

Репрезентации в науках о человеке: время переосмысления

И. Ф. Михайлов

Институт философии РАН,
Российская Федерация, 109240, Москва, ул. Гончарная, 12

Для цитирования: Михайлов И. Ф. Репрезентации в науках о человеке: время переосмысления // Вестник Санкт-Петербургского университета. Философия и конфликтология. 2019. Т. 35. Вып. 2. С. 297–307. <https://doi.org/10.21638/spbu17.2019.206>

Понятие репрезентации играло фундаментальную роль в когнитивной науке с момента ее появления в середине прошлого века. Акцент на репрезентации был сделан, поскольку когнитивная психология и лингвистика с самого начала опирались на «компьютерную метафору» и, соответственно, на классическую Тьюрингову концепцию вычислений как операций с символами, которые рассматривались как репрезентации (представления) некоторых данных или содержания. В связи с появлением многочисленных неклассических моделей вычислений (распределенные, аналоговые, генетические, квантовые и т. д.) в последние десятилетия она подвергается атакам со стороны различных антирепрезентационалистских или посткогнитивистских школ. Как показывается в статье, полный отказ от понятия репрезентации в теоретических схемах радикального коннекционизма и радикального энактивизма, равно как и иррелевантность этого понятия в когнитивной парадигме «динамических систем», означает шаг назад к традиционному естествознанию с его слабой способностью объяснения когнитивных явлений, в то время как сильный репрезентационализм, тесно связанный с когнитивным классицизмом, чреват парадоксами типа парадокса гомункула, и его интерпретация на реальных нейрофизиологических фактах составляет теоретическую проблему. Напротив, умеренный или слабый репрезентационализм исходит из того, что сложные вычислительные системы так или иначе используют репрезентации результатов вычислений своих подсистем или уровней в качестве данных для вычислений в других подсистемах или на других уровнях. Поэтому понятие репрезентации представляет операциональную ценность для когнитивных наук, но при этом не может быть универсальным ключом для когнитивных или философских проблем.

Ключевые слова: репрезентация, вычисление, когнитивная наука, символизм, коннекционизм, энактивизм, динамические системы, посткогнитивизм.

Если существуют способы, которыми философы могут помочь когнитивным исследователям, то, безусловно, один из них состоит в том, чтобы помочь этим исследователям определить уместность использования таких понятий, как репрезентация, в различных контекстах.

Уильям Рэмзи [1, р. 36].

1. Понятие репрезентации в исторической перспективе

Компьютерная революция XX в. открыла новые перспективы не только в математике и прикладных науках, но и в психологии и даже в лингвистике. Поскольку компьютер воспринимался как интеллектуальная машина, он естественным обра-

зом стал моделью человеческого интеллекта, породив «компьютерную метафору» в психологии. И, поскольку Тьюрингова модель вычислений представляла их как правилосообразные операции с символами, науки о языке также оказались включенными в новую парадигму. Необихевиоризм как непосредственный предшественник когнитивной психологии видел идеал научности в отказе от менталистских терминов. Новой психологической школе удалось восстановить их в правах без малейшего намека на метафизику. Это случилось благодаря возрождению на новой основе старой идеи Томаса Гоббса: мышление есть вычисление. Вычисления в их классическом Тьюринговом понимании суть операции над символами, которые представляют некое внешнее содержание и, следовательно, являются репрезентациями. Таким образом, *вычисления* и *репрезентации* — два опорных концепта в рамках когнитивных подходов в различных науках. Вокруг них разворачиваются основные философские дискуссии об основаниях когнитивной парадигмы.

