

## **Задача создания Общего искусственного интеллекта и проблема сознания\***

*Д.И. Дубровский*

*Институт философии РАН, Москва, Россия*

### **Аннотация**

В статье рассмотрена задача создания Общего искусственного интеллекта – системы искусственного интеллекта, которая приближается к функциональным возможностям естественного интеллекта. Сделан акцент на ведущей роли Сбербанка в актуализации и организации программы исследований в целях разработки этой проблемы, имеющей для России стратегическое значение. Успешная реализация данной программы предполагает решение принципиальных методологических вопросов, требующих участия философов – специалистов в области эпистемологии и методологии науки. Показано соотношение понятий «искусственного» и «естественного» «сильного» и «слабого», «общего» и «узкого» интеллекта. Выявлены теоретические трудности, связанные с четким определением свойств Общего интеллекта и способов его практической реализации. Автор статьи обращает внимание на значение исследований сознания для разработки Общего искусственного интеллекта и приходит к выводу о том, что на первом плане находятся вопросы использования результатов феноменологического анализа субъективной реальности, ее ценностно-смысловых и операциональных структур. Такие вопросы освещены подробно, поскольку они имеют прямое отношение для построения новых когнитивных архитектур. Последние 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. В этой связи показана ограниченность операционалистской методологии Тьюринга, которая исключает применение результатов специальных исследований сознания как субъективной реальности. Использование таких результатов связано с разработкой посттьюрин-

---

\* Работа выполнена при поддержке РФФИ, проект № 21-18-00184 «Социогуманитарные основания критериев оценки инноваций, использующих цифровые технологии и искусственный интеллект».

говой методологии, открывающей значительные возможности для создания Общего искусственного интеллекта.

**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](mailto:ddi29@mail.ru)

<https://orcid.org/0000-0003-4392-2526>

**Для цитирования:** *Дубровский Д.И.* (2021) Задача создания Общего искусственного интеллекта и проблема сознания // *Философские науки*. 2021. Т. 64. № 1. С. 13–44. DOI: 10.30727/0235-1188-2021-64-1-13-44

## **The Task of the Creation of Artificial General Intelligence and the Problem of Consciousness\***

*D.I. Dubrovsky*

*Institute of Philosophy, Russian Academy of Science, Moscow, Russia*

### **Abstract**

The article discusses the task of creation of Artificial General Intelligence (AGI), that is, an artificial intelligence system that approaches the functional capabilities of natural intelligence. Emphasis is made on the leading role of Sberbank in updating and organizing a special research program in order to develop this problem, which is of strategic importance for Russia. It has been established that the successful implementation of this program presupposes solving fundamental methodological issues that require input of philosophers – specialists in the field of epistemology and methodology of science. We show the relations of the concepts of “artificial” and “natural,” “strong” and “weak,” “general” and “narrow” intelligence. The article reveals the theoretical difficulties associated with a clear definition of the properties of general intelligence and the ways of its practical implementation. The author draws attention to the importance of studies of consciousness for the development of general artificial intelligence and comes to the conclusion that the priority issues are using the results of the phenomenological analysis of subjective reality, its value-semantic and op-

---

\* This work was supported by the Russian Science Foundation (RSF), project no. 21-18-00184 “Social and Humanitarian Foundations of Criteria for Evaluating Innovations Using Digital Technologies and Artificial Intelligence.”

erational structures. The paper discusses these issues in detail, as they are of direct relevance to the construction of new cognitive architectures. The latter make it possible, to go beyond the limits of “narrow” artificial intelligence and create AGI with a high degree of autonomy and independent solutions to a wide range of problems in different environments. It demonstrates the limitation of Turing’s operationalist methodology that excludes the use of the results of special studies of consciousness as a subjective reality. Application of such results is associated with the development of post-Turing methodology, which opens up significant opportunities for creating an AGI.

**Keywords:** natural intelligence, artificial intelligence, general intelligence, narrow intelligence, thought, consciousness, subjective reality, operational dynamic structures of subjective reality, factors of self-organization.

**David I. Dubrovsky** – D.Sc. in Philosophy, Professor, Chief Research Fellow, Department of the Theory of Knowledge, Institute of Philosophy, Russian Academy of Sciences.

[ddi29@mail.ru](mailto:ddi29@mail.ru)

<https://orcid.org/0000-0003-4392-2526>

**For citation:** Dubrovsky D.I. (2021) The Task of the Creation of Artificial General Intelligence and the Problem of Consciousness. *Russian Journal of Philosophical Sciences = Filosofskie nauki*. Vol. 64, no. 1, pp. 13–44. DOI: 10.30727/0235-1188-2021-64-1-13-44

## Введение

Основные цели и задачи развития в нашей стране искусственного интеллекта (далее – ИИ) четко сформулированы 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, т.е. к задаче создания **Общего ИИ** (в принятой международной номенклатуре – **AGI, Artificial General Intelligence**). Эта задача соответствует мировому тренду развития ИИ в условиях нарастающей конкуренции между крупнейшими научными центрами и специализированными корпорациями, – а более широком масштабе, – между Россией и такими лидирующими в области ИИ государствами, как США, Китай, Япония, страны Западной Европы. Факт наличия острой конкуренции очевиден,

поскольку Россия пока существенно отстает от своих главных конкурентов. Догонять их по основным сложившимся приоритетным направлениям развития ИИ – бессмысленно.

