

Некоторые онтологические аспекты проблемы искусственного интеллекта

© М.М. Догужиева, Э.Н. Кузнецов

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

На обширном историко-философском и научном материале проанализированы онтологические модели человеческого интеллекта, способные прояснить проблемы и трудности, которые связаны с возможностью технической реализации проекта искусственного интеллекта в версии сильного искусственного интеллекта. Показано, какие из моделей допускают создание в обозримой перспективе технологий искусственного интеллекта, а какие налагают на это принципиальный запрет. Обоснованы актуальность и значимость философской проблематики в контексте развития современных технологий.

Ключевые слова: искусственный интеллект, онтологическая модель, асимметрия живого, время, вечность, эволюция, интуиция

В настоящее время проблема искусственного интеллекта (ИИ) перешла из разряда теоретических в практическую плоскость. Многие футурологи считают реализацию этой идеи делом вполне обозримого будущего и видят проблему только в том, насколько положительными или отрицательными будут последствия широкого внедрения ИИ. Например, известный японский футуролог Мориносукэ Кавагути (консультант правительства Японии) считает, что люди напрасно боятся ИИ: «На деле мы просто становимся все более супер и супер» [1]. В противоположность этому израильский историк и философ Ю.Н. Харари — автор бестселлеров «Sapiens. Краткая история человечества» и «Homo Deus. Краткая история будущего» — обращает гораздо больше внимания на опасность внедрения современных технологий (включая ИИ) [2]. Однако не менее важной и актуальной в философском плане является другая проблема: возможно ли *в принципе* создание ИИ силами человеческого разума в условиях, когда люди до сих пор не понимают природу и происхождение интеллекта естественного? Решение этого вопроса напрямую зависит от того, какая философская парадигма лежит в основе технологического оптимизма разработчиков и пропагандистов идеи ИИ. По мнению авторов статьи, их позиция напоминает некритически принятые на веру постулаты — как у «свидетелей» в псевдорелигиозном смысле. Безусловно, такая вера вполне может привести «свидетелей ИИ»

к достижению кажущихся сегодня фантастическими результатов. Но при этом нельзя исключать и другие онтологические модели, которые раскрывают проблему ИИ в совершенно ином ключе. В настоящей статье, не претендуя на полноту охвата всех возможных интерпретаций, авторы рассмотрят три самые существенные из них.

Современные «свидетели ИИ» считают его (в версии сильного ИИ) полноценной заменой человеческого интеллекта. Важно отметить, что авторы термина «искусственный интеллект» изначально не имели столь высоких притязаний. Так, по мнению Дж. Маккарти, «проблема состоит в том, что пока мы не можем в целом определить, какие вычислительные процедуры мы хотим называть интеллектуальными. Мы понимаем некоторые механизмы интеллекта и не понимаем остальные [курс. авт. — Э.К.]. Поэтому под интеллектом в пределах этой науки понимается только вычислительная составляющая способности достигать целей в мире» [3, с. 179].

Для людей, пытающихся создать ИИ на базе информационных технологий, психофизической проблемы (так называемой *mind-body problem*, с художественной убедительностью представленной в пьесе знаменитого драматурга Тома Стоппарда «*The Hard Problem*») фактически не существует: она считается разрешенной в пользу безусловного приоритета физического. На парадигмальном уровне они уже неявно приняли для себя определенную онтологическую модель природы человеческого интеллекта, игнорируя альтернативные модели, существующие в истории философии с глубокой древности до настоящего времени. Целью данной статьи является философский анализ основных онтологических моделей феномена человеческого интеллекта, который призван прояснить реальные трудности (а в пределе — возможность и невозможность) практического создания ИИ.

Первая модель восходит к древнейшей концепции материализма, берущей свое начало в идеях Левкиппа — Демокрита на Западе и школы локаята (чарвака) на Востоке. Она представляет человеческое сознание (интеллект, духовную жизнь) как результат естественного эволюционного развития природы (от неживой — через живую — к разумной форме жизни). Материалистическая трактовка с неизбежностью предполагает включение фактора времени как длительности — но не в смысле *durée* А. Бергсона, понятия, которое В.И. Вернадский правильно передал как «дление». В этой концепции большой интервал времени необходим, чтобы сделать возможным крайне маловероятное, по всем расчетам, появление жизни и разумной жизни. Тем самым «дурная бесконечность» времени помогает преодолеть качественный скачок от неживого через живое к живому-разумному. Сегодня объяснение этих процессов предлагается, в частности, в рамках синергетики (а также в трудах И.Р. Пригожина и др.), постулирую-

щей возможность самоорганизации материи, но, по мнению авторов статьи, оно не является достаточным. По сути, первую модель в полной мере можно считать вариантом редукционизма (сведения сложного к простому) — обратного к линии эволюции движения интерпретирующего разума.