Тема вычислений как специального предмета теории — прежде всего и в основном математической — привлекла внимание в 1930-х годах, когда Алан Тьюринг, думая над проблемами гильбертовского обоснования математики, предложил теоретическую абстракцию, известную как машина Тьюринга [2], позволившую лучше обозначить класс вычислимых функций. Поскольку идеи Тьюринга легли в основу теории и практики современных компьютеров, его подход долгое время считался — а многими и по сей день считается — теорией вычислений в собственном смысле слова. Однако с появлением новых подходов — квантовые, аналоговые, генетические, параллельные вычисления — встал вопрос о границах Тьюринговой парадигмы, породив широкий спектр исследований в области нетьюринговых моделей [3]. Таким образом, проблема вычислений — их определения, онтологии, спектра применимости понятия — постепенно стала общенаучной и в какой-то степени философской. В общем виде вычислительные процессы, в отличие от простых динамических, можно охарактеризовать как алгоритмические и существенно независимые от материальных реализаций. Иными словами, любые структурные изменения, которые можно описать как последовательность состояний, смена которых подчиняется определенным правилам («если — то») и которые можно реализовать более чем в одном материальном субстрате, суть вычисления. Естественным образом появилась концепция «натуральных вычислений» [4] — вычислительных процессов в природе, таких как трансляция РНК и последующий синтез протеинов. Это дало начало таким научным дисциплинам, как вычислительная термодинамика, вычислительная биология, вычислительная астрономия, экономика и т. д. Так как в самом понятии вычислений присутствует изрядная доля антропоморфизма, дело часто представляется так, что вычисления непременно связаны с обработкой информации. Но поскольку эта тема до сих пор является в высшей степени дискуссионной, ограничимся лишь следующим соображением: об информационном аспекте в вычислениях можно говорить, если имеет место инвариантная функциональная взаимозависимость изменений в различных структурах — как в уже упомянутых РНК и протеинах. Можно с некоторыми оговорками сказать, что информация как бы добавляет телеологическое измерение в мир: изменение *a* в структуре *A* происходит для того, чтобы вызвать изменение *b* в структуре *B*. Вопрос же о том, является ли информация необходимым элементом вычислений, безусловно, заслуживает отдельного обсуждения, для которого здесь недостаточно места. Однако, как ста-

нет ясно далее, сама его постановка проливает свет на концептуальные проблемы когнитивных наук. Так, если некая структура или последовательность элементов обрабатывается в вычислительном процессе именно как информация о состоянии другой структуры, то первая как бы представляет — репрезентирует — вторую, т. е. является ее *репрезентацией* в данном вычислительном процессе.

Поскольку первые инкарнации когнитивной психологии и лингвистики в 1950-х годах существенно направлялись восходящей звездой компьютерной теории и промышленности, вычисления и репрезентации стали двумя концептуальными столпами, на которых зиждется когнитивистская парадигма. Они оставались таковыми на классическом этапе развития когнитивных наук, который в литературе принято обозначать терминами «символизм», «классицизм» и «компьютериализм»¹. В основу современных компьютеров архитектуры фон Ноймана была положена машина Тьюринга, идея которой основана на алгоритмической обработке последовательностей символов. Несмотря на то что, по некоторым свидетельствам, сам Алан Тьюринг незадолго до смерти начал разрабатывать альтернативные модели вычислений — в том числе распределенные и генетические, — его соответствующие рукописи не были вовремя опубликованы, и слава за эти открытия досталась другим людям [5]. Ирония истории состоит в том, что Тьюринг *de facto* заложил основы нетьюринговых вычислений, но об этом мало кто знает. Серийные же символьные вычисления, теория которых связана с его именем, оказались «на коне» по исключительно практическим соображениям — в те времена (а во многом и в наши тоже) любая альтернатива была более ресурсоемкой [6]. Исследования Питтса и Маккалока в области нейронов и «перцептрон» Розенблатта (соответственно, в 1940-е и в 1950-е годы) не принесли результатов, пригодных к использованию в промышленности или в обороне, и попытки реализации распределенных вычислений в нейронных сетях были приостановлены до начала 1980-х годов. Поэтому компьютер фон-ноймановской архитектуры стал на десятилетия главной метафорой для когнитивных способностей живых субъектов.

В 1986 г. выходит фундаментальный труд группы авторов под названием «Параллельный распределенный процессинг» [7]. Там были продемонстрированы теоретические разработки и результаты экспериментов с когнитивными моделями, представляющими собой эмулированные нейронные сети. В таких моделях данные, подаваемые на входной слой сети, случайным образом прокладывают себе путь через внутренний слой к выходному слою, меняя по пути количественные характеристики связей между нейронами, называемые «весами». Вес каждой связи означает сравнительную вероятность прохождения сигнала по этой связи в будущем. Данные, получаемые на выходе, интерпретируются как ответ сети на вопрос, заданный ей на входе. Если ответ неправильный с точки зрения экспериментатора-тренера, то запускается алгоритм «обратного распространения ошибки», и уже вычисленные веса связей вновь пересчитываются. Так происходит обучение сети на большом количестве данных, пока сеть не научается давать правильные ответы. В качестве данных может использоваться что угодно: изображения животных, форма прошедшего времени неправильных глаголов и т. п.

