

Философия и культура

Правильная ссылка на статью:

Грибков А.А., Зеленский А.А. — Общая теория систем и креативный искусственный интеллект // Философия и культура. – 2023. – № 11. DOI: 10.7256/2454-0757.2023.11.68986 EDN: EQVTJY URL: https://nbpublish.com/library_read_article.php?id=68986

Общая теория систем и креативный искусственный интеллект

Грибков Андрей Армович

ORCID: 0000-0002-9734-105X

доктор технических наук

ведущий научный сотрудник, НПК "Технологический центр"

124498, Россия, г. Москва, Зеленоград, площадь Шокина, 1, строение 7

✉ andarmo@yandex.ru



Зеленский Александр Александрович

ORCID: 0000-0002-3464-538X

кандидат технических наук

ведущий научный сотрудник, НПК "Технологический центр"

124498, Россия, г. Москва, Зеленоград, пл. Шокина, 1, строение 7

✉ zelenskyaa@gmail.com



[Статья из рубрики "Философия техники"](#)

DOI:

10.7256/2454-0757.2023.11.68986

EDN:

EQVTJY

Дата направления статьи в редакцию:

09-11-2023

Дата публикации:

16-11-2023

Аннотация: В статье анализируются возможности и ограничения искусственного интеллекта. Рассматривается субъектность искусственного интеллекта, представляющая

собой способность быть субъектом активности, в том числе в отношении объектов познания, определяется необходимостью субъектности для решения интеллектуальных задач в зависимости от возможности представления реального мира как детерминированной системы. Констатируются методологические ограничения искусственного интеллекта, в основу которого положено использование технологий больших данных, общая идея которых сводится к формированию информационной модели объекта познания в виде большого массива данных, упорядоченных на основе выявленных эмпирических связей, корреляции различных фрагментов, недетерминированных (по причине и механизму) трендов. Методологические ограничения искусственного интеллекта обуславливают невозможность формирования целостного представления об объектах познания и мира в целом. Рассматривается различие естественного (человеческого) и искусственного интеллекта, в том числе реализация мультисистемой интеграции интеллекта в физические, биологические, социальные и духовные системы. В качестве инструмента детерминированного описания мироздания предлагается использовать эмпирико-метафизическую общую теорию систем, являющуюся расширением существующих общих теорий систем за счет онтологического обоснования явления изоморфизма и определения ограниченной совокупности законов, правил, паттернов и примитивов форм и отношений объектов в мироздании. Необходимым условием практической реализации креативного искусственного интеллекта, сопоставимого или даже превосходящего человека в решении интеллектуальных задач, является создание адекватной общей теории систем, возможным вариантом которой является разрабатываемая эмпирико-метафизическая общая теория систем. В этом случае возможными станут дефрагментация системы знаний и детерминированное решение интеллектуальных задач как человеческим, так и искусственным интеллектом. В контексте проблемы обеспечения развития интеллекта формулируется философски обоснованный подход к обеспечению эволюционных свойств искусственного интеллекта, основанный на включении в него механизмов неравновесия. Механизмы неравновесия, через которые реализуется устойчивость, должны быть заложены в систему искусственного интеллекта при ее формировании. В этом случае она будет эволюционировать, но останется неживой – без потребностей, без эмоций и никогда не будет представлять угрозы для человечества.

Ключевые слова:

интеллект, субъектность, интеллектуальные задачи, творчество, целостность, общая теория систем, мультисистемность, неравновесная устойчивость, эволюция, детерминированное описание

Введение

Арсенал средств познания мира, формирования новых знаний и объектов в последние годы пополнился новым инструментом, возможность существования которого еще недавно было трудно представить. Этим новым инструментом являются системы искусственного интеллекта.

Для дальнейшего объективного анализа возможностей и ограничений искусственного интеллекта необходимо ответить на несколько вопросов: Что такое искусственный интеллект? Необходима ли субъектность для выполнения им своей функции решения интеллектуальных задач? Как связан искусственный интеллект с большими данными? В чем ограниченность современных технологий больших данных и как это влияет на

возможности использующего их искусственного интеллекта? В чем различие между естественным и искусственным интеллектом?

Ответы на эти вопросы должны стать основой для расширения возможностей искусственного интеллекта, а также позволят обозначить подходы к сепарации искусственного и естественного интеллекта.

