

В защиту гипотезы внутренних репрезентаций в современных исследованиях восприятия и познания¹

Сущин М.А.

В работе рассматриваются три основные представленные в современной литературе вызова идее внутренних репрезентаций, являвшей собой одно из теоретических оснований когнитивной науки на протяжении долгого времени. Автор по очереди обращается к каждому критическому доводу и заключает, что понятие репрезентации сохраняет свою значимость для исследований познания и разума. В противовес антирепрезентационистским подходам автор приводит в пример несколько проектов в современных когнитивных исследованиях, включая недавние байесовские модели восприятия и действия, а также концепцию зрительных ансамблей, в фундаментальном смысле сохраняющих приверженность понятию внутренних репрезентаций. По мнению автора, при рассмотрении зрительных репрезентаций необходимо учитывать тот факт, что они являются результатом работы несовершенной зрительной системы и служат в первую очередь для ориентации и действия в сложной динамической среде.

КЛЮЧЕВЫЕ СЛОВА: репрезентация; когнитивная наука; сенсорномоторный подход; познание; восприятие; действие.

СУЩИН Михаил Александрович – кандидат философских наук, старший научный сотрудник кафедры философии и социологии Юго-Западного государственного университета, старший научный сотрудник Института научной информации по общественным наукам Российской академии наук.

sushchin@bk.ru

Введение

Понятие репрезентации в том или ином виде являлось одним из движущих мотивов современных исследований познания и разума, начиная с 1940–1950-х гг. Так, в 1943 г. кембриджский философ и психолог Кеннет Крэйк в своей книге «Природа объяснения» подробным образом разработал предположение, что «Одним из наиболее фундаментальных свойств мысли является ее способность предсказания событий» [Craik 1967, 50]. По Крэйку, мысль моделирует или создает уменьшенные модели внешних событий и процессов, позволяющие приходить к выводам об их свойствах и функционировании таким же образом, как если бы за ними наблюдали в реальности. Работая над книгой ещё до появления первых программируемых ЭВМ, Крэйк отмечал, что в способности моделировать внешние события, или, в

¹ Работа выполнена при поддержке гранта РНФ, проект № 15-18-10013 «Социо-антропологические измерения конвергентных технологий». This work was supported by the Russian Science Foundation, project No. 15-18-10013 “Socio-anthropological dimensions of converging technologies”.

его терминологии, символизме, мысль человека в значительной степени сходна с оперированием механических устройств, «которые помогают мысли и вычислению» [Craik, 57]. Часто приводимое положение из работы Крэйка гласит, что «Если организм несет в голове небольшую модель реальности и своих собственных возможных действий, он способен проверять альтернативы, заключать о наилучшей из них, реагировать на будущие ситуации до их возникновения, использовать знание о прошлых событиях во взаимодействии с настоящим и будущим, и во всех отношениях действовать гораздо более полноценно, безопасно и компетентно при встречающихся ему непредвиденных обстоятельствах» [Craik, 61]. Крэйку, таким образом, удалось предвосхитить суть современной теории ментальных моделей [Johnson-Laird 1983; Johnson-Laird 2013] – респектабельного направления, возникшего в русле классической когнитивной науки.

Несколько позже американский психолог Эдвард Толман ввел понятие когнитивной карты [Tolman 1948]. Толман, экспериментально исследовавший процессы обучения у крыс, отмечал, что нервную систему животных отличает высокая избирательность к входящим стимулам, что было трудно объяснить с позиций доминировавшей в то время классической схемы «стимул – реакция». Вместо этого он исходил из предположения, что в ходе обучения крысы формируют внутренние карты местности, содержащие в себе маршруты и определяющие, «как в конечном счете будет реагировать организм» [Tolman 1948, 192].

Крэйк и Толман стали непосредственными предшественниками новых когнитивных дисциплин, исходивших из рассмотрения разума и познания как разновидности вычислительных процессов на основе особых информационных структур – ментальных репрезентаций. В течение почти полутора десятков лет данное видение в общих чертах составляло официальную точку зрения научного истеблишмента в самых разных отраслях исследований познания – от психолингвистики до области изучения зрительного восприятия. И хотя, как известно, отдельные исследователи энергично оппонировали вычислительному репрезентационизму классической когнитивной науки еще во времена ее безраздельного господства [Гибсон 1988], лишь во второй половине 1980-х гг. созрели условия для возникновения полномасштабного альтернативного направления, поставившего под сомнение необходимость обращения к понятию репрезентации для объяснения работы процессов познания.

