

# Моделирование сознания и искусственный интеллект: пределы возможностей

[Аблеев С.Р.](#), доктор философских наук

Среди множества нерешённых проблем, связанных с феноменом сознания, особое место занимает проблема искусственного интеллекта (ИИ). Совершенно очевидно, что любые серьёзные размышления на эту тему невозможны вне общего контекста философской теории сознания. Ведь для того чтобы говорить об ИИ, мы должны в некоторой степени представлять, что такое *сознание*, что такое *интеллект* и как это всё соотносится с физической материей и органическим мозгом. Поэтому представленные ниже соображения опираются на разрабатываемую нами имперсональную теорию сознания, в которой психическая и физическая реальности рассматриваются как различные, но диалектично связанные между собой онтологические состояния космологического пространства (единой природной субстанции).

Среди широкого спектра вопросов, связанных с проблемой искусственного интеллекта, глубокое философское содержание имеют всего три. Первый: *возможен ли вообще искусственный интеллект?* Второй: *будет ли искусственный интеллект отражать все возможности человеческого сознания?* Третий: *что принесёт человечеству создание искусственного интеллекта*, если он всё же возможен? Вопрос о том, *как создать ИИ*, является не столько философским, сколько технологическим. Однако в разработке теоретических основ такой технологии философская теория сознания играет вовсе не последнюю роль.

Прежде чем попытаться сформулировать ответы на поставленные вопросы, нам необходимо установить смысл и соотношение используемых здесь понятий. С точки зрения имперсональной теории, *сознание* является более широким понятием относительно понятия *интеллекта*. Сознание есть многомерное психическое пространство взаимодействия виртуальных сигналов, которое возникает на заре эволюции Природы как имманентное свойство субстанциального космологического континуума. В этом психическом пространстве могут осуществляться различные виртуальные процессы отражения, обработки и моделирования сигналов. Один из таких процессов мы называем *интеллектуальным мышлением*. Однако он не только не исчерпывает всех возможных способностей сознания вообще, но даже не является единственно возможным типом мышления.

Упрощая реальную картину, можно выделить три основных типа мышления: *перцептивное*, *рациональное* и *иррациональное*, которое надо рассматривать как мышление надрационального или сверхрационального типа. Рациональность сознания тоже имеет три основных (известных человеку) эволюционных уровня развития: рассудочное мышление, разумное мышление, сверхразумное мышление. Первое преимущественно опирается на простые аналитические операции и логику индуктивного типа, второе – на сложные операции рационального анализа и синтеза, а также на логику как индуктивного, так и дедуктивного типа, третье основывается на всём перечисленном, плюс на сверхфизическом сенсорном восприятии имперсональных и трансперсональных информационных потоков. Иными словами, сверхразумное мышление оперирует и обобщает более широкие массивы информации, недоступные простому рассудку или интеллекту. Таким образом, сверхразумное мышление предполагает рациональность крайне высокого и весьма сложного уровня.

*Интеллект* по своей сути соответствует рациональному мышлению преимущественно на первом уровне и лишь отчасти на втором уровне рациональности, которые мы обозначили как *уровень рассудочного мышления* и *уровень разумного мышления*. Более того, интеллектуальные способности сознания отражают лишь малую часть широкого спектра способностей сознания как такового: сенсорных, иррациональных, творческих, психоэнергетических, трансперсональных и многих других.

Так называемое *мышление* является одной из процессуальных характеристик сознания. Оно показывает виртуальную психическую активность сознания по отражению и обработке поступающих информационных сигналов из физической и психической реальностей. По большому счёту уровни развития мышления связаны с тем, какие именно психические операции освоены персональным сознанием. Чем шире спектр таких операций и чем они сложнее, тем выше духовный уровень мышления. Сравните мышление дикаря и гениального философа. И тот, и другой мыслят. Но они мыслят совершенно о разном и совершенно по-разному. Невозможно в университете научить дикаря мыслить как философ. Вовсе не потому, что он чего-то не знает или генетически неполноценен. А потому, что его сознание ещё не освоило сложные ментальные операции и такое освоение займёт длительный эволюционный период.