Субъектность искусственного интеллекта

В зависимости от уровня функциональных возможностей различают три вида искусственного интеллекта: слабый или узкий (Artificial Narrow Intelligence, ANI), сильный или общий (Artificial General Intelligence, AGI) и суперинтеллект (Artificial Super Intelligence, ASI). Узкий искусственный интеллект предназначен для узкого круга интеллектуальных задач, которые им решаются не хуже или даже лучше человека. Все существующие в настоящее время системы искусственного интеллекта относятся к узкому искусственному интеллекту. Общий искусственный интеллект будет обладать возможностями в решении интеллектуальных задач, приближенными к человеческим по разнообразию, а также сопоставимыми или превосходящими их по эффективности (скорости, достоверности получаемых результатов и т.д.). Суперинтеллект будет превосходить человека при решении любой интеллектуальной задачи.

В качестве параметра, определяющего уровень искусственного интеллекта, иногда также рассматриваются наличие или отсутствие у искусственного интеллекта самостоятельности, самосознания и эмоций. Узкий искусственный интеллект может обладать самостоятельностью (автономностью) в решении интеллектуальных задач, однако эта самостоятельность ограничена предварительно определенной (заданной разработчиком) структурой в виде методов, алгоритмов и т.д. Интеллектуальная задача не может ставиться самим узким искусственным интеллектом или корректироваться (пересматриваться) в процессе решения. Узкий искусственный интеллект не может обладать самосознанием или эмоциями (хотя может их имитировать если его запрограммировать на такую имитацию). Общий искусственный интеллект и суперинтеллект для обеспечения необходимой функциональности в решении интеллектуальных задач, сопоставимой или превосходящей человека, обязательно должны обладать самостоятельностью. Вопрос о том, необходимы ли общему искусственному интеллекту и суперинтеллекту самосознание и эмоции, требует дополнительного рассмотрения.

Наличие или отсутствие у искусственного интеллекта самосознания (наряду с эмоциями и др.) является одним из основных факторов, определяющих его субъектность [\[1\]](#). Субъектность в философии трактуют как системное качество субъекта, наличие которого делает субъект тем, чем он является [\[2\]](#). Данное определение, в рамках которого субъектность эквивалентна самости, имеет онтологический характер. Самость в этом случае понимается как способность субъекта выйти из-под власти внешних причин и создать воспроизводящийся порядок жизни, детермируемый изнутри [\[3\]](#). Более адекватным в контексте нашего исследования является гносеологическое определение субъектности как способности быть субъектом активности, в том числе в отношении объектов познания.

Ключевым показателем, определяющим функциональность искусственного интеллекта, является его способность решать интеллектуальные задачи. В этой связи возникает вопрос том, какого рода задачи следует относить к интеллектуальным. Обычно под интеллектуальной понимают слабо формализованную задачу, т.е. задачу, «для которой

не существует общепризнанного (классического) решения, поэтому чтобы решить такую задачу, необходимо придумать способ решения» [4]. Более того, в процессе решения такой задачи она может пересматриваться (корректироваться).

Приведенное описание интеллектуальной задачи соответствует ее квалификации как творческой задачи. Творческая задача (в научно-техническом творчестве) – это такая, «для решения которой в науке до сих пор нет общепринятых правил и положений, определяющих точную программу ее решения» [5, с. 39-40].

Необходима ли субъектность для решения интеллектуальных (творческих) задач? Ответ на этот вопрос зависит от того, как мы представляем себе реальный мир.

Если мы представляем себе реальный мир как недетерминированную систему, в которой вероятностные параметры являются онтологическими и в котором творчество – иррациональный процесс, то достоверного ответа на вопрос о необходимости субъектности для решения интеллектуальных задач мы получить не можем. Мы лишь знаем, что человек может решать интеллектуальные задачи и обладает субъектностью. Характер связи этих свойств нам неизвестен.

Однако у нас имеются достаточные основания полагать, что в действительности мир детерминирован [6] и в его основе лежат механизмы, имеющие рациональное объяснение. Рациональное объяснение может быть дано и для интуиции, в основе которой лежит неосознанное использование аналогий и существующих в мироздании паттернов форм и отношений. Существование аналогий и паттернов является следствием изоморфизма, через который проявляется единство бытия [7].

Если мир детерминирован, все в нем имеет причины и (теоретически) может быть объяснено, то между вычислительной или логической задачей (с которой прекрасно справляется компьютер, не обладающий искусственным интеллектом) и интеллектуальной задачей нет качественного различия. При этом, конечно, методы и алгоритмы, используемые при решении интеллектуальной задачи, будут несравнимо более сложными и разнообразными, чем при решении вычислительных или логических задач. Однако все алгоритмы и последовательность действий искусственного интеллекта могут быть определены и объяснены. Такой искусственный интеллект называют объяснимым [8]. В этом случае для выполнения искусственным интеллектом своей функции решения интеллектуальных задач субъектность не требуется.

