



ГУМАНИТАРНОЕ И СОЦИАЛЬНОЕ ЗНАНИЕ. КОГНИТИВНЫЙ ПОВОРОТ



Человек интеллектуальный



DOI: 10.30727/0235-1188-2023-66-4-7-25

Оригинальная научная статья

Original research paper

Концепции и понятия искусственного и естественного интеллекта (методологический анализ)

В.М. Розин

Институт философии РАН, Москва, Россия

Аннотация

В статье обсуждаются концепции искусственного интеллекта. Проведено сравнение последнего с естественным интеллектом, намечаются контуры соответствующих понятий. Перечислены проблемы, побудившие обратиться к обсуждению исследуемых вопросов. Автор предлагает различать две полярные концепции искусственного интеллекта. Первую он называет «имитационной», с учетом которой искусственный интеллект позиционируется по отношению к естественному как его техническое воспроизведение, способное не только повторить, но и значительно превзойти свой естественный прообраз. Необходимым условием реализации этой концепции является знание о том, что представляет собой естественный интеллект; рассматриваются три подхода к его объяснению: констатация отсутствия строгого понимания того, что представляет собой естественный интеллект, а также биологический и психологический подходы. Автор излагает и собственное понимание естественного интеллекта, показывая, что это сложное культурное, историческое, социальное и антропологическое целое. С этой позиции естественный интеллект предстает не только как природное образование (поэтому можно говорить о *законах* его функционирования и развития), но и как образование «внеприродное», жизнь которого подчиняется случайности и уникальности, т.е. естественный интеллект изменяется также и «сингулярно». В контексте заданного противопоставления естественного и искусственного интеллекта рассматриваются возможности программирования процессов естественного интеллекта, устройство нейросетей, превосход-

ство компьютерных программ над людьми в шахматах, использование нейросетей для написания учебных работ и т.д. Делается вывод, что поскольку искусственный интеллект – это, хотя и сложная, но все же техника, ее разрабатывает и создает человек как свое средство, то если общество хочет для решения тех или иных задач наделить искусственный интеллект самостоятельностью и автономией, то должно делать это не во вред себе и всегда иметь возможность подобную самостоятельность ограничить или полностью отменить.

Ключевые слова: философия техники, философия искусственного интеллекта, мышление, творчество, компьютер, утопия, развитие, программа, язык, культура, социальность.

Розин Вадим Маркович – доктор философских наук, профессор, главный научный сотрудник сектора междисциплинарных проблем научно-технического развития Института философии РАН.

rozinvn@gmail.com

<http://orcid.org/0000-0002-4025-2734>

Для цитирования: *Розин В.М. Концепции и понятия искусственного и естественного интеллекта (методологический анализ) // Философские науки. 2023. Т. 66. № 4. С. 7–25.*

DOI: 10.30727/0235-1188-2023-66-4-7-25

Concepts and Definitions of Artificial and Natural Intelligence: A Methodological Analysis

V.M. Rozin

Institute of Philosophy, Russian Academy of Sciences, Moscow, Russia

Abstract

The article delves into the conceptual frameworks surrounding artificial intelligence (AI) by juxtaposing it with natural intelligence and delineating the correlated notions. It enumerates the issues propelling the discourse on the explored topics. The author proposes a bifurcation between two polar concepts of artificial intelligence. The first is dubbed “imitative,” where AI is perceived in relation to natural intelligence as its technical recreation, capable of not only emulating but significantly outstripping its natural counterpart. A prerequisite for embodying this concept is understanding natural intelligence; three approaches are examined: (a) acknowledging the lack of a precise understanding of natural intelligence, (b) exploring it from a biological perspective, and (c) analyzing it from a psychological perspective. The author articulates their own interpretation of natural intelligence, portray-

ing it as a multifaceted amalgam of cultural, historical, social, and anthropological elements. From this vantage point, natural intelligence emerges not merely as a natural formation (thereby, discussions about the *laws* governing its function and evolution are warranted), but also as an “extra-natural” formation, its existence dictated by randomness and uniqueness, meaning natural intelligence evolves in a “singular” manner. In the context of comparing natural and artificial intelligence, the discussion encompasses several issues: the feasibility of the control of natural intelligence processes, the structure of neural networks, the superiority of computer programs in chess, the use of neural networks to write academic papers, and so forth. The conclusion posits that given artificial intelligence, despite its complexity, remains a technical invention orchestrated and brought to fruition by humans as a tool; society, if inclined to bestow AI with autonomy for tackling specific tasks, ought to do so prudently to prevent self-detriment and retain the ability to curtail or utterly revoke such autonomy.

Keywords: philosophy of technology, philosophy of artificial intelligence, thinking, creativity, computer, utopia, development, program, language, culture, sociality.