¹ Последний термин, как мы увидим дальше, не совсем точен.

Новый подход в когнитивных исследованиях долгое время после его появления был известен как *коннекционизм*². Его успехи в биологически реалистичных объяснениях когнитивных способностей поставили новые вопросы перед философами-методологами: вместо правилсообразных операций с символами мы имеем теперь многочисленные итерации пересчета весов межнейронных связей, вместо простого семантического отношения символа к значению — векторные величины активации нейронных ансамблей. Стала очевидной потребность в переосмыслении опорных концептов — вычислений и репрезентаций³, референты которых в новой когнитивной модели локализовались не на уровне процессинга, а на уровне архитектуры. Сигналом к такому переосмыслению была известная в истории науки атака Джерри Фодора и Зенона Пилишина на семантические возможности коннекционистских сетей [8]. Авторы — основоположники классицистского взгляда на когнитивную архитектуру — выдвинули аргументы, основанные на комбинаторном характере ментальных репрезентаций, следствием которого являются их продуктивность (способность порождать неопределенно большое количество мыслей из конечного числа элементов), систематичность (способность мыслить некоторые содержания, отличные от данного, но семантически связанные с ним) и композициональность (зависимость значения целостной репрезентации от значений ее составных частей). По их мнению, этот аргумент показывает, что архитектура сознания не является коннекционистской *на когнитивном уровне*, поскольку коннекционистские репрезентации не демонстрируют этих свойств. Впрочем, они допустили, что коннекционизм может давать представление о *нейронных (или «абстрактно неврологических») структурах*, на которых реализована классическая когнитивная архитектура.

Затем последовала дискуссия с участием Пола Смоленски [9], утверждавшего, что критики не учли распределенный характер коннекционистских репрезентаций и что коннекционистские исследования должны предложить *новые формализации фундаментальных вычислительных понятий*, которые до сих пор были формализованы одним определенным образом в традиционной символической парадигме — очень важное замечание в контексте нашего обсуждения основополагающих понятий когнитивной науки. Часто цитируемая статья Уильяма Рэмзи [1] многими рассматривается как важный вклад в интересующую нас дискуссию. Рэмзи показал, что однонаправленные сети с обратным распространением ошибки не нуждаются в понятии репрезентации как таковом. Это породило целый букет антирепрезентационалистских концепций, в том числе отрицающих какую-либо роль этого понятия в когнитивных науках вообще⁴.

Следующим вызовом классическому когнитивизму оказались нетождественные, но соотносимые концепции, известные как *знактивизм* и *embodied mind*. Странники этих подходов, в многом черпающие вдохновение в феноменологической философии, противопоставляют пониманию когнитивного аппарата как встроено-

² Сегодня в литературе чаще используется термин *deep learning networks* — сети глубокого обучения.

³ Многие считают, что вектор значений весов связей внутри обученной сети можно рассматривать как репрезентацию категориальной структуры пропущенных через нее данных.

⁴ Сюда относятся как новое направление, которое основывается на математической теории динамических систем [10], так и работы исследователей, сочетающих коннекционизм с подходами, восходящими к Выготскому [11].

ного компьютера, манипулирующего символами, новый подход, располагающий когнитии в пространстве между мозгом, телом и окружающей средой. Сторонники этого направления сосредоточиваются скорее на моторике и бессознательных реакциях, а не на символьных вычислениях. Как и в коннекционизме, антирепрезентационализм здесь не необходим, но возможен. Так, «радикальный энактивизм», манифест которого сформулирован в работах Дэниела Хутто [12; 13], провозглашает отказ не только от репрезентационализма, но и от компьютерационализма как такового, обозначая новую парадигму как посткогнитивизм.