Методологические ограничения технологий больших данных

Для понимания характера методов обработки информации, положенных в основу искусственного интеллекта, рассмотрим системы глубинного машинного обучения (Deep learning), которые являются одной из начальных версий узкого искусственного интеллекта.

В основе глубинного машинного обучения, как это понятно из его названия, лежит обучение: с учителем, с частичным привлечением учителя, без учителя, с подкреплением и т.д. Объектом обучения являются искусственные нейронные сети, реализуемые в настоящее время на базе различных систем высокопроизводительных вычислений (графических процессоров, программируемых пользователем вентильных матриц, интегральных схем специального назначения и др.), а также на базе нейронных процессоров. В дальнейшем основой для нейронных сетей также могут стать квантовые процессоры.

Глубинная нейронная сеть (Deep neural network) представляет собой искусственную нейронную сеть с большим числом слоев, т.е. математических преобразований от входа до выхода. Функция глубинной нейронной сети подобрать (в процессе обучения) такую последовательность математических преобразований (от слоя к слою), которая при различных значениях на входе обеспечивает на выходе правильный результат. Верификация «правильности» результата может осуществляться как внешним пользователем («учителем»), так и посредством дополнительных алгоритмов глубинной нейронной сети.

Системы искусственного интеллекта неразрывно связаны с технологиями больших данных (Big Data) [\[9\]](#). Искусственный интеллект на практике реализует не только методы глубинного машинного обучения, но и большую часть методов, используемых в технологиях больших данных, в том числе методов Data Mining, служащих для обнаружения определенных закономерностей в сырых данных с помощью интеллектуального анализа.

Общая идея технологий больших данных и задействованных в ней методов сводится к формированию информационной модели объекта познания в виде большого массива данных, упорядоченных на основе выявленных эмпирических связей, корреляции различных фрагментов, недетерминированных (по причине и механизму) трендов. Эти массивы данных используются вместо детерминированных представлений об объекте познания: описания его элементов, структуры, внутренних и внешних связей, реализуемых в объекте законов природы и т.д. Методологические подходы, являющиеся в настоящее время основой технологий больших данных (и транслируемые на создаваемые на их основе системы искусственного интеллекта), являются результатом распространения в научном сообществе идей позитивизма, несовместимых с формированием целостного представления мироздания.

Неудачи в формировании целостной картины физического мира и природы в целом породили в философии тенденцию к эмпиризму, возникли сомнения в возможности построения целостной картины мира, а вслед за этим – вопрос о том, необходимо ли такое построение. Выражением формирующихся тенденций отказа от построения целостной картины мира, ограничения познания результатами опыта, его систематизации и обобщения стала концепция позитивизма [\[10\]](#). Концепция позитивизма, таким образом, приобрела нынешнее доминирующее положение не в силу объективных философских достоинств, а как результат капитуляции философов перед сложностью и противоречивостью явлений и объектов в природе.

Поскольку в основе творчества (в том числе, решения интеллектуальных задач) лежит осознанное или неосознанное использование изоморфизма мира, являющегося выражением его целостности, существующие подходы к созданию искусственного интеллекта (на основе технологий больших данных и всех используемых в них методов) являются бесперспективными. Они могут позволить решать лишь отдельные примитивные интеллектуальные задачи, для которых достаточно выявления эмпирических связей и трендов. Узкий искусственный интеллект на основе такого подхода (с существенными ограничениями) реализуем, искусственный интеллект более высокого уровня, очевидно, нет.

Технологии больших данных уверенно расширяют свое присутствие в нашей жизни. Они стали основой машинного обучения, на их базе строятся современные системы искусственного интеллекта. Преобладающим становится мнение, что именно таким образом должны формироваться когнитивные системы, именно таким образом

искусственный интеллект сможет приблизиться к естественному (человеческому). Как мы выяснили, это мнение является глубоко ошибочным.

Для построения общего искусственного интеллекта или суперинтеллекта потребуется выйти за методологические ограничения, сформировавшиеся в рамках технологий больших данных. При этом мы имеем единственный ориентир – человеческое мышление, которое, вероятно, не имеет таких ограничений.

Опыт познания показывает, что здоровый человеческий интеллект (и это продемонстрировано всей историей развития системы знаний человечества) не ищет тренды, корреляцию данных в окружающем мире. Более того, фиксация мышления на поиске связей (часто несуществующих) часто является свидетельством психического расстройства (депрессия, шизофрения, бред и т.д.). Здоровый человеческий интеллект ищет не тренды, а смыслы. Основой творческого мышления является осмысленное или бессознательное понимание целостности мира, существования в нем изоморфизмов, т.е. подобия форм и законов в различных предметных областях, на разных уровнях мироздания. Это делает возможным использование аналогий и в целом определяет продуктивность традиционного мышления.