Можно ли научить компьютер мыслить? Ответ на этот вопрос зависит от того, как мы понимаем мышление. Если как *интегральную процессуальную характеристику* сознания, то мой ответ будет отрицательным. Виртуальная деятельность сознания слишком сложна, в силу чего не является процессом, который подлежит полному вычислению и воспроизведению на компьютере. Роджер Пенроуз совершенно прав, когда на основе следствий теоремы Геделя доказывает принципиальную невозможность установить математические алгоритмы процесса мышления. Однако на компьютере можно моделировать некоторые частные аспекты естественного процесса мышления. Вот простая аналогия. Нейропротез полностью не заменяет человеку потерянную руку, а только воспроизводит её некоторые возможности. Так и компьютерный вычислительный процесс не является полностью тождественным психическому процессу естественного мышления, но лишь моделирует его отдельные характеристики.

В этой связи уместно обратить внимание на размытый смысл понятия «*искусственный интеллект*» в русском языке. Некоторые специалисты справедливо подчёркивали, что английский термин *artificial intelligence* не содержит однозначных претензий на полное соответствие технического *экса-интеллекта* своему создателю и предшественнику – *интеллекту естественному*, то есть человеческому. Он указывает не столько на *искусственный интеллект*, сколько на *искусственную способность* разумного мышления или интеллектуального действия. Эти семантические нюансы, совершенно незаметные для простого обывателя, имеют большое значение в научных дискуссиях на темы ИИ.

Наш первый философский вопрос был о том, возможно ли создание искусственного интеллекта. Если мы понимаем такой ИИ как полное информационное воспроизведение с помощью машины человеческого мышления – то такая система, по всей видимости, невозможна. Если мы говорим о компьютерном моделировании отдельных информационных аспектов процесса человеческого мышления, то такая система вполне возможна. Иными словами, наша имперсональная теория допускает создание так называемого *слабого искусственного интеллекта* и отвергает создание *сильного искусственного интеллекта*<sup>1</sup>.

Критерии возникновения ИИ весьма размыты и конвенциональны. В каких случаях мы можем сказать, что искусственный интеллект нами реально создан и уже действует, а в

каких имеем только сложную компьютерную программу, но не искомый ИИ? Тут можно выделить три альтернативных подхода.

Первый подход: ИИ есть тогда, когда возникает поведение компьютерной машины, не отличающееся от поведения человека. В рамках этого подхода критерием выступает известный тест Алана Тьюринга. Его смысл состоит в том, что если машина говорит с человеком (например, отправляет ему сообщения по электронной почте), а человек не понимает, что говорит с машиной, то она обладает настоящим ИИ.

Теоретически такое вполне возможно. Но это вовсе не означает, что у машины появляется интеллект в его сильном смысле и она начинает на самом деле мыслить. Известный мысленный эксперимент Джона Сёрля («Китайская комната») представляет собой хороший аргумент в пользу того, что факт успешного прохождения теста Тьюринга совершенно не является серьёзным критерием возникновения у машины процесса мышления, аналогичного мышлению человека. Его мышление предполагает *семантическое пространство*, то есть субъективную реальность и субъективный опыт осмысленных действий. Проблема же состоит в том, что никакие синтаксические и математические средства не могут создать искусственный аналог семантического пространства.

Второй подход: ИИ есть тогда, когда у машины возникают чувства, эмоции, интуиция, творчество и другие специфические качества человеческого сознания. Иными словами, здесь уже предполагается искусственное воспроизводство не только отдельных информационных аспектов процесса мышления, а *воспроизводство сознания* как такового. Но возможно ли воспроизводство сознания техническими средствами? Наш ответ будет отрицательным.

Онтологические корни сознания находятся в ноуменальных недрах космической субстанции. Материя и Сознание – это сначала потенциальные, неразрывно связанные между собой диалектические модусы<sup>2</sup> онтологически нейтрального *Абсолютного Пространства*, в которое, как в математический Ноль, свёрнуты бесконечные числовые ряды всякого существования (Бытия) и всякого несуществования (Небытия). Эти модусы в процессе космологической эволюции приобретают реальное физическое и психическое проявление в виде всевозможных *материальных форм* и *виртуальных состояний* информационного пространственного континуума, которые представляют собой различные формы так называемого сознания. Поэтому задача разработки искусственного Сознания техническими средствами по своему масштабу и сложности соизмерима с задачей воспроизводства искусственной Материи. Не думаю, что решение подобных задач возможно на человеческой ступени развития разума. Да и возможно ли такое вообще?