Сюда же примыкает относительно новое направление, которое исследователи часто называют «динамическими системами», имея в виду используемый там математический аппарат [14]. Когнитивные процессы помещаются здесь в многомерное «пространство состояний», измерения которого представляют параметры системы, выраженные переменными с изменяющимися во времени значениями. Когнитивные процессы описываются как сложные траектории в этом n -мерном динамическом пространстве. Когнитивная наука должна обнаруживать инварианты в этих движениях и устанавливать функциональные связи между ними — во многом подобно тому, как действуют естественные науки довычислительной эпохи. Если попытки проникнуть в конкретные механизмы когнитивных процессов предполагают обращение к понятиям вычислений и репрезентаций, то в динамическом подходе они оказываются в общем и целом неуместны.

На другом конце спектра позиций относительно роли репрезентаций располагается, пожалуй, новейшее из конкурирующих направлений — *predictive mind* [15] — возвращающее нас, по мнению некоторых авторов, в лоно строгого и крайнего репрезентационализма [16]. «Предсказывающий разум» — именно так переводится самоназвание нового движения — основывается на модели когнитивного аппарата как иерархии статистических вычислительных машин, задача которых — предсказание входных данных. В качестве математического аппарата используется байесовская статистика. Эти машины детерминируют другие машины, расположенные рядом и ниже в иерархии, создавая антиципаторную модель мира. Входящие данные оцениваются системой по энергетической интенсивности их воздействий и передаются дальше поверхностных датчиков *только в случае их несовпадения* с построенной моделью. Причем статистические предсказания начинаются уже на уровне сенсорных модулей, преодолеть который способны только не ожидаемые ими данные. По мнению создателей и адептов этого направления, заложенная здесь методология исключительно перспективна не только в изучении традиционных тем когнитивной психологии, но и в поиске решений «трудных проблем», таких как квалиа, эмоции и даже психические расстройства. Однако в задачи настоящей статьи не входит какая-либо оценка конкурирующих концепций: их широкий спектр интересует нас в отношении к основополагающим когнитивным понятиям.

И, как мы видим, по отношению к проблеме репрезентаций существующие парадигмы могут быть выстроены по трехступенчатой шкале: (1) *сильный* репрезентационализм — (2) *слабый* репрезентационализм — (3) *антирепрезентационализм*. Точное позиционирование интересующих нас подходов в этой системе координат возможно, однако, только при условии учета глубоких концептуальных изменений, происходящих сегодня в естественных науках.

2. Современное понимание: трудности и парадоксы

В рамках классической когнитивной науки, как указывает Нир Фреско [17, р. 356], репрезентации должны обладать двумя важными свойствами — быть *физически реализуемыми* и быть *интенциональными*. Интенциональность понимается также классически: как наличие значения или содержания, т. е. того, репрезентацией чего она является. Физическая реализуемость предполагает наличие физически допустимого *транспорта* (vehicle) репрезентаций, в качестве которых могут выступать вычислительные структуры или состояния мозга. В рамках такого понимания репрезентации действительно являются физически воплощенными сущностями, обладающими семантикой, т. е. *символами*. Над ними возможны тьюринг-вычислимые операции, и вся модель когнитивных актов оказывается полностью аналогичной работе компьютера фон-ноймановской архитектуры. К несомненным достоинствам классической модели можно отнести ее непосредственную компьютерную реализуемость: один из классиков символического подхода Дж. Р. Андерсон создал компьютерную платформу АСТ-R для моделирования когнитивных функций с целью последующего экспериментального исследования поведения моделей [18]. Объяснительный принцип построен на научной абдукции: если модель показывает те же результаты, что и живой испытуемый, то с высокой степенью вероятности когнитивный аппарат испытуемого имеет ту же структуру, что и компьютерная модель.

Однако, как мы помним, исторически когнитивистская парадигма торжествует после победоносной полемики Н. Хомского против Б. Ф. Скиннера в конце 1950-х годов [19]. Теория врожденной генеративной грамматики как бы смогла объяснить продуктивность человеческого языка — способность составлять и понимать ранее не слышанные высказывания. Лингвистика победила необихевиористскую психологию в союзе с восходящей компьютерной наукой. Неудивительно, что новый когнитивный подход имел ярко выраженные лингвистические родовые черты: представление о психической жизни как о потоке вычислительных операций над символами, обладающими семантикой. Закономерным следствием этого видения явилась концепция «мысленного языка» (*Mentalese*), выдвинутая Джерри Фодором, согласно которой нашим операциям с внешними символами соответствуют внутренние манипуляции с символами-репрезентациями, такие, в которых репрезентации имеют ярко выраженную знаковую природу, а сами операции аналогичны высказываниям в естественном языке. Именно язык мысли, согласно Фодору, является основой нашего понимания языка общения [20; 21]. Такая позиция не должна непременно уводить в дурную бесконечность, тем более что Фред Эттнив еще в 1959 г. представил механистическую модель когнитивного аппарата, позволяющую избежать *парадокса гомункула* за счет перераспределения функций между уровнями организации системы [22]. Однако на концептуальном уровне, как представляется, парадокс не был преодолен, поскольку, чтобы распознать внешнее содержание в символе, необходимы когнитивные способности. Но именно их мы пытаемся объяснить с помощью этой схемы.

