



УДК 168
DOI 10.52575/2712-746X-2022-47-4-670-681

Искусственный интеллект как философская проблема и искусственные интеллектуальные системы

Швырков А.И.

Брянский государственный технический университет,
Россия, 241035, Брянская область, г. Брянск, бульвар 50 лет Октября, 7
E-mail: aishvyrkov@rambler.ru

Аннотация. Прояснена терминология, связанная с проблемой искусственного интеллекта, различных способов ее трактовки, места и отношения друг к другу ключевых понятий соответствующей сферы. Констатируется, что фокус исследований за последние 30 лет существенно сместился: если раньше суть проблемы искусственного интеллекта большинством исследователей виделась в том, можно ли создать некий искусственный аналог человеческого интеллекта, то теперь она фактически превратилась в конгломерат самых разнообразных проблем и задач, связанных, например, с моделированием психики, мышления, распознаванием образов, роботикой, автоматическим доказательством теорем и т.д. Тем не менее, такая трансформация не означает, что идея искусственного интеллекта как целостного феномена не имеет реального содержания и что мы должны оставить попытки решить проблему возможности именно такого искусственного интеллекта. Однако для того чтобы наши попытки стали более плодотворными, нам следует коренным образом пересмотреть статус данной проблемы. До сих пор она классифицировалась как «комплексная», «междисциплинарная» и т.п. Однако, по нашему мнению, гораздо плодотворнее будет рассматривать ее как имеющую принципиально философский характер, то есть как чисто философскую проблему. Также сделана попытка обосновать идею о том, что искусственные интеллектуальные системы могут рассматриваться как феномен, достаточно независимый от идеи искусственного интеллекта. Соответственно, их анализ может осуществляться без постоянного соотношения с идеей искусственного интеллекта.

Ключевые слова: искусственный интеллект, проблема искусственного интеллекта, искусственные интеллектуальные системы

Для цитирования: Швырков И.А. 2021. Искусственный интеллект как философская проблема и искусственные интеллектуальные системы. НОМОТНЕТКА: Философия. Социология. Право, 47(4): 670–681. DOI: 10.52575/2712-746X-2022-47-4-670-681

Artificial Intelligence as a Philosophical Problem and Artificial Intellectual Systems

Alexander I. Shvyrkov

Bryansk State Technical University,
7 50 years of October boulevard, Bryansk, Bryansk region 241035, Russian Federation
E-mail: aishvyrkov@rambler.ru

Abstract. The paper analyzes terminology associated with the problem of artificial intelligence, various ways of its interpretation, place and relation to each other of the key concepts of the field. It is stated that the focus of research over the last 30 years has shifted significantly: if earlier most researchers saw the essence of the problem in whether an artificial analogue of the human intellect can be created then now it has actually become a conglomerate of a wide range of problems connected, for example, with psyche and thinking modeling, pattern recognition, robotics, automatic proof of theorems, etc. Nevertheless, such a transformation does not mean that the idea of artificial intelligence as a holistic phenomenon has no real

sense and that we should abandon attempts to solve the problem of the possibility of *such* an artificial intelligence. However, in order to make our efforts more fruitful, we should radically reconsider the status of the problem. Until now, it has been classified as “complex”, “interdisciplinary”, etc. However, according to the author of the paper, it will be much more fruitful to consider it as having a fundamentally philosophical character, that is, as a purely philosophical problem. Also an attempt to justify the idea that artificial intellectual systems can – and even should – be regarded as a phenomenon rather independent of the idea of artificial intelligence has been made. Accordingly, they can be analyzed without permanent referring to the idea of artificial intelligence.

Key words: artificial intelligence, artificial intelligence problem, artificial intelligence systems

For citation: Shvyrkov A.I. 2021. Artificial Intelligence as a Philosophical Problem and Artificial Intellectual Systems. NOMOTNETIKA: Philosophy. Sociology. Law, 47(4): 670–681 (in Russian). DOI: 10.52575/2712-746X-2022-47-4-670-681

Введение

Наверно, трудно найти сегодня проблему более неоднозначную, чем проблема искусственного интеллекта (ИИ). С момента появления в 1956 года термина «искусственный интеллект» представления об ИИ и том, можно ли его создать, постоянно менялись и продолжают меняться. Причем разные трактовки ИИ и проблемы ИИ – не редко слабо совместимые друг с другом – сосуществовали и сосуществуют в одном предметном поле, внося изрядную путаницу в теоретические построения. В следствие этого в одной из своих статей попытался несколько прояснить эти вопросы [Швырков, 2009].

Приведем некоторые основания для принятия следующих двух тезисов.

Первый состоит в том, что для успешного анализа проблемы ИИ эту проблему следует рассматривать как такую, которая имеет принципиально философский характер.

Второй гласит, что *искусственные интеллектуальные системы* необходимо рассматривать в качестве феномена, достаточно независимого от идеи искусственного интеллекта.

Однако сначала следует прояснить несколько терминологических моментов (кратко повторяя основные идеи упомянутой статьи). В различных работах по проблеме термин «искусственный интеллект» употребляется как минимум в *трёх* основных значениях.

