

П.Н. Барышников

«ЯЗЫКОВОЙ БАРЬЕР» В ТЕОРИЯХ СОЗНАНИЯ И ОГРАНИЧЕНИЯ ВЫЧИСЛИТЕЛЬНОГО ПОДХОДА*

Барышников Павел Николаевич – доктор философских наук, доцент, профессор кафедры исторических и социально-философских дисциплин, востоковедения и теологии. ФГБОУ ВО «Пятигорский государственный университет». Российская Федерация, 357532, г. Пятигорск, пр. Калинина, д. 9; e-mail: pnbaryshnikov@pgu.ru

В данной работе в особом ракурсе рассматривается проблема методологических ограничений вычислительного подхода в философии сознания и эмпирических науках. Основная цель – последовательно обосновать зависимость философских «метафорических словарей» от успехов в области компьютерных наук, исторических контекстов теории познания, формальных и методологических ограничений алгоритмических и вычислительных процедур. Ключевая идея состоит в том, что, несмотря на успехи вычислительных моделей в эмпирических исследованиях, их концептуальный уровень не позволяет корректно сформулировать вопрос об онтологии сознания. Компьютериализм в философских теориях сознания представлен как практика словоупотребления, которая помещает системы непротиворечивых описаний информационных свойств сознания и когнитивных процессов в рамки тех или иных методологических правил. Ограничения вычислительного подхода связываются с отсутствием научной теории субъективности, принципиальной несводимостью внешних свойств сознания к внутренним состояниям, фундаментальными ограничениями на полноту и непротиворечивость вычислительной математики. Проведена классификация нескольких антикомпьютериалистских программ. В статье обсуждаются различные ограничения вычислительных подходов, такие как сложность формализации нелинейных и скрытых процессов, зависимость эмоций и интуиции от ситуационного контекста и индивидуальных различий, сложность прогнозирования возникающих свойств автономных систем, ограничения, связанные с нестандартными вычислительными процессами квантовых и динамических явлений. Особая роль отводится проблеме теоретико-множественного редуктивного реализма в рамках математического структурализма и некоторым аспектам р-адических числовых систем как возможной альтернативе «универсального языка» для вычислительной философии сознания.

Ключевые слова: вычислительный подход, философия сознания, научная метафора, «разрыв в объяснении», эмпирические теории сознания, методологические ограничения

* Исследование выполнено за счет гранта Российского научного фонда № 24-28-00540, <https://rscf.ru/project/24-28-00540/>.

Для цитирования: Барышников П.Н. «Языковой барьер» в теориях сознания и ограничения вычислительного подхода // Философский журнал / Philosophy Journal. 2024. Т. 17. № 2. С. 122–136.

1. Введение

Перед изложением основных рассуждений необходимо дать введение в используемый в данной статье терминологический аппарат и описать предметное поле. Прежде всего, стоит указать на то, что наработки последних лет по проблеме компьютеризации в философских теориях сознания в качестве объекта исследования полагают не столько рациональные свойства ума (*mind*), феноменальное сознание (*consciousness*) или когнитивные процессы, сколько системы непротиворечивых описаний перечисленных объектов в рамках тех или иных методологических правил. Можно утверждать, что современная западная философия сознания – это своеобразная академическая культура анализа и толкования философских и научных текстов и прояснения значений языковых выражений. Вопрос о том, чем является сознание, по значимости сопоставим с вопросом о структуре концептуальной системы, правила которой позволяют непротиворечиво выразить значение предложения «Сознание является X». Предлагаемые рассуждения – это рассуждения о вычислительном подходе как своеобразном «универсальном словаре», состав которого позволяет формулировать высказывания о сознании. Поэтому сопоставление линий аргументов, описание методологических возможностей и ограничений компьютеризации указывают не на свойства объекта теории, а на свойства самой теории как объяснительной системы. Вычислительный подход здесь понимается как способ употребления слов, на основании «словаря», в статьях которого значения знаков объясняются при помощи других знаков.

Чем является процесс вычисления с точки зрения компьютерных наук? В качестве исходного утверждения возьмем определение из работы С. Вольфрама: «Вычисление – это операция, которая начинается с некоторых начальных условий и дает результат, следующий из определенного набора правил. <...> Набор правил, используемых для выполнения вычислений, известен как алгоритм»¹. В основе вычислительного подхода лежат представления о формах механической рациональности, реализуемой информационными процессами: каждое действие агента (и состояние типа убеждений, мышления, эмоций, мотивов, желаний, феноменального опыта) имеет под собой ментальную причину, вызванную, в свою очередь, алгоритмической (эволюционной) обработкой статистических данных для достижения адаптивных целей. Из приведенного определения понятия вычисления следует, что процедура вычисления – это изменение состояний вычислителя по определенным правилам. Изменение состояний за счет особенностей архитектуры вычислителя преобразует структуру начальных условий. Правила преобразования, примененные к структуре начальных условий, определяют результаты вывода (структуру конечных условий). Собственно, задача эмпирических исследований сознания – найти корреляцию между функциональными свойствами преобразователя и структурой входящих и исходящих данных.

¹ Wolfram S. A new kind of science. Champaign (IL), 2019. P. 367.

2. Классификации ограничений вычислительного подхода

Работа по систематизации слабых сторон вычислительных подходов проделана не одним десятком авторов. Антикомпьютериализм, как правило, укрепляет свои позиции не за счет собственных творческих прорывов, а за счет поиска слабых мест и скрупулезного анализа положений компьютеризации и многообразных его форм.

