

ОТЗЫВ

официального оппонента Финна Виктора Константиновича на диссертацию Ковалева Максима Александровича на тему «Роль символьных методов в задачах общего искусственного интеллекта», представленную к защите на соискание ученой степени кандидата философских наук по научной специальности 5.7.6. Философия науки и техники

Искусственный интеллект (ИИ) как научно-практическая область исследования имитации и поддержки познавательной деятельности человека посредством компьютерных систем стал существенным явлением цивилизационных и культурных аспектов современного общества, характеризуемого как **информационное** общество. Это означает применение ИТ-технологий во многих его областях – управлении, производстве, науке, медицине, обороне, правовой и гуманитарной сферах, и в сфере услуг. Важно отметить, что весьма значимыми ИТ-технологиями являются те, которые применяют методы ИИ, продуктами которых являются интеллектуальные системы, интеллектуальные роботы. Это означает, что информатизация жизни современного общества не сводится к так называемой цифровизации (цифровизации лишь **технологический** аспект информатизации), а использует **интеллектуализацию** программных продуктов, что делает последние, как говорят популяризаторы, «умными».

Технологии не являются гермафродитами – они не рождают сами себя: они имеют родителей, которыми являются научные методы. Как говорил академик Л.Д. Ландау – «метод важнее результата». В связи с этим следует заметить, что понимание природы интеллекта и методов его имитации, формализации и автоматизации является необходимым условием создания технологий (в том числе, ИТ-технологий) и их **развития**. Отсюда следует необходимость методологии исследований в области ИИ и философского понимания возможностей и ограничений в применении методов ИИ. Это весьма естественно, ибо эвристика, логика рассуждений и представление

знаний имеют прямое отношение к философии науки и техники, востребованным информационным обществом.

Проблемам осмысления роли ИИ в современном обществе, роли различных подходов к созданию методов имитации интеллекта для практических приложений посвящена диссертация М.А. Ковалева. Очевидно, что это означает её значительную актуальность и практическую значимость, которые подтверждены добросовестным и систематическим исследованием диссертанта большого литературного массива философских размышлений об ИИ отечественных и зарубежных авторов.

Отмечу основные достоинства диссертационного исследования М.А. Ковалева.

(1) В диссертации аргументировано показано, что проблемы ИИ не могут быть сведены к решению их только в рамках **коннекционистского** понимания интеллекта, реализуемого **нейронными сетями**. В диссертации утверждается необходимость применения, так называемых, **символьных методов**, которые, фактически, являются реализацией логики рассуждений, которая есть средство формализации эвристик получения нового знания (knowledge discovery), а это – цель применения методов ИИ. Как отмечал создатель ИИ Д. Маккарти с соавтором Д. Хейесом (статья 1969 года): имеются три главных проблем развития ИИ: уточнение идеи интеллекта, создание принципов эпистемологии получения нового знания (knowledge discovery) и разработка эвристик для получения нового знания в компьютерных системах. Поэтому создание таких эвристик – проблема, решаемая «символьными» методами. В этом состоит актуальность и ценность оправдания необходимости применения символьных методов, ибо логика рассуждений – главное средство ИИ.

(2) М.А. Ковалев в своем исследовании обоснованно формулирует главную проблему практического применения идей ИИ – создание **партнерских человеко-машинных систем**, называя их, согласно некоторой терминологии, гибридными системами. Естественно эти компьютерные

системы называть «интеллектуальными системами», если их архитектура состоит из баз фактов, баз знаний, Решателя задач, когнитивного интерфейса, последние два модуля реализуют в автоматическом и интерактивном режимах **имитацию** и **усиление** интеллектуальных способностей, среди которых имеются способности к рассуждению, аргументации, синтезу познавательных процедур, объяснению (ответ на вопрос «почему?»), рефлексии (оценки результатов), обучения на примерах.

(3) Важным достоинством диссертационной работы М.А. Ковалева является систематическое исследование истории развития идей искусственного интеллекта. Эта история включает не только историю технологических достижений, но и историю эпистемологических идей, необходимых для понимания формирования нового знания для последующей его автоматизации.

Известна книга Г. Шпета «История как проблема логики», аналогично: ИИ есть проблема логики рассуждений и точной эпистемологии, а связь ИИ с логикой и эпистемологией есть предмет интеллектуальной истории.

Возможно, следует предложить автору диссертации создать книгу об истории развития идей искусственного интеллекта и их технологических реализаций: от эпистемологии, логики и компьютерной науки – к технологиям применения идей ИИ и машинного обучения.