Во-первых, как название для некоторого научного направления, объединяющего большое количество достаточно разнородных исследований [Байдык, 2001, Искусственный интеллект, 1990; Russell, Norvig, 2009].

Во-вторых, как некий целостный феномен, артефакт по своим возможностям так или иначе соотносящийся или даже превосходящий человеческий интеллект (но не обязательно аналогичный последнему) [Haugeland, 1985; Bostrom, 2014]. Под эту категорию, очевидно, подпадают и «сильный», и «слабый» ИИ [Searle, 1997]. В данной статье термин «искусственный интеллект» будет употребляться именно в этом третьем смысле.

В-третьих, как совокупность уже существующих или только разрабатываемых так называемых интеллектуальных систем, используемых для решения различных задач (системы распознавания, слежения, автоматического перевода, автоматизации различных производственных процессов и т.п. [Войскунский, 2001; Глазунов, 2002; Искусственный интеллект, 1991; Уинстон, 1980; Bruce, 2005]). Последние 10–15 лет можно наблюдать то, как это третье значение термина «ИИ» вытесняет все остальные. В большинстве случаев сегодня, когда говорят об ИИ (в том числе в СМИ), имеют в виду именно это значение.

Теперь о том, какой смысл может вкладываться в словосочетание «проблема ИИ». Здесь я вижу две основные возможности.



Первая: проблема ИИ – это название для некоторой совокупности проблем, феноменов, артефактов. Какого-то особого содержания с соответствующим словосочетанием в этом случае не связывается.

Вторая: проблему ИИ в наиболее общем виде можно сформулировать так: возможен ли ИИ и если возможен, то каким образом. В дальнейшем, говоря о проблеме ИИ, я буду иметь в виду именно эту её формулировку.

Все исследования и разработки, так или иначе связываемые с ИИ, можно условно разделить на следующие два класса.

1. Исследования, проводящиеся тогда, когда понятие ИИ так или иначе определено (во всяком случае, у исследователей есть некоторое представление об этом ИИ), и направленные на поиск путей создания такого ИИ. Этому классу, очевидно, соответствует третье значение термина ИИ.

2. Исследования и разработки, возникшие из неких практических потребностей и, как кажется, первоначально с ИИ особенно не связывавшиеся. Во всяком случае, создание ИИ (как некоего целостного феномена, артефакта) отнюдь не являлось целью этих исследований. По крайней мере, ближайшей практической целью.

Важно подчеркнуть, что большинство проводившихся в последние несколько десятков лет исследований и разработок относилось именно к этому, второму, классу. В первую очередь речь идет об исследованиях и разработках, связанных с созданием экспертных систем, роботикой, автоматическим доказательством теорем, распознаванием образов и т.д. [Миклошко, 1991; Russell, Norvig, 2009].

Объединение всех этих разнородных исследований под лозунгом создания ИИ достаточно условно и, вероятно, именно поэтому сам термин «искусственный интеллект» в связи с ними исследователями употреблялся сравнительно редко. Показательно, что в книге Б. Гейтса «Бизнес со скоростью мысли», посвященной использованию информационных технологий в различных областях экономики, образования, производства и т.д., термин ИИ встречается только *один* раз [Гейтс, 2002]. Однако, как уже было сказано, в последние годы данные разработки (и совокупность артефактов, созданных в ходе них) все чаще называют именно ИИ.

Искусственный интеллект как философская проблема

Как писали В.Т. Пушкин и А.Д. Урсул в 1989 г., вместо «единой теории ИИ существует ряд теоретических дисциплин, которые должны изучаться теми, кто выбрал ИИ своей специальностью» [Пушкин, 1989, с. 130]. Отмечу, что за прошедшие 30 лет ситуация мало изменилась.

Каковы же причины того, что такая теория до сих пор не создана (или, по крайней мере, не создаётся)?

Прежде чем говорить о том, способны ли мы в настоящее время создать общую теорию ИИ, необходимо определиться с тем, что следует понимать под такой теорией.

Все теории вообще можно разделить на два вида. К первому относятся такие, которые призваны объяснить некоторую совокупность фактов. Учёных и философов, создающих их, не особенно интересует их практическое приложение. По крайней мере, оно имеет второстепенное значение. Такие теории, как правило, имеют свой вполне определённый, реально существующий предмет, вполне доступный восприятию (возможно, с помощью приборов) или, хотя бы воображению. В качестве примера подобных теорий прежде всего нужно указать математические, физические, биологические, вообще естественнонаучные теории. Можно утверждать, что учёных и философов, создающих теории первого вида, интересует истина *сама по себе*.

Ко второму виду относятся теории, которые условно могут быть названы практически ориентированными или целеориентированными. Как следует из названия, они созда-

ются под конкретную цель, результат. Те или иные принципы, составляющие их, как правило, носят подчинённый, вспомогательный характер: если бы данную цель можно было достичь каким-либо иным путём, такие теории, скорее всего, просто не возникли бы. Часто подобные теории могут быть модификациями уже существующих теорий или объединением их.

Для целеориентированных теорий чрезвычайно важен *образ цели*, результата, поскольку именно этот образ определяет отбор компонентов будущей теории. Чем он яснее, детальнее, тем меньше в соответствующей теории будет лишних элементов, тем адекватнее будут методы достижения цели. Следует отметить, что в качестве цели может выступать и некий материальный *продукт*, и некое *желаемое состояние*.