Если говорить о персональном вкладе в развитие современных антирепрезентационных настроений, то здесь, прежде всего, выделяются такие исследователи, как робототехник Родни Брукс [Brooks 1999], выступивший одним из основоположников поля «ситуативной робототехники», психолог Кевин О’Риган [O’Regan 1992; O’Regan, Noë 2001; O’Regan 2011], один из пионеров исследований феномена «слепоты к изменению» и соавтор сенсорномоторного подхода к зрительному сознанию, а также философ Тим ван Гелдер, представивший получившую достаточно широкий отклик динамическую концепцию познания – версию теории динамических систем в приложении к когнитивным явлениям [van Gelder 1995]. Именно на

сформулированной этими авторами (а также аналогичной более ранней) критике идеи внутренних репрезентаций будет сосредоточено преимущественное внимание в данной работе.

Необходимо отметить, что критические аргументы современных противников репрезентационной перспективы, безусловно, не могли остаться без ответа [Grush 2003]. Тем не менее стремительное развитие когнитивных дисциплин и возникновение все новых исследований и подходов позволяет более полно и исчерпывающе оценить представленную критику. Таким образом, мы видим необходимость в том, чтобы еще раз кратко и систематично привести, а затем ответить на основные вызовы, брошенные идее внутренних репрезентаций в современной литературе сторонниками альтернативных антирепрезентационных проектов изучения познания. Мы собираемся показать, что идея внутренних репрезентаций не утратила своей объяснительной ценности, будучи одним из ключевых элементов целого ряда современных подходов к исследованию восприятия, познания и мозга.

Вместе с тем мы не собираемся здесь просто отмахиваться от представленной критики. Мы считаем, что работы в рамках той же программы «ситуативного и воплощенного познания» или же исследования феноменов «слепоты к изменению» и «слепоты по невниманию», выступившие в качестве платформы для современных критиков идеи внутренних репрезентаций, содержат в себе множество ценных замечаний относительно того, каким образом в особенности зрительные репрезентации должны пониматься. Будущие адекватные репрезентационные проекты и модели познания, мы будем утверждать, не должны упускать из виду то обстоятельство, что, к примеру, зрительные репрезентации (человека и других существ) являются результатом работы весьма несовершенного зрительного аппарата, необходимым в первую очередь для взаимодействия со сложным динамически меняющимся окружением, подчас в условия дефицита времени и возможностей принятия решений.

Во избежание возможного недопонимания необходимо также внести ясность, что в рамках данной работы мы собираемся исходить из, по сути, общепринятого среди философов и когнитивных ученых определения репрезентации как объекта (физического или психического), замещающего другой объект [Haugeland 1998], несущего о нем определенные сведения, не обязательно обладая при этом теми же свойствами или организацией. Развитие любого более детального взгляда на репрезентации потребовало бы значительно больше места, чем доступно здесь с учетом поставленных целей.

Три вызова идее внутренних репрезентаций

Итак, какова же была теоретическая мотивация тех исследователей, которые ратовали за отказ от идеи внутренних репрезентаций? Мы считаем, что сумма претензий адептов антирепрезентационизма сводится к трем главным пунктам, а именно (1) неудовлетворительности идеи внутренних репрезентаций в силу того, что она имплицитно предполагает существование в мозге некоего

централизованного механизма наподобие пресловутого гомункулюса, который, как предполагается, и должен эти репрезентации каким-то образом «воспринимать» [Гибсон 1988; O'Regan, Noë 2001; O'Regan 2011]; (2) отсутствию необходимости во внутренних репрезентациях, поскольку агенты, как утверждается, способны полагаться на использование непосредственно самого мира, а не его копий или моделей – мира как его собственной модели [Brooks 1999] или внешней памяти [O'Regan 1992]; (3) возможности установления в фундаментальном смысле отличной от репрезентационизма теоретической перспективы, более подходящей для описания взаимодействия агентов со средой в реальном времени, в рамках которой применение понятий репрезентации, вычисления и т.п. было бы лишено своей объяснительной силы [van Gelder 1995]. Рассмотрим каждый из этих доводов по очереди.

