

ОТЗЫВ

официального оппонента, доктора философских наук, ведущего научного сотрудника кафедры философии и методологии науки философского факультета МГУ имени М.В. Ломоносова Алексеева Андрея Юрьевича на диссертационную работу Михайлова Игоря Феликсовича «Когнитивные основания социальности», представленной к защите на соискание ученой степени доктора философских наук по научной специальности 09.00.01 - Онтология и теория познания в совет по защите диссертаций на соискание ученой степени кандидата наук, на соискание ученой степени доктора наук Д 002.015.03 на базе федерального государственного бюджетного учреждения науки Института философии Российской академии наук

Диссертационная работа непосредственно вписывается в современные междисциплинарные исследования искусственного интеллекта.

Актуальность исследования обусловлена разработкой философской методологии, объединяющей на концептуальном фундаменте основные проекты современных исследований искусственного интеллекта. Эти проекты ориентированы на компьютерную имитацию когнитивных феноменов различного содержания. Проект «искусственной жизни» моделирует витальные феномены. Проект «искусственного мозга» (или «искусственного сознания», в зависимости от ориентации психофизического вектора) изучает возможности компьютерной репликации, репрезентации, репродукции ментальных феноменов. Проект «искусственной личности» ориентирован на моделирование персонологических феноменов креативности, самости, смысла, Я. Наконец, проект «искусственного общества» включает методологию программной имитации социокультурных феноменов. Все классы обозначенных когнитивных феноменов в работе Игоря Феликсовича явно представлены и достаточно подробно исследованы. Поэтому *важность работы* определена стратегией построения т. н. общего искусственного интеллекта, а её народно-хозяйственная значимость подчеркивается прямым соответствием Указу Президента Российской Федерации № 490 от 10 октября 2019 г. «О развитии искусственного интеллекта в Российской Федерации».

Автор принимает общую когнитивную метафору о том, что способ реализации разнородных когнитивных феноменов имеет вычислимый характер и для выработки сквозной интегральной методологии он решает следующие задачи: анализирует основные направления когнитивной науки, их онтологические и концептуальные основы; проводит анализ онтологического и концептуального базиса биологии, психологии, социологии и, отчасти, культурологии, а также уточняет понимание вычислений для определения возможностей компьютерного моделирования обозначенных когнитивных феноменов; определяет концептуально-онтологические пересечения указанных наук и перспективы их междисциплинарного синтеза.

Научная новизна исследования состоит в выработке комплексной методологии междисциплинарной интеграции научных направлений, исследующих когнитивные основания человеческой социальности; в разработке концептуальной схемы, позволяющей применять однородные формальные средства для описания когнитивных и социальных явлений; в демонстрации применимости вычислительных процессов в исследовании биологических, когнитивных и социальных систем; во введении в отечественный научный оборот проблематики, широко обсуждаемой в многочисленных зарубежных публикациях по философским проблемам когнитивных наук и искусственного интеллекта.

Логическая структура работы подчеркивает последовательность рассуждений автора, которые начинаются от формулировки т. н. «рабочей гипотезы» и в последующем отчетливо раскрываются в ходе её доказательства.

Первая глава посвящена анализу философско-методологических оснований исследования. В этой главе вводится и защищается понятие научных (предметных, доменных) онтологий, которые по содержанию и функциям отличаются как от теоретических, так и от эмпирических компонентов научного знания. Предлагается философское исследование