Таким образом, основные компоненты целеориентированных теорий следующие: образ цели, теоретические положения, объясняющие некоторый набор фактов и каким-то образом связываемые с целью, практические методы, основанные на этих положениях и призванные обеспечить достижение цели. Прекрасный пример целеориентированных теорий дают различные психологические теории. В первую очередь здесь следует упомянуть психоанализ З. Фрейда и аналитическую психологию К.Г. Юнга.

К какому из двух вышеописанных видов теорий мы должны отнести теорию ИИ? Очевидно ко второму, поскольку главная цель подобной теории – не познание какой-то абстрактной истины, а создание конкретного «продукта». Однако, что же это за продукт?

Ответ на этот вопрос зависит, очевидно, от того, какой смысл мы вкладываем в понятие об ИИ. Действительно, как уже говорилось, в качестве ИИ мы можем рассматривать и некие системы, в той или иной мере берущие на себя работу интеллекта, и некий целостный феномен. В первом случае теория ИИ будет чем-то вроде *теории машин и механизмов*. Создание такой теории вполне возможно. Более того, можно, вероятно, считать, что она уже существует или, по крайней мере, появится в ближайшем будущем.

В случае, когда под ИИ понимается некий целостный феномен, дело, очевидно, обстоит сложнее. Какой конкретно смысл следует вкладывать в понятие «целостный феномен» совершенно не понятно. То есть не понятно, какова конкретная цель изысканий. Другими словами, отсутствует сколько-нибудь чёткий *образ цели*. А раз так, то нет критерия, позволяющего оценить те или иные методы и подходы, нет руководящего правила для их генерации.

Очевидно, что создание общей теории ИИ и выработка общего подхода к решению проблеме ИИ это не одно и то же. Действительно, говорить о некой единой теории ИИ можно только в том случае, если ИИ признаётся в принципе возможным. В случае же, когда мы говорим о проблеме ИИ, подразумевается, что этот последний может и не существовать, во всяком случае его возможность должна быть доказана.

Итак, предположим, что мы все-таки захотели выработать тот или иной вариант общего подхода к решению проблемы искусственного интеллекта. С какими трудностями мы бы столкнулись?

Во-первых, мы бы столкнулся с обилием «белых пятен», «лакун» в наших знаниях об интеллекте, психике, мозге, принципах их функционирования и организации. Причем, как правило, эти «лакуны» настолько велики, что каким-либо образом их *обойти* или просто *игнорировать* не представляется возможным [Симонов, 1992].

Во-вторых, серьезные проблемы возникли бы вследствие необходимости того или иного *однозначного* ответа на те основополагающие вопросы нашего мировоззрения, на которые самым непосредственным образом выводит проблема ИИ. Условно эти трудности могут быть названы *философскими*.

Следует отметить, что если, к примеру, в физике философские проблемы нужно *выискивать*, то в случае проблемы ИИ мы сталкиваемся с ними практически сразу же, как только начинаем ею заниматься. Материя и сознание, дискретное и непрерывное и т.д., и т.п. – все эти проблемы сразу же предстают перед нами со всей своей непосредственно-



стью, конкретностью и философской напряженностью. Причем для того, чтобы их увидеть, вовсе не обязательно быть философом.

Необходимо также отметить, что именно на основе правильного понимания проблем второго типа возможно правильное понимание того, в чём именно состоят и откуда происходят ограниченность и недостаточность тех методов, которые когда-либо рассматривались как методы создания ИИ.

Возникает отнюдь не праздный вопрос: а нельзя ли как-то *обойти* эти проблемы?

Единственная возможность для этого, по моему мнению, состоит в том, чтобы ***вывести проблему ИИ в чисто философскую область, то есть, перейти от рассмотрения философских проблем искусственного интеллекта к рассмотрению искусственного интеллекта как философской проблемы.***

Развернутое обоснование этого положения выходит далеко за рамки данной статьи. Да его, наверное, и невозможно провести в принципе. Однако надеюсь, что сказанное на предыдущих страницах достаточно естественно на него выводит. В любом случае, полагаю, нам ничто не мешает рассматривать положение о принципиальном философском характере проблемы ИИ как рабочую гипотезу.

Каковы же следствия принятия подобной гипотезы? Вот лишь несколько основных.

Во-первых, становится не так уж важно, как именно мозг обрабатывает поступающие в него данные – в аналоговой или какой-то другой форме, на каких элементах построен тот или иной компьютер и т.д., то есть от всего этого можно абстрагироваться. В этой связи особенно важно отказаться от термина «машина».

Вообще, как кажется, в настоящее время компьютерам уделяется слишком много внимания. Даже вопрос о том, возможен ли ИИ, фактически свёлся к вопросу, может ли мыслить *компьютер*. Особенно характерно это для англо-американской философии (см. [Dennet, 1987; Searle, 1990]).

На самом же деле компьютеры – суть нечто *случайное*, преходящее. Их или подобные им устройства ни в коей мере нельзя рассматривать как *единственно возможное* средство создания ИИ.