(1) *«Проблема гомункулюса»*. Рассуждая схематично, аргументация тех, кто апеллирует к проблеме такого характера, может быть представлена следующим образом: внутренние репрезентации в физиологической (сетчаточные изображения, ретинотопические карты в зрительной коре мозга) или психической (перцепты, ментальные образы) форме предполагают существование некоего централизованного нейронного/психического механизма, который должен их каким-то образом «воспринимать» (декодировать, «интерпретировать» и т.д.). Однако же попытка постулирования такого рода механизма является чрезвычайно сомнительной как с философской, так и с научной точки зрения. Следовательно, отказавшись от рассмотрения познания сквозь призму идеи репрезентаций (по крайней мере, в некоторых их формах) в пользу альтернативных проектов (теории извлечения информации, сенсорномоторного подхода), проблема снимается сама собой.

В частности, именно к такому типу аргументации имплицитно прибегал Гибсон в контексте критики им теорий, связывающих восприятие с обработкой данных и интерпретацией сетчаточных изображений [Гибсон 1988, 101–102]. Если предполагается, что сетчаточные изображения являются чем-то, на что можно «смотреть», то должен быть кто-либо, кто их «воспринимает» или декодирует (крошечный человечек, разум, существующий отдельно от тела). Следовательно, отказываясь от такого рода заманчивого, но ошибочного представления и понимая восприятие как активный процесс извлечения информации из среды, в котором мозг является лишь элементом большей системы глаз – голова – мозг – тело, проблема, по мысли Гибсона, устраняется сама собой.

В более ухищренной форме данное затруднение получило обсуждение в совместной статье О'Ригана и Ноэ [O'Regan, Noë 2001] и в особенности в книге самого О'Ригана [O'Regan 2011]. Так, в своей совместной работе О'Риган и Ноэ предложили так называемый сенсорномоторный подход к пониманию зрения и зрительного сознания, который они противопоставили в первую очередь распространенной точке зрения, что зрительное восприятие заключается в построении внутренней репрезентации воспринимаемого мира, активация которой каким-то образом порождает зрительный опыт. Основная проблема теорий, которые пытаются установить причину возникновения зрительного

сознания на основе исследования специфических механизмов или процессов в мозге, по О’Ригану и Ноэ, заключается в том, что они оставляют открытым фундаментальный вопрос о том, «каким образом нейронный код в принципе способен привести к возникновению сознания» [O’Regan, Noë 2001, 940]. Для любого конкретного предполагаемого физического механизма в мозге, будь согласованные колебания в диапазоне 40–70 Гц, квантовые явления в тубулиновых микротрубочках нейронов и т.д., всегда, утверждали авторы, можно задать вопрос, что же именно заключено в этих механизмах и процессах, что позволяет им порождать сознание? И что именно позволяет активности в определенных областях мозга порождать, к примеру, специфически зрительный опыт, в отличие от опыта в других сенсорных модальностях?

С целью разрешения указанных и иных трудностей авторы предлагают исходить из понимания зрения как особого рода активности, опосредованной практическим знанием (навыками) воспринимающих субъектов правил и особенностей изменения сенсорного входа (восприятия) в зависимости от их собственной двигательной активности: «Центральная идея нашего нового подхода, – пишут они, – состоит в том, что *зрение является способом исследования мира, опосредованным знанием того, что мы называем сенсорномоторными зависимостями* (курсив авторов – М.С.)» [Ibid.]. Каждая сенсорная модальность предполагает наличие связанных с ней отличных правил изменения сенсорного входа в зависимости от двигательной активности, и чтобы обладать соответствующим перцептивным опытом субъект восприятия должен, по О’Ригану и Ноэ, во-первых, применять соответствующие навыки и знания в данный момент времени (а не просто пассивно хранить их в памяти), и, во-вторых, соединить текущее использование этих навыков с высокоуровневыми когнитивными процессами, включающими способности рассуждения, планирования, речь и т.д. [Ibid., 944]. Только в этом случае, полагают авторы, можно утверждать, что субъект зрительно осознает нечто (например, как они замечают, кто-либо может вести машину и одновременно разговаривать по телефону, пропуская мимо своего внимания, т.е. фактически не видя, множество деталей и объектов встречающихся ему зрительных сцен – цвет автомобиля спереди, цвет неба и т.п.).