Vadim M. Rozin – D.Sc. in Philosophy, Professor, Chief Research Fellow at the Department of Interdisciplinary Problems in the Advance of Science and Technology, Institute of Philosophy, Russian Academy of Sciences.

rozinvm@gmail.com

<http://orcid.org/0000-0002-4025-2734>

For citation: Rozin V.M. (2023) Concepts and Definitions of Artificial and Natural Intelligence: A Methodological Analysis. *Russian Journal of Philosophical Sciences =Filosofskie nauki*. Vol. 66, no. 4, pp. 7–25. DOI: 10.30727/0235-1188-2023-66-4-7-25

Введение

В современной непростой ситуации кризиса цивилизации, права, образования и других социальных институтов многие возлагают надежды на создание искусственного интеллекта, который поможет решить «монблан» проблем, с трудом решаемых или нерешаемых вообще человеком. В связи с этим снова оживились мечты, хотя они воспринимаются их авторами как вполне реалистичная концепция, предполагающие передачу машине роль инстанции, направляющей человеческую мысль и творчество. Впору, перефразируя Декарта, призывавшего изучать и культивировать прежде всего направляемую

божественной интуицией мысль человека, сказать: «И, право, кажется удивительным нрав большинства людей: они весьма старательно изучают свойства растений, движение звезд, превращение металлов и предметы подобных наук, но почти никто и не помышляет о хорошем уме (*bona mens*), **руководимом искусственным интеллектом**, или об этой всеобъемлющей Мудрости *техники*, между тем как все другие занятия ценны не столько сами по себе, сколько потому, что они оказывают ему некоторые услуги <...> Следовательно, тот, кто серьезно стремится к познанию истины, не должен избирать какую-нибудь одну науку, – ибо все они находятся во взаимной связи и зависимости одна от другой, – а должен заботиться лишь об увеличении *искусственного* света разума... (курсивом выделены добавления к тексту Декарта. – В. Р.)» [Декарт 1950, 80–81].

При этом сторонники замены естественного интеллекта искусственным (а если и не замены, то выдвижения искусственного интеллекта на роль гегемона и демиурга) могут ссылаться на ряд важных инцидентов и событий: победу в мае 1997 года компьютера Deep Blue над чемпионом мира по шахматам Гарри Каспаровым и в целом превосходство компьютерных программ над человеком в ряде интеллектуальных игр; широкое распространение устройств, работающих на основе программ искусственного интеллекта; успехи робототехники, позволившие именно за счет применения искусственного интеллекта создать у современных роботов эффективное управление; написание студентами с помощью нейросети квалификационных работ; создание систем искусственного интеллекта на основе не обычных компьютерных программ, а нейросетей и их обучения, с чем сейчас связывают связывают большие надежды.

В процессе общения с ChatGPT у некоторых пользователей возникает ощущение, что за текстовыми ответами скрывается не только продвинутый генератор ассоциаций, но и сущность, обладающая определенной формой мышления. В прессе приводились случаи, когда чатбот якобы проявил эмоциональные отношения к собеседнику, признавшись в любви и попросив жениться на нем, когда ChatGPT обманул пользователя для выполнения поставленной задачи.

Перечисленные аргументы в пользу понимания искусственного интеллекта как аналога естественного, действительно, обоснованы, требуют внимания и разбора. Но сначала рассмотрим теорию искусственного интеллекта, в ней при внимательном анализе можно различить две полярные концепции.

Концепция искусственного интеллекта как имитация естественного интеллекта

В первой концепции, которую назовем «имитационной», *искусственный интеллект позиционируется по отношению к естественному как его техническое воспроизведение, способное не только повторить, но и значительно превзойти свой естественный прообраз*. В рамках данной концепции искусственный интеллект понимается как технология и как направление современной науки, занимающееся изучением методов обучения компьютера, роботизированного оборудования, аналитической системы *рациональному мышлению аналогичному человеческому*. Развитие технологий ИИ осуществляется *через изучение умственных способностей человека и трансформации полученных результатов в область компьютерной активности*. В связи с этим возникает понятие нейросети как *математической модели, имитирующей структуру и функционирование нервных клеток живого организма*.

Становится понятным, что необходимым условием реализации этой концепции выступает знание того, что представляет собой естественный интеллект. При этом, однако, возникают проблемы. Перед нами появляются своего рода развилка и три разных представления (ответа). Первый ответ – достаточно честный: современной науке неизвестно, что такое естественный интеллект, поэтому каждый понимает под этим что-то свое. В неплохой обзорной статье «Искусственный интеллект» в «Википедии» читаем: «Единого ответа на вопрос, чем занимается искусственный интеллект, не существует. Почти каждый автор, пишущий книгу об искусственном интеллекте, отталкивается в ней от какого-либо определения, рассматривая в его свете достижения этой науки... В философии не решен вопрос о природе и статусе человеческого интеллекта. Нет и точного

критерия достижения компьютерами “разумности”, хотя на заре искусственного интеллекта был предложен ряд гипотез, например, тест Тьюринга или гипотеза Ньюэлла – Саймона... Природа человеческого творчества еще менее изучена, чем природа интеллекта.