Вторая модель наиболее ярко представлена в философии жизни (Бергсон) и в концепциях ноосферы Э. Леруа, П. Тейяра де Шардена и Вернадского. Эти мыслители считали феномен жизни так или иначе вечным в космическом времени и отрицали возможность ее самозарождения в неорганической природе. В то же время человеческий разум они рассматривали как этап естественной эволюции жизни (образно выражаясь, как один из побегов на вечном древе жизни). Вот как пишет об этом Тейяр де Шарден в своей знаменитой книге «Феномен человека»: «...первичная материя представляет собой нечто большее, чем кишение частиц... Под этим первичным слоем следует представить себе до крайности тонкий, но абсолютно необходимый для объяснения состояния космоса в последующие времена “биологический” слой. ...В цельной картине мира наличие жизни неизбежно предполагает существование до нее беспредельно простирающейся преджизни. Если бы органическое не начало существовать на Земле с первого возможного момента, то оно так никогда бы и не возникло» [4, с. 164, 165, 180].

К этой же модели отчасти можно отнести и концепцию психической энергии *libido* у З. Фрейда, для которого оно было энергией жизни, определяющей — как базис — надстройку в виде разнообразных видов человеческой деятельности (от поэтического творчества до политической активности). Биологизаторские тенденции в творчестве Фрейда также порой приводили его к предельному редукционизму.

Третья модель органически связана с высоким идеализмом Платона, неоплатоников и христианской теологии. Особенно отчетливо это можно видеть на примере концепции Плотина, у которого ум (греч. νοῦς — нус) — отнюдь не продукт эволюции природы: он появляется как первая эманация божественного Единого (которая объединяет мыслящее и мыслимое). Второй эманацией является душа, и только третьей — природа (включающая в себя жизнь). А материя у Плотина вообще является последней — низшей — эманацией Единого, в которой божественное полностью угасает (превращаясь в «низшее ничто»): «...мировая, универсальная Душа произвела, вдохнув в них жизнь, все живые существа..., она же произвела божественные звезды и солнце, да и всю красоту форм необъятного неба... Душа, нисходя в эту громадную инертную массу, сообщает ей движение и жизнь и превращает ее в мир, который, вечно движимый ее разумной

силой, стал живым вседвоящим существом... Душа придала ему смысл, ценность и красоту, между тем как без нее все это было бы мертво, земля да вода, или даже нечто еще худшее — темная бездна и небытие...» [5].

По мнению авторов статьи, именно первая онтологическая модель лежит в основе представлений о человеческом интеллекте у большинства современных «свидетелей ИИ». Ведь если природа за миллиарды лет эволюции создала — как результат слепого случая — жизнь и разум, то человек разумный вполне в состоянии сжать этот процесс во времени и на элементной базе косной материи создать не только искусственную жизнь (возможность чего допускал уже Тейяр де Шарден: «Химики сумеют воспроизвести этот феномен в лаборатории» [4, с. 207]), но и ИИ. При таком подходе предполагается абсолютность именно физического времени и пространства (И. Ньютон — Л. Эйлер) или пространства-времени (А. Эйнштейн), что связано с представлением о линейной одноправленности времени.

Между тем сторонники второй модели считали необходимым указать на специфику биологического пространства-времени (Л. Пастер, Бергсон, П. Кюри, Вернадский): хиральность (левизна-правизна) в живой природе (что доходчиво объясняет в своей лекции «Асимметрия биологических молекул» профессор Бостонского университета доктор физико-математических наук биофизик М.Д. Франк-Каменецкий [6]), время-длание (*durée* Бергсона) и т. п. Они отмечали непреодолимые трудности, с которыми сталкиваются теории, объясняющие возникновение жизни (асимметричной) из «симметричной причины». Как об этом писал Вернадский (ко всему прочему, один из основоположников кристаллографии, воплощающей в себе симметрию неживого): «Диссимметрическое явление... всегда должно вызываться такой же диссимметрической причиной» [7, с. 31].