Если широко резюмировать типичные ответы антикомпьютериалистов, сформулированные в период второй половины XX в., то можно привести следующий список:

- ветви аргумента «Китайская комната» Дж. Сёрла;
- аргумент от тривиальности;
- аргумент от Гёделява предложения;
- аргумент от ограниченности компьютерного моделирования;
- аргумент от темпоральности;
- аргумент от телесности².

Подробно каждый из аргументов рассмотрен в книге «Вычислительные модели разума: от кода к смыслу»³. Однако голоса антикомпьютериалистов сегодня стихают (или по крайней мере привлекают к себе не так много внимания) на фоне успехов нейросетевых моделей (в частности, предобученных генеративных трансформеров). Отметим также, что с 2017 г. статья М. Рескорлы «Вычислительная теория сознания» в Стэнфордской энциклопедии не редактировалась. Значит ли это, что коннекционистский метод «грубой силы» и «нейросетевая революция» опровергли все аргументы, которые так тщательно оттачивались в течение полувека? Полагаем, что тактика поиска слабых сторон компьютеризации себя не изжила.

М. Милковски предлагает свою классификацию объяснительных ограничений вычислительного подхода⁴. Эта классификация детализирует «Аргумент от ограниченности компьютерного моделирования». Каждый из пунктов нуждается в комментарии.

1. *Ограничения при объяснении нелинейных и скрытых процессов, таких как эмоции и интуиция.* Этот тип ограничений связан с тем, что эмоции и интуиция очень зависят от ситуативного контекста и индивидуальных различий, что затрудняет формализацию. Кроме того, они связаны с эмпирически нерегистрируемыми бессознательными процессами. Эмоции тесно связаны с телесными реакциями. Тело в этом случае может выступать как источник данных для обучения моделей. Но для строгих корреляций необходимо установить строгую казуальную зависимость эмоциональных реакций (от репрезентаций или от среды), что пока вызывает затруднения⁵.
2. *Ограничения при объяснении процессов, связанных с автономией и аутопоэзисом.* Вычислительные модели плохо предсказывают зарождение эмерджентных свойств автономных и аутопоэтических

² Rescorla M. The Computational Theory of Mind // The Stanford Encyclopedia of Philosophy. 2017. URL: <https://plato.stanford.edu/archives/spr2017/entries/computational-mind> (дата обращения: 21.01.2023).

³ Барышников П.Н. Вычислительные модели разума: от кода к смыслу. М., 2022.

⁴ Milkowski M. Explaining the Computational Mind. Cambridge (Mass.), 2013.

⁵ Bosse T., Jonker C.M., Treur J. Formalisation of Damasio's theory of emotion, feeling and core consciousness // Consciousness and Cognition. 2008. Vol. 17. No. 1. P. 94–113.

систем, таких как самоорганизация и эволюция структуры. При больших объемах данных и сложности вычислений в самих вычислительных процедурах могут зародиться непредсказуемые эмерджентные свойства. Наиболее яркий пример – при увеличении количества параметров особенное поведение LLM (LLM – Large Language Models (большие языковые модели)) при взаимодействии с человеком⁶. Сегодня нет модели, которая вычислила бы точное «место перехода» количества данных и параметров в качество ответов.

3. *Проблема рамки (frame problem)* заключается в том, что вычислительные модели не могут учитывать все возможные взаимодействия между объектами в окружающей среде. Чтобы предсказать и задать новое состояние системы, необходимо, чтобы модель «на вход» получила данные обо всех возможных состояниях окружающей среды. Масштабирование здесь уходит в бесконечность, и не существует универсального алгоритма, который определял бы достаточные параметры точности для каждой задачи.
4. *Ограничения, связанные с нестандартными вычислительными процессами, такими как квантовые вычисления или моделирование динамических систем.* Этот тип ограничений – результат сложности задач по непротиворечивому описанию квантовых процессов или нелинейных динамических процессов. Универсальный квантовый компьютер сегодня не сконструирован. В инженерном поле квантовый компьютинг осуществляется за счет сложного распараллеливания процессов на классической архитектуре. В философии сознания эти ограничения послужили причиной долгого спора о редукции волновой функции и роли сознания наблюдателя⁷.
5. *Проблема формализации и представления в вычислительных системах чувства «здорового смысла».* Человеческая интеллектуальная вовлеченность в собственную деятельность и во взаимодействие с остальными людьми составляет основу для так называемого практического разума. Практический разум способен к построению нормативных суждений и инференциальных связей на основе опыта при минимальном наборе данных. Для систем ИИ это пока недостижимый уровень. Бытие существа и бытие вещи разнятся по своим онтологическим основаниям. Отождествить или максимально приблизить эти основания – нетривиальная инженерная задача. Несмотря на колоссальные объемы обучающих данных, сегодня не существует универсального вычислительного представления «опыта жизни» или «здорового смысла»⁸. Приходится прибегать к методу ручной разметки элайнмента (alignement – выравнивание модели по усредненным ценностным и мировоззренческим установкам), чтобы встроить «поведение» машин в контексты человеческой жизни.

⁶ Bubeck S., Chandrasekaran V., Eldan R., Gehrke J., Horvitz E., Kamar E., Lee P., Lee Y.T., Li Y., Lundberg S., Nori H., Palangi H., Ribeiro M.T., Zhang Y. Sparks of Artificial General Intelligence: Early experiments with GPT-4. 2023. DOI: 10.48550/arXiv.2303.12712.

⁷ Сторожук А.Ю. Философские основания моделирования квантовых компьютеров // Сибирский философский журнал. 2019. Т. 17. № 4. С. 85–92.