Естественный (человеческий) интеллект не использует методологию технологий больших данных. Парадоксально, но человеческий разум, реализующийся в реальном мире, полотно которого в системе знаний соткано из эмпирического опыта, при этом является дедуктивным, тяготеет к метафизическому представлению мира. И это неизбежно, поскольку мир един (обладает целостностью) и человек, как часть мира, должен это принимать (пусть даже неосознанно). Ведь успешность человека, и это общеизвестно, в огромной мере определяется его интегрированностью в мир, т.е. встраиванием как элемента в целостную систему мира.

Возможности общей теории систем

Характерной чертой современных научных и философских представлений является практическая невозможность их объединения в целостную систему знаний, в рамках которой возникающие в процессе развития противоречия и несоответствия могут находить свое решение согласно логике целостности мироздания.

Опасность фрагментации знаний заключается в неопределенности перспектив научного, технологического и социального развития. Эта неопределенность порождается ограниченностью представлений рамками отдельных малых предметных областей, вследствие чего широкое понимание контекста происходящих событий, долгосрочные тенденции как развития, так и деградации оказываются за пределами поля зрения ученых, философов и общественных деятелей.

Потенциальной способностью взять на себя функцию продвижения концепции целостности мира обладает общая теория систем. В своем настоящем виде она сильно подвержена эмпиризму, однако этот изъян может быть устранен при должном развитии теории, в том числе за счет выстраивания ее онтологического фундамента, позволяющего определить генезис явления изоморфизма, через который эмпирически проявляется целостность мироздания. В настоящее время ведется работа над созданием общей теории систем (эмпирико-метафизической общей теорией систем), свободной от выявленных недостатков и дополненной необходимой онтологической частью [\[11\]](#).

Эмпирико-метафизическая общая теория систем использует два взаимодополняющих подхода к познанию, первый из которых можно условно назвать «онтологическим», а

второй – «эпистемологическим». Согласно «онтологическому» подходу, из первичных свойств бытия (которые можно квалифицировать как априорные знания) логически могут быть выведены формы и законы мироздания. Эпистемологический подход основан на использовании для описания объектов наборов паттернов и примитивов, а также выявляемых в практике познания законов и свойств (вторичных законов и свойств) без определения их генезиса и внутренних механизмов. Понятия паттернов и примитивов могут быть определены следующим образом: паттерны – репрезентативный, но ограниченный по количеству набор шаблонов форм и отношений элементов внутри системы, широко распространенных в различных предметных областях; примитивы – множество элементов, лежащих в основе формирования всех систем.

Эмпирико-метафизическая общая теория систем включает в себя внутренние механизмы развития, детализации и повышения достоверности результатов познания. Сформулированные первичные и вторичные законы бытия, выявленные группы паттернов и примитивов не являются однозначно и окончательно зафиксированными. Они могут уточняться, детализироваться и расширять свою допустимую интерпретацию. При этом логика выстраивания эмпирико-метафизической общей теории систем должна сохраниться.

Методология общей теории систем недостаточна для окончательного решения задач. Общая теория систем, представляющая собой часть теории познания, служит лишь для выбора направления и области поиска решений, определения возможных вариантов решения, допустимости тех или иных решений и т.д., т.е. является лишь одним из необходимых инструментов в арсенале средств решения интеллектуальных задач. Однако при решении интеллектуальных задач (особенно сложных, соответствующих, согласно современной классификации, общему искусственному интеллекту или суперинтеллекту) этот инструмент незаменим, поскольку формируемые им знания не могут быть получены из собранных эмпирических данных об объекте познания, а следуют из представлений о целостности мира, формализованных в виде первичных свойств, базовых и первичных законов бытия, известных паттернов и примитивов, а также вторичных законов и свойств бытия. До настоящего времени в качестве инструмента творческого осмысления на основе целостности мира выступал человеческий интеллект. Развитие и формализация методологии общей теории систем позволит вначале частично, а в дальнейшем полностью ликвидировать зависимость от человека в решении интеллектуальных задач. И в этом нет угрозы для интересов человечества также как и не было угрозы в появлении вычислительных машин для решения расчетных и логических задач.

В связи с оценкой возможностей общей теории систем имеет смысл рассмотреть вопрос о суперинтеллекте. Мы уверенно говорим об этом третьем виде искусственного интеллекта, как о чем-то, что обязательно рано или поздно будет создано. На самом деле, появление суперинтеллекта необязательно. И причина не в технологических сложностях, а в ограниченности качественной сложности мира.