(4) М.А. Ковалев в своем диссертационном исследовании рассматривает достижения в применении идей ИИ как нового направления в технике, посредством которой ИТ-технологии стали инструментарием современного информационного общества. Автор диссертации развивает идею «подручности» М. Хайдеггера, которая для продуктов ИИ фактически является следствием понимания интеллектуальных (гибридных) систем как **партнерских человеко-машинных систем**, управляемых человеком и являющихся его помощниками и «усилителями умственных способностей» согласно У.Р. Эшби («Введение в кибернетику», 1959). В силу сказанного «подручность» по М.А. Ковалеву выводима из его понимания

интеллектуальных систем как партнерских человеко-машинных систем, а, следовательно, идея М. Хайдеггера лишь согласие, а не источник явления.

(5) Достоинством диссертационного исследования является также рассмотрение возникшего влияния идей и методов ИИ на проблемы культуры – технологизация обучения, влияние на когнитивные исследования (в том числе в психологии), создание формальных и алгоритмических средств для социальных и гуманитарных наук. Последнее обстоятельство фундаментально, ибо создает возможность этих дисциплин быть **экспериментальными**, что порождает возможности точного исследования и реализации в них критерия демаркации в смысле К.Р. Поппера. Это означает опровержение идеи Г. Риккерта о принципиальном различии наук о природе и наук о культуре. Главной областью применения ИИ являются науки о жизни и социальном поведении, так как методы ИИ применимы там, где знания слабо формализованы, а данные могут быть структурированы для определения сходства фактов. Свидетельством этого являются успешное применение методов ИИ и машинного обучения в социологии, преобразовании текстов на естественных языках, криминалистике и медицине.

(6) К числу достижений автора диссертации следует отнести содержательный и глубокий анализ проблемы **понимания** текстов естественных языков, которая не решается модными большими языковыми моделями, весьма эффективными как средства порождения новых текстов. Следствием формализации понимания текстов было бы автоматизированное формирование баз фактов посредством извлечения релевантных знаний из текстов естественного языка. Решение этой проблемы создало бы новый класс интеллектуальных информационных систем.

(7) В диссертации содержится убедительная критика многих неясных идей, включающих такие как «сильный ИИ» Д. Сёрля и саму идею «общего интеллекта».

Следует отметить также содержательный и обширный обзор по философским проблемам ИИ, о чем уже упоминалось в настоящем отзыве на диссертацию.

Из перечисленных достоинств диссертационного исследования М.А. Ковалева следует его актуальность, новизна и практическая значимость.

В диссертации М.А. Ковалева обсуждаются и исследуются актуальные и трудные проблемы методологии искусственного интеллекта – научно-практической области имитации и усиления познавательной деятельности человека, порождающей новое знание (knowledge discovery). В силу сказанного ИИ не является **программным продуктом**, ибо он – наука о создании программных продуктов. Программные продукты ИИ используют представление знаний и данных (фактов), логику рассуждений и процедуры, реализующие эвристические принципы порождения нового знания.

Для ИИ существенно использование именно **знаний**, а не **информации**, ибо знание в компьютерных системах, реализующих идеи и методы ИИ, имеет определенную организацию и средства выражения: предикаты, представляющие отношения, термы, обозначающие понятия, высказывания, образованные предикатами, кванторами и логическими связками. Эти средства образуют высказывания, которые связаны логическими зависимостями, а некоторые из них порождаются заданными процедурами. Таким образом, знание в компьютерной системе ИИ является декларативным и процедурным. В случае наличия когнитивного интерфейса в программной системе, который реализует управление рассуждениями и высказываниями, используется и концептуальное знание, представляющее соответствующую методологию.

Таким образом, из сказанного следует, что методология и философия ИИ должна учитывать особенности представления знаний (**не информации!**) в компьютерной системе, средства логики рассуждений и особенности реализации процедур, порождающих новые знания посредством логики рассуждений и используемых вычислений.

Из сказанного следует, что используются формальные языки логики рассуждений и представление знаний, а также алгоритмические языки программирования для осуществления процедур. Следовательно, использование символов **очевидный научный факт**, не являющийся специфическим средством, т.н., «символьных методов» ИИ, ибо нейронные сети (продукт коннекционизма) также используют формальный язык, в котором выразимы логические и другие функции. Отсюда следует неудачность терминологии, отличающей символьные методы и коннекционистский подход к имитации интеллектуальных способностей.

Более существенно различие методов ИИ и методов машинного обучения, что связано с отличием интеллектуального анализа данных (преобразование знаний) от анализа данных (преобразование данных). Это различие имеет глубокий смысл, объяснимый с помощью фундаментальной эпистемологической триады И.Канта – «чувственное созерцание – рассудок – разум». Эта триада помогает понять необходимость создания интеллектуальных систем как **партнерских человеко-машинных систем**, применяемых в автоматическом режиме для реализации рассудка и в интерактивном режиме для поддержки реализации разума. Отсюда следует невозможность создания «сверхразума» – систем без участия человека (т.н., принцип сингулярности).