Нам нужен прорыв! Прорыв на принципиально новый уровень развития ИИ, о чем говорил Президент России В.В. Путин. Для всех является понятным, в чем заключается стратегическое значение прорыва, от которого зависит наше экономическое развитие, решение проблем безопасности, многих важных социальных вопросов. Но возможно ли это? Теоретически прорыв (несмотря на наше положение в настоящее время) возможен с учетом того, что проблемное поле ИИ в высшей степени многомерно. В нем назревает сегодня множество мыслимых и немыслимых пока новаций разного рода и разного порядка. Вопрос состоит 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). Наша страна обладает высокоразвитой научно-технологической базой, сильной математической школой для подобных 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, предусматривает разработку новых подходов и методов, которые способствовали бы созданию AGI, т.е. системы, приближающейся по функциональным возможностям к естественному интеллекту (далее – ЕИ).

### **О соотношении понятий общего и узкого, сильного и слабого, естественного и искусственного интеллекта**

В актуализации, организации, стимулировании исследований в области AGI ведущую роль сыграла активная позиция Сбербанка России, которому удалось создать за счет собственных средств мощную научно-исследовательскую команду и вычислительную инфраструктуру в области ИИ и робототехники. В рамках этой структуры, объединяющей ведущих специали-

стов, созданы экспертные советы, на которых подробно обсуждают ключевые проблемы ИИ и AGI, ежегодно проводится одна из крупнейших мировых конференций по искусственному интеллекту «Artificial Intelligence Journey» (8–9 ноября 2019 года, 4–5 декабря 2020 года), где постоянно выступает Президент России В.В. Путин.

В 2021 году вышла первая в нашей стране книга, посвященная основательному обзору проблематики AGI, анализу состояния ее разработки, основных трудностей и возможных способов их преодоления. В предисловии к ней президент, председатель правления Сбербанка России Герман Греф оценивает современное 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]. 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]. Указанные три составляющие ввиду их неотъемлемой взаимообусловленности должны быть в центре внимания государственных органов, призванных содействовать реализации программы AGI.

Далее рассмотрим теоретико-методологические вопросы философского плана. Для создания AGI в первую очередь следует преодолеть ограниченности традиционной методологии разработки ИИ, основы которой заложены А. Тьюрингом. Согласно этой методологии, понятие «интеллект» трактуется в сугубо функционально-операциональном смысле, исключая роль сознания, вся сложная проблематика которого выносится за скобки. Исследование сознания представляется излишним в данном случае: существует когнитивная или практическая задача, которая формулируется и решается посредством сугубо операциональных методов с помощью компьютерных программ.

Именно методология Тьюринга открыла компьютерную эру, обеспечила выдающиеся успехи ИИ, развитие информационных технологий, качественное изменение земной цивилизации. Однако сегодня мы переходим к новому этапу, нуждающемуся в существенных теоретико-методологических обновлениях. Парадигма функционализма, в формирование которой А. Тьюрингом внесен первостепенный вклад, естественно, остается действующей, но она требует более широкой интерпретации с учетом ее роли в объяснении сознания и способов использования новейших результатов исследований сознания для развития ИИ.

Тенденция «реабилитации» сознания и потребность коррективов методологии Тьюринга отчетливо проявлялась в разных формах за последние два десятилетия. Но в этом направлении нам предстоит глубокая теоретико-методологическая работа. К тому же немало специалистов в области ИИ опираются на традиционные методы, позволяющие и в настоящее время успешно решать многие практические задачи, и убеждены, что для реализации намечаемых больших задач достаточно совершенствования, улучшения существующих методов. Они скептически относятся к привлечению проблемы сознания, считая, что она влечет за собой множество неопределенностей, «сплошной туман», что нет смысла в разработке т.н. трудной проблемы сознания, только еще более усложняющей ситуацию. Подобные убеждения проявились у ряда некоторых видных

специалистов, выступавших на экспертных обсуждениях проекта разработки AGI, предложенного Сбербанком России. Соответствующее мнение сложилось у меня, поскольку я участвовал в этих обсуждениях полемизировал по этим вопросам.

Соответствующие настроения присущи, конечно, лишь некоторым специалистам. Существуют разные взгляды, разные направления разработки ИИ. Среди них – концептуальные подходы, в которых внимание обращено на изучение ЕИ, моделирование его специфических свойств, создание новых когнитивных архитектур с целью сближения функциональных возможностей ИИ и 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 всесторонне анализирует функциональные когнитивные структуры ЕИ и, главное, придает огромное 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 значительные результаты [Финн 2016]. Развитие этого направления может быть перспективным и для разработки AGI. Существуют и многие другие подходы, применяющие в этом аспекте 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 для прорыва нужны колоссальные творческие и волевые усилия, связанные с преодолением привычных ментальных клише, выбором (и созданием) наиболее перспективных направлений исследований, концентрацией сил и средств для разработки принципиально новых теоретических и методологических концепций развития ИИ, новых подходов к решению междисциплинарных проблем.