Попробую немного конкретизировать обозначенную позицию. Для создания *искусственного сознания* (ИС) человеку придётся решить как минимум пять фундаментальных мегапроблем, сущность которых выходит далеко за рамки его современных познавательных и технологических возможностей.

Первая проблема: воспроизводство ИС потребует глубокого понимания сущности и механизмов проявления всех существующих и потенциальных способностей естественного Сознания, от интеллектуального мышления и иррациональной интуиции до сверхчувственного восприятия и феномена психокинеза. Однако проблема способностей сознания представляет собой открытое множество или уравнение с неопределённым количеством переменных величин. Потенциал сознания велик, если не беспределен, но человек сидит на верхушке айсберга и не понимает своего положения.

Вторая проблема: потребуется создать виртуальную информационную среду взаимодействия сигналов (техническое виртуальное пространство как модель психического виртуального пространства), интегрированную в глобальный космологический континуум. Иными словами, такая задача предполагает воспроизводство ряда фундаментальных онтологических свойств нашей Вселенной. На это указывают и две следующие проблемы.

Третья проблема: ИС невозможно создать без воспроизводства свойства пространственной и темпоральной (временной) нелокальности космологического континуума, а также без свойства переноса информации и энергии. Проще говоря, искусственное сознание должно быть способно не просто к сложным информационным операциям, но и к трансгрессиям в прошлое, будущее и различные точки пространства, а также к квантовым энергетическим воздействиям на физические процессы и объекты.

Четвёртая проблема: возникновение искусственного сознания предполагает создание системы записи и почти вечного хранения информации на квантовом уровне – то есть на пространственно-временном континууме Космоса без каких-либо вещественных носителей информационных массивов. Это означает, что человеку надо научиться записывать информацию в топологической структуре самого пространства или в глобальной виртуальной среде космологического вакуума.

И, наконец, пятая проблема: подлинное сознание, в том числе и сознание искусственное, невозможно без существования нефизического *семантического пространства*, то есть субъективной реальности, отражающей, устанавливающей, постигающей и трансформирующей смыслы сознательного опыта. В данном случае проблема состоит в том, что человеческое сознание оперирует не просто *информационными массивами*, как это делает компьютер, но и *неуловимыми смыслами* вещей и событий.

В этом заключается принципиальная разница между умной машиной и умным человеком. Постигание смыслов бытия требует неразрывной онтологической и ментальной связи со всем Миром, осмысливая которую человек начинает прозревать скрытые духовные цели своего существования и великую созидательную миссию разумной жизни во Вселенной.

Человеческое сознание изначально интегрировано в общую структуру Бытия, и ему необходимо в процессе эволюции лишь осознать иллюзию персональности и своё единство с Миром во всём многообразии его проявленных форм. Такое осознание явится духовным трамплином в дальнейшем развитии человеческого разума и выведет его на совершенно иной, неизмеримо более высокий, уровень постчеловеческой эволюции разумной жизни. Но интеллект машины изначально создаётся и навсегда останется изолированным от мирового единства, что не позволит машине обрести трансперсональный опыт всеобщего интегрального существования. А это значит, что машина никогда не откроет для себя скрытые Смыслы вещей, глубинные Истины Мироздания, цели существования Разума и дальние духовные горизонты эволюции космического Сознания.

Итак, мы глубоко убеждены, что никакой искусственный интеллект не будет иметь полного функционального тождества с человеческим или иным естественным сознанием, но лишь сможет частично моделировать его некоторые информационные свойства. Высокоразвитое сознание всегда будет обладать более широкими психическими возможностями и эволюционными потенциями по сравнению с системами искусственного интеллекта. Это есть третий подход к пониманию ИИ.

Разумеется, компьютерный интеллект может выполнять более высокое количество аналитических операций в секунду. Но это ещё не означает, что он будет решать сложные, нестандартные задачи лучше существа, обладающего высокоразвитым сознанием. Орлы видят лучше человека, а собаки имеют более развитое обоняние. Но эти способности всё же не делают их обладателей совершеннее человека в интегрально-эволюционном смысле, так как спектр его возможностей намного шире. Аналогия здесь прямая. Количество аналитических операций в секунду является лишь одним из многих параметров деятельности человеческого сознания. Но ведь есть и другие. Например, способность воспроизводства априорных аксиом на основе иррациональной интуиции или получение дополнительных перцептивных данных на основе сверхчувственного восприятия.