⁸ Hasselberger W., Lott M. Where lies the grail? AI, common sense, and human practical intelligence // Phenomenology and the Cognitive Sciences. 2023. DOI: 10.1007/s11097-023-09942-x.

Резюмируя изложенное выше, можно утверждать, что все ограничения вычислительных подходов в основании имеют уровни сложности объектов (на которые они направлены), которые всегда будут сложнее, чем их вычислительные модели, т.к. их сложность увеличивается при применениях к ним вычислительных моделей. Иными словами, *вычисление есть ограничение*.

Тем не менее все меньше остается областей когнитивной и интеллектуальной деятельности, которые невозможно было бы представить в виде компьютерной модели. Сегодня вопрос в стиле Х. Дрейфуса «Чего не могут вычислительные машины?» звучит старомодно, с учетом того, что данные для обучения машин формируются на основании информации, порождаемой интеллектуальной и когнитивной деятельностью человека. То есть не будет большим преувеличением, если сказать на уровне обыденного языка: «Чего не могут машины (в когнитивном или интеллектуальном поле), того не может и человек, но не наоборот».

Здесь возникает парадоксальная ситуация, когда революционное развитие глубокого машинного обучения и повсеместное внедрение нейросетевых моделей приводит к так называемой «consciousness winter»⁹ – стагнации в развитии научных теорий сознания. С одной стороны, как отмечает И.Ф. Михайлов, вычислительный подход – «оптимальное методологическое решение проблемы сознания, предполагающее описание предмета как комбинации вычислительных примитивов, которые позволяют реализовывать алгоритмы, порождающие сознательные состояния»¹⁰. С другой стороны, вычислительные методы, претендуя на универсализм, «наследуют» трудные проблемы той предметной области, в которой применяются. В связи с чем на основании некоторых работ строятся не вполне логичные высказывания типа: «Компьютеризация не универсальна, потому что мы не знаем, что такое коллапс квантовой волновой функции»¹¹, или «Компьютеризация не универсальна, потому что мы не знаем, что такое математика»¹², или «Компьютеризация ложна, потому что мы не знаем, вычисляет ли мозг» и т.п.

Например, если сознание определять как функцию головного мозга, то необходимо примирить объекты нейробиологии с методами компьютерных наук. Прежде чем рассуждать о роли вычислительного подхода, необходимо выбрать «методологический домен», чтобы увидеть ограничения только одной предметной области. Как только мы утверждаем, что сознание – продукт живой материи, все затруднения, связанные с вычислительными методами в биологии, присовокупляются к философским критическим программам компьютеризации. Д. Брэй подробно описывает ограничения вычислительных методов в биологических исследованиях (речь идет не о высших мозговых функциях, а лишь о хемотаксисе¹³ кишечной палочки *Escherichia coli*):

⁹ Hoel E. The risk of another consciousness winter // The Intrinsic Perspective. 2023. URL: <https://www.theintrinsicperspective.com/p/the-risk-of-another-consciousness> (дата обращения: 21.04.2023).

¹⁰ Михайлов И.Ф. Предметы и методы эмпирических исследований сознания // Философский журнал / Philosophy Journal. 2024. Т. 17. № 2. С. 92–109.

¹¹ Penrose R. Shadows of the mind: a search for the missing science of consciousness. Oxford; New York, 1994.

¹² Целищев В.В. Априорные структуры как представление знания. Новосибирск, 2013.

¹³ Хемотаксис – двигательная реакция микроорганизмов по направлению увеличения концентрации питательных веществ (аттрактантов) и уменьшения концентрации токсинов (репеллентов), простейшая реакция на внешнюю информацию.

Существует определенный уровень молекулярной неопределенности, особенно тот, который отвечает за тонкую настройку и адаптацию ко множеству внешних условий, который мы в настоящее время не можем разрешить или воспроизвести на компьютере. Даже для этого относительно простого процесса мы не можем перечислить все гены, влияющие на его работу. Мы также не можем сказать для каждого мыслимого условия окружающей среды, как регулируются эти гены. У нас мало надежды выявить все комбинации модификаций белков, присутствующие в клетке в любой момент времени¹⁴.

Помимо того что отношения организма и среды порождают неперечислимое количество параметров, сами вычислительные методы накладывают методологические ограничения на исследуемые материальные объекты и их статистическое представление. Отвечая с помощью эффективных компьютерно-аналитических методов на вопрос о том, как работает (или устроен) мозг, мы не можем сформировать естественнонаучный словарь, чтобы сформулировать вопрос «Что есть сознание?». Если придерживаться мнения, что мозг не является необходимым условием существования сознания, то вычислительную методологию будет не к чему применить. Одновременно сильная и слабая сторона компьютеризации – прикладная направленность. Чтобы вычислять, нужно что-то, на что можно указать пальцем (даже в платоновском мире): необходимы операциональные правила, грамматики, структуры данных, алфавиты, множества, переменные или геометрия переносимых колес, логика полупроводника, костяшки, камни и т.п.

В недавней работе Э. Хоэла перечислены границы применимости научных методов, которые в особом качестве проявляются при попытках построить строгое научное знание о сознании. К таким ограничениям автор относит:

- 1) отсутствие научной теории субъективности;
- 2) принципиальную несводимость внешних свойств сознания (поведение, язык, нейронные корреляции, телесность, социально-коммуникативный слой и т.п.) к внутренним состояниям;
- 3) фундаментальные ограничения на полноту и строгость универсального языка науки – математики (стоит добавить, что к проблеме соотношения дескрипции и действия присоединяется проблема реализации математических сущностей в архитектуре цифровых вычислителей¹⁵);
- 4) проблему теоретико-множественного редуктивного реализма¹⁶.