В фокусе внимания исследователей оказываются проблемы теории, методологии, междисциплинарности, понятия интеллекта, мышления, познания и сознания, проблемы и гипотезы,



объяснения и понимания, вопрос о соотношении эмпирического и теоретического знания, характера связи между ИИ и ЕИ. Первостепенная роль в их осмыслении принадлежит философам – специалистам в области эпистемологии (теории познания), методологии науки, феноменологии, изучения сознания в широком смысле включая те направления, которые в западной номенклатуре именуются аналитической философией сознания.

Однако нужен *особый вид специализации философа*, работающего на стыке философии и науки, требующий глубоких методологических знаний и теоретических навыков, аналитического опыта и, конечно, достаточной компетенции в соответствующей области научного знания. В связи с этим центральное место занимают вопросы о понятии интеллекта и соотношении 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 как стадии продвижения к полноценному AGI. Все это получило широкое освещение в указанной книге [Сильный ИИ 2021], которая, подчеркнем еще раз, вносит исключительно большой вклад в разработку AGI.

Под сильным интеллектом понимают человеческий интеллект, наделенный сознанием, т.е. ЕИ. Все существующие формы ИИ чаще относят к слабому интеллекту. Но все, что касается Общего интеллекта (AGI) и его отношения к ЕИ, к обычному ИИ, сопровождается вопросами, ответы на которые истолкованы неоднозначно. Термин «Общий интеллект» появился в 1997 году, когда компьютер *Deep Blue* обыграл чемпиона мира в шахматы. На базе 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 определенную задачу). В 2007 году состоялась первая конференция, посвященная проблемам развития Общего ИИ, которые впоследствии стали предметом возрастающего внимания специалистов в области ИИ. Тем более что подавляющее большинство систем ИИ характеризуются как узкоспециализи-



рованные, способные качественно решать определенную узкую задачу. Поэтому критики не раз говорили о том, что компьютер, запрограммированный на решение тысячи задач, не способен самостоятельно научиться решать тысячу первую.

В указанной нами книге о сильном ИИ подробно рассказано об истории развития ИИ, ее значимых этапах, о том, что такое развитие происходило в направлении снижения зависимости технологий ИИ от человека, их способности к решению более широкого круга задач. Но достижения по-прежнему далеки от решения проблемы AGI. Авторами издания дана общая оценка 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]. Авторы книги продолжают: «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 символической информации (рассуждений на основе знаний) и т.д., то есть интегративными не являются» [Сильный ИИ... 2021, 52]. Тем самым четко обозначены главные, пока отсутствующие свойства, требуемые для построения действительного AGI.

Интерес представляет то, как решен в книге вопрос об определении AGI, его степени общности и отношении к ЕИ. Авторы анализируемого издания ссылаются на ряд определений: «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) [Сильный ИИ... 2021, 28]. Находит отражение и еще одно утверждение: «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...», в том числе «на неизведанных планетах» [Сильный ИИ... 2021, 29].

Как видим, 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 о том, почему сразу отодвигается на задний план «сильный» интеллект, наш ЕИ как высшая форма из всех известных форм интеллекта, наделенная сознанием, благодаря которому мы создаем все искусственные интеллекты. Ведь именно в ЕИ присутствуют 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. 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»

[Сильный ИИ... 2021, 30]. В подтверждение о «зыбкости» дано определение интеллекта из Википедии, далеко не надежного в научном отношении источника. В нем, наряду с указанием на то, что он является «качеством психики», перечислены в несистематизированном виде его многочисленные «способности», от приспособления к новым ситуациям до ощущений, мышления, воображения, внимания, воли и рефлексии.

Приведенное определение трудно принять в полной мере, поскольку непонятно, каким образом упорядочить, соотнести между собой эти способности и использовать их для целей AGI. В итоге предложено не вполне ясное заключение: «Определение интеллекта в области AGI не включает эти способности, но и не отвергает их» [Сильный ИИ... 2021, 31]. Заслуживает внимания и 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» [Сильный ИИ... 2021, 32]. Под «сверхчеловеческим» имеются в виду действительно достижения «узкого» ИИ: например, победа компьютера над чемпионом мира по игре в го, многие другие результаты, связанные с задачами анализа больших данных, быстрого действия, использования ИИ в производственных и коммуникативных процессах.

Но далее читаем:

Можно предположить, что *некоторые аналоги некоторых «сильных» качеств у действительно общего ИИ должны быть* (здесь и далее курсив и выделение полужирным мои. – Д. Д.). 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 удивле-

ния. 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 [Сильный ИИ... 2021, 32–33].