Современная философия и наука исходят из представлений о неограниченной (в контексте позитивизма – неопределяемой) сложности мироздания. Если это было бы так, то, действительно, для познания мира требовался бы все более обширный искусственный интеллект, но при этом его возможностей никогда бы не хватило для познания мира.

По нашему мнению, задача познания мира является одновременно и более сложной, и более простой. Сложность заключается в том, что достоверное представление реального

мира в системе знаний невозможно в силу непереводаемости с языка бытия на язык теории познания и науки [\[11\]](#). Теория познания и наука строятся на обобщенных понятиях и вероятности событий, а в реальности их не существует – все конкретно: существует в единственном экземпляре в данный момент времени. Система знаний о мире является моделью реальности, соответствующей ей лишь по ограниченному набору признаков, принимаемых (обоснованно или необоснованно – не всегда понятно) существенными. С другой стороны, хотя мы не можем представить мир в системе знаний точно, однако его качественное описание возможно через имеющийся в общей теории систем ограниченный набор методов и средств. Даже при текущем уровне ее развития мы способны адекватно описывать мир, в дальнейшем детализация описания будет повышаться. Парадокс заключается в том, что качественно мир не очень сложен и для его качественного описания суперинтеллект не требуется. При этом точное количественное описание практически невозможно, а там, где мы к нему стремимся (с большим или меньшим успехом), сложности связаны с большим объемом вычислений (по строго определенным алгоритмам), а не с их интеллектуализацией.

Итак, представляется логичным исключить суперинтеллект из вариантов искусственного интеллекта. Как мы выяснили, вычислительная мощность искусственного интеллекта не является основным определяющим его критерием. Возможно ли, что разделение искусственного интеллекта на узкий и общий также не является оптимальным?

Принципиальное отличие между узким и общим искусственным интеллектом заключается в «ширине» решаемых интеллектуальных задач. При этом, как показали наши размышления, общий искусственный интеллект должен опираться (наряду с существующими инструментами анализа данных) также на общую теорию систем. Это придает общему искусственному интеллекту способность верифицировать поставленную перед ним интеллектуальную задачу и получаемые варианты решения в рамках логики целостности мира (корректная или некорректная постановка вопроса, имеются аналоги или нет, согласуется ли решение с существующими паттернами, примитивами и законами или нет и т.д.). Это дает общему искусственному интеллекту способность корректировать постановку интеллектуальной задачи и осуществлять выбор допустимого решения из нескольких имеющихся. При этом его определяющей характеристикой становится не способность решать широкий спектр интеллектуальных задач, а креативность, заключающаяся в выходе не только за пределы заданных алгоритмов в решении самой задачи, но и за пределы самой задачи.

По нашему мнению, обоснованным с точки зрения определения специфики искусственного интеллекта является его деление на два вида: системы глубинного машинного обучения и креативный искусственный интеллект.

Мультисистемная интеграция

Для определения систем искусственного интеллекта, как это видно из их названия, используется системный подход. Экспликация интеллекта в общем случае (т.е. как искусственного, так и человеческого) также, по мнению авторов, может быть построена на основе системного подхода. Поскольку, вероятно, человеческий интеллект обладает существенно более сложной организацией, чем искусственный, при определении структуры и составляющих интеллекта будет ориентироваться на человеческий интеллект.

На первый взгляд обоснованным выглядит отождествление интеллекта с когнитивной системой. Когнитивная система (от лат. *cognito* — познание, узнавание, ознакомление),

с точки зрения философии, – это многоуровневая система, осуществляющая функции распознавания и запоминания информации, принятия решений, хранения, объяснения, понимания и производства новых знаний [\[12\]](#). Когнитивная система – это информационная система (реализующая обработку информации), дополняемая некоторыми расширениями, т.е. элементами, обеспечивающими поступление входных данных. Примером таких дополнительных элементов у человека являются органы чувств, у искусственного интеллекта – сенсоры и другие интегрированные с когнитивной системой измерительные устройства.

Когнитивная система является важнейшей составляющей интеллекта, однако, по крайней мере в случае человеческого интеллекта, не тождественна ему. Человеческий интеллект не может существовать как исключительно информационная система. Он интегрирован в большое число систем: в физический мир, в биологическую систему мыслящего индивида и систему биологического вида, к которому индивид принадлежит, в личность (социального индивида) и социальную систему. Кроме того, интеллект как информационная система интегрирован в систему знаний человечества, включая ее парадигмальные (этические, религиозные, различные научные и т.д.) и локальные (страновые, национальные, классовые и т.д.) подсистемы.