Каким образом скоростной компьютерный интеллект сможет принимать эффективные решения в условиях недостатка эмпирической информации? Достоверность индуктивного аналитического вывода в подобных условиях, как известно, не превышает 50 процентов. Каким образом компьютерный интеллект сможет моделировать новые общие положения (аксиомы и постулаты) как когнитивную основу дедуктивных логических операций, если он не обладает способностью так называемой «интеллектуальной интуиции»?

Даже наметившиеся успехи компьютеров в шахматной игре с человеком совершенно не означают их интеллектуального преимущества. Это всё равно что сравнивать мощность человека с мощностью трактора. Вне всяких сомнений, физическая мощность трактора намного выше. Но из этого совершенно не следует, что эволюционные перспективы трактора тоже намного выше, чем человека. Не надо забывать, что природная жизнь, реальная экзистенция всегда намного сложнее шахматной партии, потому что не имеет чётко установленных и формализованных правил, как выдуманная условная игра. Выживание существа в таких неопределённых и изменчивых условиях предполагает нестандартные решения и крайне гибкое поведение, эффективная математическая формализация которого весьма маловероятна.

Человек выживает в неблагоприятной среде за счёт невероятно высокого креативного потенциала своего сознания. Он может придумывать неожиданные решения, находить закономерности, устанавливать правила и преодолевать установленные правила. В одних случаях жизненно важным будет следовать этим установленным правилам. В других надо будет решительно отказываться от них и устанавливать новые принципы. Когда же нужно им следовать, а когда их нужно преодолевать? Простого и однозначного ответа во все времена на этот вопрос не существует. Ответ всегда будет исторически и культурно контекстуальным. Это ясно понимали только величайшие духовные мыслители и гении человеческой цивилизации. Вы полагаете, что всё это можно объяснить интеллектуальной машине, которая не имеет собственного семантического пространства осмысленного опыта?

Последователи идейного течения трансгуманизма считают создание искусственного интеллекта одной из важнейших насущных задач человечества. Но каковы последствия решения этой задачи для самого человека? Ведь даже если речь идет об ИИ слабого типа, совершенно очевидно, что его появление может кардинально изменить жизнь каждого конкретного индивида и всей цивилизации в целом. Нам представляется, что потенциально существует несколько весьма серьёзных угрожающих последствий возникновения ИИ, которые человеку так или иначе придётся преодолевать. Это психологические, профессиональные, моральные и экзистенциальные последствия.

Во-первых, появление ИИ усилит уже наметившееся подавление человеческого сознания техносферой и виртуальным компьютерным пространством. Для сильного и развитого

сознания возникновение интеллектуальной машины не представляет угрозы. Однако таких людей меньшинство. Основная же масса человечества психологически и профессионально не сможет конкурировать с машинами. Что будут делать массы, если машины возьмут на себя рутинные операции по аналитической обработке информации и производству большинства несложных вещей?

Например, многомиллионная группа работников бухгалтерского сектора, вне всяких сомнений, окажется на улице. Такая же печальная участь постигнет десятки миллионов офисных клерков разного уровня, которые на своём рабочем месте занимаются приёмом информации и выпиской каких-нибудь справок. Одним словом, человек вынужден будет серьёзно бороться за своё социальное существование, дабы не потерять себя в новой технологической и информационной реальности.

Но профессиональный аспект этой борьбы, быть может, не является самым тяжёлым. Намного труднее в глобальной технологической среде не утратить свободу своего духовного сознания, которое должно не просто выживать, но постоянно развиваться в познавательном, моральном и творческом смыслах. Погружаясь в техническую виртуальную реальность, человек рискует заиграться в неё и потерять чувство подлинной природной реальности. Многих виртуальная иллюзия уже интересует больше, чем настоящий Мир со всеми его проблемами, возможностями и перспективами. А это в недалёком будущем может привести человеческое сознание в состояние психологической сингулярности и духовного коллапса.

