множества компьютеров (и роботов), человека и роботов,

автономной системы роботов и т.п. Проблема такого рода решается в сложной биологической системе *гораздо эффективнее*, чем в случаях взаимоотношений человека и общества, человека и технических устройств и в самих технических системах. Это требует внимания и изучения, когда речь идет о новых разработках в области искусственного интеллекта и интеллектуальной робототехники, связанных с аналогичными задачами эффективного управления.

Возникновение у животных субъективной реальности стало исторически первой формой *виртуальной реальности*, открывавшей по мере развития все более широкий диапазон таких способностей как абстрагирование, обобщение, прогнозирование, планирование, пробные действия в виртуальном плане и других подобных операций (в форме своего рода «мысленных экспериментов»), повышавших приспособляемость к среде. Важно обратить особое внимание на то, что именно это качество виртуальности во многом определяет способность формирования новых навыков, приспособления к изменившейся среде, т.е. по существу то главное свойство, которое требуется для AGI – способность самостоятельно решать задачи «в широком диапазоне сред».

4. В ходе антропогенеза произошло качественное развитие субъективной реальности, возникает сознание (как высшая форма субъективной реальности), а вместе с ним язык. Особенностью сознания по сравнению с животной психикой является то, что психическое отображение и управление сами становятся объектом отображения и управления. Возникает способность, по существу, неограниченного производства информации об информации, возможность высокой степени абстрагирования, резкого расширения диапазона мысленных действий, моделирования ситуаций, проектирования, целеполагания и волеизъявления, фантазирования, творческого решения задач.

Качество субъективной реальности человека открыло принципиально новый этап развития информационных процессов. Оно представляет собой два основных свойства психической деятельности: 1) *способность индивида иметь информацию*

«в чистом» виде – в том смысле, что ее мозговой носитель для нас элиминирован: когда я переживаю образ дерева, мне дана информация об этом предмете и отображение мной этой информации, т.е. знание о том, что именно я вижу это дерево; но я ничего не знаю, не чувствую, что при этом происходит в моем головном мозгу; а в нем происходят очень сложные, многоступенчатые сетевые процессы переработки информации, которые предшествуют, например, переживанию мной ощущения красного или образа бегущего ко мне человека, т. е. мне дан **итоговый результат** сложных информационных процессов именно в той **интегральной** форме («простой» и понятной), которая необходима для адекватного поведения и решения задач в данной ситуации. 2) Вместе с этим в явлениях субъективной реальности: нам дана также **способность оперировать этой информацией в «чистом» виде по своей воле** в довольно широком диапазоне (контролировать собственное Я, переключать внимание, направлять движение своей мысли и концентрировать ее на определенном объекте, давать волю своему воображению, ставить определенные цели и решать мысленно разные задачи). Это означает, что наше Я является самоорганизующейся структурой, способной к эффективной саморегуляции. Здесь возникает острый вопрос о свободе воли, который встречает большие трудности в нейронаучных исследованиях сознания, но может получить объяснение с позиций информационного подхода (см. [Dubrovsky 2019a]).

Именно данность информации в «чистом» виде и способность управлять ею выражают специфические черты субъективной реальности.

5. Динамическая структура субъективной реальности с ее многообразными способами продления и смены переживаемого контента (наличного «содержания») включает такие функциональные регистры как единство *рефлексивного–арефлексивного* и *актуального–диспозиционального*. Рефлексивное – это длящееся *сейчас* осознаваемое переживание данного контента, которое в следующий момент «уходит» в память, становится арефлексивным (заменяется другим рефлексивным контентом, но потом может снова «возвращаться»),

становится рефлекслируемым). Актуальное - это рефлексивное, но взятое в его отношениях и зависимости от *диспозиционального*, которое представляет собой ценностно-смысловые структуры, хранящиеся в памяти или на более глубоком психическом, а затем и на еще более глубоком допсихическом (генетическом) уровне неосознаваемых информационных процессов.