Во-вторых, мы также можем абстрагироваться от времени – в том смысле, что для нас не должно быть особенно важно, какие успехи достигнуты (или какие неудачи постигли) исследователей к настоящему моменту. Создание ИИ, чтобы ни понималось под этим понятием, должно рассматриваться как *существенным образом ориентированное на неопределённое будущее*, точно так же, как решение фундаментальных проблем нашего мировоззрения. Для нас – если уж мы решим подойти к проблеме ИИ с таких максимально широких позиций – должно быть важно одно: существуют искусственные, созданные человеком системы, способные осуществлять такую деятельность, которую ранее мог осуществлять только человек. Важен сам факт этого, *прецедент*. Из этого и нужно исходить. И дело здесь, пожалуй, даже не в том, сможет ли философия решить проблему ИИ, а в том, что мы должны *относиться* к последней как к философской проблеме.

Насколько плодотворной окажется выдвинутая здесь гипотеза покажет время, однако, думаю, уже сейчас можно сказать, что подобный *эксперимент* стоит того, чтобы его провести.

Проблема искусственного интеллекта и искусственные интеллектуальные системы

Рассмотрим вопрос об отношении проблемы, *идеи* ИИ к тем проблемам, которые традиционно с ней связываются, а также к тем *искусственным интеллектуальным системам* (далее – ИИ-системы), которые существуют в настоящее время или появятся в ближайшем будущем.

Под искусственными интеллектуальными системами я понимаю любые программно-аппаратные комплексы, способные, как считается, осуществлять такую деятельность, ко-

торию осуществляет человеческий интеллект. Что касается этих последних, то хотя, как я уже говорил, сегодня в англоязычной литературе искусственным интеллектом называют именно ИИ-системы, для всех понятно, что даже в своей *совокупности* они *не являются* ИИ (в том смысле, в котором этот термин употребляется в данной статье). Из этого непосредственно следует ряд вопросов. Например, создавая все эти системы, приближаемся ли мы к созданию ИИ? Действительно ли эти системы можно рассматривать как обладающие *элементами интеллекта*¹?

Примерно то же самое касается тех проблем и задач, которые традиционно связываются с проблемой ИИ. Действительно ли решение всех этих проблем приближает нас к созданию ИИ, или это не более чем очередная «тупиковая ветвь»? Тупиковая не потому, что, решая эти проблемы, мы ничего не познаём, а потому, что путь, ведущий непосредственно к ИИ, лежит *в другой стороне, не через решение этих проблем*.

В связи с этими соображениями и учитывая вышесказанное, хочу высказать следующее предложение. Может быть стоит, хотя бы на время, *развести* проблемы, связанные с формализацией, моделированием мышления и т.д., и *идею* ИИ, не рассматривать эти проблемы в контексте последней (в очередной раз замечу, что в данной статье ИИ понимается в *идеальном* смысле и что в это понятие вкладывается максимум содержания)? Слишком уж велик разрыв между нашей целью (ИИ) и теми средствами, с помощью которых, как мы предполагаем, её можно достичь.

Однако рассмотрим этот вопрос более строго.

Итак, как правило, большинство сложных конкретно-научных и философских проблем представляют собой, или, точнее сказать, *распадаются на* комплекс более мелких проблем и задач. Эти сложные проблемы по отношению к входящим к ним более мелким можно разделить на два класса.

К первому относятся те проблемы, для которых возможно дать *сущностную формулировку*, то есть формулировку, четко и ясно говорящую о сущности данной проблемы, сущности тех подпроблем, которые в неё входят и из неё следуют. Выгоды из наличия такой формулировки очевидны. Сущностная формулировка позволяет определить общие методы решения проблемы, методы решения входящих в неё подпроблем; отбросить те из них, которые имеют к данной сложной проблеме лишь косвенное отношение; объединить несколько подпроблем в одну, подведя их под общее основание. Вообще, сущностная формулировка позволяет оперировать с некоторой сложной проблемой как с единым целым, то есть на максимально высоком уровне абстракции. Это, кстати, особенно важно для философии.

Ко второму классу относятся те проблемы, для которых сущностную формулировку дать нельзя. Поэтому, когда по отношению к некоторой проблеме из этого класса мы говорим «проблема такая-то», мы на самом деле можем иметь в виду только то, что это не более чем просто общее *название* для некоторого комплекса проблем. В этом втором случае может возникнуть сомнение на счёт того, справедливо ли вообще объединять некий комплекс проблем под одним названием, подводить их под одну рубрику. Может быть, лучше было бы *распределить* их каким-то другим образом, попытаться подвести под какие-то *другие* проблемы. Однако этот вопрос должен решаться в каждом конкретном случае отдельно.

В связи с вышесказанным возникает вопрос: является ли та формулировка проблемы ИИ, которая была дана в предыдущем разделе, сущностной? Если да, то для каких проблем?

¹ «Одна из основных предпосылок создания ИИ заключается в том, что интеллектуальное поведение можно моделировать на цифровых ЭВМ» [Штанько, 1993, с. 246].



Поскольку в настоящее время не существует общепринятых концепций интеллекта и ИИ, решать проблему ИИ, исходя лишь из той формулировки, которая была дана в предыдущей главе просто невозможно. То есть её нельзя решать, *всё время оставаясь на максимально высоком уровне абстракции*.