Последняя проблема заслуживает отдельного рассмотрения, т.к. ее логическая структура и свойства «словаря» коррелируют с проблемой множественной реализации в философии сознания. (Очень сложным выглядит вопрос, почему проблемы основания математики так схожи с фундаментальными проблемами философии сознания.) Здесь же кратко опишем проблему и укажем на важные для дальнейших рассуждений ее структурные свойства. «Существовать – значит быть свободным от противоречия»¹⁷, –

¹⁴ *Bray D.* Limits of computational biology // *In Silico Biology*. 2015. Vol. 12. No. 1 (2). P. 1–7.

¹⁵ *Ющенко Р.* Что нужно знать про арифметику с плавающей запятой. 2011. URL: <https://habr.com/ru/articles/112953> (дата обращения: 21.01.2023).

¹⁶ *Hoel E.* The world behind the world: consciousness, free will, and the limits of science. New York; London, 2023. P. 120–133.

¹⁷ *Bosse T., Jonker C.M., Treur J.* Op. cit.

этот математический лозунг дает надежду компьютероналистам-платонистам на то, что, если подобрать непротиворечивый язык описания для феноменального сознания и реализовать его через вычислимую функцию, то «проблема сознания» исчезнет сама собой, т.к. сформируется «полная дескрипция» свойств. Если мы сможем вычислительными методами доказать существование неоткрытых математических истин, то они (истины) существуют, но при этом не формулируются на языке доказательств. То есть компьютеры многое могут нам дать при описании свойств сознания, но не могут доказать существование сознания. В этом случае сознание существует, но необъяснимо во всей своей полноте компьютероналистскими методами. Хоэл выражает последнюю мысль (со ссылкой на математический структурализм в духе П. Бенацерафа) следующим образом: «...умы могут быть неалгоритмичными, и это указывает на то, что трудности вокруг теории сознания являются примером неполноты науки как способа познания»¹⁸. Даже если сознание – это расширение конечного автомата и оно действительно является лишь определенной моделью машины Тьюринга, то в этом сознании не будет ни одного состояния вычислителя, указывающего на свойства самого сознания.

Характерно, что П. Бенацераф отвергает возможность существования абстрактных объектов математики (платонического толка), указывая на то, что онтология формальной системы должна быть связана с непротиворечивостью самой системы, а не со значением ее языковых выражений. Структуралисты указывают на противоречивость теоретико-множественного редуцированного реализма, т.к. ни один математический объект не представляет собой уникальное и строго определенное множество. Цель Бенацерафа состоит в том, чтобы показать возможность приложений математики без утверждений о существовании абстрактных объектов. Можно предположить, что структуралистский подход в математике подсказывает решение для методологических затруднений при изучении сознания инструментами современных эмпирических наук. Решение это таково: возможны эвристические приложения для компьютеронализма без необходимости строить онтологию сознания. То есть можно познавать свойства объекта без необходимости создавать язык для обоснования существования объекта. В структуралистской модели оснований математики числа не способны обладать индивидуальностью, т.к., будучи представленными знаками цифр (любой из возможных систем счисления с любым алфавитом), зависят от правил, предписываемых системой, для управления и различения знаков.

Немаловажным является и то, что помимо выбора правил управления знаками играют роль физические принципы работы вычислителя. Можно привести пример из области компьютерной математики (точнее, арифметики с типом float):

```
>>> 0.1 + 0.1 + 0.1 == 0.3
# False
>>> 0.1 + 0.1 + 0.1 == 0.30000000000000004
# True
```

¹⁸ Hoel E. The risk of another consciousness winter // The Intrinsic Perspective. 2023. URL: <https://www.theintrinsicperspective.com/p/the-risk-of-another-consciousness> (дата обращения: 21.04.2023).

На современных машинах числа с плавающей запятой отображают с использованием двоичной дроби: числитель использует первые 53 бита, начиная с самого старшего бита, а знаменатель – как степень двух:

$$\ggg 0.1 \approx 3602879701896397/2^{**} 55$$

Проблема представления вещественных чисел за ограниченное время в конечной памяти компьютера решалась еще в 1970-е гг. XX столетия, но, на наш взгляд, не потеряла своего философского содержания. Континуум вещественных чисел приходится представлять сначала в нормализованном экспоненциальном виде посредством переменного типа float (число одинарной точности, занимающее 32 бита в памяти компьютера) и потом в итоге на конечном наборе нулей и единиц, выбирая между точностью и полнотой диапазона принимаемых значений. Сегодня как аппаратные, так и программные способы представления операций над числами с плавающей точкой почти совершенны. Однако это не ликвидирует полностью такие аномалии, как некоммутативность и неассоциативность арифметических операций, ноль со знаком, разность неравных чисел, дающая ноль, и прочее¹⁹.

3. Универсальные свойства компьютеризации и «языковой барьер»

Возвращаясь к идеям Бенацераффа, можно утверждать, что онтология чисел в основаниях математики в чем-то схожа с проблемой онтологии сознания. Существует множество систем, предписывающих правила описания и указания на корреляты сознания (компьютеризация – одна из эффективных систем). Но ни одна из этих систем не претендует на формирование высказываний, способных превратить квантор существования в предикат.

...если мы утверждаем «Существование является предикатом», используя «предикат» в грамматическом смысле, то наша пропозиция – это пропозиция об определенных словах в том смысле, что они часто используются определенным способом, но, как это ни странно, это – пропозиция не о самом слове «существование»²⁰.