Мы еще будем иметь возможность вернуться к этому интригующему аспекту сенсорномоторного подхода. Пока же заострим интересующую нас сейчас тему – проблематичность идеи внутренних репрезентаций в силу необходимости постулирования некоего централизованного агента, гомункулюса, которому надлежит их каким-то образом «воспринимать». Обсуждая отношение своего подхода к эмпирическим исследованиям, О’Риган и Ноэ уделили внимание нескольким особенностям зрительного восприятия – кажущейся стабильности видимого мира, несмотря на движения глаз, тем фактам, что мы не замечаем размытые очертания во время движений глаз или же не отдаем себе отчета о наличии слепой области прямо в центре зрительного поля каждого глаза, а также неоднородности зрительного поля в силу неравномерного распределения фоторецепторов сетчатки – которые склоняли часть исследователей к предположению о наличии определенных

компенсирующих механизмов, необходимых, по их мысли, для сглаживания соответствующих несовершенств в работе и устройстве зрительного аппарата.

Эта линия рассуждения в более развернутом виде была представлена в работе самого О'Ригана [O'Regan 2011]. К примеру, как он замечает, допуская наличие компенсирующих зрительных механизмов, «исследователи непреднамеренно впали в образ мышления, в рамках которого они имплицитно предполагают наличие чего-то вроде шишковидной железы по Декарту: то есть существование внутри головы того, что философы называют «гомункулюсом» («крошечным человечком»), который воспринимает информацию, предоставляемую зрительной системой» [Ibid., 12]. Так, чтобы объяснить, почему мы не замечаем слепую область в центре зрительного поля каждого глаза, вызванную отсутствием фоторецепторов на соответствующем участке сетчатки ввиду того, что окончания нервных клеток там собираются в пучок, формируя зрительный нерв, многие исследователи прибегают к так называемой гипотезе «заполнения»: считается, что недостающая информация «заполняется» самим мозгом (либо концептуально, в виде «знания», что в данном месте что-то находится или же в форме вполне реального «зарисовывания»). Эта гипотеза, согласно О'Ригану, может влечь за собой, что после того, как мозг произвел «заполнение» соответствующей области, «создается некоторое совершенное изображение мира. Но кто или что, – спрашивает он, – тогда будет смотреть на это изображение? Не могли ли ученые попасться в ту же самую ловушку, в которую попадают обычные люди, когда идет речь о перевернутом сетчаточном изображении: ловушку представления, что в мозге есть крошечный человечек, гомункулюс, смотрящий на внутренний экран, куда проецируется внешний мир?» [Ibid., 16].

Аналогичным образом попытка объяснить, почему другие известные несовершенства работы и устройства зрительной системы не находят отражения в нашем субъективном опыте восприятия мира, при помощи иных соответствующих компенсирующих механизмов, по О'Ригану, обладает опасностью все той же пресловутой «ошибки гомункулюса», задачей которого было бы воспринимать «отретушированные» изображения мира.

Но если постулирование компенсирующих механизмов является сомнительным в концептуальном плане, то что предлагается в рамках сенсорномоторного подхода? Коротко говоря, ответ О'Ригана состоит в том, что, на первый взгляд, нуждающиеся в компенсирующих механизмах несовершенства функционирования и организации зрительного аппарата являются частью и условием нормальных направленных вовне активных процессов исследования мира. Так, «слепое пятно может быть использовано, чтобы видеть: если бы стимуляция сетчатки не изменялась существенно, когда объект попадает на слепое пятно, то мозг должен был бы заключить, что объект не воспринимался зрительно, а был плодом галлюцинации» [O'Regan, Noë 2001, 951]. К тому же, полагает О'Риган, на слепое пятно и другие несовершенства устройства зрительного аппарата типичным образом тяжело обратить внимание, потому что мозг, используя определенные статистические нейронные механизмы [O'Regan 2011, 50], в первую очередь ориентирован на

восприятие объектов во внешнем мире, а не на обнаружение артефактов собственного устройства.