Диспозициональное большей частью выражает весьма устойчивые, ценностно-смысловые образования нашей психики (усвоенные нормы культуры, убеждения, проверенные опытом знания, привычки, умения, интересы, предпочтения и т.п.), которые задают направленность актуализуемых в данном интервале желаний, мыслей и действий; в некоторых же случаях такого рода направленность становится для личности непреодолимой или крайне трудно преодолимой (как например, при наркотической зависимости).

В этом плане рассмотрения регистров рефлексивное-арефлексивное и актуальное-диспозициональное для разработки ИИ представляет интерес работа нашей памяти в связи с тем, что компьютерам присуща фон-неймановская архитектура, основанная на постоянном разделении памяти и вычисления. Субъективная реальность демонстрирует принципиально другое устройство памяти, а, соответственно, способов обработки информации (среди них есть и такие, которые можно назвать вычислениями), способов ассоциативной, контекстуально значимой выборки определенного контента из памяти и замены его другим (в активированных динамических контурах рефлексивного-арефлексивного, актуального-диспозиционального). Но эти специфические для нашей психики формы информационных процессов определяются основными системно-структурно-функциональными свойствами сознания (субъективной реальности) с учетом непрерывной связи с бессознательными процессами.

6. Рассмотрим далее основные функциональные структуры субъективной реальности и присущие ей *операциональные способности*, которые могут иметь

непосредственное отношение к моделированию когнитивных архитектур, соответствующих уровню AGI.

Уже элементарный анализ раскрывает фундаментальное структурно-операциональное свойство всякого явления субъективной реальности – его *двумерность*, т.е. способность одновременного отображения в нем некоторого объекта и вместе с тем самого себя, т.е. *единство иноотображения и самоотображения*. Субъективная реальность в любых ее интервалах и проявлениях обладает в той или иной степени способностью самоотображения на уровне нашего Я. Эта **фундаментальная** способность отчетливо проявляется на уровне зрительного восприятия, что было показано в работах Дж. Риццолатти и его сотрудников в исследовании зеркальных нейронов. Зеркальные нейроны одновременно осуществляют "кодирование сенсорной информации в моторных терминах», т.е. создают единство зрительного восприятия и его вероятной моторной репрезентации. Это свидетельствует о том, что в то время как индивид наблюдает за выполнением действий другими, в его мозге выполняется потенциальный моторный акт, по всем своим характеристикам сходный с тем, который спонтанно активируется при подготовке и выполнении реального действия. Различие лишь в том, что в одном случае это действие остается потенциальным, а в другом переводится в последовательность конкретных движений " [Риццолатти., Синигалья 2012, 97]

Благодаря нашим зеркальным системам, мы способны незамедлительно распознавать определенные типы действий, производимых другими, достигать *понимания* другого. Это указывает на особенности организации зрительного образа (объединения в нем отображения предмета и вероятного способа действия с ним, а также санкционирующего акта понимания, «принятия» смысла этого действия), Такое единство трех операций способно обеспечить весьма эффективное взаимодействие между людьми. Особенно важно иметь в виду, что именно на моторных репрезентациях наблюдаемых действий, создаваемых зеркальными нейронами, *«основана способность обучения при помощи подражания»*

[Риццолатти Джакомо, Синигалья Коррадо 2012, 97]. Это обстоятельство может служить стимулом для поиска способов построения моделей саморасширения диапазона «навыков» интеллектуального робота, его способности *умножать свои «навыки»* в данном классе действий, способности столь существенной для построения AGI.

Единство иноотображения и самоотображения выражается в многообразных формах – от развитой рефлексии до элементарного *чувства принадлежности мне переживаемого мной в данный момент состояния* субъективной реальности. Зачастую чувство принадлежности находится вне фокуса осознания, слабо выражено; требует для его актуализации специального внимания. Оно особенно остро переживается лишь во время его нарушения, в патологических случаях *деперсонализации*, когда больному кажется, что переживаемые им ощущения, мысли «не его», и он впадает в паническое состояние. В более тяжелых патологических случаях (например, при параноидальном бреде) больной уже не воспринимает нарушений процесса самоотображения, «не замечает» своей деперсонализации.