В настоящее время существует целый ряд проблем, феноменов, традиционно связываемых с проблемой ИИ. Поэтому вопрос о том, может ли формулировка проблемы ИИ рассматриваться как сущностная, можно конкретизировать так: является ли формулировка проблемы ИИ сущностной для *этих* проблем, *объемлет* ли она их, выражает ли их сущность? Как я полагаю, в настоящее время дать ответ на этот вопрос также невозможно. И причина этого опять-таки состоит в том, что нет определенности с понятием интеллекта, понятием ИИ, нет определенности с *конечной целью*. Из-за этой неопределённости между формулировкой проблемы ИИ и множеством данных проблем существует *разрыв*, сократить который в ближайшее время вряд ли удастся.

Таким образом, вследствие всего вышесказанного, мы, вероятно, должны заключить, что в настоящее время проблема ИИ действительно может рассматриваться только лишь как *название* некоторой совокупности проблем. Вкладывать в неё какое-то большее содержание мы просто не имеем права.

Каковы же должны быть наши дальнейшие шаги в том случае, если мы это признаем? По-другому, как мы должны относиться к данной совокупности проблем и артефактов. Я вижу две возможности.

1. Можно попытаться найти сущностную формулировку для всех этих проблем. Из чего же мы должны исходить, чтобы её найти? Из самого её определения. Действительно, если сущностная формулировка есть такая, которая четко и ясно говорит о сущности некоторой проблемы, сущности входящих в неё проблем, то отсюда сразу же вытекает, что для того, чтобы её найти, необходимо в первую очередь найти ту *идею*, которая объединяет интересующий нас комплекс проблем (вообще, *феноменов*), найти то, что у них всех есть общего.

Однако не всё здесь так просто.

Нельзя, например, исключать тот случай, когда мы сможем найти общие идеи только для каких-то отдельных *групп* проблем. Эти же идеи, в свою очередь, не будут иметь между собой ничего общего. В этом случае нам останется только попытаться как-то *минимизировать* их количество.

Возможно также, что все проблемы (или какую-то их часть) удастся подвести сразу под *несколько* совершенно *разных* идей. Таким образом, может получиться, что для всего данного комплекса проблем просто не окажется одной общей сущностной формулировки. Либо их окажется несколько, причем таких, среди которых ни одну нельзя будет предпочесть другим. Вполне возможно также, что ни в одной из таких формулировок для термина «искусственный интеллект» просто не найдется места.

Ясно далее, что многие из этих идей должны быть философскими: проблемы, составляющие проблему ИИ, настолько разноплановы, что нам просто не удастся отыскать какую-либо объединяющую их *конкретно-научную* идею.

Однако не возвращаемся ли мы таким образом к тому, уход от чего был провозглашен выше, то есть, не подменяем ли мы анализ проблемы ИИ как таковой, анализом конкретных подпроблем, связываемых с проблемой ИИ?

Для того чтобы это понять, необходимо разобраться в том, что конкретно представляет собой философский анализ той или иной научной проблемы, в чем конкретно он состоит. Он состоит в том, что мы *подводим* эту проблему под ту или иную философскую идею или теорию и затем рассматриваем её как некий *частный случай* этой идеи или теории. Другими словами, мы *выбираем* из множества философских идей и теорий ту, которая может быть так или иначе *приведена в связь* с данной научной проблемой. Таким образом, мы имеем *одну* конкретно-научную проблему и *множество* философских идей. И именно в выборе

среди последних *подходящей* и состоит суть философского анализа любой конкретно-научной проблемы. Во всяком случае, суть *первого шага* такого анализа. Всё это, очевидно, справедливо и по отношению к философскому анализу проблем, обозначаемых в настоящее время термином «проблема ИИ».

Вернёмся теперь к тому подходу, который был предложен выше. Очевидная суть этого подхода, как это следует из его описания, состоит в том, что исходить необходимо не из какой-то отдельной конкретно-научной проблемы, а брать сразу множество проблем (лучше *все*) и, разлагая их, искать такую философскую идею, на основе которой их в конце концов можно было бы объединить. Если это удастся, то мы таким образом сможем подвести множество *конкретно-научных* проблем (а также множество феноменов, *артефактов*, созданных в результате попыток решения этих проблем) под одну *философскую* идею. Анализ проблемы в данном случае является не самоцелью, а лишь вспомогательным средством для дальнейшего *синтеза*. Другими словами, мы не заикливаемся на анализе, а идем дальше, к синтезу.