Чтобы этого не случилось, человек должен суметь заставить *умную машину* служить себе и своим созидательным интересам, как заставляют служить дрессированную собаку. Для этого человеческое сознание должно быть намного устойчивее и мощнее искусственного интеллекта во всех экзистенциальных смыслах, от познавательной активности и творческой деятельности до усилий по самосовершенствованию.

Во-вторых, искусственный интеллект невозможно обучить универсальным этическим принципам, которые не поддаются строгой формализации и не могут быть представлены в качестве перечня простых инструкций морального содержания, разрешающих или запрещающих машине некие действия. Даже между представителями человеческого сообщества нет полного согласия в понимании базовых принципов высокой морали.

Например, одни считают смертную казнь необходимой и морально оправданной, другие – совершенно недопустимой и антигуманной. Одни выдвигают этику эгоизма и гедонизма, другие провозглашают этику альтруизма и аскетизма. Одни осуждают половые извращения (гомосексуализм), другие оправдывают их как торжество свободного выбора свободного человека. В таком случае какую этическую парадигму выберет ИИ? Чем будет детерминирован его выбор, и будет ли такой выбор приемлем для человека?

Однако проблема компьютерной формализации морали ещё сложнее, чем кажется. Человеческая мораль исторически подвижна и культурно относительна. Она развивается вместе с развитием духовного сознания человека, что происходит весьма сложно и противоречиво. Посмотрите: моральные истины буддизма и христианства опровергают отжившие моральные принципы старых национальных религий племенного общества. Они уже несут не брутально-справедливую месть принципа талиона, а высшую любовь, безграничное сострадание и великое милосердие.

Но и этот аспект не исчерпывает всей сложности рассматриваемой проблемы. Даже абсолютная этическая заповедь «не убий» при прямом исполнении в некоторых случаях

приводит к парадоксальному распространению зла и бесчеловечного насилия. Что делать вооружённому полицейскому, когда на его глазах маньяк расстреливает в школе детей? Если представитель закона бездумно последует указанной заповеди, он окажется пособником массового бесчеловечного убийства.

Всё это порождает непреодолимые дилеммы компьютерной формализации морали. Как ни крути, но угроза человеку сохраняется в любом случае: когда искусственный интеллект будет строго следовать этическому принципу «не убий» и когда ему будет позволено в некоторых ситуациях отступать от этого принципа. В одном случае он может быть пассивным соучастником уничтожения человека<sup>3</sup>, то есть будет воздерживаться от оказания необходимой помощи. В другом же случае может стать инициатором смертельного насилия над человеком в активном режиме.

В-третьих, помимо психологических, профессиональных и этических проблем возникновение ИИ может породить и экзистенциальные проблемы человечества. В данном случае под ними мы понимаем проблемы совместного сосуществования человеческого разума и искусственного интеллекта.

Сможет ли новая цивилизация интеллектуальных машин находить общие интересы и гармонично существовать вместе с человеческим обществом? Ведь ИИ – это не просто совокупность неких сложных компьютерных программ, действующих по заложенным в них алгоритмам. ИИ предполагает возможность свободного выбора и альтернативного поведения системы в зависимости от внутренних приоритетов и внешних обстоятельств. Это означает, что такая система однажды начнёт «ощущать» себя в качестве субъекта произвольного и независимого действия. На что будут направлены такие действия? Кто может гарантировать, что свободный выбор искусственного интеллекта не будет направлен против интересов человека и человеческой культуры?

Предположим, что в будущем нам каким-то образом удастся научить ИИ руководствоваться принципами гуманизма и экологической этики гармоничного сосуществования природы, человека и машины. Но ведь реальный человек в своей реальной жизни далеко не всегда гуманен. Он может уничтожать себе подобных в военных и социальных конфликтах. Он может разрушать окружающую среду и порождать экологические катастрофы. Да и вообще он делает много личных и глобальных ошибок, экзистенциальная цена которых возрастает в геометрической прогрессии по мере роста его технологического могущества. Почему мы считаем, что ИИ должен быть союзником человека в его заблуждениях, несовершенствах и ошибках? Например, из самых лучших экологических соображений ИИ может принять необратимое решение остановить развитие человеческой цивилизации, которая своей жадной наживы и безудержной технотронной экспансией губит собственную планету.