Данные психопатологии и психоневрологии предоставляют исключительно ценный материал для исследования самоотображения, убедительно раскрывают *взаимозависимость иноотображения и самоотображения*, демонстрируют различные варианты и стадии его нарушения, особенно ярко в тех случаях, когда состояния деперсонализации вызывает явления *дереализации*, нарушения отображения внешних объектов и собственного тела ([Меграбян 1962], [Рамачандран 2014], [Литвак 2007], [Назлюян 2002] и др).

7. Взаимозависимость иноотображения и самоотображения проявляется в функционировании памяти (в регистрах рефлексивного и арефлексивного, актуального и диспозиционального), но особенно полно она выражена на уровне многомерной динамической структуры нашего Я, в том, что можно назвать *базисной динамической структурой субъективной реальности*. Она представляет собой *единство и переменное соотношение противоположных модальностей «Я» и*

«не-Я». Это единство представлено в каждом наличном интервале субъективной реальности, оно формирует его ценностно-смысловой каркас и деятельно-волевые векторы. В динамическом биполярном контуре «Я»–«не-Я» совершается движение «содержания» (контента) субъективной реальности. Это «содержание» способно переходить из модальности «Я» в модальность «не-Я» и наоборот (например, когда «содержание», относящееся к модальности «Я», мои личностные свойства, мои оценки себя становятся для меня объектом внимания, анализа и оценки, а, значит, выступают в данном интервале уже в модальности «не-Я» и т.п.). Соответственно свойства другого человека или определенной социальной общности могут «присваиваться» мной, переходя в модальность «Я». Такого рода взаимопреобразования, перемена модальности переживаемого «содержания» – *механизм эффективного отображения и освоения действительности*, в том числе самой субъективной реальности, ее саморегуляции, а вместе с тем механизм освоения социального опыта и осуществления творческой деятельности.

Взаимопереходы модальностей «Я» и «не-Я» постоянно сохраняют биполярную структуру субъективной реальности в любом ее интервале, не нарушая идентичности персонального Я. (исключая психопатологию) *Каждая из модальностей определяется лишь через противопоставление другой и соотнесение с ней*. Поэтому в самом общем виде «Я» есть то, что противопоставляется «не-Я» и соотносится с ним; и, наоборот, «не-Я» есть то, что противопоставляется «Я» и соотносится с ним. Первостепенный интерес для понимания операциональных структур системы Я, Самости представляют исследования выдающегося нейропсихолога Антонио Дамасио; особенно это относится к его анализу автобиографической самости [Дамасио 2018].

Базисная динамическая структура субъективной реальности раскрывается конкретнее, когда в ходе феноменологического анализа выясняются основные виды противопоставления и соотнесения «Я» с «не-Я». Если взять за систему отсчета модальность «Я», то в первом приближении «Я» выступает по отношению к «не-Я»

как отношение «Я»: 1) к внешним объектам, процессам (внешней предметности); 2) к собственному телу; 3) к самому себе; 4) к другому «Я» (другому человеку); 5) к «Мы» (той социальной общности, группе, с которой «Я» себя идентифицирует, к которой оно себя в том или ином отношении причисляет, ценности которой разделяет); 6) к «Они» (той общности, социальной группе, которой «Я» себя противопоставляет или, по крайней мере, от которой оно себя отделяет, ценности которой резко отрицает или равнодушно не принимает, не понимает); 7) к «Абсолютному» («Мир», «Бог», «Космос», «Природа» и т.п.).