2. Идеи, которые мы находим в соответствии с первым подходом, по своему *объёму* должны либо полностью *совпадать* с объёмом проблем, подводимых под них, либо незначительно его *превышать*. Но что если попытаться отыскать такую философскую идею, по отношению к которой частным случаем будет не только *каждая* из этих проблем, но и *все они в совокупности*. Что если попытаться отыскать такую идею, объём которой будет в значительной степени *превышать* объём проблем, объединяемых названием «проблема ИИ»? При этом совсем не обязательно, чтобы эта идея была какой-то *новой* идеей в философии. Вполне возможно, что это будет идея хорошо известная, хорошо разработанная. Не обязательно также и то, чтобы с помощью какой-либо такой идеи можно было провести исчерпывающий анализ каждой проблемы, феномена, чтобы она *полностью* исчерпывала их сущность. Однако должно быть возможно усмотреть наличие тех или иных аспектов её в *каждой* проблеме (феномене, артефакте). Может оказаться, что ни одну из подобных идей нельзя будет предпочесть другим. Но в любом случае, посредством таких идей мы сможем как бы делать *срезы, сечения* интересующего нас комплекса проблем и феноменов. При этом одни идеи в большей мере будут показывать, выявлять сущность одних проблем, другие – других. Чем больше мы проведём подобных *сечений*, тем точнее и полнее будет наше представление о сущности данных проблем.

Отмечу, что при всей своей абстрактности эти идеи, очевидно, должны быть достаточно *специфичны*.

В смысле своего *начала* путь, который был только что описан, более труден, чем первый; он более труден хотя бы потому, что отысканию подобной идеи практически невозможно предпослать какую-либо рациональную схему, *алгоритм*. Многое здесь зависит и от общей философской культуры исследователя, и от простого случая. Однако этот путь более предпочтителен из-за тех выгод, которые он сулит. Он более плодотворен в плане следствий, приложений, менее кропотлив, да и, в конечном счете, более *реален*.

То, что справедливо относительно проблем, очевидно, будет справедливо и относительно тех продуктов, которые получаются в результате попыток их решения, то есть тех искусственных интеллектуальных систем, которые существуют в настоящее время.

Таким образом, задача в данном случае состоит в том, чтобы сформулировать такие философские идеи, *объём и содержание которых максимально соответствовали бы объёму и содержанию проблем, составляющих проблему ИИ (или, точнее, проблем, объединяемых под соответствующей рубрикой)*.

Также представляется интересным отыскание общих идей и для совокупности существующих в настоящее время искусственных интеллектуальных систем. В определенном смысле эта последняя задача может иметь и вполне самостоятельное значение.



В заключение этого раздела отмечу, что попытка реализовать второй подход была осуществлена в моей кандидатской диссертации [Швирков, 2006]. Насколько удачно – судить, разумеется, не мне.

Проблема ИИ и философия техники

Интересно рассмотреть отношение между проблематикой, связанной с ИИ и ИИ-системами, и проблематикой, относящейся к так называемой *философии техники*. В зависимости от того, как понимать ИИ и ИИ-системы, здесь могут быть как сходства, так и отличия.

В настоящее время под техникой чаще всего понимают:

- совокупность технических устройств, артефактов – от отдельных простейших орудий до сложнейших технических систем;
- совокупность различных видов технической деятельности по созданию этих устройств – от научно-технического исследования и проектирования до их изготовления на производстве и эксплуатации, от разработки отдельных элементов технических систем до системного исследования и проектирования;
- совокупность технических знаний – от специализированных рецептурно-технических до теоретических научно-технических и системотехнических знаний [Степин, 1996].

Что следует понимать под «проблемой техники»? Здесь также, как и с проблемой ИИ, ясности не много. Однако можно попытаться кое-что вывести из того контекста, в котором обычно употребляется это словосочетание в литературе.

Итак, под проблемой техники понимают, во-первых, выявление социальных, экономических и других последствий существования и функционирования техники, определение её влияния на общество, его структуру, на человека; во-вторых, определение места техники в нашей культуре; в-третьих, выявление того, что есть техника сама по себе, какова сущность техники (Э. Капп, М. Хайдеггер). Другими словами, какова идея, выражающая сущность той совокупности артефактов, которую принято называть техникой (и деятельности по ее созданию).

Для нас проблема выяснения сущности техники представляет интерес в первую очередь потому, что по своему характеру она оказывается сходна с той проблемой, на которую мы вышли в предыдущем разделе.

Действительно, и в том и в другом случае, речь, очевидно, идёт о том, чтобы найти такую идею, под которую можно было бы подвести некое множество феноменов, идею, которая бы позволяла рассматривать это множество как нечто единое.

Такое *сходство* проблем может во многом оказаться выгодным. Например, благодаря ему мы можем получить *подтверждение правомерности* многих из тех подходов, которые в настоящее время используются при анализе и оценке искусственных интеллектуальных систем.

С другой стороны, если мы будем рассматривать совокупность ИИ-систем не как самостоятельный феномен, а в контексте его возможной связи с созданием ИИ, то ситуация приобретет несколько иную окраску.

Как мы определили выше, техника в первую очередь есть совокупность неких артефактов. Эти артефакты, как правило, создаются для решения конкретных задач, достижения определённых целей. Для разных задач существует своя специфическая техника – как в смысле артефактов, так и в смысле действий, операций. Другими словами, техника не универсальна. То, что можно сделать с помощью одних технических устройств, нельзя сделать с помощью других.

Можем ли мы утверждать, что, создавая различные технические устройства, мы имеем своей целью создание некоего универсального технического устройства, которое

соединило бы в себе все их свойства и функции? Очевидно, нет. У нас нет такой цели, ни явной, ни мнимой. Хотя конечно, неплохо, если одно и то же устройство, одна и та же техника годится для различных целей.