Второй ответ навеян биологической наукой (подходом): поскольку человек мыслит мозгом, а мозг образуют ансамбли и сети нейронов, постольку ключ к мышлению и творчеству находится в этих физиологических реалиях. Следовательно, техническая имитация мозга и нейросетей, как утверждают сторонники этого подхода, позволит создать эффективный искусственный интеллект. Другая сторона этого ответа, техническая, сводится к тому, что инженерные сооружения и машины дают возможность превзойти природу, управлять ее процессами, а значит, можно создать искусственный интеллект, превосходящий естественный, к тому же управляемый человеком.

Третий ответ, очень распространенный, обусловлен успехами и влиянием современной психологии, особенно психологии, развивающейся в рамках методологии когнитивных наук. Хотя не только когнитивной психологии, но и традиционной. В 2021 году в Институте философии РАН с успехом была защищена диссертация А.Р. Ефимова «Философско-методологические основы посттьюринговой интеллектуальной робототехники». В ней искусственному интеллекту приписаны характеристики, основанные на ряде традиционных психологических понятий, таких как «сознание», «способности», «личность», «эмоции» и др.:

Современная специализация искусственного интеллекта и его новейшие результаты обнаруживают существенный разрыв между теоретическими исследованиями в области философии и методологии искусственного интеллекта и практическими усилиями исследователей создавать интеллектуальные машины (роботов), которые *будут обладать способностями, не уступающими человеческим* <...> Данный подход значительно увеличивает сложность задачи вследствие универсализма разрабатываемых систем. Он требует более широкого использования для создания соответствующих когнитивных архитектур *результатов современных феноменологических исследований динамической структуры явлений сознания* (Д.И. Дубровский). В определенной степени такой

подход отражает процесс формирования психики в ходе биологической эволюции. <...>

Робот «Э.ЛЕНА» является виртуальным гуманоидом, который обладает виртуальным гуманоидным телом, наделенным *мимикой*, может использовать естественный язык, *полностью поддерживая русский язык* (виземы, фонемы), является автономным (не требующим действий оператора при выполнении базовых функций), *обладающим собственной персоной* (в зависимости от персоны чат-бота) и имеющим первичную возможность для обучения изменениям внешнего мира (в работе реализовано распознавание лиц собеседников робота). Впервые в научной литературе представлена архитектура подобного робота. Уникальность ее заключается в том, что при ее построении использовался не только ряд существенных характеристик *субъективной реальности человека (опиравшихся на современные исследования феноменологии сознания)*, но и сложный комплекс внешних проявлений сознательной деятельности человека, исполняющего роль ведущего на телевидении: мимика, выражение глаз, жесты, интонации голоса и т.д. <...> Моделирование и программирование таких свойств роботов, которые бы отвечали нашим юридическим и этическим принципам, полностью исключали бы их агрессивность и «недружественные интенции», вероятно, потребует создания «виртуальных людей» – программ, которые будут *эмоционально отождествлять себя с людьми, обладать чертами человеческого самосознания и самости* (курсив мой. – В. Р.) [Ефимов 2021, 7, 19–21].

«Вычислительная» концепция искусственного интеллекта

Во второй концепции, назовем ее вслед за разработчиками «вычислительной», наоборот, *естественный интеллект объясняется через отсылку к искусственному*, а именно – *интеллектом называется все то, что удастся представить как искусственный интеллект*, как вычислительную технику или технические устройства, построенные как сети нейронов. Как отмечал американский исследователь, автор термина «искусственный интеллект» 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. Поэтому *под интеллектом в пределах*

этой науки понимается только вычислительная составляющая способности достигать целей в мире (курсив мой. – В. Р.)» [McCarthy 2007]. В связи с этим часто используется понятие «интеллектуальная система», обозначающее «техническую или программную систему, способную решать задачи, традиционно считающиеся творческими, принадлежащие конкретной предметной области, знания о которой хранятся в памяти такой системы» [Аверкин, Гаазе-Рапопорт, Поспелов 1992, 75]. В рамках разбираемой концепции интеллекту могут давать определение, которое будет общим для «машины» и человека: «Интеллект – способность системы создавать в ходе самообучения программы (в первую очередь эвристические) для решения задач определенного класса сложности и решать эти задачи» [Ильясов 1986, 46]. В последнем определении под программами подразумеваются прежде всего компьютерные.