Таков один из мыслимых способов выделения основных видов «содержания» (контента) «не-Я», а, следовательно, самого «Я», ибо оно полагает и раскрывает себя лишь посредством своего «не-Я». Иными словами, таковы основные смысловые (когнитивные и ценностные) измерения нашего «Я» (*категории контентов Я*). И чтобы сравнительно полно раскрыть одно из выделенных отношений «Я», нужно рассмотреть его не только само по себе, но и сквозь призму всех остальных. Так, нельзя основательно понять отношения «Я» к самому себе, если оставить в стороне отношение «Я» к предметному миру, к собственной телесности, к другому «Я», к «Мы», к «Они» и к «Абсолютному». В этой биполярной динамической контентной многомерности «Я» и полагает себя как свое «не-Я», которое выступает в форме «знания», «оценки» и «действия». В ней непрестанно совершаются процессы самоотображения и самоорганизации структуры субъективной реальности, формируется и реализуется ее активность, ее *деятельно-волевые функции*. Особо важный вопрос о связи отображения с действием, как в форме самопознания и самопреобразования сознания, так и в форме познания и преобразования внешнего мира, требует отдельного специального обсуждения.

8. Приведенные результаты феноменологического анализа динамической структуры субъективной реальности для специалистов в области ИИ, особенно для тех, кто нацелен на разработку AGI, могут быть полезны во многих отношениях. Прежде всего, это касается ясного понимания той огромной дистанции, которая

отделяет нынешний уровень ИИ от ЕИ, от тех конкретных специфических функциональных свойств естественного процесса мышления, к моделированию и реализации которых у нас пока нет технологического доступа, Но занимая реалистическую позицию, мы сохраняем веру в наши *творческие* способности добиваться шаг за шагом сближения с ЕИ, т.е. **добиваться все более глубокого самопознания и самопреобразования, ибо эти функции исходно присущи нашему ЕИ.**

На каждом этапе мы используем методологические установки, позволяющие корректно сокращать чрезвычайную многомерность субъективной реальности (в ее динамических операциональных и творческих проявлениях), выбирать и осваивать лишь некоторые специфические ее функциональные свойства, опираться на приемлемые аналогии в операциональных действиях мышления человека и интеллектуального робота. Лучше говорить в данном отношении не об ИИ в общем смысле, а об интеллектуальном роботе, поскольку он обладает не только определенной когнитивной архитектурой, но совершает физически реальные действия во внешней среде, подлежащие строгой проверке на их эффективность.

Возьмем рассмотренную выше двумерность явлений сознания, *единство иноотображения и самоотображения – необходимое условие эффективности мыслительных и практических действий человека.* Интеллектуальный робот обладает в известной мере аналогичными функциями. У него есть четко заданное иноотображение (и соответствующий набор внешних действий). Есть и нечто подобное самоотображению, которое может быть заложено в надежности его программного обеспечения и технического устройства, в механизмах слежения за их исправностью. Но здесь мы сразу видим качественное различие. У человека функция самоотображения обладает способностью *саморегуляции*, которая постоянно стимулируется, корректируется, совершенствуется процессами иноотображения, как и наоборот: последнее стимулируется, корректируется, повышает свою

результативность под воздействием процессов самоотображения, т.е. также обладают в этом смысле способностью саморегуляции.

В этом и заключается свойственное сознанию, структуре СР единство функций иноотображения и самоотображения, их органическая взаимообусловленность, которая, как было показано, особенно ярко проявляется на уровне базовой динамической бимодальной структуры субъективной реальности – в единстве и переменном соотношении модальностей «Я» и «не-Я».

Когда говорят об автономности роботов, то, конечно, хорошо осознают ее ограниченность: узость заданного диапазона восприятий и действий, неспособность к самостоятельному его расширению при существенном изменении внешних условий. Еще сильнее эта ограниченность касается функции самоотображения. В программном обеспечении робота и в его техническом устройстве нет специальных регистров текущего непрерывного самоотображения его подсистем и процессов и соответственной самонастройки, подверженных непрерывной коррекции со стороны иноотображения и внешних действий, как это имеет место у человека; Функция иноотображения и то, что условно можно было бы отнести к функции самоотображения, фактически *разделены*.