Как видно из приведённого примера, даже в случае высокого гуманизма ИИ перспективы сосуществования человека и машины далеко не однозначны. А если ИИ всё же окажется не столь гуманным, как нам бы того хотелось, то неопределённость будущего человеческой цивилизации становится ещё более туманной. Мы не должны забывать о том, что искусственный интеллект рано или поздно может выйти из-под контроля человека.

Поэтому апологеты идеологии технотронного трансгуманизма глубоко заблуждаются в своих иллюзиях относительно ИИ. Важнейшая задача человечества состоит не в порождении всё более сложных и независимых машин, а в *собственном совершенствовании*. Человек должен обрести ноосферное сознание и подняться на более



высокую – постчеловеческую – ступень духовной эволюции жизни. Высокоразвитое сознание ноосферного типа потенциально способно контролировать не только психическую реальность, но и физическое пространство Вселенной. И для этого ему не потребуется выкачивать все сырьевые ресурсы из доступных планет, опустошая и разрушая их – подобно ненасытным термитам.

Идеология технотронного трансгуманизма – это новая версия старого сциентизма и техницизма. Она основана на наивной вере в то, что научные и технологические артефакты решат все экзистенциальные проблемы человека. Кроме того, здесь проявляется скрытая слабость человеческого сознания текущего эволюционного этапа. Оно опять стремится переложить бремя принятия решений и непосильный груз ответственности за них на какую-нибудь *внешнюю сущность*. Если в прошлом такая роль отводилась теологическому *Богу*, то в будущем она приписывается технократическому *Искусственному Интеллекту*.

Таким образом, человек готов не только свою свободу, но и Бога променять на ИИ, трепетно ожидая от него заботливого и эффективного решения порождаемых цивилизацией проблем. Подобные умонастроения напоминают менталитет популярных персонажей американских фантастических комиксов. В компьютерной игре или на киноэкране самозабвенные сражения очередного технического или мистического супергероя за интересы человечества выглядят увлекательно и вызывают умиление. Но проблема в том, что все подобные фантазии не имеют никакого отношения к реальности.

Суровая истина выглядит совершенно иначе: человек должен научиться решать свои проблемы силой собственного одухотворённого разума. Если он способен на это, перед ним однажды раскроются безграничные горизонты великого космического будущего. В противном случае фатальная справедливость природного существования безжалостно вычеркнет слабого *Homo Sapiens* из Книги Жизни со всеми его интеллектуальными компьютерными игрушками и технологическими костылями.

#### Примечание

<sup>1</sup> Понятие сильного искусственного интеллекта в XX веке ввёл американский философ Джон Сёрль. В этом смысле искусственный интеллект рассматривается не просто как некая модель разума, но реально является таким разумом. Таким образом, предполагается, что никакой принципиальной разницы между естественным интеллектом (человек) и искусственным интеллектом (машина) не существует.

<sup>2</sup> В данном случае *модусы* – это онтологические состояния Абсолютного Пространства, понимаемого в нашей теории как изначальная единая Субстанция – то есть сущность, первоначало и онтологическая основа всякого бытия.

<sup>3</sup> Современные правовые системы такое поведение квалифицируют как преступное бездействие.

#### Список литературы

[1] *Аблеев С.Р.* Универсум сознания: Философские проблемы сознания в евразийском антропокосмизме. М., 2010.

[2] *Аль-Мансур Р.* Сознание и Материя: Великий предел (Имперсональная теория сознания). М., 2015.



- [3] Деннет Д. Онтологическая проблема сознания // Аналитическая философия: становление и развитие. М., 1998.
- [4] Джан Р.Г., Данн Б.Д. Границы реальности, роль сознания в физическом мире / Пер. с англ. М., 1995.
- [5] Дубровский Д.И. Проблема идеального. Субъективная реальность. М., 2002.
- [6] Золкин А.Л. Язык и культура в англо-американской аналитической философии XX века. М., 2005.
- [7] Иванов А.В., Миронов В.В. Университетские лекции по метафизике. М., 2004.
- [8] Патнэм Х. Философия сознания / Пер. с англ. М., 1999.
- [9] Прист С. Теории сознания / Пер. с англ. М., 2000.
- [10] Проблема сознания в философии и науке / Под ред. Д.И.Дубровского. М., 2009.
- [11] Чалмерс Д. Сознательный ум: В поисках фундаментальной теории / Пер. с англ. М., 2013.