Приведенные суждения показательны. Они свидетельствуют о том, что специфические свойства и функции сознания, присущие нашему мыслительному процессу, должны привлекаться для построения AGI. Но возможно ли это за вычетом состояния *субъективной реальности – неотъемлемого качества сознания?*

Теоретически возможно, даже необходимо, но при условии адекватного вычленения и описания оперативных аналогов определенных состояний субъективной реальности. 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 (исследуемого нами ЕИ). Они являются весьма сложными, многомерными, взаимодействующими друг друга по содержанию. В когнитивном плане каждое из них включает в себя возможность множества аналитических и синтетических оперативных действий. Поэтому не вполне ясно, о каком аналоге идет речь. Приведенные понятия представлены слишком абстрактно. Следовательно, указания на подобные аналогии вряд ли могут иметь предполагаемое значение для моделирования когнитивных структур AGI.

Несмотря на это, установка на поиск аналогий заслуживает поддержки, т.к. обращена к задаче использования специфических свойств сознательной деятельности для разработки AGI. Нужны 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 опреде-

ляется без отсылки к человеческому интеллекту как ИИ, способный решать широкий круг задач» [Сильный ИИ... 2021, 27]. И она противостоит упомянутым выше случаям прямого отрицания роли сознания в разработке проблем ИИ и AGI. Специалист в области ИИ, занимающий эту позицию, демонстрирует парадоксальную ситуацию: он *как бы забывает* о том, что является носителем ЕИ и обладает искомыми свойствами AGI, постоянно применяет их в своей деятельности, несмотря на то, что они скрыты для него в хитросплетениях собственного сознания.

Большинство специалистов в области ИИ занимаются решением прикладных задач. Они не могут быть компетентны в философско-методологических вопросах в контексте проблем сознания, анализа ценностно-смысловых и деятельно-волевых структур субъективной реальности; им «хватает» своих дел, трудностей и проблем. Но сегодня важно понимать, что профессионалы, занимающиеся *теорией познания, методологией науки, феноменологией субъективной реальности*, способны быть активными помощниками в общем деле, в частности в вопросах теоретического осмысления роли сознания на современном этапе развития ИИ. Разумеется, многие специалисты в области ИИ отдают себе в этом 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 неотъемлемом качестве субъективной реальности, позволяют сформулировать основные системные, структурные, функциональные, операциональные характеристики естественной мыслительной деятельности человека, а значит, представляют собой важный ресурс для моделирования когнитивных структур, отвечающих задачам построения AGI. С этой целью актуальным становится дальнейшее рассмотрение основных философских и научных положений о сущности самого качества субъективной реальности, ранее изложенных (эти вопросы

подробно освещались в ряде моих работ: [Дубровский 2002; Дубровский 2020].

Качество нередко выносятся 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 большинстве описанных случаев нет необходимости выделять и анализировать качество субъективной реальности, поскольку внимание сосредоточено на поставленной задаче и наличном процессе деятельности; мы будто не замечаем это качество, и даже когда напрягаем, корректируем собственное мышление; оно подобно воздуху, который мы тоже «не замечаем» во многих интервалах жизни. Качество субъективной реальности становится слишком «заметным», когда предметом исследования оказывается мышление и сознание. При этом возникает целый ряд принципиальных вопросов. Рассмотрим каждый из них.

1. Именно качество субъективной реальности создает главные теоретические трудности проблемы сознания. Они связаны в первую очередь с тем, что явлениям субъективной реальности нельзя приписывать физические свойства (массу, энергию, пространственные характеристики). Возникает вопрос о том, как, не обращаясь к идеалистическим постулатам о наличии особой духовной субстанции, построить теоретически корректное объяснение сознания с материалистических позиций, учитывая его очевидную связь с физическими процессами. Вопросы, относящиеся к требуемому объяснению, принято объединять под общим названием «Трудная проблема сознания». Она более 60 лет находится в центре внимания западной аналитической философии, служит предметом острых дискуссий. Ей посвящены тысячи публикаций. Однако общепринятое решение отсутствует. В большей части 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]). В отличие от последних мною предложен нередукционистский способ решения глав-

ных теоретических вопросов «Трудной проблемы сознания» [Дубровский 1971; Дубровский 2015; Dubrovsky 2019b], который может представлять определенный интерес и для разработки ряда вопросов ИИ.

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

1) в **онтологическом плане** поставлены вопросы объяснения связи явлений субъективной реальности с деятельностью мозга [Koch, Massimini, Boly, Tononi 2016], их способности управлять телесными процессами, а также более широкие вопросы происхождения субъективной реальности, форм ее существования, роли в физическом мире (при условии, что ей нельзя приписывать физические свойства);

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

Вместе с тем для проблематики AGI важное значение имеют вопросы онтологического плана, т.е. 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, объ-



ективно присущих любой среде, в которой он должен учиться ориентироваться и действовать; для агента значимо умение различать физические причины и информационные, способность «понимать» специфику информационной причинности, видом которой служит психическая причинность. Для этого и требуется решение «Трудной проблемы сознания».