рассматривать как онтологический аспект науки, избегая каких-либо утверждений на её теоретическом и эмпирическом уровнях. В случае с социальными науками существенную помощь может оказать номиналистическая метаонтология, разработанная некоторыми философскими школами и признающая существование множества простых объектов, все свойства которых или сводятся к отношениям, или обеспечивают возможность вступать в отношения. Применительно к интересующим нас группам наук это означает предпочтение взгляда на их предметы как на сетевые структуры, состоящие из узлов (нейронов, агентов) и рёбер (соединений, коммуникаций) между ними. Простота и лаконичность этой онтологии означает, что объяснительная функция почти целиком ложится на собственно теоретический уровень соответствующих наук, т.е. на формальные описания законов (правил) и алгоритмов. В этой же главе обосновывается точка зрения, что правильно построенная когнитивная наука, способная к междисциплинарной интеграции с социальным знанием, должна быть очищена от неоправданных философских презумпций. Показано, что сознание онтологически соотносится, но логически не зависит от наличия осознанных состояний. Последние, в свою очередь, не предполагают с необходимостью самоосознание. Также в этой главе показано, что, во-первых, интенциональность как свойство сознательных актов быть «о чём-то» не исчерпывает всех явлений сознания, так как то, что делает квалиа субъективными и невыразимыми в перспективе третьего лица, не является интенциональным. Во-вторых, имеется необходимость обратиться к коллективной интенциональности, т.е. к интенциональным свойствам тех явлений сознания, которые имеют основу в социальном образе жизни людей.

Вторая глава посвящена анализу концептуальных основ когнитивных наук. Здесь устанавливается, что известные сомнения Канта в том, что психология как наука возможна, имели основанием особенности качественно-субстанциального (умозрительного) этапа её развития. Переход на количественно-функциональные основания, подкреплённые процедурами

эмпирической верификации, позволил в онтологическом плане расширить сферу психического до пространства межсубъектной коммуникации, а в методологическом — продуктивно применить «вычислительную метафору» к этой науке. На этом основании делается предположение, что разработка концепции вычислений, адекватной предмету когнитивных наук, и перевод этих наук на компьютерно-функционалистские основания поможет преодолеть ограничения классических естественно-научных методологий, не справляющихся со сложностью предмета. Когнитивная наука нуждается в опоре на «вычислительную метафору», но последняя должна выйти за тесные тьюринговы рамки, избавиться от других частных, связанных с особенностями отдельных дисциплин, и стать общей (родовой) теорией, объединяющей предметные дисциплины, изучающие процессы в сложных системах. Особо анализируется понятие репрезентации, вызывающие многочисленные дискуссии в публикациях по философским проблемам когнитивных наук. Из анализа понятия вычислений следует, что репрезентации являются необходимым элементом вычислительных процессов. В качестве такового элемента репрезентации имеют место также и в когнитивных актах. Однако именно поэтому понятие репрезентации, будучи родовым понятием, не выражает видовой специфики когнитивного, и вряд ли поэтому может быть методологической основой собственно когнитивных наук. В этой же главе рассматриваются различные антикомпьютерно-функционалистские подходы и, в частности — концепция «биологического мозга», популярная среди отечественных нейробиологов. Анализ этой позиции показывает, что аргумент «биологического мозга» недостаточен для дискредитации вычислительного, и, в частности нейросетевого, подхода в когнитивной нейронауке. В целом во второй главе показывается, что психология, исторически преодолевшая качественно-субстанциалистские основания в пользу количественно-функционального подхода, естественным образом приняла «вычислительную метафору», став когнитивной психологией не только по предмету, но и по методу. Трудности

с адаптацией этой метафоры свидетельствуют о необходимости нового понимания вычисления. Результатом этого этапа исследования является онтологическая концепция вычисления как потенциально многоуровневого процесса, предполагающего представленность одних процессов в других с помощью механизмов репрезентации. В свою очередь, репрезентации могут пониматься как необходимый элемент многоуровневых вычислительных процессов в когнитивных системах.