Это указывает на те проблемы, которые требуют разработки при создании интеллектуальных роботов нового поколения, т.е. достигающих уровня AGI, Думаю, для этого есть определенные ресурсы и уже наметились соответствующие тенденции, связанные с критическим взглядом на ограниченность классической методологии А. Тьюринга признанием важной роли учета феноменов сознания в разработке ИИ.

В моделировании интеллектуальных роботов могут использоваться результаты исследования сознания, что ведет к расширению «мира» робота, его «техно-умвельта», тем самым и диапазона его успешных действий (концепция известного нашего специалиста по интеллектуальной робототехнике А. Е. Ефимова [Efimov 2020]). В этом направлении, повидимому, будет возможность использовать

ресурсы, связанные с методами глубоких нервных сетей для создания программ обозначенной выше *двумерной функциональности* в поведении интеллектуального робота.

Разумеется, пунктирно намеченные возможности новых подходов нуждаются в дальнейшей конкретизации, в критическом обсуждении, но они указывают на широкую творческую перспективу использования феноменологических исследований операциональной специфики сознательной деятельности.

9. К этому следует добавить, что изложенные выше результаты феноменологических исследований субъективной реальности обращают внимание на чрезвычайную сложность «содержания» (множества различных *контентов*) нашей сознательной деятельности, взаимосвязь, взаимовлияние различных по своему виду (классу) контентов. Они были приблизительно определены выше в качестве категориальных ценностно-смысловых структур, отображающих области внешней предметности и процессуальности, телесности человека, собственного Я, понимания Я другого человека, социально и мировоззренчески обусловленных ценностно-смысловых и деятельно-волевых структур (см. их перечень выше).

Эти различные категориальные структуры *диспозиционально* укоренены в нашем сознании и на бессознательном уровне психики, они обычно вовлечены в той или иной степени во все мыслительные процессы и задают широкий по их «содержанию» *спектр операций абстрагирования, анализа и синтеза*. Эти операции существенно различаются по своим когнитивным свойствам в зависимости от того или иного категориального контента. Одно дело операции абстрагирования, анализа и синтеза при исследовании предметной области внешнего мира, другое – при исследовании нашего Я или социокультурных реалий.

Различие «онтологий» обуславливает соответственные особенности эпистемологических подходов. Тем не менее, при решении сложных теоретических и особенно практических задач естественным интеллектом обычно используется весьма широкий спектр этих операций, многие из которых протекают слитно,

включают *интуитивный уровень*, связанный нередко с новыми, находящимися в стадии формирования творческими результатами. Из всего этого многообразного спектра в системах ИИ используется лишь небольшая часть когнитивных операций, правда, в достаточной степени определенных. В этом состоит одно из главных отличий ЕИ от ИИ. Все эти обстоятельства также должны служить поводом для размышлений при разработке AGI, поскольку для него не требуется способность решения *всего разнообразия* человеческих задач (к примеру, таких как объяснение природы темной энергии и т.п.), но важно уточнять те их классы, которые соответствуют его целям и возможностям на нынешнем этапе. Среди них особую актуальность в последнее время приобретают этические проблемы развития систем ИИ, требующие специального основательного исследования.

10. Наконец, в многомерных планах проблемы сознания надо выделить еще один вопрос. Он касается не просто когнитивных операций абстрагирования, обобщения и т.д. в различных категориальных контентах, указанных выше, а учета множества *экзистенциально значимых* состояний сознательной деятельности, которые, хотя и включены в процесс познания и играют в нем существенную роль, но не могут быть четко выражены в принятых когнитивных терминах. Все они, однако, требуют учета и охвата для наличия целостного образа сознания и для полноты анализа познавательного процесса и оценки его целей, ибо в настоящее время, в условиях нарастания глобального кризиса нашей потребительской цивилизации, первостепенное значение приобретают вопросы о подлинных смыслах личной и социальной деятельности, о противостоянии нарастающему абсурду в нынешних процессах социального развития.