Мультисистемная интеграция человеческого интеллекта в множество различных систем обуславливает формирование потребностей: биологических, духовных (интеллектуальных) и социальных. Потребности, в свою очередь, являются необходимым условием формирования эмоций. Эмоции – это психофизическая реакция субъекта (человека) на предметы чувственного восприятия в контексте удовлетворения или неудовлетворения его потребностей [\[13, том 4, с. 436-437\]](#).

Интеграция человеческого интеллекта во множество взаимосвязанных систем не является пассивной. Она предполагает реализацию механизмов обратной связи (как отрицательной, так и положительной), взаимное влияние между интеллектами и другими акторами взаимосвязанных систем и т.д. В результате человеческий интеллект в процессе своего существования воспроизводит целостность мира. Поэтому человеку в существенной степени доступно осознание целостности мира и без ее формализации в виде общей теории систем. И поэтому человек обладает интуицией и способен на творчество.

Мультисистемная интеграция искусственного интеллекта, очевидно, никогда не будет такой, как у естественного интеллекта. Она не будет спонтанной, она не предполагает механизмов, обуславливающих формирование потребностей и эмоций. И для реализации своей функции решения интеллектуальных задач искусственному интеллекту это не требуется. Вместо эмпирической неполной модели целостности мира в виде мультисистемной интеграции, присущей человеческому интеллекту, искусственный интеллект должен быть наделен непосредственным систематизированным знанием о целостности мира в виде общей теории систем (в своем завершенном окончательном виде).

Сможет ли искусственный интеллект, ограниченный в интеграции в некоторые существующие системы, эволюционировать? Человеческий интеллект, как живая система, обеспечивает свои устойчивость и развитие посредством механизма устойчивого неравновесия, неизбежным следствием которого является эволюция [\[14, с. 321\]](#). Искусственный интеллект – не живой, однако эволюция возможна не только для «живых» (биологических, социальных, экономических) систем, но и многих динамических

неживых систем (например, химических). Механизмы неравновесия, через которые реализуется устойчивость, должны быть заложены в систему искусственного интеллекта при ее формировании. В этом случае она будет эволюционировать, но останется неживой – без потребностей, без эмоций и никогда не будет представлять угрозы для человечества.

Необходимым условием практической реализации креативного искусственного интеллекта, сопоставимого или даже превосходящего человека в решении интеллектуальных задач, является наличие адекватной общей теории систем, возможным вариантом которой является указанная выше Эмпирико-метафизическая общая теория систем [\[11\]](#). Только в этом случае возможными станут дефрагментация системы знаний и детерминированное решение интеллектуальных задач как человеческим, так и искусственным интеллектом.

Сегодня лишь человеческий разум в силу биологической, социальной и интеллектуальной интеграции человека в мир обладает способностью его целостного восприятия, интуитивного поиска и нахождения в нем логически необнаружимых закономерностей и связей. Если формы и законы, определяющие целостность мира, будут формализованы, человек утратит свое преимущество перед искусственным интеллектом. Последний конечно не станет при этом тождественным человеческому, который неразрывно связан с биологическими, социальными и духовными потребностями, которых у машины никогда не будет, однако интеллектуальная монополия человека будет разрушена.

Будущее человечества трудно прогнозировать, однако, вероятно, развитие знаний не имеет альтернативы. Искусственный интеллект, если в основу его совершенствования будет положена дефрагментация знаний посредством общей теории систем, не будет угрозой для человечества. Напротив, такой креативный искусственный интеллект станет перспективным инструментом расширения системы знаний и значимым фактором развития человеческой цивилизации.

Человеческий интеллект – носитель субъектности и самости человека. Цель существования любого живого существа заключается в том, чтобы жить. Применительно к человеческому интеллекту эта цель – осознавать все многообразие своих потребностей и решать задачу по их удовлетворению. Тот факт, что какие-то интеллектуальные задачи искусственный интеллект может решать лучше человека, не имеет для отдельного человека принципиального значения. Это важно для науки, технологий, экономики и через них, конечно, оказывает влияние на человека. Это расширяет возможности удовлетворения человеком своих биологических, социальных и интеллектуальных потребностей, но не формирует их.

Вопрос о замене человека искусственным интеллектом лишен смысла. Допустим, достижения науки позволили клонировать человека и записать клону память оригинала. Для других людей клон может быть заменой оригинала, но не для того, кого клонировали. Ему безразлично, что есть другой такой же, и он не согласится, чтобы его заменили клоном. Человек и человеческий интеллект ценны сами для себя и неважно, можно их заменить на лучшую или просто альтернативную версию или нет.