Третья глава посвящена выработке понимания социальной реальности как метасетевого расширения когнитивных способностей индивидов. Здесь показывается, что внедрению номиналистической онтологии в науки об обществе мешают традиционные философские предубеждения, в соответствии с которыми общество состоит из крайне сложно организованных личностей. Однако, именно когнитивные теории, построенные на той же номиналистической онтологии, способны не только справиться с интенциональностью и относительным индетерминизмом социальных акторов, но и способствовать интеграции когнитивных и социальных наук. Для этого когнитивные науки должны рассмотреть сеть нейронов головного мозга и социальную сеть как параллельные распределённые вычислительные устройства, способные рассчитывать вероятностные зависимости между переменными окружения на основе обработки больших массивов данных. Язык, в свою очередь, понимается в рамках такого подхода как интерфейс между этими двумя типами компьютеров сетевой архитектуры, который сам, напротив, реализуется как серийная последовательность команд. Язык не только «переодевает мысли», по словам Витгенштейна, но и переформатирует под себя мозговую деятельность и социальные отношения. Социальная сеть, внешняя по отношению к нейросети мозга и связанная с нею языковым интерфейсом, образует метасеть со сложной комплексной вычислительной архитектурой. Последние исследования в области концептуальных и методологических проблем биологии показывают, что мир органических молекул и живых

организмов представляется как многоуровневая структура вложенных сетеподобных систем, осуществляющих вычислительную обработку информации на основе, скорее всего, статистических алгоритмов. Этот результат позволяет предположить, что когнитивные способности, в основе которых лежат те же механизмы, по сути являются продолжением и усложнением базисного устройства жизни, а биологические, когнитивные и социальные науки закономерным образом должны образовать единый комплекс знаний, связанный предметно и методологически. Таким образом, как предполагает автор, социальная организация возникает и существует как система параллельно разделяемого доступа индивидов к когнитивным ресурсам друг друга, а также к общим (распределённым) когнитивным ресурсам. Мультиагентные суперкомпьютерные системы (как современный результат эволюции программных средств информационных технологий) моделируют социальные взаимодействия, основанные на правилах эпистемических логик. Мультиагентная система, состоящая из обучаемых нейросетевых агентов, создающих и использующих символический язык для коммуникации, была бы наиболее реалистичной моделью человеческого общества. Когнитивная наука об обществе понимается в этой главе как наука, основанная на вычислительном подходе. Когнитивно обоснованная социальная теория необходимо должна быть вычислительной и видеть свой предмет как результат эволюции нейроцеребральных вычислений и «коммуникативном функционализме». Развитие когнитивных наук и искусственного интеллекта в последние десятилетия свидетельствует в пользу распределённой (сетевой) архитектуры когнитивных систем, которая реализует вероятностные алгоритмы для поиска ассоциативных связей. Эволюция, идущая в направлении повышения эффективности вычислений при снижении их энергозатратности, приводит к объединению когнитивных возможностей живых организмов, что выражается в появлении их социальных организаций, также воплощающих сетевую архитектуру. Интерфейсом для такого объединения выступают сигнальные системы,

позже — символические системы (языки). Последние в какой-то степени перестраивают познание, основанное на нём социальное поведение в направлении от ассоциативно-вероятностного к вербально-категориальному (рациональному). Таким образом, во-первых, когнитивные способности, выработанные в ходе эволюции, имеют ассоциативно-вероятностный характер, в то время как вербально-категориальное мышление как концептуальное основание классической философии и науки ограничено сферой коммуникации и предназначено для обеспечения сетевых взаимодействий когнитивных агентов. Во-вторых, сама социальная метасеть возникает как расширение когнитивных возможностей живых организмов за счёт распределения вычислительных задач между ними.

К содержанию работы уместен ряд вопросов и замечаний.

Во-первых, не совсем понятна роль т. н. «рабочей гипотезы» диссертационной работы. Она сформулирована во Введении, с. 10-11: *«если когниция (это по сути совокупность субъективных ментальных процессов, исходя из глоссария стр. 360) может быть понята как вычислительный процесс той или иной архитектуры, то, учитывая некоторую общность методов когнитивных и социальных дисциплин, общественная жизнь и социальные взаимодействия также могут быть объяснены как вычислительные процессы в самом широком смысле слова. Тогда язык и — шире — символические системы, будучи важной формой социализации индивидов, могут быть поняты как своего рода программный интерфейс между когнитивными и социальными процессами».*