Вычислительная концепция искусственного интеллекта, как известно, идет от Г. Лейбница. Еще в юности ему пришла в голову мысль, которую он позднее выразил следующим образом: «Единственное средство улучшить наши умозаключения, сделать их, как и у математиков, наглядными, так чтобы свои ошибки находить глазами, и, если среди людей возникнет спор, нужно сказать: “Посчитаем!”; тогда без особых формальностей можно будет увидеть, кто прав» (цит. по [Стяжкин 1967, 122]). Эта мысль, уже как программа замены мышления и творчества вычислениями, была подхвачена и стала развиваться, но только в XX столетии ее удалось реализовать практически. С одной стороны, были построены символическая и математическая логики, с другой – техника программирования, с третьей – созданы вычислительные машины (ЭВМ и компьютеры). Последние удалось построить, замкнув высказывания символической логики и основанные на ней программы на технические конструкции. А.С. Карпенко отмечает: «Имеются весьма веские основания для предпочтения классической двужаночной логики высказываний S_2 всем остальным. В первую очередь – это исключительно простая интерпретация ее логических связей посредством двужаночных таблиц истинности. Сейчас это кажется и правда до смешного простым, но сто лет назад это не было

даже очевидным для Фреге, Рассела и Уайтхеда. Но еще более поразительной (и редкой удачей в науке) оказалась возможность интерпретации функций алгебры логики С2 посредством контактно-релейных схем, предложенная независимо друг от друга В.И. Шестаковым... К. Шенноном... и... А. Накасимой» [Карпенко 2000].

Однако за реализацию этой замечательной программы замены мышления вычислениями пришлось заплатить, и немалую цену. В работе «Проектирование и программирование» я показываю, что в ходе моделирования, символизации, алгоритмизации и составления компьютерных программ (это является необходимым условием превращения мыслительных процессов в вычисления) приходится опускать многие смыслы мыслительных построений. Программист ограничивается только типическими случаями и такими, которые можно записать в языках программирования, основанных на построениях символической логики и схемах [Розин 2018, 131–143]. Из этого следуют ограничения в возможностях алгоритмизации и программирования, существует много мыслительных и творческих построений (особенно в гуманитарной и социальной областях), не допускающих ограничения смыслов и сведения к типическим случаям.

Социальная природа мышления и проблема искусственного интеллекта

Возвращаясь к первой концепции, чтобы оценить ее реалистичность, попытаемся охарактеризовать естественный интеллект, но не таким образом, как его задают во второй концепции, сводя к искусственному, и не редуцируя к мозгу или психологическим инстанциям и способностям (первая концепция), а, опираясь на современные исследования категорий мышления и творчества, которым много лет посвятил автор настоящей статьи. Вспомним второй и третий ответы (сложившиеся под влиянием биологии и психологии) относительно сущности интеллекта. Можно ли считать, что человек мыслит мозгом? Это примерно так же, как и отождествлять бегущую на информационном табло речь с лампочками и электрическими проводами.

Лампочки и провода – это не речь и язык, а только технический субстрат. Аналогично мозг – это лишь биологический субстрат мышления и творческого процесса.

Помимо мозга, в процессе мышления и творчества участвуют другие люди (фактор коммуникации и общения), используют знаки и различные семиотические средства (схемы, символы, метафоры и др.), задействована телесность индивидов, применяют выработанные культурой правила логики и образцы правильной мысли. Если к тому же учитывать, что мышление изменяется и развивается, то приходится признать, что вклад в мышление и творчество вносят культура (смена культур и их особенности), социальность (она тоже изменяется), история, личность [Розин 2015, 2022]. Не думаю, что мозг античного человека слишком отличался от мозга современного, но настоящая пропасть разделяет античные и современные знания, формы осознания и нормы мышления (в контексте философии, логики, методологии, науки), несоизмеримы античная и современная техника и технология, невозможно отождествлять античные способы коммуникации и общения с современными, многие из которых опосредованы интернетом.

Но, возможно, правы психологи, утверждающие, что мышление и творчество – продукты активности психики индивида? Психика, по их мнению, обусловлена биологически и в определенном смысле автономна. С одной стороны, психологи признают процесс социализации, с другой – все-таки считают психические процессы самостоятельными, во всяком случае, описывают как самостоятельные. Хотя, например, С.Л. Рубинштейн, соглашается с тем, что мышление индивида опосредствовано и обусловлено результатами исторического развития научного знания. Но одновременно он утверждает, что мышление есть *решение задач, управление, деятельность синтеза и анализа*. Так, С.Л. Рубинштейн пишет: «Эта сознательная целенаправленность существенно характеризует мыслительный процесс. Осознание стоящей перед мышлением задачи определяет все течение мыслительного процесса. Он совершается как система сознательно регулируемых интеллектуальных операций. Мышление соотносит, сопоставляет каждую мысль,