Стало общим местом многочисленных «круглых столов» безосновательное утверждение о том, что современным ученым удалось синтезировать жизнь. На этом основании строятся смелые утверждения о возможности создания ИИ — несмотря на все технологические и прочие трудности, которые считаются преодолимыми в перспективе времени. Однако биологам в реальности удалось лишь модифицировать некоторые виды живой материи (клетки, белковые соединения и т. д.) — оказалось, что жизнь «прикуривается» от жизни. Неявные сторонники второй модели иногда предлагают создавать ИИ на базе биологических носителей. Это как минимум обеспечивает включение в ИИ биологического пространства-времени (выявляющего нелинейные процессы типа фолдинга белка, связанного с проблемой зарождения и функционирования живых организмов).

При этом адепты второй модели игнорируют возможные различия между биологическим временем и временем разумного духа, представляя его всего лишь переходным этапом в эволюции живой природы: «*Homo sapiens* не есть завершение создания, он не является обладателем совершенного мыслительного аппарата. Он служит промежуточным звеном в длинной цепи существ, которые имеют прошлое и, несомненно, будут иметь будущее» [5, с. 75]. В настоящее время это будущее раскрывается, в частности, в концепции трансгуманизма, которая предполагает скорее не создание ИИ, а некий апгрейд биологического вида *Homo sapiens* с использованием всех технологических возможностей, созданных силами естественного интеллекта.

Третья онтологическая модель подчеркивает именно принципиальное различие между физическим и биологическим уровнями реальности и феноменом разумной жизни, который не только существует во времени и временном, но и открыт вечности (понимаемой не как «дурная бесконечность» во времени, но как особая вневременная реальность). В этом отношении человеческий ум (с греческого языка слово «ум» — νοῦς — не случайно переводится на латинский как *intellectus*) можно считать неким «перпендикуляром» к оси времени, устремленным к вечности.

Одним из проявлений этой открытости-в-вечность является такая способность человеческого интеллекта, как интуиция, без которой невозможно творчество. Но одной из целей «свидетелей ИИ» как раз и является создание «сильного ИИ», который сможет обладать некими творческими способностями. В чем видят их проявление адепты ИИ? Среди исследователей ИИ есть общая договоренность о том, что сильный ИИ обладает следующими свойствами:

- принятие решений, использование стратегий, решение головоломок и действия в условиях неопределенности;
- представление знаний, включая общее представление о реальности;
- планирование;
- обучение;
- общение на естественном языке;
- объединение всех этих способностей для достижения общих целей.

Компьютеры, воплощающие в себе сильный ИИ, «могут приобрести способность мыслить и осознавать себя как отдельную личность (в частности, понимать собственные мысли)» [8]. Технологической базой для реализации подобных задач становятся нейронные сети с возможностями обработки больших баз данных (*big data*) и сложными алгоритмами самообучения. При этом упускается из виду,

что все это разворачивается в исключительно физическом времени, так как информация обрабатывается физическим носителем: это микропроцессоры (тактовая частота), базы данных (многомерный массив) и многоуровневые взаимодействия между ними. Хотя носителем человеческого мышления тоже является мозг, отчасти подобный искусственным нейронным сетям, оно в полной мере не сводится к ним (все-таки пресловутая *mind-body problem* не разрешена до сих пор!). Это является одним из аргументов против первой онтологической модели.

Сторонники третьей модели указывают на то, что человеческий интеллект напрямую связан с интуицией как непосредственным схватыванием информации, без временного дискурсивного развертывания (что означает открытость интеллекта вечности как некой вневременной реальности). Подобные рассуждения присутствуют, например, у Плотина в его описании Ума как второй довременной эманации Единого, что было отмечено выше. Человеческий разум как его воплощение тоже мыслит интуитивно, схватывая все непосредственно, сразу, вне времени. В противоположность этому связанная с материальным космосом душа мыслит как раз дискурсивно, дискретно, развернуто, что с неизбежностью требует линейного времени [9]. Именно такие возможности открывает перед человеком ИИ в его слабом варианте, который и реализуется на современном этапе развития технологий.