То есть компьютеризация в современной философии сознания и философии когнитивных наук, обладая разнообразием направлений и подходов, – это лишь способ употребления слов. Несмотря на общий «зонтичный» термин «вычислительный» (computational), пока невозможно говорить не только о единообразии исследовательской доктрины, но и о единой онтологии исследовательских объектов, находящихся в фокусе данного направления. Причин такому положению вещей существует множество. Но все эти причины прямо или косвенно связаны с проблемой непроясненного значения языковых выражений. Перечислим некоторые из них:

- 1) зависимость философских «метафорических словарей» от успехов в области компьютерных наук;

¹⁹ Юценко Р. Что нужно знать про арифметику с плавающей запятой. 2011. URL: <https://habr.com/ru/articles/112953> (дата обращения: 21.01.2023).

²⁰ Мур Дж.Э. Является ли существование предикатом? // Вестник Томского государственного университета. Философия. Социология. Политология. 2009. Т. 8. № 4. С. 111.

- 2) зависимость философских «метафорических словарей» от исторических контекстов теории познания;
- 3) существование формальных и методологических ограничений в вычислительных архитектурах;
- 4) концептуальное замещение философских категорий понятиями из эмпирических исследований.

Основной наш тезис сводится к тому, что, несмотря на успехи вычислительных моделей в эмпирических исследованиях, их концептуальный уровень не позволяет корректно сформулировать вопрос об онтологии сознания. Попытки обойти этот «языковой барьер» приводят к тому, что в язык философии проникают технические гипотезы из инженерных прикладных задач. В настоящее время цифровые данные и компьютерные алгоритмы стали своеобразным универсальным языком науки вместе с тем материалом, из которого создан непреодолимый «языковой барьер» для философии.

Сегодняшний энтузиазм и надежды, связанные с революционным ростом мощности вычислительных систем, объема обрабатываемых данных и сложности алгоритмов, напоминают о схожих настроениях в 1940–1950-е гг. XX в., когда абстрактные правила работы конечных автоматов были приняты как модель функционирования мозга. Идея о существовании универсальных правил, описываемых универсальным алгоритмом, наконец, нашла свое технологическое воплощение. Упомянутый выше «Принцип наследственности», согласно которому различные версии компьютеризации наследуют слабые стороны подходов из других областей, проявился и в ранних версиях компьютеризации. В качестве примера можно привести пошаговую «биологическую корректировку», внесенную в машинный функционализм Х. Патнэмом, когда стало очевидно, что состояния абстрактной машины Тьюринга не соответствуют значению вычисляемых функций. То есть ментальные виды как «композиционно», так и «вычислительно» пластичны так, что невозможно обнаружить корреляции между пропозициональными отношениями и свойствами мозга²¹. Шаг первый известен как аргумент множественной реализуемости. Шаг второй указывал на то, что нет никаких шансов обнаружить психофизические эквиваленции вне интенционального содержания. О. Шагрин позже аргументированно указал на то, что все попытки реабилитировать машинный функционализм сводились либо к видоизмененным формам бихевиоризма, либо к физикализму²².

Приходится признать, что вся эволюция «Вычислительной теории сознания» представляет собой перевод концептуальной системы одной теории в концептуальную систему другой теории. Так, попытки лишить машинный функционализм его врожденных противоречий приводят к тому, что это направление теряет свою обособленность и принимает иные формы (бихевиоризм или физикализм), наследуя их слабые стороны.

Собственно, такое положение вещей мы будем получать при любой попытке «высказаться о сознании» при помощи «словарей» эмпирических наук. К.В. Анохин в работе «Когнитом: в поисках фундаментальной нейронаучной теории сознания», классифицируя большинство научных и философских подходов и разрабатывая методологические требования, выдвигаемые

²¹ Putnam H. Representation and reality. Cambridge (Mass.), 2011.

²² Shagrir O. The Rise and Fall of Computational Functionalism // Hilary Putnam. Cambridge, 2005. P. 220–251.

к эмпирическим объяснениям сознания, указывает на так называемую циркулярную ловушку теории сознания. В результате автор получает «замкнутый круг» из пяти суждений, начинающихся со слова «невозможно»²³. Суть «циркулярной ловушки» состоит в том, что для понимания субъективного опыта необходимо раскрыть принципы развития нервной системы в онтогенезе и принципы эволюционных изменений в филогенезе. Но при этом крайне важна роль субъективного опыта в развитии закономерностей эволюции нервной системы в филогенезе. Все существующие теории сознания (как правило, страдающие от дефицита объяснений природы носителя субъективных состояний, «кто» – сознания) можно подразделить на два класса: 1) корреляционные – ведущие поиск нервных коррелятов сознания «на языке анатомии и физиологии мозга»; 2) эпифеноменалистские – ведущие «поиск специфических процессов в высокоуровневой организации мозга, которые составляют фундаментальную нередуцируемую сущность когнитивных систем»²⁴.

Собственно, итоговый вывод К.В. Анохина, указывающий на путь ликвидации «разрыва в объяснении», необходимо понимать как попытку преодоления «языкового барьера». «Словарь» традиционной анатомии и нейрофизиологии накладывает ограничения на возможность переноса представлений о мозге на антиредукционистский уровень представлений о разуме. В статье дается пошаговый рецепт перехода от коннектома (физиологических процессов нейронной сети) к когнитому (когнитивным процессам нейронной гиперсети). Пока без ответа остается вопрос о том, какая структура какой формально-аксиоматической системы способна непротиворечиво, полно и строго представить «мозг в его максимально нередуцируемой сущности как особый когнитивный орган, обладающий особыми связями с когнитивным агентом (mind), реализованными через функционирование нейронной гиперсети»²⁵.