2. Понятие субъективной реальности означает всякое осознаваемое состояние, в отличие от бессознательных процессов, которые существуют одновременно с ним и служат его непрерывной 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). Речь идет о непрерывном процессе, который временно прерывается в условиях глубокого сна без сновидений, общего наркоза, обморока, в иных случаях временной потери сознания (в этих интервалах я для себя не существую). 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 свидетельствуют 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. Речь идет о *допсихическом уровне информационной реальности*. Следующим этапом ее развития стало зарождение психики, обладающей 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 представляют собой самоорганизующиеся системы со своими жесткими программами, «отработанными» эволюцией в течение многих сотен миллионов лет. Но теперь последние должны были согласовываться с общеорганизменной программой 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. меры централизации управления, которая не нарушает фундаментальные программы отдельных клеток, и 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 актуальна и для ИИ, если появляются 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. Данная проблема решается в сложной биологической системе *гораздо эффективнее*, чем в случаях взаимоотношений человека и общества, человека и технических устройств, в технических системах. Это требует внимания и изучения, когда речь идет о новых разработках в

области ИИ и интеллектуальной робототехники, связанных с аналогичными задачами эффективного управления.

Возникновение у животных субъективной реальности стало исторически первой формой **виртуальной реальности**, открывающей по мере развития все более широкий диапазон способностей к абстрагированию, обобщению, прогнозированию, планированию, пробным действиям в виртуальном плане, иным схожим операциям (в форме «мысленных экспериментов»), повышающим приспособляемость к среде. Это качество виртуальности во многом определяет способность формирования новых навыков, приспособления к изменившейся среде, т.е., по существу, главное свойство, необходимое 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 с ним язык. Особенностью сознания по сравнению с животной психикой является то, что психическое отображение и управление становятся объектом отображения и управления. Появляется способность, по сути, 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 данной 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 мозге; а в нем происходят очень сложные, многоступенчатые сетевые процессы переработки информации,

предшествующие, например, переживанию мной ощущения красного или образа бегущего ко мне человека, т.е. мне дан **итоговый результат** сложных информационных процессов в той **интегральной** форме («простой» и понятной), которая необходима для адекватного поведения и решения задач в данной ситуации; 2) вместе с тем в явлениях субъективной реальности нам дана и **способность оперировать этой информацией в «чистом» виде по своей воле** в 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. Возникает острый вопрос о свободе воли, который вызывает немало трудностей в нейронаучных исследованиях сознания, но может получить объяснение с позиций информационного подхода [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 в себя функциональные регистры, т.е. единство *рефлексивного и арефлексивного, актуального и диспозиционного*. Рефлексивное – это длящееся *сейчас осознаваемое* переживание данного контента, которое в следующий момент «уходит» в память, становится арефлексивным (заменяется другим рефлекслируемым контентом, но затем может снова «возвращаться», становится рефлекслируемым). Актуальное – это рефлексивное, но взятое в его отношениях и зависимости от *диспозиционного*, 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. Диспозиционное в большей степени выражает 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 ряде случаев такая направленность становится для личности непреодолимой или труднопреодолимой (как, например, при наркотической зависимости).

В аспекте рассмотрения регистров рефлексивное и арефлексивное, актуальное и –диспозициональное для разработки ИИ представляет интерес работа нашей памяти в связи с тем, что компьютерам присуща архитектура фон Неймана, основанная на постоянном разделении памяти и вычисления. Субъективная реальность демонстрирует принципиально иное 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 и арефлексивного, актуального и диспозиционального). Но эти специфические для нашей психики формы информационных процессов определяются главными системно-структурно-функциональными свойствами сознания (субъективной реальности) с учетом непрерывной связи с бессознательными процессами.

6. Охарактеризуем ключевые функциональные структуры субъективной реальности и присущие ей *операциональные способности*, которые могут иметь непосредственное отношение к моделированию когнитивных архитектур, соответствующих уровню AGI.

Элементарный анализ раскрывает фундаментальное структурно-операциональное свойство всякого явления субъективной реальности – его *двумерность*, т.е. способность одновременно отображения в нем 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 самоотобра-

жения на уровне нашего Я. Эта *фундаментальная* способность отчетливо проявляется на уровне зрительного восприятия, как показано у Дж. Ризцоллатти 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 активизируется 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» [Ризцоллатти, 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). Единство трех операций способно обеспечить эффективное взаимодействие между людьми. Нельзя не учитывать, что на моторных репрезентациях наблюдаемых действий, создаваемых зеркальными нейронами, *«основана способность обучения при помощи подражания»* [Ризцоллатти, Синигалья 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 выражено 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 внимания. Оно особенно остро переживается 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 состояние деперсонализации вызывает явления *дереализации*, нарушения отображения внешних объектов и собственного тела (см., например: [Меграбян 1962; Рамачандран 2014; Литвак 2007; Назлоян 2002]).

7. Взаимозависимость иноотображения и самоотображения проявляется в функционировании памяти (в регистрах рефлексивного и арефлексивного, актуального и диспозиционного). Полнее всего она выражена на уровне многомерной динамической структуры нашего Я, т.е. 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» (*категории контентов Я*). Такого рода взаимопреобразования, перемена модальности переживаемого «содержания» – это *механизм эффективного отображения и освоения действительности*, в том числе 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 соотношение 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, прежде всего его анализ автобиографической самости [Дамасио 2018].