Постепенно укрепляется убеждение в том, что работа человеческого интеллекта сводится по преимуществу к работе с массивом информации, разворачивающейся в линейном времени на основе физических носителей информации и средств ее обработки. Обработка больших баз данных с экспоненциально нарастающей скоростью, объединение вычислительных средств в нейронные компьютерные сети (и даже способность их к самообучению на основе алгоритмов, разрабатываемых ими же самими и уже не совсем понятных для создающих их инженеров) — все это остается в рамках дискурсивного линейного времени и однонаправленной причинно-следственной цепи событий. При этом за рамками остается подлинно творческая способность человеческого интеллекта, имеющая надвременной характер (на языке неоплатоников — высшая реальность Единого и Ума как его первой эманации, а на языке авраамических религий — духовное начало человека).

Именно в контексте третьей онтологической модели становятся понятны опасения футурологов, связанные с утратой человеком его собственной идентичности. Он начинает воспринимать себя как несовершенную машину, а свой интеллект — как несовершенный компьютер: передавая все большую часть своих способностей ма-

шине, он обедняет эмоционально-ценностную сторону своей личности, ослабляет свой творческий потенциал, утрачивает интерес к искусству и т. п. Таким образом, рвется нить, связывающая человека с божественным Единым, и гаснет «искра Божья» в его душе.

Подводя итоги работы, можно сделать следующие выводы. Первая из рассмотренных онтологических моделей не накладывает принципиальных ограничений на создание ИИ в его сильной версии — более того, предвидит конструктивный синтез искусственного и естественно-го интеллекта как двигатель прогресса человеческой цивилизации. Вторая модель допускает возможность создания ИИ на базе биологических носителей (киборгов и т. п. в рамках идей трансгуманизма). Но третья модель налагает принципиальный запрет на саму возможность реализации сильного ИИ — ни на физической, ни на биологической основе. Более того, она предостерегает от главной опасности на пути в светлое технологическое будущее: вместо прогресса, обеспечиваемого с помощью сильного ИИ, человек рискует деградировать до уровня слабого ИИ — некой бесстрастной вычислительной машины, которой неведомы любовь, творчество, свобода... Только время покажет, какого будущего заслужит интеллект в версии *Homo sapiens*.

ЛИТЕРАТУРА

- [1] Искусственный интеллект не представляет угрозы, уверен японский футуролог. *РИА Новости*. URL: <https://ria.ru/20210103/tehnologii-1591869736.html> (дата обращения 02.03.2021).
- [2] «Вы можете улыбаться и хлопать в ладоши, но уже завтра окажетесь в ГУЛАГе»: автор Sapiens напугал Давос. *Forbes*. URL: <https://www.forbes.ru/tehnologii/391789-vy-mozhete-ulybatsya-i-hlopat-v-ladoshi-no-uzhe-zavtra-okazhetes-v-gulage-avtor> (дата обращения 03.03.2021).
- [3] Осипова И.Н. Искусственный интеллект — угроза или помощник человека? *Экономика. Общество. Человек. Материалы Всероссийской научно-практической конференции с международным участием*. Белгород, Изд-во Белгородского государственного технологического университета им. В.Г. Шухова, 2019, с. 179–182.
- [4] Тейяр де Шарден П. *Феномен человека. Божественная среда*. Москва, АСТ, Астрель, 2011, 446 с.
- [5] Плотин. *Энеады. Трактат. V.1, V.2*. URL: <https://classics.nsu.ru/bibliotheca/ploti01/index.htm> (дата обращения 03.03.2021).
- [6] Франк-Каменецкий М. *Асимметрия биологических молекул*. URL: <https://www.youtube.com/63pDaP8TQaTBTxID> (дата обращения 03.03.2021).
- [7] Вернадский В.И. *Философские мысли натуралиста*. Москва, Наука, 1988, 520 с.
- [8] Сильный и слабый искусственные интеллекты. *Википедия*. URL: https://ru.wikipedia.org/wiki/Сильный_и_слабый_искусственные_интеллекты (дата обращения 03.03.2021).

- [9] Бугай Д.В. *Аристотель и философия поздней античности. Неоплатонизм (продолжение)*. URL: <https://www.youtube.com/watch?v=rJHE4jHKmnQ&t=2383s> (дата обращения 03.03.2021).