Предположим, можно найти техническое объяснение, каким образом когнитивная система распознает синтаксические свойства внутренних символов. Но откуда берется содержание, делающее их интенциональными? И кто это содержа-

ние считает? Одним словом, возникает серьезное подозрение, что объясняемое содержится в объяснении. Как замечает Фреско, «внешние репрезентации зависят от внешнего субъекта (knower): субъект приписывает объектную семантику структурам данных, строчками и символам» [17, p. 358]. Не случайно поэтому исследования в рамках символической (классицистской) парадигмы более всего преуспели в объяснении языковых способностей и языковой деятельности. Что же касается объективно регистрируемых состояний когнитивного аппарата, по всей видимости, такое понимание репрезентации играет в классицизме скорее нормативную роль: все исходит из того, что такие репрезентации должны быть, поскольку они предусмотрены в используемых компьютерных моделях.

В рамках различных проектов *brain reading* («чтение мозга») имеются результаты эмпирических исследований, демонстрирующие возможность построения функциональных отношений между паттернами активации определенных отделов мозга и внешними стимулами. Так, известна работа [23], где была сделана попытка с помощью математического моделирования продемонстрировать функциональную связь между произнесенным словом и паттерном активации верхней височной извилины, ответственной за высокоуровневую обработку семантически нагруженной акустической информации. Пациентам, оперируемым в связи с эпилепсией или опухолями мозга, в эту область были имплантированы датчики, с помощью которых стало возможно реконструировать структуру нейронных возбуждений, возникающих при восприятии пациентом реальных или специально придуманных слов. Далее были построены несколько математических моделей, описывающих функциональные связи между паттернами активации и волновой формой звучащего слова. Затем наиболее релевантная модель использовалась для обратной реконструкции акустического образа из нейронных импульсов. Был получен неоднозначный результат: звуковой образ «восстановленных» слов не распознавался слушателями, но визуально тем не менее волновая форма полученного сигнала соответствовала таковой реально произносимых человеком слов. Исследователи предположили, что по мере совершенствования технических и математических средств станет возможной разработка средств коммуникации с пациентами, лишенными речи, например, вследствие паралича.

Очевидно, что для конкретной научной области этот результат является промежуточным. Но в рамках концептуального исследования можно предположить, что эмпирический поиск увенчался успехом и найден способ трансляции как звукового образа слова в нейронные ансамбли, так и обратно. Тогда мы должны принять, что структура ансамбля возбужденных нейронов есть в точном смысле слова объективно регистрируемая репрезентация звука произносимого слова. И это, скорее всего, будет справедливо. Другое дело, что так интерпретируемая «репрезентация» не является достаточно операционализированным понятием в контексте когнитивной проблематики и не предоставляет достаточных концептуальных средств для решения философских и когнитивно-психологических проблем, связанных с сознанием и его многочисленными загадками.

Тому есть несколько причин. Во-первых, в таком случае паттерн нейронного возбуждения в такой же мере является репрезентацией звука произносимого слова, в какой, наоборот, звуковые колебания суть репрезентация структуры возбуждения нейронного ансамбля. И уже при этом соображении в самом понятии репре-

зентации нет ничего специфически когнитивного. Во-вторых, подобное расширенное, или «слабое» понимание репрезентации ведет к позиции, которую можно было бы назвать *панрепрезентационализмом*, по аналогии с панкомпьютериализмом. Мозговые структуры могут считаться репрезентациями внешних событий на тех же основаниях, на каких синтезированный протеин может считаться репрезентацией цепочки нуклеотидов РНК — или наоборот, что не принципиально. Таким образом, в основу объяснения мира когнитивных явлений кладется понятие, охватывающее широкий спектр некогнитивных явлений. С точки зрения логики объяснений и определений такое понятие может в лучшем случае служить родовым, ничего не говоря о видовых отличиях объясняемого феномена. Иными словами, для теории сознания — философской или психологической — это понятие в такой степени разработанности не может быть достаточным.