4. Универсальный язык для теории сознания и пределы выразимости

Как уже упоминалось, вопросы об онтологии математических структур и объектов необъяснимо напоминают вопросы об онтологии сознания. Говоря об объективном существовании содержания понятия числа или множества, мы проваливаемся в сложнейшие вопросы об основаниях математики. Известно, что иерархия чисел неполна. Она расширяется процедурой Кэли – Диксона сколь угодно много раз. Эта процедура позволяет построить из действительных чисел последовательно их расширения: комплексные числа, кватернионы, октонионы, седенионы и т.д. В последних параграфах работы «Почему вообще существует философия математики?» Я. Хакинг раскрывает свой взгляд на неразрешимые вечные вопросы математики, косвенно указывая на свою очарованность тайной соотношения математики и природы. Философия математики, как и философия сознания, обращена

²³ Анохин К.В. Когнитом: в поисках фундаментальной нейронаучной теории сознания // Журнал высшей нервной деятельности им. И.П. Павлова. 2021. Т. 71. № 1. С. 52.

²⁴ Там же. С. 63.

²⁵ Там же. С. 66.

к зыбкой границе между созерцанием и артикуляцией²⁶. При этом артикуляция не может содержать в себе автореферентную дескрипцию.

При попытках разорвать границы «циркулярной ловушки», «трудной проблемы» или преодолеть «разрыв в объяснении» и при этом остаться в поле эмпирических наук о сознании неизбежно возникает запрос на изобретение новой концептуальной системы. Вычислительный подход при всех его недостатках обеспечивает строгость, непротиворечивость и «методологическую разрешимость». То есть состояния вычислителя отобразят то, под отображение чего они спроектированы. Идеи перехода от вещественного континуума к особому классу p -адических числовых полей, обладающих рядом парадоксальных свойств и пригодных для описания ментальных процессов, схожи с идеями перехода от коннектома к когнитому, т.е. к построению новой онтологии сознания²⁷.

p -адические числа представляют собой своеобразные «бесконечноразрядные числа» с целочисленным основанием. Это специфичный способ записи числа в обратном порядке; способ записи, который порождает целый континуум с особенными свойствами. Простым примером p -адического числа является периодическая дробь $1/3$, представленная как бесконечная 10 -адическая дробь: $1/3 = 0.333333333... = 0.(3) = ...66666667 = (6)7_{10}$, где в скобках представлена периодическая часть. Этот способ записи позволяет перевести рациональные числа в целые. По идее, такой подход должен был бы решить проблему ошибок округления при цифровых вычислениях и снять технические ограничения вычислительных подходов. Однако здесь возникает вопрос о применимости цифровых компьютеров при работе с классом объектов, представляющим отдельную алгебраическую систему. С одной стороны, p -адические числа порождают вычислительную устойчивость, которой так не хватает при работе с числами с плавающей точкой в 64 -разрядных системах. С другой стороны, очень необычной выглядит попытка оперировать p -адическими числами методами функционального программирования²⁸.

Не вдаваясь в технические детали p -адической числовой системы, укажем на высказывание А.Ю. Хренникова о том, что поле вещественных чисел (с обобщениями в виде комплексных, интервальных чисел) является не единственно возможным расширением поля рациональных чисел. Разум математика способен отклониться от «естественного» пути и расширить рациональные числа в иное измерение p -адических чисел. Это отклонение убеждает автора в том, что разум способен породить систему символов, описывающих без противоречий и парадоксов сам разум. При попытках описать мир ментальных процессов наборами символов неизбежна постановка вопроса «о смене словарей», о переводе или наложении понятийных систем. Можно ли корреляты ментальных событий искать не на физических событиях декартова пространства R^3 , а на событиях гильбертова пространства H ? Можно ли корреляционные теории сознания развернуть от биологических свойств мозга к объективным свойствам математических структур?²⁹

²⁶ Хакинг Я. Почему вообще существует философия математики? М., 2020.

²⁷ Хренников А.Ю. Моделирование процессов мышления в p -адических системах координат. М., 2004.

²⁸ Самойленко С. Знакомство с p -адическими числами. Часть 2, практическая. 2022. <https://habr.com/ru/articles/646143> (дата обращения: 21.01.2023).

²⁹ Хренников А.Ю. Указ. соч.

Итак, при построении эмпирического объяснения сознания невозможно обойти вопрос о концептуальном аппарате новой универсальной теории. Применение вычислительных дескрипций на уровне физических или физиологических процессов сталкивается с трудностями, но у них есть неоспоримое преимущество: несмотря на неполноту и известные аппроксимации компьютерного анализа данных, когда система выполняет вычисления, эти вычисления объясняют поведение системы. Вычислительное объяснение представляет собой особый род механистического объяснения, которое применимо к системам, манипулирующим символами состояний, при соблюдении соответствующих правил⁵⁰. Вычисление примиряет физические эмпирически проверяемые процессы с математическими структурами. Какие вычислительные процедуры и какая архитектура аппаратных средств и структуры данных понадобятся для моделирования процессов когнитива (или для любого иного универсального эмпирического объяснения сознания) – крайне актуальный вопрос, требующий дополнительных исследований и всесторонней проработки.