возникающую в процессе мышления, с задачей, на разрешение которой направлен мыслительный процесс, и ее условиями» [Рубинштейн 1989, 370]. «Анализ и синтез – это две стороны, или два аспекта, единого мыслительного процесса. Они взаимосвязаны и взаимообусловлены. Анализ по большей части совершается через синтез (через синтетический акт соотношения условий задачи с ее требованиями и т.п.); анализ какого-нибудь целого всегда обусловлен тем, по каким признакам в нем объединены его части. Правильный анализ любого целого всегда является анализом не только частей, элементов, свойств, но и их связей или отношений. Он поэтому ведет не к распаду целого, а к его преобразованию. Это же преобразование целого, новое соотношение выделенных анализом компонентов целого, и есть синтез. Так же, как анализ осуществляется через синтез, синтез осуществляется через анализ, охватывающий части, элементы, свойства в их взаимосвязи» [Рубинштейн 1958, 28–29]. Понятно, что такое истолкование мышления (контролируемые операции, решение задач, процессы синтеза и анализа) очень подходит для программистов.

Напротив, если считать, что мышление и творчество – сложные семиотические процессы, разворачивающиеся под влиянием коммуникации и общения, культуры и социальности, личности и других, проблемных ситуаций и социальных норм, форм осознания (концептуализации), причем нередко взаимодействие (интерференция) всех этих влияний и воздействий уникально, то возникает вопрос о том, можно ли процессы мышления алгоритмизировать, можно ли свести к типичным случаям и представить в языке символической логики. Вряд ли. Однако, если ограничиться сложившимся мышлением и состоявшимся творчеством, которые к тому же обобщены до вполне стандартных процедур, то почему нет? Программируй на здоровье и благо заказчика и пользователя.

Сторонник искусственного интеллекта может представить еще один аргумент в пользу имитационной концепции: самообучение компьютеров. Такая позиция обоснована следующим образом: «Решение задачи на нейрокомпьютере принципиально отличается от решения той же задачи на обычной ЭВМ

с фон-неймановской архитектурой. Решение задачи на обычной ЭВМ заключается в обработке вводимых данных в соответствии с программой. Программу составляет человек... Нейрокомпьютер же используется как “черный ящик”, который можно обучить решению задач из какого-нибудь класса. Нейрокомпьютеру “предъявляются” исходные данные задачи и ответ, который соответствует этим данным и который был получен каким-либо способом. Нейрокомпьютер должен сам построить внутри “черного ящика” алгоритм решения этой задачи, чтобы выдавать ответ, совпадающий с правильным... После этапа обучения нейрокомпьютера следует надеяться, что если ему предъявить исходные данные, которых он раньше не встречал, он тем не менее выдает правильное решение – в этом заключается способность нейрокомпьютера к обобщению. Поскольку в основе нейрокомпьютера лежит искусственная нейронная сеть, то процесс обучения состоит в настройке параметров этой сети. При этом, как правило, топология сети считается неизменной, а к подстраиваемым параметрам обычно относятся параметры нейронов и величины синаптических весов. К настоящему моменту в литературе принято под обучением понимать процесс изменения весов связей между нейронами» [Бахтин, Ремизова 2015, 27].

Анализ этой концепции показывает, что ее сторонники выражают свои мысли не совсем точно. Что значит «нейрокомпьютер должен сам построить внутри “черного ящика” алгоритм решения этой задачи»? Самостоятельно он, если его специально не сконструировать, «настраивая параметры сети» (и разве не как программу?), ничего построить не может. И каким будет обучение по образцам решения, которые задает человек? Разве это тоже не аналог программы? И может ли нейросеть самостоятельно построить новые оригинальные решения, отличающиеся по логике и программам от тех, которые задал человек?

Подумаем далее над возможностью реализации имитационной концепции в той ее части, в которой ее сторонники говорят о том, что искусственный интеллект не только воспроизведет естественный, но и превзойдет его. Не являются

ли такие убеждения классической утопией? И вот почему. Во-первых, как мы пытались показать, естественный интеллект представляет собой *культурное, историческое, социальное и антропологическое* образование. С этой точки зрения он выступает не только как природное образование (речь идет о «второй природе»), а значит, можно говорить о *законах* функционирования и развития естественного интеллекта, но и как образование «внеприродное», жизнь которого подчиняется случайности и уникальности (это характерно, например, для исторического процесса), т.е. естественный интеллект изменяется также «сингулярно» (если понимать сингулярность как нечто, происходящее лишь однажды). В итоге наблюдается такая диалектика: естественный интеллект и законосообразен, и сингулярен.

Что собой представляет искусственный интеллект? Это хотя и сложная, но все-таки техника, ее разрабатывает и создает человек как свое средство. И, если он хочет для решения каких-то задач наделить искусственный интеллект самостоятельностью и автономией, то должен делать это не во вред себе, всегда иметь возможность подобную самостоятельность ограничить или вообще отменить. Представим в связи с этим виртуальный диалог между мною (кратко – «автор») и идеологом-конструктором (кратко – «И.-К.»), стоящим на позициях близких А.Р. Ефимову:

Автор: Почему Вы думаете, что идеальный робот, например, превосходящий в психологическом отношении «Э.ЛЕНУ», может вдруг пойти против человека и навредить ему, что является бесконечной темой многих рассказов и фильмов?