Напротив, рассматривая создание ИИ-систем в контексте создания ИИ, мы тем самым как бы видим, предполагаем за их разработкой некую *идеальную цель*. Правда, как уже было сказано, возможность этой идеальной цели должна ещё быть доказана. Доказать нужно и то, что создание всех этих систем действительно ведёт непосредственно к созданию некоего универсального по своим функциям артефакта, что разработка их есть в то же время и *подготовка* к его созданию.

Таким образом, основное отличие проблематики, связываемой с техникой, от проблематики, связываемой с ИИ, состоит в том, что техника как совокупность артефактов, техника *как таковая* есть нечто вполне самостоятельное, самодовлеющее. Техника не предполагает некоего универсального технического устройства, аккумулирующего в себе все её функции. Техника *уже есть*, техника – это не какая-то *идеальная цель*, а *факт*. Вследствие этого *разрыв*, аналогичный тому, который существует между *идеями* ИИ и существующими в настоящее время ИИ-системами, а также методами, рассматриваемыми, как методы создания ИИ, здесь просто невозможен. Нам нет нужды доказывать, что те машины и те технологии, которые существуют в настоящее время – это техника, то есть, нам нет нужды доказывать, что они *возможны*.

Заключение

Таким образом, главные результаты проведенного анализа сводятся к двум интуициям:

1. Проблему ИИ и ИИ-системы следует разделить – хотя бы временно, и начать рассматривать последние как феномен в достаточной степени независимый от идеи ИИ, проблемы ИИ. Возможно, подобная *операция* позволит по-новому взглянуть на эти системы.

2. Саму же проблему ИИ следует рассматривать как проблему *философскую* (а не просто «междисциплинарную», «комплексную» и т.п.). Для анализа так понимаемой проблемы ИИ необходимо привлекать самую серьезную, самую абстрактную философскую теорию.

Надеюсь, что подобные меры будут способствовать прогрессу в проводимых исследованиях.

Список литературы

- Байдык Т.Н. 2001. Нейронные сети и задачи искусственного интеллекта. Киев, Наукова думка, 264 с.
- Войскунский А.Е. 2001. Метафоры Интернета. Вопросы философии, 11: 64–79.
- Гейтс Б. 2002. Бизнес со скоростью мысли. М., Изд-во Эксмо, 480 с.
- Глазунов В.А. 2002. Робототехника в постнеклассической науке. Вопросы философии, 11: 135–148.
- Искусственный интеллект: В 3 кн. 1990. М., Радио и связь.
- Искусственный интеллект: применение в интегрированных производственных системах. 1991. М., Машиностроение, 544 с.
- Миклошко Й. 1991. Искусственный интеллект, параллельные ЭВМ и параллельные вычисления. Будущее искусственного интеллекта. М., Наука: 244–269.
- Пушкин В.Г., Урсул А.Д. 1989. Информатика, кибернетика, интеллект: Философские очерки. Кишинёв, Штиинца, 295 с.
- Рузавин Г.И. 1987. Человек и робот: (О некоторых философских и социальных проблемах роботизации). Вопросы философии, 2: 67–79.
- Симонов П.В. 1992. Мозг и творчество. Вопросы философии, 11: 147–162.
- Степин В.С., Горохов В.Г., Розов М.А. 1996. Философия науки и техники. М., Гардарики: 340 с.
- Тихомиров О.К. 1991. «Искусственный интеллект»: взгляд психолога. Будущее искусственного интеллекта. М., Наука, 132–150.



- Уинстон П. 1980. Искусственный интеллект. М., Мир, 519 с.
- Швырков А.И. 2009. Проблема искусственного интеллекта: возможности методологического подхода. Эпистемология и философия науки, 1: 111–122.
- Швирков О.І. 2006. Проблема штучного інтелекту та людиновимірність штучних інтелектуальних систем. Дис... канд. філос. наук. Київ, 170 с.
- Штанько В.И. 1993. Философский анализ информационной парадигмы исследования мышления: возможности и проблемы. Дис... д-ра филос. наук. Харьков, 279 с.
- Bostrom, N. 2014. Superintelligence: Paths, Dangers, Strategies. Oxford, UK, Oxford University Press, 352 p.
- Bruce, B.G. 2005. A (Very) Brief History of Artificial Intelligence. AI Magazine, 26(4): 53–60.
- Charniak E., McDermott, D. 1985. Introduction to Artificial Intelligence. Reading, MA, Addison Wesley, 155 p.
- Dennet D (with Haugeland J.). 1987. Intentionality. The Oxford Companion to the Mind. Oxford, University Press, 385 p.
- Haugeland, J. 1985. Artificial Intelligence: The Very Idea. Cambridge, MA, MIT Press, 287 p.
- Russell S., Norvig P. 2009. Artificial Intelligence: A Modern Approach, 2nd edition. Saddle River, NJ, Prentice Hall, 1112 p.
- Searle I. 1990. Minds, Brains, and Programs. The Philosophy of Artificial Intelligence. Oxford, University Press, 347 p.
- Searle J. 1997. The Mystery of Consciousness. New York, New York Review of Books, 224 p.