Базисная динамическая структура субъективной реальности раскрывается конкретнее, когда в ходе феноменологического анализа выявляются основные виды противопоставления и соотношения «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», поскольку оно полагает и раскрывает себя посредством своего «не-Я». Иными словами, таковы главные смысловые (когнитивные и ценностные) измерения нашего «Я» (*категории контентов Я*). Чтобы полнее раскрыть одно из выделенных отношений «Я», нужно рассмотреть его не только само по себе, но и сквозь призму всех остальных. 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», выступающее в форме «знания», «оценки» и «действия». В ней непрерывно совершаются процессы самоотображения и самоорганизации структуры субъективной реальности, формируется и реализуется ее активность, ее **деятельно-волевые функции**. Наиболее значимый вопрос о связи отображения с действием, как в форме самопознания и самопреобразования сознания, так и в форме познания и преобразования внешнего мира, заслуживает отдельного обсуждения в дальнейшем.

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, речь идет о ясном понимании той огромной дистанции, которая отделяет нынешний уровень ИИ от ЕИ, от 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 для того, чтобы добиваться шаг за шагом сближения с ЕИ, т.е. **добиваться все более глубокого самопознания и самопреобразования, т.к. 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 некоторые ее специфические функциональные свойства, опираться на приемлемые аналогии в операциональных действиях мышления человека и интеллектуального робота. Лучше говорить в этом отношении не об ИИ в общем смысле, а об интеллектуальном роботе, поскольку он обладает определенной когнитивной архитектурой и, кроме того, совершает физически реальные действия во внешней среде, подлежащие строгой проверке на их эффективность.

Обратимся к рассмотренной выше двумерности 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). Существует и нечто, подобное самоотображению, которое может быть заложено в надежности его программного обеспечения и технического устройства, в механизмах слежения за их исправностью. Но сразу при этом видим и качественное различие. У человека функция самоотображения обладает способностью *саморегуляции*, которая постоянно стимулируется, корректируется, совершенствуется процессами иноотображения, как и наоборот: стимулируется, корректируется, повышает результативность под воздействием процессов самоотображения, т.е. обладает способностью саморегуляции. В этом и заключается свойственное сознанию, структуре субъективной реальности единство функций иноотображения и самоотображения, их органическая взаимообусловленность, которая 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 в полной мере осознают ее ограниченность: узость заданного диапазона восприятий и действий, неспособность к самостоятельному его расширению при существенном изменении внешних условий. Еще сильнее такая ограниченность связана с функцией

самоотображения. В программном обеспечении робота и в его техническом устройстве нет специальных регистров текущего непрерывного самоотображения его подсистем, процессов и соответственной самонастройки, подверженных непрестанной коррекции со стороны иноотображения и внешних действий, что, однако, наблюдается у человека. Функция иноотображения и то, что условно можно было бы отнести к функции самоотображения, фактически *разделены*.

Это указывает на проблемы, требующие разработки при создании интеллектуальных роботов нового поколения, т.е. достигающих уровня AGI. Полагаем, определенные ресурсы уже существуют, и наметились 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, согласно концепции А.Р. Ефимова, известного отечественного специалиста в области интеллектуальной робототехники [Efimov 2020]. В этом направлении, видимо, будет возможность использовать ресурсы, связанные с методами глубоких нервных сетей для создания программ указанной выше *двумерной функциональности* в поведении интеллектуального робота.

Разумеется, пунктирно намеченные возможности новых подходов нуждаются в дальнейшей конкретизации, критическом обсуждении. Но они свидетельствуют о широкой творческой перспективе применения феноменологических исследований операциональной специфики сознательной деятельности.

9. Дополним, что изложенные выше результаты феноменологических исследований субъективной реальности заставляют обратить внимание на чрезвычайную сложность «содержания» (множества различных *контентов*) нашей сознательной деятельности, взаимосвязь, взаимовлияние различных по своему виду (классу) контентов. Они приблизительно определены нами

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, рассмотренных нами ранее в статье.

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, задают широкий по их «содержанию» *спектр операций абстрагирования, анализа и синтеза*. Такие операции существенно различаются по когнитивным свойствам в зависимости от того или иного категориального контента. С одной стороны, операции абстрагирования, анализа и синтеза при исследовании предметной области внешнего мира, с другой – 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, как правило, используется широкий спектр этих операций. Многие из них протекают совместно, включают *интуитивный уровень*, связанный порой с новыми, находящимися в стадии формирования творческими результатами. Из многообразного спектра в системах ИИ используется лишь небольшая часть когнитивных операций, но в достаточной степени определенных. В этом состоит одно из главных отличий ЕИ от ИИ. Все изложенные обстоятельства должны служить поводом для размышлений при разработке AGI, поскольку в отношении него не требуется способность решения *всего разнообразия* человеческих задач (например, объяснения 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, которым следует посвятить основательное самостоятельное исследование.