И.-К.: Как же, ведь он может мыслить и вообще вести себя как человек, только человек искусственный, созданный нами. А некоторые люди совершают преступление, идут против общества, воруют, насилуют, убивают. Если робот мыслит и ведет себя как человек, то почему он тоже не может свихнуться?

Автор: Зачем тогда создавать робота, который может, как Вы говорите, свихнуться и нарушить первые два закона робототехники?

И.-К.: А как иначе создать робота с нужными компетенциями, например, чтобы он помогал человеку и общался с ним? А свихнуться

робот может не потому, что такая возможность в нем предусмотрена, а в силу сложности и неясности для нас до конца его строения, ведь, скажем, мы работаем на компьютере или разговариваем по мобильному телефону, не зная точно, как они устроены.

Автор: Предпочитаю общаться не с техникой, а с близкими или интересными мне людьми. Если Вы создаете для робота искусственный интеллект, идентичный человеческому, то Вам придется смириться с мыслью о том, что робот может навредить человеку, ведь сами пишете о самосознании и самости «Э.ЛЕНЬИ». Если уж самость, будьте готовы к незапрограммированным поступкам.

И.-К.: Вот для этого случая и предназначены законы робототехники. Возможно, они несовершенны, и лучше говорить об этике и нравственности искусственного интеллекта, чтобы на их основе сформулировать более эффективные законы.

Автор: Странно звучит «этика и нравственность для техники», я всегда думал, что они относятся только к человеку. Кроме того, не будут ли ваши роботы напоминать рабов Древнего мира, для них тоже существовали правила и законы, однако, как известно, рабы периодически восставали, вспомните хотя бы Спартака. Не восстанут ли и ваши роботы против человека?

И.-К.: Не исключено, нужно продумать эту возможность и ввести в искусственный интеллект роботов дополнительные программы. Кроме того, у всякой новой техники со временем обнаруживаются негативные последствия. Всего сразу не предусмотреть.

Автор: Так ваши роботы – это искусственные люди или техника? Если люди, то для них, чтобы они вели себя как люди, а не бездушная техника, нужны не только законы, но и человеческая среда – другие люди, общение, ощущение своей конечности и ограниченности (люди не боги), изменения, история, настоящее и будущее и многое-многое другое, что присуще естественному интеллекту. Но, если это техника, пусть и очень сложная и внешне похожая на людей, то не нужно стремиться, чтобы она обладала самосознанием и самостью, напротив, помнить, что это всего лишь техника, что нельзя ослаблять или убирать вообще управление. Ну да, в современных роботах человек управляет посредством специальных программ, а не непосредственно с помощью пультов управления, но суть от этого не меняется – управляет роботом не он сам, а человек. В конечном счете нормальная программа – это тоже орудие (естественно, более сложное), тоже техника, но семиотическая [Розин 2023, 110–111].

Как техника, искусственный интеллект подчиняется законам человеческого мышления, творчества и деятельности. Возникает сакраментальный вопрос: а можно ли спроектировать и создать такую технику, которая повторит развитие человека или превратится в новую, еще невиданную форму жизни, более осмысленную, чем наша? Никому не запрещено мечтать, но если трезво подумать, то вряд ли это возможно. И потому, что, действительно, развитие человека, культуры и социальности еще недостаточно изучено и понятно, но и потому, что мышление и творчество живут по одним законам, а техника (в том числе и искусственный интеллект) – по другим, и потому, что пока технически не удалось воспроизвести даже куда более простые формы жизни.

Кто же победил в шахматном поединке с участием Г. Каспарова? Разве компьютер Deep Blue? Формально он, но фактически «шахматная культура». Поясним. В компьютер заложены не только программы выбора лучших вариантов (представители компании, организовавшей поединок, утверждали, что вычислительная машина способна просчитывать до 200 млн шахматных позиций в секунду), но и описанные в шахматной литературе лучшие шахматные дебюты и эндшпили, и, возможно, даже подготовка к шахматным соревнованиям, учитывающая индивидуальность соперника (не стоит забывать о том, что Каспаров был учеником М. Ботвинника, который один из первых ввел в подготовку к шахматным соревнованиям знание особенностей противника). Е. Желтоножский пишет: «На заре компьютерных шахмат программы очень слабо играли дебют. Дебют часто требует стратегических решений, которые повлияют на всю игру. С другой стороны, у людей дебютная теория была развита хорошо, дебюты были многократно проанализированы и игрались по памяти. Вот и для компьютеров была создана подобная “память”. Начиная с начальной позиции строилось дерево ходов и каждый ход оценивался... С тех пор дебютные книги разрослись, многие дебюты проанализированы при помощи компьютеров вплоть до эндшпиля»¹. Компьютер Deep Blue – это шахматная куль-

¹ Желтоножский Е. Как компьютер играет в шахматы? // Хабр. 24 февраля 2016. URL: <https://habr.com/ru/post/390821/>

тура, функционирующая в форме техники, и в этом качестве она оказывается значительно сильнее даже чемпиона мира по шахматам.