Вот, что пишет по этому вопросу известный феноменолог М. Мерло-Понти: «Жизнь сознания – познающая жизнь, жизнь желания, или жизнь перцептивная – скрепляется «интенциональной дугой», которая проецирует вокруг нас наше прошлое, будущее, наше житейское окружение, нашу физическую, идеологическую и моральную ситуации или, точнее, делает так, что мы оказываемся вовлечены во

все эти отношения. Эта интенциональная дуга и создает единство чувств, единство чувств и мышления, единство чувствительности и двигательной функции» [Мерло-Понти 1999, 182]. Здесь *понятие интенциональной дуги* выполняет важную интегративную роль, как средства охвата и концептуального учета всего диапазона «участников» познавательного процесса от перцептивного уровня до идеологических, моральных и других экзистенциально значимых составляющих познавательного процесса, часть которых нередко остается в тени.

М. Мерло-Понти не назвал в приведенном перечне эстетическую составляющую, *эстетическую интенцию*, которая в той или иной степени всегда присутствует в познавательных процессах (часто в неявной форме, но и тогда дающая о себе знать в виде равнодушия познающего субъекта к красоте и безобразию, к тому, что называют возвышенным и низменным, трагическим и комическим, гармоничным, изящным и т.п.). Все это приобретает высокую степень выражения и ценности в профессиональном художественном познании, в творчестве выдающихся представителей литературы и искусства. Если обратиться к интересующей нас проблеме сознания в широком плане, то надо признать, что наиболее глубокие и разносторонние знания о сознании нам даны в произведениях великих писателей и поэтов, таких как Пушкин, Толстой, Чехов и др.

Особенно хочется сказать, хотя бы кратко, о значении поэзии для эпистемологического анализа не только обыденного, но и научного мышления. Великие поэтические образы и метафоры – мощный творческий стимул развития интеллекта, так как разрушают устоявшиеся комфортные клише, решительно и убедительно пересекают границы привычных противоположных смыслов, обнаруживают неизвестные ранее степени свободы в нашем мышлении, создают в нем новые ценностно-смысловые образования и векторы активности, просветляют наш духовный мир.

Эстетическая интенция в познавательных процессах не менее важна, чем этическая интенция, и это справедливо не только для ЕИ, но и при обсуждении

проблем разработки AGI. При этом необходимо иметь в виду не только когнитивный аспект, способ решения определенной задачи, но также оценку допустимости самой задачи с позиций этических норм и принятых социальных критериев.

Особо важной и трудной темой в проблематике развития ИИ является выяснение и использование для целей AGI операций бессознательно-психического уровня и допсихического, генетически заданного уровня, на котором осуществляется управление сложнейшими процессами в организме. Биологическая эволюция выработала многочисленные информационные компетенции эффективного управления сложными системами, которые унаследованы человеком, но остаются пока крайне слабо исследованными. Яркий пример: в нашем теле смонтировано 100 000 километров кровеносных русел, и эта колоссальная динамическая система весьма эффективно управляется при постоянном изменении внешних физических условий.

В заключение хотелось бы вернуться к вопросу о чрезвычайно важной роли образования в системе мер, необходимых для создания AGI, о чем говорил в обращении к читателям книги Герман Греф. Подготовка специалистов в области ИИ обнаруживает серьезные недостатки. Как отмечают авторы книги: «Современные образовательные программы готовят инженеров, умеющих работать с существующими технологиями, а не творчески мыслящих исследователей, способных совершать новые прорывы» [Сильный ИИ 2021, 55]. В этом как раз проявляется слабость, формализм, в ряде отношений интеллектуальное убожество, нынешнего образования. Чтобы воспитывать творческую устремленность, оно должно делать акцент на фундаментальных теоретических вопросах развития ИИ, формировать у молодого инженера широкий духовный кругозор, интерес к философским и социальным проблемам познания и самопознания, к поэзии, гениальной художественной литературе, как источникам и стимулам творческого мышления.