Список литературы

- Анохин К.В. Когнитом: в поисках фундаментальной нейронаучной теории сознания // Журнал высшей нервной деятельности им. И.П. Павлова. 2021. Т. 71. № 1. С. 39–71.
- Барышников П.Н. Вычислительные модели разума: от кода к смыслу. М.: URSS, 2022.
- Михайлов И.Ф. Предметы и методы эмпирических исследований сознания // Философский журнал / Philosophy Journal. 2024. Т. 17. № 2. С. 92–109.
- Мур Дж.Э. Является ли существование предикатом? / Пер. англ. В.Е. Мельникова под ред. В.А. Суrowцева // Вестник Томского государственного университета. Философия. Социология. Политология. 2009. Т. 8. № 4. С. 111–122.
- Самойленко С. Знакомство с р-адическими числами. Ч. 2, практическая. 2022. URL: <https://habr.com/ru/articles/646143> (дата обращения: 21.04.2023).
- Сторожук А.Ю. Философские основания моделирования квантовых компьютеров // Сибирский философский журнал. 2019. Т. 17. № 4. С. 85–92.
- Хакинг Я. Почему вообще существует философия математики? / Пер. с англ. В.В. Целищева. М.: Канон+, 2020.
- Хренников А.Ю. Моделирование процессов мышления в р-адических системах координат. М.: ФИЗМАТЛИТ, 2004.
- Целищев В.В. Априорные структуры как представление знания. Новосибирск: Манускрипт – СИАМ, 2013.
- Целищев В.В. Онтология математики: объекты и структуры. Новосибирск: Нонпарель, 2003.
- Юценко Р. Что нужно знать про арифметику с плавающей запятой. 2011. URL: <https://habr.com/ru/articles/112953> (дата обращения: 21.01.2023).
- Bosse T., Jonker C.M., Treur J. Formalisation of Damasio's theory of emotion, feeling and core consciousness // Consciousness and Cognition. 2008. Vol. 17. No. 1. P. 94–113.
- Bray D. Limits of computational biology // In Silico Biology. 2015. Vol. 12. No. 1 (2). P. 1–7.
- Bubeck S., Chandrasekaran V., Eldan R., Gehrke J., Horvitz E., Kamar E., Lee P., Lee Y.T., Li Y., Lundberg S., Nori H., Palangi H., Ribeiro M.T., Zhang Y. Sparks of Artificial General Intelligence: Early experiments with GPT-4. 2023. DOI: 10.48550/arXiv.2303.12712.
- Hasselberger W., Lott M. Where lies the grail? AI, common sense, and human practical intelligence // Phenomenology and the Cognitive Sciences. 2023. DOI: 10.1007/s11097-023-09942-x.

⁵⁰ Piccinini G. Physical computation: a mechanistic account. Oxford, 2015. P. 275.

- Hoel E.* The risk of another consciousness winter // *The Intrinsic Perspective*. 2023. URL: <https://www.theintrinsicperspective.com/p/the-risk-of-another-consciousness> (дата обращения: 21.04.2023).
- Hoel E.* The world behind the world: consciousness, free will, and the limits of science. New York; London: Avid Reader Press, 2023.
- Milkowski M.* Explaining the Computational Mind. Cambridge (Mass.): The MIT Press, 2013.
- Penrose R.* Shadows of the mind: a search for the missing science of consciousness. Oxford; New York: Oxford University Press, 1994.
- Piccinini G.* Physical computation: a mechanistic account. Oxford: Oxford University Press, 2015.
- Putnam H.* Representation and reality. Cambridge (Mass): MIT Press, 2011.
- Rescorla M.* The Computational Theory of Mind // *The Stanford Encyclopedia of Philosophy* / Ed. by E.N. Zalta. 2017. URL: <https://plato.stanford.edu/archives/spr2017/entries/computational-mind> (дата обращения: 21.01.2023).
- Shagrir O.* The Rise and Fall of Computational Functionalism // *Hilary Putnam* / Ed. by Y. Ben-Menahem. Cambridge: Cambridge University Press, 2005. P. 220–251.
- Wolfram S.* A new kind of science. Champaign (IL): Wolfram Media, 2019.

“Language barrier” in theories of mind and limitations of the computational approach *

Pavel N. Baryshnikov

Pyatigorsk State University. 9 Kalinin Av., Pyatigorsk, 357532, Russian Federation; e-mail: pnbaryshnikov@pgu.ru

This paper examines from a special perspective the problem of methodological limitations of the computational approach in the philosophy of mind and empirical sciences. The main goal is to consistently substantiate the dependence of philosophical “metaphorical dictionaries” on advances in the field of computer science, historical contexts of epistemology, formal and methodological limitations of algorithmic and computational procedures. The key idea is that despite the success of computational models in empirical research, their conceptual level does not allow us to correctly formulate the question of the ontology of consciousness. Computationalism in philosophical theories of consciousness is presented as a practice of word usage, which posits systems of consistent descriptions of the information properties of consciousness and cognitive processes within the framework of certain methodological rules. The limitations of the computational approach are associated with the lack of a scientific theory of subjectivity, the fundamental irreducibility of external properties of consciousness to internal states, and fundamental restrictions on the completeness and consistency of computer mathematics. A classification of several anti-computational programs has been carried out. The article discusses various limitations of computational approaches, such as: the complexity of formalizing nonlinear and hidden processes, the dependence of emotions and intuition on the situational context and individual differences, the difficulty of predicting the emerging properties of autonomous and autopoietic systems, and the limitations associated with non-standard computational processes of quantum and dynamic phenomena. A special role is given to the problem of set-theoretic reductive realism within the framework of mathematical structuralism and some aspects of p-adic number systems as a possible alternative to a “universal language” for the computational philosophy of consciousness.