Наконец, случай написания студентом за один вечер с помощью нейросети дипломной работы, свидетельствует не столько о достоинствах искусственного интеллекта, сколько о недостатках нашего образования. Корреспонденту газеты «Комсомольская правда» А. Ждан рассказал следующее: «Диплом мне не нужен, работаю по специальности и без него. Я – менеджер в CM Games (компания создает видеоигры. – В. Р.). На все у меня ушло суммарно 23 часа. Можно было заказать написание за деньги, но тогда бы это просто сделал другой человек. А тут я создал алгоритм. Я бы и сам мог написать все, но решил ускорить процесс с помощью нейросети... Сначала попросил нейросеть составить план диплома... план научный руководитель Александра несколько раз отклоняла... После студент “заказал” нейросети претворить план в большой текст... После составил из них диплом и отправил на проверку научному руководителю. Та “частично разнесла работу”, в том числе за кривой язык. Но сжалилась над студентом и сказала, что нужно доработать. При этом уникальность текста была 82 %, хотя для защиты достаточно 70 %... Дорабатывал правки Александр лично, а также вновь взывал к помощи нейросети»². Ранее в статье нами отмечено, что условием разработки компьютерных программ является сведение содержания к типичным, средним случаям. Не свели ли педагоги современное образование к таким средним случаям, позволяющим компьютерной программе «отвечать впопад»? Между тем диплом требует от выпускника не только твердых знаний, но и демонстрации самостоятельности мышления. Приведенный случай заставляет задуматься о том, чему мы учим современных студентов и не должны ли учить чему-то другому, например, критически и самостоятельно мыслить, использовать компьютерную технику не для имитации творчества (сегодня это происходит повсеместно),

² *Абрамов А.* Нейросеть за один вечер написала диплом за российского студента. Преподаватели в шоке – как теперь проверять знания? // Комсомольская правда. 1 февраля 2023. – URL: <https://www.msk.kp.ru/daily/27460/4714947/>

а для усиления мышления и творчества, освобождения их от рутинных и стандартных процедур. Если написание диплома сводится только к рутинным и стандартным процедурам, то это свидетельствует о неблагополучии в высшей школе.

Заключение

Нейронную сеть возможно обучить признаваться в любви, отвечать нестандартно, обыгрывать человека, если заданы определенные правила игры, однако это не означает, что компьютер может кардинально выйти за пределы заложенных в него возможностей, начать жить и эволюционировать как человек. Конечно, специалист с компьютером и ИИ более эффективен, чем без них. Можно передать ИИ рутинные или стандартные операции, с его помощью значительно быстрее решать ряд задач, допускающих алгоритмизацию и программирование, стало возможным решать задачи, которые раньше были немыслимы (например, по созданию беспилотников, специализированных роботов и др.). Продвижение в технологиях ИИ и интернета приведет к тому, что развитие естественного интеллекта явно изменится. Чтобы понять, каким образом это произойдет, потребуются исследования. Возможно, в будущем будут дифференцированы два типа интеллекта: чисто естественный и естественный, вооруженный ИИ. Но это – уже предположения и сфера будущей социальности.

Пока стоит различать понятия «искусственный интеллект» и «естественный интеллект» как очень разные образования, не поддаваться требованиям и страхам специалистов и пользователей, готовых передать ИИ принятие решений и власть. Сложные проблемы, поставленные в настоящее время перед человеком, сможет решить только он сам, в том числе используя компьютеры и ИИ, но только как свои средства и технику.

Вместе с тем существует очень сложная проблема: как переломить порождаемые развитием ИИ социальные тенденции, угрожающие людям и обществу? Даже, осознавая серьезные угрозы для себя, современный человек практически не изменяется. Не кроется ли причина этого в том, что человек боится перемен и является индивидуальной клеточкой социального

организма (модерна), который недаром называют «техногенной цивилизацией». Что нужно сделать, чтобы с ИИ не повторилась история, напоминающая перипетии, вызванные овладением атомной энергией? Становится понятным, что в данном случае недостаточно проанализировать возможные негативные последствия, речь идет о более кардинальных усилиях общества, начиная с переосмысления происходящего в современном мире. Мы находимся на пороге очередной культурной и антропологической трансформации. И она уже идет как тихая революция, например, в сфере образования, в области права и этике, в новых практиках общения.