Непонятно, что подразумевается под понятием «архитектура вычислительного процесса» и нет ли возможности привести некоторые примеры «архитектур вычислительных процессов»? Так же просьба уточнить квантификатор «некоторая общность» методов когнитивных и социальных дисциплин. Какого рода «общность» возможна между субъективными и интерсубъективными феноменами? Наконец, следует усомниться в логической целесообразности «рабочей гипотезы». В ней четко

прослеживается три компьютерные метафоры. Две метафоры связаны с «вычислимостью» ментальных и социальных феноменов, а третья метафора — со взаимодействием этих метафор посредством своеобразного API, программного интерфейса приложений, посредством которых одна компьютерная программа (программой реализации ментального) может взаимодействовать с другой программой (программой реализации социального). Однако метафоры не стоит доказывать, а именно на доказательство «рабочей гипотезы» направлено все диссертационное исследование: риторические фигуры призваны убеждать в истине или лжи, но не выражать собственно логические значения.

Во-вторых, автор иногда бывает чересчур категоричным, выделяя и подробно раскрывая содержание противоречий, не замечая возможной контрарности утверждений. Например, при изучении философии сознания он выделяет две исследовательские установки — феноменологическую, в «которой принимающее свободные решения «я» есть непосредственная данность» (с.58) и редукционистскую, в которой «ментальные состояния, как более сложные, являют собой функциональные или эмерджентные свойства не-ментальных элементов» (там же). Но, помимо этих полярных направлений, существует и установки антиредукционистского содержания. Они не замыкаются во внутреннем мире интроспективно схватываемых очевидностей сознания, но и не отдают свое приватное переживание обезличенным отчетам от третьего лица. Например, к ним относятся проекты 2D- и 3D-дименсионализма, которые выражают «смысловое» единство суждений о приватных ментальных феноменах с научными и компьютерно-программными суждениями. Собственно, машинный функционализм Х. Патнэма изначально и ориентировался на такую антиредукционистскую установку: в работе 1960 г. «Minds and machines» он, решая проблемы субъективного и интерсубъективного переживания боли (то есть «когнитивные» и «социальные» проблемы) предлагал использовать машину Тьюринга в роли функционального инварианта реального мозгового

процесса возбуждения «болевых нейронов», с одной стороны, и поведенческой диспозиции боли, с другой стороны. К сожалению, И.Ф. Михайлов проигнорировал первоисточники машинного функционализма и в своих исследованиях оснований компьютеризма сослался на весьма незначительную работу (с. 48), в которой ошибочно указывается, что машинный функционализм — это программа бихевиоризма, впервые предложенная в работе Х. Патнэма 1967 г. Это не совсем верно. Это третье замечание.

Наконец, в-четвертых, и с сожалением, следует отметить, что в коннекционизме И.Ф. Михайлова не нашлось места для машины Корсакова. Машина Корсакова — это первый параллельный компьютер, принцип действия которого основан на связях (коннекциях) между ассоциативно связанными «идеями», формально представленных на перфокартах, накладываемых друг на друга, в отверстия которых проваливаются иголки при наличии общности признаков. Машина Корсакова — это чистый коннекционистский компьютер, в то время как машина Тьюринга является чистым символьным компьютером. Машина Корсакова была предложена в 1832 г. С.Н. Корсаковым. В свою очередь, прообразом машины Тьюринга является машина Ч. Бэббиджа, первую программу к которой составила А. Лавлейс в 1843 г. Коннекционистскую (параллельную распределенную вычислимость) последовательными машинами Тьюринга не реализовать. К сожалению, концепция вычислимости И.Ф. Михайлова при рассмотрении вопросов коннекционизма приобрела редукционистский характер, сводимый к работе последовательной машины Тьюринга.