Необходимо также отметить некоторые методологические недостатки классического репрезентационализма. На наш взгляд, классическое понимание репрезентации исходит из примитивной схемы когнитивного субъекта, окруженного объектами, которые отражаются в представлениях. На самом деле все гораздо сложнее. Вся концепция семантически нагруженных внешних репрезентаций объектов недостаточно обоснована и убедительна. Подобно субсимвольным вычислениям, на которые опирается коннекционизм, было бы правильно говорить о субсимвольных репрезентациях — как внешних, так и внутренних. Тогда, конечно, репрезентации имеют смысл только в контексте вычислений. Так, например, структура *A* может быть понята как репрезентация структуры *B* в рамках некоторого вычисления, тогда и только тогда, когда в этих же самых рамках *A* и *B* связаны устойчивой инвариантной функцией.

Принимая такое «слабое» определение репрезентации, мы оказываемся дальше от конечной цели когнитивной теории, поскольку так понятые репрезентации не обязательно позволяют нам объяснить процесс получения знания в его окончательном виде. Но это единственный способ избавиться от «гомункула» и увидеть познание как процесс в сложной многоуровневой вычислительной системе.

Другая трудность может быть связана с излишне антропоморфным пониманием «семантического» отношения репрезентации к тому, что репрезентируется. Подобно тому, как в человеческом мире знаки и их значения связаны между собой конвенционально, здесь также люди должны *знать*, а следовательно, быть *обучены* этим связям. Если мы позаимствуем эту семантическую теорию для классической версии когнитивной науки, мы рискуем столкнуться с другой версией парадокса гомункула: для того чтобы символическое вычисление было семантически эффективным, когнитивная система должна «знать» семантические отношения между символами и их референтами. Но любое знание представляет собой (или имеет в своей основе) репрезентацию. Тогда каждая репрезентация требует другую, поддерживающую ее репрезентацию, и так далее до бесконечности.

В случае чисто синтаксического вычисления мы избегаем этого парадокса, но оставляем механизм, благодаря которому психические состояния вообще имеют содержание (т. е. интенциональность) необъясненным. Но если, потерпев неудачу с компьютерными подходами, мы отступим к естественно-научным объяснениям репрезентации, то мы в конечном итоге упустим смысл всего когнитивного предприятия — представить экономные теоретические средства для

такого объяснения эмерджентности сознания и познания, которое оставалось бы в целом в рамках натуралистической картины мира. Таким образом, это еще один аргумент в пользу более слабого понимания репрезентации, с тем чтобы она оставалась в когнитивной науке как полезный — но не универсальный — объяснительный принцип.

3. Эскиз «слабого» репрезентационализма

В контексте современных теорий вычислений было бы разумно предположить, что репрезентации являются не только объектом, но прежде — результатом вычислительных процессов. Вычислительные процессы, осуществляемые на более высоком системном уровне, принимают выходные данные процессов более низкого уровня в форме репрезентаций, но это репрезентативное отношение не является семантическим — оно представляет собой особый тип причинной связи между уровнями, что позволяет повысить эффективность системных взаимодействий, например, понизив их термодинамическую стоимость [24]. В этом смысле, скажем, ощущение теплоты представляет собой репрезентацию броуновского движения молекул, а зрительные образы — репрезентацию результатов нейробиологических вычислений, осуществляемых в сетчатке и зрительных нервах, при этом обе репрезентации подаются в качестве данных на вход сенсомоторных процессоров.