Статья поступила в редакцию 09.04.2021

Ссылку на эту статью просим оформлять следующим образом:

Догужиева М.М., Кузнецов Э.Н. Некоторые онтологические аспекты проблемы искусственного интеллекта. *Гуманитарный вестник*, 2021, вып. 2.
<http://dx.doi.org/10.18698/2306-8477-2021-2-713>

Догужиева Марина Муратовна — канд. филос. наук, доцент кафедры «Философия» МГТУ им. Н.Э. Баумана. e-mail: aniramd@list.ru

Кузнецов Эдуард Николаевич — старший преподаватель кафедры «Философия» МГТУ им. Н.Э. Баумана. e-mail: ed0404@yandex.ru

Some ontological aspects of the problem of artificial intelligence

© M.M. Doguzhieva, E.N. Kuznetsov

Bauman Moscow State Technical University, Moscow, 105005, Russia

The article uses extensive historical, philosophical and scientific material to analyze ontological models of human intelligence that can clarify the problems and difficulties associated with the possibility of technical implementation of an artificial intelligence project in the version of a strong AI. The authors show which of them allow the creation of artificial intelligence technologies in the foreseeable future, and which impose a fundamental ban on this. The article aims to substantiate the relevance and significance of philosophical issues in the context of the development of modern technologies.

Keywords: artificial intelligence, ontological model, asymmetry of the living, time, eternity, evolution, intuition

REFERENCES

- [1] Iskusstvenny intellekt ne predstavlyaet ugrozy, uveren yaponskiy futurology [Artificial intelligence does not pose a threat, the Japanese futurist is sure]. *RIA Novosti* [RIA News]. Available at: <https://ria.ru/20210103/tehnologii-1591869736.html> (accessed March 2, 2021).
- [2] «Vy mozhetе ulybatsya i khlopat v ladoshi, no uzhe zavtra okazhetes v GULAGe»: avtor Sapiens napugal Davos [“You can smile and clap your hands, but tomorrow you will find yourself in the Gulag”: the author of Sapiens scared Davos]. *Forbes*. Available at: <https://www.forbes.ru/tehnologii/391789-vy-mozhete-ulybatsya-i-hlopat-v-ladoshi-no-uzhe-zavtra-okazhetes-v-gulage-avtor> (accessed March 3, 2021).
- [3] Osipova I.N. Iskusstvenny intellekt — ugroza ili pomoshchnik cheloveka? [Artificial Intelligence — Threat or Human Assistant?] *Ekonomika. Obschestvo. Chelovek. Materialy Vserossiiskoy nauchno-prakticheskoy konferentsii s mezhdunarodnym uchastiem* [Economy. Society. Human. Materials of the All-Russian scientific-practical conference with international participation]. Belgorod, BSTU Publ., 2019, pp. 179–182.
- [4] De Chardin P.T. *Le Phénomène humain*. Le Seuil, 1970, 320 p. [In Russ.: De Chardin P.T. Fenomen cheloveka. Bozhestvennaya sreda. Moscow, AST, Astrel Publ., 2011, 446 p.].
- [5] Plotinus. *Enneady* [The Enneads]. Vol. 1, vol. 2. Available at: <https://classics.nsu.ru/bibliotheca/ploti01/index.htm> (accessed March 3, 2021).
- [6] Frank-Kamenetskiy M. *Asimmetriya biologicheskikh molekul* [Asymmetry of biological molecules]. Available at: <https://www.youtube.com/63pDaP8TQaTBTxlD> (accessed March 3, 2021).
- [7] Vernadsky V.I. *Filosofskie mysli naturalista* [Philosophical thoughts of a naturalist]. Moscow, Nauka Publ., 1988, 520 p.
- [8] Silny i slabyi iskusstvennye intellekty [Strong and weak artificial intelligence]. *Wikipedia*. Available at: https://ru.wikipedia.org/wiki/Сильный_и_слабый_искусственные_интеллекты (accessed March 3, 2021).
- [9] Bugay D.V. *Aristotel i filozofiya pozdney antichnosti. Neoplatonizm (prodolzhenie)* [Aristotle and the philosophy of late antiquity. Neoplatonism (continued)]. Available at: <https://www.youtube.com/watch?v=rJHE4jHKmnQ&t=2383s> (accessed March 3, 2021).

Doguzhieva M.M., Cand. Sc. (Philos.), Assoc. Professor, Department of Philosophy, Bauman Moscow State Technical University. e-mail: aniramd@list.ru

Kuznetsov E.N., Assist. Professor, Department of Philosophy, Bauman Moscow State Technical University. e-mail: ed0404@yandex.ru