Таким образом, отрицая подобные фотографиям перцепты, О'Риган все же не проводит последовательно антирепрезентационный курс, предполагая, во-первых, наличие статистических механизмов и детекторов свойств в мозге (необходимых, по его мысли, для отсеивания идущих из внешнего мира сигналов от того, что относится к деятельности самой зрительной системы). А, во-вторых, значительную степень влияния на зрительное внимание (а соответственно, и зрительный опыт) нашего предшествующего знания и такого рода репрезентаций, как перцептивные ожидания [O'Regan 2011, 60]. (Менее двусмысленно о своем неполном отрицании репрезентаций О'Риган писал в более ранней работе, отмежевываясь от радикального антрирепрезентационизма Гибсона и полагая, что мы можем воспринимать только то, что может быть интегрировано в нашу «ментальную структуру» (mental framework) при помощи определенных когнитивных процессов и операций [O'Regan 1992, 473].)

Однако можно ли считать действенной его аргументацию о несуществовании перцептов (скажем, по причине того, что должен быть некий гомункулус, который их обязан воспринимать)? Нужно начать с того, что не подлежит сомнению, что мозг в процессе восприятия никоим образом не формирует фотографические репрезентации мира, и именно наглядную демонстрацию этого обстоятельства можно считать одним из главных результатов области исследований феномена «слепоты к изменению», одним из пионеров которой был сам О'Риган. Но следует ли из этого, что мозг не формирует перцептов вообще? Вовсе нет! Вот что пишут по этому поводу лидер исследований феноменов «слепоты по невниманию» и «слепоты к изменению» Дэниел Саймонс и его соавтор Брайан Шолль: «Слепота к изменению совместима с идеей, что у нас нет внутренних детальных репрезентаций: в отсутствии таких внутренних репрезентаций слепота к изменению бы имела место. Однако даже существование слепоты к изменению логически не требует отсутствия репрезентации (Simons 2000b). Репрезентации нужны, чтобы обнаружить изменение, но они также могли бы присутствовать, несмотря на слепоту к изменению. Например, наблюдатели могли репрезентировать оригинальную и измененную сцены, однако оказались не в состоянии сопоставить их прямо.... Наличие слепоты к изменению не позволяет сделать решающий вывод о существовании или отсутствии внутренних репрезентаций. Все, что оно говорит нам, так это то, что, если мы обладаем такими репрезентациями, то мы не получаем спонтанно сознательный доступ к различиям между ними» [Scholl, Simons 2001, 1004–1005].

Забегая вперед, мы можем отметить, что существует возможность, что мозг формирует не только репрезентации нескольких объектов на зрительной сцене в пределах области высокого разрешения (плюс «размытый фон»), а, как предполагается в недавней работе М. Коэна, Д. Деннета и Н. Кэнвишер [Cohen, Dennett, Kanwisher 2016], некоторые усредненные (статистически суммированные), но в достаточной степени аккуратные репрезентации

множества объектов на данной сцене за пределами центральной области как часть единого зрительного ансамбля.

Наконец, как быть непосредственно с проблемой гомункулюса, некоего «внутреннего агента», который, как полагается, в случае наличия внутренних репрезентаций должен их воспринимать? Безусловно, к анализу этого затруднения необходимо подходить с осторожностью, поскольку данная тема за редкими исключениями была и продолжает оставаться фактически табуированной для обсуждения в исследовательской литературе в серьезном ключе. Как отмечал психолог Фред Аттнев [Attneave 1961], это табу имеет в качестве причины, во-первых, опасения по поводу нематериальной природы гомункулюса, возможно, обитающего где-то в регионе шишковидной железы и соединяемого с телом тонкими нитями эктоплазмы, что, конечно, в этом случае делает его недоступным для исследования научными методами. А во-вторых, связываемую с ним проблему порочного регресса: воспринимающий мои репрезентации гомункулюс будет, в свою очередь, нуждаться в таком же крошечном человечке в своей голове и так до бесконечности.