Литература

- Дамасио 2018 – Дамасио А. Так начинается «Я». Мозг и возникновение сознания. М.: Карьера-Пресс, 2018. – 384 с.
- Дубровский 1971 – Дубровский Д. И. Психические явления и мозг. Философский анализ проблемы в связи актуальными задачами нейрофизиологии, психологии и кибернетики. М.: Издательство «Наука», 1971. – 386 с.; издание 2-е, доп., М.: ЛЕНАНД, 2020. - 400 с
- Дубровский 2002 – Дубровский Д. И. Проблема идеального. Субъективная реальность. Изд. 2-е, доп., М.: Канон+, 2002. - 368 с.
- Дубровский 2019 – Дубровский Д. И. Субъективная реальность // Философская Энциклопедия. М.: ИФ РАН, 2019.
- Дубровский 2015 – Дубровский Д. И. Проблема "Сознание и мозг": Теоретическое решение. М.: Канон+, 2015. - 208 с.
- Дубровский 2020 – Дубровский Д.И. Субъективная реальность и мозг: к вопросу о полувековом опыте разработки проблемы сознания в аналитической философии // В книге: Дубровский Д.И. Проблема сознания. Теория и критика альтернативных концепций. М.:ЛЕНАНД, 2020.С.140 – 183.
- Кузнецов 2016 – Кузнецов О. П. Избранные труды: автоматы, языки и искусственный интеллект. М., 2016. – 464 с
- Литвак 2007 – Литвак Л.М. «Жизнь после смерти»: предсмертные переживания и природа психоза. Опыт самонаблюдения и психоневрологического исследования. Изд. 2-е, переработанное и дополненное / Под ред. и со вступительной статьей Д.И. Дубровского. М.: Канон+, 2007. – 672 с.
- Меграбян 1962 – Меграбян А.А. Дегерсонализация. Армянское госуд. издательство Ереван, 1962.. – 355 с.
- Мерло-Понти 1999 – Мерло-Понти М. Феноменология восприятия. СПб.: Ювента, 1999,
- Назлюян 2002 – Назлюян Г. М. Концептуальная психотерапия. Портретный метод. М.: ПЕР СЭ, 2002. – 239 с.
- Рамачандран 2014 – Рамачандран В.С. Мозг рассказывает. Что делает нас людьми. М.: Карьера Пресс. 2014.- 422 с.
- Риццолатти, Синигалья 2012 – Риццолатти Джакомо, Синигалья Коррадо. Зеркала в мозге. О механизмах совместного действия и переживания. М.: Языки славянских культур, 2012. - 208 с.
- Сильный ИИ 2021 – Сильный искусственный интеллект: На подступах к сверхразуму / Александр Ведяхин [и др.]. М.: Интеллектуальная Литература, 2021. - 232 с.
- Финн 2011 – Финн В.К. Искусственный интеллект: методология, применения, философия. М.: КРАСАНД, 2011.
- Финн 2016 – Финн В.К. О классе ДСМ-рассуждений, использующих изоморфизм правил индуктивного вывода // Искусственный интеллект и принятие решений. – 2016. №3. С. 48-61.
- Dubrovsky 2019 – David I. Dubrovsky. “The Hard Problem of Consciousness”. Theoretical solution of its main questions //AIMS Neuroscience, 6(2): 85–103. DOI: .3934/Neuroscience.2019. 2.8.

Dubrovsky 2019 a – Dubrovsky D. I. The Problem of Free Will and Modern Neuroscience //Neuroscience and Behavioral Physiology, Vol. 49, No. 5, June, 2019 DOI 10.1007/s11055-019-00781-5

Efimov 2020 – Efimov A. Post-Turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence // Artificial General Intelligence (AGI). 2020. Vol. 12177. P. 83–94.

Koch 2016 – Koch C. and al. Neural correlates of consciousness: progress and problems // Nature / Reviews / Neuroscience/ Vol. 17 | May 2016 | 307.