10. В многомерных планах проблемы сознания целесообразно выделить еще один вопрос. Он касается не просто когнитивных операций (абстрагирования, обобщения и др.) в различных категориальных контентах, указанных выше, а учета множества *экзистенциально значимых* состояний сознательной деятельности, которые, хотя и включены в процесс познания, 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, полноты анализа познавательного процесса и оценки его целей, т.к. в настоящее время, в условиях нарастания глобального кризиса нашей потребительской цивилизации, особое значение приобретают вопросы о подлинных смыслах личной и социальной деятельности, о противостоянии нарастающему абсурду в нынешних процессах социального развития.

Неслучайно известный феноменолог М. Мерло-Понти писал: «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]. В данном случае понятие *интенциональной дуги* выполняет интегративную роль, т.е. роль средства охвата и концептуального учета диапазона «участников» познавательного процесса, от перцептивного уровня до идеологических, моральных и иных, 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 тем не менее дает 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 смысле, то, полагая, справедливо признать, что наиболее глубокие и разносторонние сведения о сознании даны в произведениях Пушкина, Толстого, Чехова, других писателей и поэтов.

Велико значение поэзии для эпистемологического анализа обыденного и научного мышления. Поэтические образы и метафоры – мощный творческий стимул развития интеллекта, т.к. разрушают устоявшиеся комфортные клише, решительно и убедительно пересекают границы привычных противоположных смыслов, обнаруживают ранее не известные 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 значима, чем этическая, что находит отражение и в контексте ЕИ, и при обсуждении проблем разработки AGI. Следует иметь в виду не только когнитивный аспект, способ решения определенной задачи, но и оценку ее допустимости с позиций этических норм и принятых социальных критериев.

Особенно важная и трудная тема в проблематике развития ИИ – 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 малоисследованными. Подтверждением служит следующий пример. В нашем теле присутствует 100 000 километров кровеносных русел, и эта колоссальная динамическая система эффективно управляется при постоянном изменении внешних физических условий.



### **Заключение**

В завершение еще раз обратим внимание на ключевую роль образования в системе мер, необходимых для создания AGI, о чем говорил в обращении к читателям первой книги об AGI Герман Греф. Подготовка специалистов в области ИИ обнаруживает существенные недостатки. Авторы книги убеждены в следующем: «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 проявляется слабость, формализм, интеллектуальное убожество (в ряде отношений) современного 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.

### **ЦИТИРУЕМАЯ ЛИТЕРАТУРА**

Дамасио 2018 – Дамасио А. Так начинается «Я». Мозг и возникновение сознания. – М.: Карьера-Пресс, 2018.

Дубровский 1971 – Дубровский Д.И. Психические явления и мозг. Философский анализ проблемы в связи актуальными задачами нейрофизиологии, психологии и кибернетики. – М.: Наука, 1971.

Дубровский 2002 – Дубровский Д.И. Проблема идеального. Субъективная реальность. – М.: Канон+, 2002.

Дубровский 2015 – Дубровский Д.И. Проблема «Сознание и мозг». Теоретическое решение. – М.: Канон+, 2015.

Дубровский 2019 – Дубровский Д.И. Субъективная реальность и мозг: к вопросу о полувековом опыте разработки проблемы сознания в аналитической философии // Дубровский Д.И. Проблема сознания. Теория и критика альтернативных концепций. – М.: Ленанд, 2019. С. 140–183.

Дубровский 2020 – Дубровский Д.И. Субъективная реальность // Философская электронная энциклопедия. – М.: Институт философии РАН, 2020. – URL: [https://iphlib.ru/library/collection/elphilenc/document/dubrovskii\\_subektivnaya\\_realnost](https://iphlib.ru/library/collection/elphilenc/document/dubrovskii_subektivnaya_realnost).

Кузнецов 2016 – Кузнецов О.П. Избранные труды. Автоматы, языки и искусственный интеллект. – М.: Наука, 2016.

Литвак 2007 – Литвак Л.М. «Жизнь после смерти»: предсмертные переживания и природа психоза. Опыт самонаблюдения и психоневрологического исследования. – М.: Канон+, 2007.

Меграбян 1962 – Меграбян А.А. Деперсонализация. – Ереван: Армгосиздат, 1962.

Мерло-Понти 1999 – Мерло-Понти М. Феноменология восприятия. – СПб.: Ювента; Наука, 1999.

Назлован 2002 – Назлован Г.М. Концептуальная психотерапия. Портретный метод. – М.: ПЕР СЭ, 2002.

Рамачандран 2014 – Рамачандран В.С. Мозг рассказывает. Что делает нас людьми. – М.: Карьера Пресс, 2014.

Риццолатти, Синигалья 2012 – Риццолатти Дж., Синигалья К. Зеркала в мозге. О механизмах совместного действия и сопереживания. – М.: Языки славянских культур, 2012.

Сильный ИИ... 2021 – Сильный искусственный интеллект. На подступах к сверхразуму / А.А. Ведяхин и др.; науч. ред. А.С. Потапов. – М.: Интеллектуальная литература, 2021.

Финн 2011 – Финн В.К. Искусственный интеллект: методология, применения, философия. – М.: URSS, 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 2019a – Dubrovsky D.I. “The Hard Problem of Consciousness.” Theoretical Solution of Its Main Questions // AIMS Neuroscience. 2019. Vol. 6. No. 2. P. 85–103.