* The research was funded by Russian Science Foundation No. 24-28-00540, <https://rscf.ru/en/project/24-28-00540/>

Keywords: computational approach, philosophy of mind, philosophy of consciousness, scientific metaphor, “explanatory gap”, empirical theories of consciousness, methodological limitations

For citation: Baryshnikov, P.N. “‘Yazykovoi bar’er’ v teoriyakh soznaniya i ogranicheniya vychislitel’nogo podkhoda” [“Language barrier” in theories of mind and limitations of the computational approach], *Filosofskii zhurnal / Philosophy Journal*, 2024, Vol. 17, No. 2, pp. 122–136. (In Russian)

References

- Anokhin, K.V. “Kognitom: v poiskakh fundamental’noi neironauchnoi teorii soznaniya” [Cognitome: in search of a fundamental neuroscientific theory of consciousness], *Zhurnal vysshei nervnoi deyatel’nosti im. I.P. Pavlov*, 2021, Vol. 71, No. 1, pp. 39–71. (In Russian)
- Baryshnikov, P.N. *Vychislitel’nye modeli razuma: ot koda k smyslu* [Computational Models of the Mind: From Code to Meaning]. Moscow: URSS Publ., 2022. (In Russian)
- Bosse, T., Jonker, C.M. & Treur, J. “Formalisation of Damasio’s theory of emotion, feeling and core consciousness”, *Consciousness and Cognition*, 2008, Vol. 17, No. 1, pp. 94–113.
- Bray, D. “Limits of computational biology”, *In Silico Biology*, 2015, Vol. 12, No. 1 (2), pp. 1–7.
- Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y.T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M.T. & Zhang, Y. *Sparks of Artificial General Intelligence: Early experiments with GPT-4*, 2023, DOI: 10.48550/arXiv.2303.12712.
- Hacking, I. *Pochemu voobshche sushchestvuet filosofiya matematiki?* [Why is there philosophy of mathematics at all?], trans. by V.V. Tselishchev. Moscow: Kanon+ Publ., 2020. (In Russian)
- Hasselberger, W. & Lott, M. “Where lies the grail? AI, common sense, and human practical intelligence”, *Phenomenology and the Cognitive Sciences*, 2023, DOI: 10.1007/s11097-023-09942-x.
- Hoel, E. “The risk of another consciousness winter”, *The Intrinsic Perspective*, 2023 [https://www.theintrinsicperspective.com/p/the-risk-of-another-consciousness, accessed on 21.04.2023].
- Hoel, E. *The world behind the world: consciousness, free will, and the limits of science*. New York; London: Avid Reader Press, 2023.
- Khrennikov, A.Yu. *Modelirovanie protsessov myshleniya v r-adicheskikh sistemakh koordinat* [Modeling of thinking processes in p-adic coordinate systems]. Moscow: FIZMATLIT Publ., 2004. (In Russian)
- Mikhailov, I.F. “Predmety i metody ehmpiricheskikh issledovaniy soznaniya” [Subjects and methods of empirical studies of consciousness], *Filosofskii zhurnal / Philosophy Journal*, 2024, Vol. 17, No. 2, pp. 92–109. (In Russian)
- Milkowski, M. *Explaining the Computational Mind*. Cambridge, Mass.: The MIT Press, 2013.
- Moore, G.E. “Yavlyaetsya li sushchestvovanie predikatom?” [Is existence a predicate?], trans. by V.E. Mel’nikov, ed. by V.A. Surovtsev, *Vestnik Tomskogo gosudarstvennogo universiteta: Filosofiya, Sotsiologiya, Politologiya*, 2009, Vol. 8, No. 4, pp. 111–122. (In Russian)
- Penrose, R. *Shadows of the mind: a search for the missing science of consciousness*. Oxford; New York: Oxford University Press, 1994.
- Piccinini, G. *Physical computation: a mechanistic account*. Oxford: Oxford University Press, 2015.
- Putnam, H. *Representation and reality*. Cambridge, Mass.: MIT Press, 2011.
- Rescorla, M. “The Computational Theory of Mind”, *The Stanford Encyclopedia of Philosophy*, ed. by E.N. Zalta, 2017 [https://plato.stanford.edu/archives/spr2017/entries/computational-mind, accessed on 21.01.2023].
- Samoilenko, S. *Znakomstvo s p-adicheskimi chislami. Chast’ 2, prakticheskaya* [Introduction to p-adic numbers, Pt. 2, practical], 2022 [https://habr.com/ru/articles/646143, accessed on 21.04.2023]. (In Russian)

- Shagrir, O. “The Rise and Fall of Computational Functionalism”, *Hilary Putnam*, ed. by Y. Ben-Menahem. Cambridge: Cambridge University Press, 2005, pp. 220–251.
- Storozhuk, A.Yu. “Filosofskie osnovaniya modelirovaniya kvantovykh komp’yuterov” [Philosophical foundations for modeling quantum computers], *Sibirskii filosofskii zhurnal*, 2019, Vol. 17, No. 4, pp. 85–92. (In Russian)
- Tselishchev, V.V. *Apriornye struktury kak predstavlenie znaniya* [A priori structures as knowledge representation]. Novosibirsk: Manuscript – SIAM Publ., 2013. (In Russian)
- Tselishchev, V.V. *Ontologiya matematiki: ob’ekty i struktury* [Ontology of mathematics: objects and structures]. Novosibirsk: Nonparel Publ., 2003. (In Russian)
- Wolfram, S. *A new kind of science*. Champaign, IL: Wolfram Media, 2019.
- Yushchenko, R. *Chto nuzhno znat’ pro arifmetiku s plavayushchei zapyatoi* [What you need to know about floating point arithmetic], 2011 [<https://habr.com/ru/articles/112953>, accessed on 21.01.2023]. (In Russian)