Развитие систем искусственного интеллекта, пусть даже лишенных субъектности, не мешает дать им возможность имитации человеческого поведения. И, очевидно, спрос на такой искусственный интеллект будет увеличиваться. Как и любой самообман, это вредное явление, однако оно не может представлять значимой угрозы для людей.

Создание искусственного интеллекта, наделенного субъектностью, – цель труднодостижимая и при этом бесполезная с точки зрения решения интеллектуальных задач. Тем не менее, даже если работы в данной области будут запрещены законом, такие системы, вероятно, будут созданы. Остается надеяться, что сравнительно малая распространенность таких систем (ограниченная масштабами лабораторий) позволит избежать тех потенциально возможных опасных последствий, которые возникают при господстве интеллектуально более слабых субъектов (людей) над предположительно более сильными субъектами (искусственным интеллектом, наделенным субъектностью).

Выводы

Приведенные в статье исследования позволяют сделать следующие основные выводы:

1. Одним из ключевых вопросов, возникающих при определении возможностей и опасностей, связанных с развитием искусственного интеллекта, является вопрос его субъектности. Развитие методологии познания за счет формирования адекватной общей теории систем, предполагающее ограниченность качественной сложности мироздания, позволит сделать субъектность систем искусственного интеллекта необязательной для выполнения ими своей функции решения интеллектуальных задач.
2. В зависимости от уровня функциональных возможностей в настоящее время различают три вида искусственного интеллекта: слабый или узкий, сильный или общий и суперинтеллект. Ввиду ограниченности качественной сложности мироздания представляется логичным исключить суперинтеллект из вариантов искусственного интеллекта. Обоснованным с точки зрения определения специфики искусственного интеллекта является его деление на два вида: системы глубинного машинного обучения и креативный искусственный интеллект.
3. Существование естественного интеллекта предполагает его мультисистемную интеграцию в физические, биологические, социальные, духовные системы. Через эту интеграцию реализуется локализованное в естественном интеллекте моделирование целостности мира, осознанное или неосознанное выстраивание на основе этой модели системы аналогий и паттернов. Следствием этого становится обретение человеческим интеллектом способности к творчеству, в том числе к решению интеллектуальных задач.
4. Мультисистемная интеграция искусственного интеллекта будет существенно ограниченной. В частности, она не должна порождать механизмов, обуславливающих формирование потребностей и эмоций. При этом искусственный интеллект должен быть наделен непосредственным систематизированным знанием о целостности мира в виде общей теории систем (в своем завершенном окончательном виде).
5. Придание искусственному интеллекту, ограниченному в интеграции в некоторые существующие системы, эволюционных свойств возможно, если в него при формировании заложить механизмы неравновесия, через которые реализуется устойчивость. В этом случае искусственный интеллект будет эволюционировать, но останется неживым – без потребностей, без эмоций и никогда не будет представлять угрозы для человечества.

Библиография

1. Алексеева И.Ю. Субъектность искусственного интеллекта: старые вопросы в новых контекстах // Информационное общество, 2020, №6. С. 2-6
2. Радченко Е.В., Ранг К.А. Понимание субъектности в философии и языкознании //

- Вестник ЮУрГУ, № 2, 2012. С. 74-78
3. Гиренок Ф.И. Самость. Энциклопедия фонда знаний "Ломоносов". URL: <http://www.lomonosov-fund.ru/enc/ru/encyclopedia:0127733>
 4. Пруцков А.В. Современные проблемы систем искусственного интеллекта // Cloud of Science. 2020, Т. 7, №4. С. 897-904
 5. Рапацевич Е.С. Словарь-справочник по научно-техническому творчеству. Мн.: ООО «Энтоним», 1995. 384 с.
 6. Грибков А.А. Простейшие материальные структуры – следствия первичных свойств бытия // Общество: философия, история, культура, 2023, №6, с. 23-29
 7. Грибков А.А. Паттерны и примитивы эмпирико-метафизической общей теории систем // Общество: философия, история, культура, 2023, №5, с. 15-22
 8. Райков А.Н. Субъектность объяснимого искусственного интеллекта // Философские науки. 2022, Т. 65, № 1. С. 72–90
 9. Асеева И.А. Искусственный интеллект и большие данные: этические проблемы практического использования. (Аналитический обзор) // Социальные и гуманитарные науки. Отечественная и зарубежная литература. Сер. 8: Науковедение. 2022, №2. С. 89-98
 10. Султанов К.В. и др. Философия науки позитивизма: эволюция предмета и понятий / Султанов К.В., Корольков А.А., Пую Ю.В., Рабош В.А., Стрельченко В.И. // Общество. Среда. Развитие (Terra Humana), №1 (42), 2017. С. 25-32
 11. Грибков А.А. Эмпирико-метафизический подход к построению общей теории систем // Общество: философия, история, культура, 2023, №4. С. 14–21
 12. Философия: Энциклопедический словарь / Под ред. А.А. Ивина. М.: Гардарики, 2004. 1072 с.
 13. Степин В.С., Гусейнов А.А., Семигин Г.Ю., Огурцов А.П. и др. Новая философская энциклопедия: В 4 томах. М.: Мысль, 2010
 14. Бауэр Э.С. Теоретическая биология. М.-Л.: Изд. ВИЭМ, 1935. 151 с