Выше было отмечено важное достижение диссертации М.А. Ковалева, состоящее в необходимости разработки и развития гибридных систем ИИ (удачнее говорить о партнерских человеко-машинных системах, их синоним – «интеллектуальные системы»).

Итак, из сказанного следует, что «общий ИИ», «сильный ИИ» - неясные идеи, не имеющие практического значения. Повторю, что ИИ – не программный продукт, программными же продуктами ИИ являются системы ИИ (программные системы, реализующие некоторые процедуры или их комбинации, но использующие лишь способности прединтеллекта – распознавание, обучение на примерах, память; в основном это системы

машинного обучения), интеллектуальные системы (партнерские человеко-машинные системы, имитирующие и усиливающие способности теоретического интеллекта – преобразователя знаний), интеллектуальные роботы. ИИ-роботы есть интеллектуальная система, сенсорный блок (динамическое формирование баз фактов) и мехатроника.

В силу сказанного формулирую первое замечание:

1. В диссертации нет четкого разграничения ИИ как научной дисциплины и продуктов ИИ, что, однако, распространенное явление в литературе об ИИ (см. сочинения Д. Сёрля). В диссертации используется неясный термин «система ИИ».

2. В диссертации последовательно утверждается, что результатом работы «систем ИИ» являются новые знаки (символы), понятия и правила. Это утверждение неудачно по следующим причинам.

2а. Если порождаются понятия, то, следовательно, возникает имя (согласно треугольнику Г. Фреге).

2б. Новые понятия иногда могут порождаться с помощью интеллектуальных систем, но только в интерактивном режиме и притом весьма редко.

2с. Цель действий интеллектуальной системы порождения новых знаний (knowledge discovery), а, следовательно, не является обязательным порождение правил. Новым знанием могут быть гипотезы, эмпирические закономерности, множество утверждений между которыми могут существовать логические зависимости. Поэтому эта тройка терминов – знак, понятие, правило не является реальной и обязательной.

3. В диссертации М.А. Ковалева утверждается, что системы ИИ осуществляют обучение и самообучение. Действительно, системы, реализующие методы машинного обучения, осуществляют сравнение примеров, обобщают результаты этих сравнений и в этом смысле обучаются. Возможно при этом и участие «учителя». Соответственно, имеются системы обучения с подкреплением. Но для интеллектуального анализа данных

(то-есть, применения интеллектуальных систем) необходимо решить классическую проблему индукции. А, значит, формализовать проблему **доверия** к результатам рассуждений, дополнив её формализацией абдукции – логического средства принятия гипотез, объясняющих их принятие.

Таким образом, для методологии ИИ существенно рассмотрение **проблемы индукции**, решение которой порождает возможность обучения на примерах. **Синтез познавательных процедур** – индукции (порождение гипотез о причинах), аналогии (использование причин для предсказания) и абдукции (средства принятия гипотез) является одной из **главных проблем ИИ**, а обучение на примерах является её следствием.

4. В исследовании М.А. Ковалева недостаточное внимание уделено **логике рассуждений** и её роли для формализации эвристик knowledge discovery. Однако, ИИ есть прежде всего наука о представлении знаний и логике рассуждений, применяемой вместе с вычислениями к данным и знаниям в компьютерных системах.

5. В содержательном и систематическом обзоре становления идей и методов ИИ отсутствует информация об известном японском проекте компьютеров 5^{ого} поколения, сыгравшим важную роль в применении суперкомпьютеров и понимании необходимости синтеза познавательных процедур (в том числе, индукции, абдукции и дедукции).

6. В упомянутом выше обзоре, по моему мнению, следовало бы выделить следующую последовательность мыслителей и научных сообществ, сыгравших роль в потенциальном и реальном смысле порождения идей, необходимых для развития эпистемологии и методологии ИИ: Аристотель, стоики (Хризипп), мегарская школа (Эвбулид), Ф. Бэкон, Д. Юм, И. Кант, Д. Гершель и Д.С. Милль, Ч.С. Пирс, К. Ясперс, логический позитивизм (в том числе, Г. Рейхенбах), Э. Кассирер, Д. Поппер, К.Р. Поппер, Д. Маккарти, Д.А. Поспелов.

7. В диссертации М.А. Ковалева имеется тенденция обратить внимание на роль философии для ИИ (думаю, что здесь важно влиять на эрудицию

специалистов в области ИИ). Однако почти отсутствует объяснение роли ИИ для философии, что важно для создания точной (эволюционной) эпистемологии – методологического основания эвристик для knowledge discovery.