«Слабый» репрезентационализм в нашей версии предполагает одно важное философское следствие: понятие репрезентации в том виде, в котором оно сохраняется в когнитивном тезаурусе, недостаточно для выражения таких глобальных тем, как познавательное отношение субъекта к миру. Это понятие уместно при интерпретации когнитивной «фабрики» как иерархии вычислительных процессов для обозначения формы детерминации одних процессов другими. Последовательные репрезентации образуют многоуровневый путь прохождения данных. Общий результат, данный нам в виде единой упорядоченной картины мира, является результатом сложного взаимодействия взаимно детерминированных когнитивных процессов, которые обмениваются данными и результатами вычислений в приемлемых для них репрезентативных формах. За пределами рассмотрения в этой статье остаются социальные вычисления, которые — в виде языка, производного от него рационального дискурсивного мышления, норм и культурных сценариев — достраивают уникальную когнитивную сферу *homo sapiens*.

* * *

Представленные здесь соображения, на наш взгляд, делают более правдоподобными следующие выводы:

- 1) репрезентации, как они концептуализированы в классическом когнитивизме, более всего приспособлены к непосредственной компьютерной реализации на современном уровне развития технологий. Но на концептуальном уровне они чреваты парадоксами и слишком зависимы от лингвосемантической объяснительной схемы;
- 2) радикально-антикомпьютационалистские программы в когнитивной науке частично или полностью приносят в жертву вычислительный дискурс

в пользу классической математики, нейрофизиологии и биологии, что приближает их к естественно-научным взглядам на эволюцию и прижизненное обучение живых существ. Но этот же шаг лишает их главного преимущества когнитивного подхода — расширения эмпирической сферы за счет компьютерного моделирования психических процессов;

- 3) и потому с этой точки зрения наиболее перспективной выглядит слабая версия репрезентационализма, в рамках которой репрезентации понимаются как необходимый элемент многослойных вычислительных процессов в сложных когнитивных системах, но однако недостаточный для окончательного объяснения психической жизни.

Литература / References

1. Ramsey, W. (1997), "Do Connectionist Representations Earn their Explanatory Keep?", in *Mind & Language*, vol. 12, no. 1, pp. 34–66.
2. Turing, A. M. (1936), "On computable numbers, with an application to the Entscheidungsproblem", in *Proceedings of the London Mathematical Society*, vol. 2–42, no. 1, January 1, 1937, pp. 230–265.
3. Hogarth, M. (1994), "Non-Turing Computers and Non-Turing Computability", in *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 1: Contributed Papers, pp. 126–138.
4. MacLennan, B. J. (2004), "Natural computation and non-Turing models of computation" *Theoretical Computer Science*, vol. 317, pp. 115–145.
5. Eberbach, E., Goldin D. and Wegner P. (2004), "Turing's Ideas and Models of Computation", in Teuscher C., ed., *Alan Turing: Life and Legacy of a Great Thinker*, Springer Berlin — Heidelberg: Berlin; Heidelberg, pp. 159–194.
6. Copeland, M. (2016), "What's the Difference Between Artificial Intelligence, Machine Learning, and Deep Learning?", available at: <https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/> (accessed 5 August, 2018).
7. Rumelhart, D. E., McClelland, J. L. and Corporate PDP Research Group (eds) (1986), *Parallel distributed processing: explorations in the microstructure of cognition*, vol. 1: Foundations, MIT Press; Cambridge, MA, USA.
8. Fodor, J. A. and Pylyshyn, Z. W. (1988), "Connectionism and cognitive architecture: A critical analysis", in *Cognition*, no. 28, pp. 3–71.
9. Smolensky, P. (1988), "On the Proper Treatment of Connectionism", in *Behavioral and Brain Sciences*, no. 11, pp. 1–74.
10. Thelen, E. and Bates, E. (2003), "Connectionism and dynamic systems: are they really different?", in *Developmental Science*, vol. 6, no. 4, pp. 378–391.
11. O'Brien, G. and Opie, J. "Radical connectionism: thinking with (not in) language", in *Language & Communication*, vol. 22, no. 3, pp. 313–332.
12. Hutto, D. D. (2011), "Representation Reconsidered", in *Philosophical Psychology*, vol. 24, no. 1, pp. 135–139, <https://doi.org/10.1080/09515089.2010.529261>
13. Hutto, D. D. and Myin, E. (2013), *Radicalizing enactivism: basic minds without content*, MIT Press, Cambridge, MA.
14. Smith, L. B. and Thelen, E. (2003), "Development as a dynamic system", in *Trends in Cognitive Sciences*, vol. 7, no. 8, pp. 343–348, [https://doi.org/10.1016/S1364-6613\(03\)00156-6](https://doi.org/10.1016/S1364-6613(03)00156-6)
15. Hohwy, J. (2013), *The Predictive Mind*, Oxford Press, Oxford, UK.
16. Gärtner, K. and Clowes, R. W. (2017), "Enactivism, Radical Enactivism and Predictive Processing: What is Radical in Cognitive Science?", in *Kairos. Journal of Philosophy & Science*, no. 18, pp. 54–83.
17. Fresco, N. (2012), "The Explanatory Role of Computation in Cognitive Science", *Minds & Machines* no. 22, pp. 353–380, <https://doi.org/10.1007/s11023-012-9286-y>
18. Anderson, J. R. (1983), *The architecture of cognition*, Harvard University Press, Cambridge, MA, USA.
19. Chomsky N. (1959), "Review of B. F. Skinner, Verbal Behavior", in *Language*, no. 35, pp. 26–58.
20. Fodor, J. A. (1975), *The Language of Thought*, Thomas Y. Crowell Company, New York, USA.
21. Fodor, J. A. (2008), *LOT 2: The Language of Thought Revisited*, Oxford University Press, New York, USA.