Еще одно небольшое замечание. На с. 126 - 127 автор не совсем корректно полагает, что «вычислимость» и «алгоритмичность» синонимичны по причине того, что *«формальное описание алгоритмического процесса было дано в знаменитой статье Алана Тьюринга, посвящённой математической проблеме вычислимых чисел. Поэтому «алгоритм» прочно ассоциируется с «вычислением» в математическом смысле, хотя, например,*

проезд перекрёстка на автомобиле в соответствии с правилами дорожного движения является алгоритмическим действием, несмотря на то что никакие математические объекты в ходе его не вычисляются» (с. 127). Это не совсем корректно. Дело в том, что программа абстрактно представима целочисленным рядом шагов ее выполнения. «Вычислимость» понимается именно в этом смысле, в смысле преобразований пронумерованных шагов ее выполнения, а не в смысле вычисления каких-то, скажем, дифференциальных, то есть явно выраженных, «числовых» функций.

Эти незначительные замечания ни в коей мере не снижают высокой оценки диссертации. Работа богата эвристическим материалом философского и научного содержания. Она чрезвычайно полезна философам, ученым и инженерам, вовлеченным в междисциплинарные исследования искусственного интеллекта. Автор демонстрирует эрудированность и креативность. Диссертация актуальна и важна; научные положения, выносимые на защиту, являются обоснованными и новыми; заключения и выводы аргументированы, рекомендации значимы и затрагивают сферы философии науки, когнитивных наук и социальных наук. Основные результаты исследования в полной мере отражены в 55 публикациях автора (33 увидели свет в изданиях, рекомендованных ВАК, 15 — в журналах, индексируемых в базах Web of Science и Scopus). Автореферат адекватно отражает содержание диссертации.


В целом, можно констатировать, что диссертационная работа Михайлова Игоря Феликсовича «Когнитивные основания социальности» является самостоятельным оригинальным исследованием актуальной проблемы с достаточной доказательной базой и собственной концепцией автора. Выполненное И.Ф. Михайловым диссертационное исследование является завершённой научно-квалификационной работой, характеризуется научной новизной, теоретической и практической значимостью, полностью соответствует требованиям пп. 9-11, 13, 14 Положения о присуждении ученых степеней, утверждённого Постановлением Правительства Российской

Федерации от 24 сентября 2013 г. № 842¹, предъявляемым к диссертациям на соискание ученой степени доктора философских наук по научной специальности 09.00.01 – Онтология и теория познания.

Автор диссертации «Когнитивные основания социальности» Михайлов Игорь Феликсович заслуживает присуждения ему искомой ученой степени доктора философских наук по научной специальности 09.00.01 – Онтология и теория познания.

Официальный оппонент

доктор философских наук
(09.00.08 – Философия науки и техники),
ведущий научный сотрудник
кафедры философии и методологии науки
философского факультета
Московского государственного
университета имени М.В. Ломоносова



Алексеев
Андрей Юрьевич

10 сентября 2021 года

Подпись ведущего научного сотрудника Алексеева А.Ю. заверяю:

И.о. декана философского факультета
МГУ имени М.В. Ломоносова



А.П. Козырев

ФГБОУ ВО «Московский государственный университет имени М.В.Ломоносова», философский факультет (119991, Москва, Ленинские горы, МГУ учебно-научный корпус «Шуваловский», тел.: + 7 (495) 939-19-25, e-mail: info@philos.msu.ru, сайт: www.philos.msu.ru)

С перечнем публикаций официального оппонента Алексеева Андрея Юрьевича можно ознакомиться на следующих Интернет-ресурсах:

- https://www.elibrary.ru/author_items.asp?authorid=76136
- <https://istina.msu.ru/profile/AlekseevAY/>.

¹ С изменениями, внесенными Постановлениями Правительства РФ от 30.07.2014 № 723, от 21.04.2016 № 335, от 02.08.2016 № 748, от 29.05.2017 № 650, от 28.08.2017 № 1024, от 01.10.2018 № 1168, от 20.03.2021 № 426, Решением Верховного Суда РФ от 21.04.2014 N АКПИ14-115, Постановлением Правительства РФ от 26.05.2020 № 751.