8. В диссертации М.А. Ковалева упоминается необходимость уточнения идеи интеллекта, важной для понимания архитектуры и логических средств интеллектуальных систем и интеллектуального анализа данных. Однако эта потребность не нашла развития в тексте.

Теоретический интеллект познающего субъекта – это система знаний, множество интеллектуальных способностей, включающее способности к рефлексии, рассуждению, аргументации, объяснению, интеграции знаний, распознаванию, обучению и т.п., а также высшие психические функции – интуиция, интенция, инициатива, рефлексия (отличная от аргументированной рефлексии), воображение. Из такого понимания интеллекта следует необходимость интерактивного режима интеллектуальных систем и, следовательно, «подручность» о которой говорит автор диссертации.

Имеются также редакционные замечания, сообщенные автору диссертации.

В заключение отмечу важное следствие диссертационного исследования М.А. Ковалева – оно порождает комплекс серьезных исследовательских проблем, которые свидетельствуют о содержательности и практической значимости проведенных исследований автором диссертации.

1. В связи с пониманием роли обучения на примерах возникает оправдание полученных обобщений, средствами теории индукции. Отсюда следует целесообразность рассмотрения истории развития идей и процедур индукции у следующих мыслителей от Аристотеля до Ф. Бэкона, Д. Гершеля и Д.С. Милля, У. Хьюила, Ч.С. Пирса, Р. Карнапа, Г. Рейхенбаха, К.Р. Поппера. Эту историю индукции полезно сравнить с современными

методами индукции, включая и ДСМ-метод автоматизированной поддержки исследований.

2. Следует рассмотреть три проблемы Д. Маккарти-Д. Хейеса об уточнении идеи интеллекта, эпистемологии получения нового знания и соответствующих эвристик в интеллектуальных системах.

3. Существенно рассмотреть кантианские мотивы в проблемах ИИ: логику рассуждений как трансцендентальную логику, триаду «чувственное созерцание – рассудок – разум», синтетические суждения a priori и a posteriori как схемы для knowledge discovery.

4. Следует обсудить необходимость при интеллектуальном анализе данных применения трех теорий истины – когерентной (для гипотез), корреспондентной (для верификации результатов), прагматической (для принятия результатов с точки зрения интерпретатора).

5. Следует также охарактеризовать интеллектуальные системы как особый вид артефактов **несводимых только к инженерному статусу**, ибо они **партнеры** человека, ведущие с ним **диалог** в когнитивном интерфейсе посредством интерактивного режима («инженерность» имеет место только в применении компьютера как технического устройства).

6. Богатой темой являются цивилизационные и культурные аспекты применения методов ИИ в информационном обществе и соответствующие социальные последствия для трудовой занятости, коммуникации, управления, образования, правовой сферы, науки, обороны.

Необходимость рассмотрения перечисленных проблем является свидетельством плодотворности исследований автора диссертации, у которого имеется реальная возможность продолжить исследование философских проблем ИИ. Добавлю, что имеющиеся замечания, не влияют на высокую оценку диссертационного исследования М.А. Ковалева в весьма актуальной и сложной области науки и техники.

Диссертация отвечает требованиям, установленным Московским государственным университетом имени М.В.Ломоносова к работам подобного рода. Содержание диссертации соответствует специальности 5.7.6. Философия науки и техники (по философским наукам), а также критериям, определенным пп. 2.1-2.5, 3.1 Положения о присуждении ученых степеней в Московском государственном университете имени М.В.Ломоносова. Диссертация оформлена согласно приложениям № 8, 9 Положения о совете по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук Московского государственного университета имени М.В. Ломоносова.

Таким образом, соискатель Ковалев Максим Александрович заслуживает присуждения ученой степени кандидата философских наук по специальности 5.7.6. Философия науки и техники.

Официальный оппонент:

доктор технических наук, профессор,
заслуженный деятель науки Российской Федерации,
главный научный сотрудник отдела интеллектуального анализа данных
федерального государственного учреждения «Федеральный
исследовательский центр «Информатика и управление» Российской
академии наук»

ФИНН Виктор Константинович



6 мая 2024 года

Контактные данные:

тел.: +7 (495) 250-63-29; e-mail: kmlis@rggu.ru

Специальность, по которой официальным оппонентом
защищена диссертация:

05.13.17 - Теоретические основы информатики

Адрес места работы:

119333, Москва, Вавилова, д. 44, корп. 2,

ФИЦ ИУ РАН, отдел интеллектуального анализа данных, главный научный
сотрудник.

Подпись В.К. Финна заверяю
Ученый секретарь ФИЦ ИУ РАН
Д.т.н. В.Н. Захаров
06 мая 2024 г.