Результаты процедуры рецензирования статьи

В связи с политикой двойного слепого рецензирования личность рецензента не раскрывается.

Со списком рецензентов издательства можно ознакомиться [здесь](#).

Рецензируемая статья представляет собой отражение глубоких размышлений автора о возможностях конструирования различных систем искусственного интеллекта (автор придерживается различия «узкого интеллекта», «общего интеллекта» и «суперинтеллекта»), взаимосвязи проблемы достижения искусственным интеллектом статуса субъекта и решения им творческих задач, о значении технологий «больших данных» для понимания природы и ограниченности возможностей существующих систем искусственного интеллекта, а также о взаимосвязи проблемы решения искусственным интеллектом творческих задач и структуры реальности (исчерпывается ли она рациональными связями, и «раскладывается» ли в последовательности причинно-следственных связей). Философские составляющие «проблемы искусственного интеллекта» связаны, главным образом, с пониманием возможностей его эволюции в направлении «естественного интеллекта». Сегодняшние образы искусственного интеллекта относятся к классу разработок «узкого интеллекта», в их рамках задача, а также методология её решения, принимаются искусственным интеллектом извне, от человека. Самый простой способ развития технологий искусственного интеллекта, как

представляется, состоит в том, чтобы попытаться сконструировать «искусственную субъектность», наделять искусственный интеллект эмоциями и волей и, в конце концов, самосознанием. Автор полагает, однако, что подобная постановка вопроса бессмысленна в теоретическом плане и бесполезна в практическом отношении. С нашей точки зрения, возможности конструктивного обсуждения этой проблемы остаются пока крайне ограниченными, и мы не можем достичь должной ясности в понимании самого концептуального наполнения выражений «эмоции», «воля» или «самосознание» в отношении искусственного интеллекта. Далее автор сосредотачивает внимание на более частном вопросе, который, с его точки зрения, уже может получить достаточно определённое решение в современных условиях, а именно, на вопросе о том, необходимо ли предполагать «субъектность» искусственного интеллекта для решения им творческих задач. Ответ на этот вопрос, полагает автор, зависит от того, как мы представляем себе реальный мир. Если предположить, что мир в целом является, как выражается автор, «детерминированным», или, точнее, если в его основе лежат исключительно рациональные отношения, то решение творческих задач, полагает автор, оказывается возможным для искусственного интеллекта даже безотносительно к проблеме достижения цели конструирования его «субъектности». В этом случае «творческий потенциал» искусственного интеллекта совсем не обязательно должен быть бесконечным (как у самого человека), а значит, можно представить ситуацию, в которой «творческое начало» в определённых границах будет поддаваться программированию. Касается автор и интересного вопроса о значимости технологий «больших данных» для разработки искусственного интеллекта, и он (на наш взгляд, оправданно) оценивает роль этих технологий критически: «Поскольку в основе творчества ... лежит осознанное или неосознанное использование изоморфизма мира, являющегося выражением его целостности, существующие подходы к созданию искусственного интеллекта (на основе технологий больших данных и всех используемых в них методов) являются бесперспективными. ... Узкий искусственный интеллект на основе такого подхода (с существенными ограничениями) реализуем, искусственный интеллект более высокого уровня, очевидно, нет»; «Для построения общего искусственного интеллекта или суперинтеллекта потребуется выйти за методологические ограничения, сформировавшиеся в рамках технологий больших данных. При этом мы имеем единственный ориентир – человеческое мышление, которое, вероятно, не имеет таких ограничений». Рецензируемая статья представляет собой оригинальное исследование, она может быть интересна широкому кругу читателей. Отдельные погрешности стилистического характера «...связей (часто несуществующих) часто является...» и т.п.) могут быть исправлены в рабочем порядке. Рекомендую принять статью к публикации в научном журнале.