References

- Baydyk T.N. 2001. Neyronnyye seti i zadachi iskusstvennogo intellekta [Neural networks and tasks of artificial intelligence]. Kiev, 264 p.
- Voyskunskiy A.Ye. 2001. Metafory Interneta [Metaphors of the Internet]. *Voprosy filosofii*, 11: 64–79.
- Gates B. 2002. Biznes so skorost'yu mysli [Business at the speed of thought]. Moscow, 480 p.
- Iskusstvennyy intellekt: V 3 kn. 1990. [Artificial intelligence: In 3 books]. Moscow.
- Iskusstvennyy intellekt: primeneniye v integrirovannykh proizvodstvennykh sistemakh [Artificial intelligence: using in integrated production systems]. 1991. Moscow, Mashinostroyeniye, 544 p.
- Mikloshko Y. 1991. Iskusstvennyy intellekt, parallel'nyye EVM i parallel'nyye vychisleniya [Artificial intelligence, parallel computers and parallel computing]. *Budushcheye iskusstvennogo intellekta* [The future of Artificial Intelligence]. Moscow, Publ. Nauka: 244–269. (in Russian)
- Pushkin V.G., Ursul A.D. 1989. Informatika, kibernetika, intellekt: Filosofskiye ocherki [Informatics, cybernetics, intellect: Philosophical essays]. Chisinau, 295 p. (in Russian)
- Ruzavin G.I. 1987. Chelovek i robot: O nekotorykh filosofskikh i sotsial'nykh problemakh robotizatsii [Man and robot: On some philosophical and social problems of robotics]. *Voprosy filosofii*, 2: 67–79.
- Simonov P.V. 1992. Mozg i tvorchestvo [Brain and Creativity]. *Voprosy filosofii*, 11: 147–162.
- Stepin V.S., Gorokhov V.G., Rozov M.A. 1996. Filosofiya nauki i tekhniki: ucheb. posobiye dlya vuzov [Philosophy of Science and Technology: textbook for universities]. M., Gardarika: 340 p.
- Tikhomirov O.K. 1991. «Iskusstvennyy intellekt»: vzglyad psikhologa [“Artificial intelligence”: the view of a psychologist]. *Budushcheye iskusstvennogo intellekta* [The future of Artificial Intelligence]. Moscow, Publ. Nauka: 132–150.
- Winston P. 1980. Iskusstvennyy intellekt [Artificial Intelligence]. Moscow, Mir, 519 p. (in Russian)
- Shvyrkov A.I. 2009. Problema iskusstvennogo intellekta: vozmozhnosti metodologicheskogo podkhoda [The problem of Artificial Intelligence: the possibilities of the methodological approach]. *Epistemologiya i filosofiya nauki*, 1: 111–122.
- Shvyrkov O.I. 2006. Problema shtuchnogo intelektu ta lyudinovimirmist' shtuchnikh intelektual'nikh system [The problem of artificial intelligence and human measurements of artificial intellectual systems]. Candidate's thesis, Kiev, 170 p. (in Ukrainian)
- Shtan'ko V.I. 1993. Filosofskiy analiz informatsionnoy paradigmy issledovaniya myshleniya: vozmozhnosti i problem [Philosophical analysis of the information paradigm of thinking research: opportunities and problems]. Candidate's thesis, Khar'kov, 279 p.
- Bostrom N. 2014. Superintelligence: Paths, Dangers, Strategies. Oxford, UK, Oxford University Press, 325 p.
- Bruce B.G. 2005. A (Very) Brief History of Artificial Intelligence. AI Magazine, 26(4): 53–60.



- Charniak E., McDermott D. 1985. *Introduction to Artificial Intelligence*, Reading, MA, Addison Wesley, 155 p.
- Dennet D (with Haugeland J.). 1987. *Intentionality*, in: *The Oxford Companion to the Mind*. Oxford, *University Press*: 245–267.
- Haugeland J. 1985. *Artificial Intelligence: The Very Idea*. Cambridge, MA, MIT Press, 287 p.
- Russell S., Norvig P. 2009. *Artificial Intelligence: A Modern Approach*, 2nd edition. Saddle River, NJ: Prentice Hall, 1112 p.
- Searle I. 1990. *Minds, Brains, and Programs*, in: *The Philosophy of Artificial Intelligence*. Oxford, *University Press*: 132–154.
- Searle J. 1997. *The Mystery of Consciousness*, New York, NY, *New York Review of Books*, 224 p.

Конфликт интересов: о потенциальном конфликте интересов не сообщалась.

Conflict of interest: no potential conflict of interest has been reported.

Поступила в редакцию 18.09.2021

Received September 18, 2021

Поступила после рецензирования 18.12.2021

Revised December 18, 2021

Принята к публикации 09.02.2022

Accepted February 9, 2022

ИНФОРМАЦИЯ ОБ АВТОРЕ

INFORMATION ABOUT THE AUTHOR

Швырков Александр Иванович, канд. филос. наук, доцент кафедры гуманитарных и социальных дисциплин, Брянский государственный технический университет, г. Брянск, Россия

Alexander I. Shvyrkov, PhD in Philosophy, Associate Professor of the Department of Humanities and Social Studies, Bryansk State Technical University, Bryansk, Russia