Dubrovsky 2019b – Dubrovsky D.I. The Problem of Free Will and Modern Neuroscience // Neuroscience and Behavioral Physiology. 2019. Vol. 49. No. 5. P. 629–639.

Efimov 2020 – Efimov A. Post-Turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence // Artificial General Intelligence. 13<sup>th</sup> International Conference, AGI 2020 (Lecture Notes in Computer Science. Vol. 12177) / ed. by B. Goertzel, A. Panov, A. Potapov, R. Yampolskiy. – Cham: Springer, 2020. P. 83–94.

Koch, Massimini, Boly, Tononi 2016 – Koch C., Massimini M., Boly M., Tononi G. Neural Correlates of Consciousness: Progress and Problems // Nature Reviews Neuroscience. 2016. Vol. 17. No. 5. P. 307–321.

#### REFERENCES

Damasio A. (2010) *Self Comes to Mind: Constructing the Conscious Brain*, Pantheon. New York: Pantheon Books (Russian translation: Moscow: Karyera-Press, 2018).

Dubrovsky D.I. (1971) *Mental Phenomena and the Brain. Philosophical Analysis of the Problem in Connection with the Urgent Tasks of Neurophysiology, Psychology, and Cybernetics*. Moscow: Nauka (in Russian).

Dubrovsky D.I. (2002) *The Problem of the Ideal. Subjective Reality*. Moscow: Kanon+ (in Russian).

Dubrovsky D.I. (2015) *The Problem “Consciousness and the Brain.” Theoretical Solution*. Moscow: Kanon+ (in Russian).

Dubrovsky D.I. (2019a) “The Hard Problem of Consciousness”. Theoretical Solution of Its Main Questions. *AIMS Neuroscience*. Vol. 6, no. 2, pp. 85–103.

Dubrovsky D.I. (2019b) The Problem of Free Will and Modern Neuroscience. *Neuroscience and Behavioral Physiology*. Vol. 49, no. 5, pp. 629–639.

Dubrovsky D.I. (2019c) Subjective Reality and the Brain: On the Issue of the Semi-Centennial Experience in the Development of the Problem of Consciousness in Analytical Philosophy. In: Dubrovsky D.I. *The Problem of Consciousness. Theory and Criticism of Alternative Concepts* (pp. 140–183). Moscow: Lenand (in Russian).

Dubrovsky D.I. (2020) Subjective Reality. In: *Philosophical Electronic Encyclopedia*. Moscow: RAS Institute of Philosophy. Retrieved from [https://iphlib.ru/library/collection/elphilenc/document/dubrovskii\\_subektivnaya\\_realnost](https://iphlib.ru/library/collection/elphilenc/document/dubrovskii_subektivnaya_realnost)

Efimov A. (2020) Post-Turing Methodology: Breaking the Wall on the Way to Artificial General Intelligence. In: Goertzel B., Panov A., Potapov A., & Yampolskiy R. (Eds) *Artificial General Intelligence. 13<sup>th</sup> International Conference, AGI 2020* (Lecture Notes in Computer Science. Vol. 12177) (pp. 83–94). Cham: Springer.

Finn V.K. (2011) *Artificial Intelligence: Methodology, Applications, Philosophy*. Moscow: URSS, Krasand (in Russian).

Finn V.K. (2016) On the Class of JSM-Reasoning Using Isomorphism of Inductive Inference Rules. *Artificial Intelligence and Decision Making*. No. 3, pp. 48–61 (in Russian).

Koch C., Massimini M., Boly M., & Tononi G. (2016) Neural Correlates of Consciousness: Progress and Problems. *Nature Reviews Neuroscience*. Vol. 17, no. 5, pp. 307–321.

Kuznetsov O.P. (2016) *Selected Works. Automata, Languages and Artificial Intelligence*. Moscow: Nauka (in Russian).

Litvak L.M. (2007) “Life after Death”: Near-Death Experiences and the Nature of Psychosis. *The Experience of Self-Observation and Neuropsychiatric Research*. Moscow: Kanon+ (in Russian).

Megrabyan A.A. (1962) *Depersonalization*. Yerevan: Armgosizdat (in Russian).

Merleau-Ponty M. (1945) *La Phénoménologie de la perception*. Paris: Gallimard (Russian translation: Saint Petersburg: Yuventa; Nauka, 1999).

Nazloyan G.M. (2002) *Conceptual Psychotherapy. Portrait Method*. Moscow: PER SE (in Russian).

Ramachandran V.S. (2011) *The Tell-Tale Brain: A Neuroscientist's Quest for What Makes Us Human*. New York: W.W. Norton & Company (Russian translation: Moscow: Karyera-Press, 2014).

Rizzolatti G. & Sinigaglia C. (2008) *Mirrors in the Brain: How Our Minds Share Actions and Emotions*. Oxford: Oxford University Press (Russian translation: Moscow: LRC Publishing House, 2012).

Vedyakhin A.A. et al. (2021) *Strong Artificial Intelligence. On the Approaches to the Supermind*. Moscow: Intellektual'naya literatura (in Russian).