22. Attneave, F. (1961), "In defense of homunculi", in: Rosenblith, W. A., ed., *Sensory communication: Contributions to the symposium on principles of sensory communication*, Endicott House, M.I.T., Cambridge, MA, pp. 777–782.

23. Pasley, B. N., David, S. V., Mesgarani, N., Flinker, A., Shamma, S. A., et al. (2012), "Reconstructing Speech from Human Auditory Cortex", *PLOS Biology*, vol. 10, no. 1, <https://doi.org/10.1371/journal.pbio.1001251>

24. Kempes, C. P., Wolpert, D., Cohen, Z. and Pérez-Mercader J. (2017), "The thermodynamic efficiency of computations made in cells across the range of life", *Philosophical Transactions of The Royal Society A*, no. 375: 20160343, <https://doi.org/10.1098/rsta.2016.0343>

Статья поступила в редакцию 6 августа 2018 г.;
рекомендована в печать 7 февраля 2019 г.

Контактная информация:

Михайлов Игорь Феликсович — канд. филос. наук, ст. науч. сотр.; ifmikhailov@gmail.com

Representations in human sciences: time for reconsideration

I. F. Mikhailov

Institute of Philosophy, Russian Academy of Sciences,
12, ul. Goncharnaya, Moscow, 109240, Russian Federation

For citation: Mikhailov I. F. Representations in human sciences: time for reconsideration. *Vestnik of Saint Petersburg University. Philosophy and Conflict Studies*, 2019, vol. 35, issue 2, pp. 297–307. <https://doi.org/10.21638/spbu17.2019.206> (In Russian)

The concept of representation has played a fundamental role in cognitive science since its inception in the 1950s. Representation gained this emphasis insofar as cognitive psychology and linguistics since their beginnings were based on a "computer metaphor" and, accordingly, on the classical Turing concept of computations as operations with symbols construed as representations of some data or content. With the emergence of numerous non-classical models of computing in recent decades (distributed, analog, genetic, quantum, etc.), representation has been attacked by various anti-representationalist or "post-cognitive" schools. This article shows that the complete rejection of representations in theoretical schemes of radical connectionism and radical enactivism, as well as the irrelevance of this concept in the cognitive paradigm of "dynamic systems," means a step back to traditional natural science with its weak ability to explain cognitive phenomena. Meanwhile, the strong representationalism closely related to computational symbolism is fraught with paradoxes, such as the "homunculus paradox," and its interpretation on real neurophysiological facts is a theoretical problem. On the contrary, moderate, or weak, representationalism proceeds from the fact that complex computing systems use representations of computational outcomes in their subsystems or at their particular levels as input for computations in other subsystems or at other levels. Therefore, the concept of representation is of operational value to cognitive science, but it cannot be a universal clue to cognitive or philosophical problems.

Keywords: computation, representation, cognitive science, symbolism, connectionism, enactivism, dynamic systems, predictive mind.

Received: August 6, 2018
Accepted: February 7, 2019

Author's information:

Igor F. Mikhailov — PhD, Senior Researcher; ifmikhailov@gmail.com