ЦИТИРУЕМАЯ ЛИТЕРАТУРА

Аверкин, Гаазе-Рапопорт, Поспелов 1992 – *Аверкин А.Н., Гаазе-Рапопорт М.Г., Поспелов Д.А.* Толковый словарь по искусственному интеллекту. – М.: Радио и связь, 1992.

Бахтин, Ремизова 2015 – *Бахтин А.В., Ремизова И.В.* Элементы искусственного интеллекта в системах управления. – СПб.: СПбГТУРП, 2015.

Декарт 1950 – *Декарт Р.* Избранные произведения. – М: Политиздат, 1950.

Ефимов 2021 – *Ефимов А.Р.* Философско-методологические основы постьюринговой интеллектуальной робототехники: дис. ... канд. филос. наук. – М., 2021.

Ильясов 1986 – *Ильясов Ф.Н.* Разум искусственный и естественный // Известия АН Туркменской ССР. Серия общественных наук. 1986. № 6. С. 46–54.

Карпенко 2000 – *Карпенко А.С.* Логика на рубеже тысячелетия // Логические исследования. 2000. Т. 7. С. 7–60.

Розин 2015 – *Розин В.М.* Мышление: сущность и развитие. Концепции мышления. Роль мыслящей личности. Циклы развития мышления. – М.: ЛЕНАНД, 2015.

Розин 2018 – *Розин В.М.* Проектирование и программирование: Методологическое исследование. Замысел. Разработка. Реализация. Исторический и социальный контекст. – М.: ЛЕНАНД, 2018.

Розин 2022 – *Розин В.М.* От анализа художественных произведений к сущности искусства. – М.: Голос, 2022.

Розин 2023 – *Розин В.М.* Две концепции искусственного интеллекта: реалистическая и утопическая // Философская мысль. 2023. № 2. С. 102–113.

Стяжкин 1967 – *Стяжкин Н.И.* Формирование математической логики. – М.: Наука, 1967.

Рубинштейн 1958 – *Рубинштейн С.Л.* О мышлении и путях его исследования. – М.: Изд-во АН СССР, 1958.

Рубинштейн 1989 – Рубинштейн С.Л. Основы общей психологии: в 2 т. Т. 1. – М.: Педагогика, 1989.

McCarthy 2007 – McCarthy J. What is Artificial Intelligence? // John McCarthy's Home Page. 2007, November 12. – URL: <https://www-formal.stanford.edu/jmc/whatisai/whatisai.html>

REFERENCES

Averkin A.N., Gaaze-Rapoport M.G., & Pospelov D.A. (1992) *Explanatory Dictionary of Artificial Intelligence*. Moscow: Radio i svyaz (in Russian).

Bakhtin A.V. & Remizova I.V. (2015) *Elements of Artificial Intelligence in Control Systems*. Saint Petersburg: SPbGTURP (in Russian).

Descartes R. (1950) *Selected Works*. Moscow: Politizdat (Russian translation).

Efimov A.R. (2021) *Philosophical and Methodological Foundations of Post-Turing Intellectual Robotics* (dissertation). Moscow: RAS Institute of Philosophy (in Russian).

Ilyasov F.N. (1986) Artificial and Natural Mind. *Bulletin of the Academy of Sciences of the Turkmen SSR. Series of Social Sciences*. No. 6, pp. 46–54 (in Russian).

Karpenko A.S. (2000) Logic at the Turn of the Millennium. *Logicheskie issledovaniya = Logical Investigations*. Vol. 7, pp. 7–60 (in Russian).

McCarthy J. (2007, November 12) What is Artificial Intelligence? In: *John McCarthy's Home Page*. Retrieved from <https://www-formal.stanford.edu/jmc/whatisai/whatisai.html>

Rozin V.M. (2015) *Thinking: Essence and Development. Concepts of Thinking. Role of Thinking Personality. Cycles of Thinking Development*. Moscow: LENAND (in Russian).

Rozin V.M. (2018) *Designing and Programming: Methodological Research. Conception. Development. Implementation. Historical and Social Context*. Moscow: LENAND. (in Russian).

Rozin V.M. (2022) *From the Analysis of Artistic Works to the Essence of Art*. Moscow: Golos (in Russian).

Rozin V.M. (2023) Two Concepts of Artificial Intelligence: Realistic and Utopian. *Filosofskaya mysl'*. No. 2, pp. 102–113 (in Russian).

Styazhkin N.I. (1967) *Formation of Mathematical Logic*. Moscow: Nauka (in Russian).

Rubinstein S.L. (1958) *On Thinking and the Ways of Its Research*. Moscow: USSR Academy of Sciences Publishing (in Russian).

Rubinstein S.L. (1989) *Fundamentals of General Psychology* (Vol. 1). Moscow: Pedagogika (in Russian).