Тем не менее, понимая опасность связанных с этой темой затруднений, все же представляется возможным выделить два основных варианта. Так, прежде всего, гомункулюсом может быть сам мозг в целом. Как писал Дэниел Деннет, критиковавший позицию «картезианского материализма», в мозге нет какой-либо иной внутренней штаб-квартиры (внутреннего наблюдателя), поскольку как таковой штаб-квартирой является мозг в целом [Dennett 1991]. Добавим, что в этом случае впечатление некой центральной воспринимающей и иницирующей команды инстанции может возникать благодаря когерентной работе множества взаимосвязанных систем мозга. Другая позиция, которую, в частности, разделял тот же Аттнев и к которой впоследствии присоединились Фрэнсис Крик и Кристоф Кох [Crick, Koch 2000], заключается в том, что можно допустить существование определенного типа (нейронного, бессознательного) гомункулюса без впадения в ошибку бесконечного регресса. По Аттневу, проблема бесконечного регресса решается в том случае, если мы перестанем думать, что гомункулюс должен отвечать за весь спектр действий, и свяжем его с каким-либо конкретным набором функций. С точки зрения Крика и Коха, гомункулюс является бессознательным, «и только его репрезентация поступает в сознание» [Ibid., 8], благодаря чему, как они утверждают, проблема регресса не возникает. В целом симпатизируя второй альтернативе, нам не остается ничего, кроме как констатировать, что в настоящий момент отсутствуют какие-либо решающие свидетельства, способные склонить чашу весов в ту или иную сторону. Однако это обстоятельство, конечно, не может использоваться как аргумент против идеи внутренних репрезентаций.

(2) *Гипотеза мира как его собственной модели/мира как внешней памяти.* Данная линия аргументации, главными выразителями которой стали робототехник Брукс [Brooks 1999], а также выступивший с иных, близких озвученным в предыдущей части позиций все тот же О'Риган [O'Regan 1992; O'Regan, Noë 2001; O'Regan 2011], имеет под собой иную мотивацию. В первом приближении суть этой позиции можно выразить таким образом, что,

по крайней мере, в тех ситуациях, когда агенты имеют возможность в полной мере обследовать нужные им аспекты и детали среды, создание внутренних копий воспринимаемого ими будет просто излишним, поскольку лучшей моделью мира является сам мир.

Так, получившая наибольшую известность формулировка этого принципа принадлежит Бруксу [Brooks 1999]. Брукс, ставший одной из наиболее влиятельных фигур в мире современной робототехники, во второй половине 1980-х гг. выступил с радикально отличной от господствовавших в то время в области искусственного интеллекта подходов программой, целью которой было создание по-настоящему мобильных и поведенчески автономных роботов. Мы уже неоднократно достаточно подробно обсуждали основные принципы предложенной им программы [Суцин 2014; Суцин 2016], оказавшей заметное влияние на ход развития робототехники и когнитивной науки в 1990-х гг., поэтому сейчас постараемся сосредоточиться на том, что может представлять для нас наибольший интерес в рамках настоящей дискуссии.

В общих чертах концептуальное ядро доктрины Брукса, обозначенной им «поведенчески основанной робототехникой», заключалось в применении новаторского метода программирования роботов, главная особенность которого заключалась в том, что системы восприятия и действия с его помощью связывались в роботах напрямую, без посредничества каких-либо центральных блоков или модулей обработки информации. Связь между системами восприятия и действия производилась в рамках некоторого программного слоя, который был ответственен за порождение определенного типа поведения (например, нижний слой мог отвечать за простое передвижение по территории, слой выше за обход встречающихся препятствий, слой выше за идентификацию объектов, еще выше – за отслеживание изменений и т.д.). Вся система управления, соответственно, складывалась из минимум двух таких слоев, выстраивавшихся, как указывалось выше, в определенной иерархии, от наиболее общего к последовательно более специализированным типам поведения. Общее согласованное (и при этом, как правило, весьма нетривиальное!) поведение робота («мобота» или «создания», как их называл Брукс) достигалось такой своеобразной самоорганизацией слоев поведения, каждый из которых выступал, когда того требовала одна ситуация, и уходил в тень, передавая «бразды правления» в другой ситуации.

Уточним позицию Брукса в отношении репрезентаций: «... посредством наблюдения мы отмечаем, что общая тема в том, как наш слоенный и распределенный подход помогает нашим Созданиям достигать целей, заключается в том, что *центральных репрезентаций не существует* (здесь и далее курсив мой – М.С.)» [Brooks 1999, 89]. «Так же, как нет центральных репрезентаций, не существует даже центральной системы. Каждый производящий активность слой соединяет восприятие с действием прямо. Это только наблюдающий за Созданием, кто приписывает центральную репрезентацию или центральный контроль. Само Создание ими не обладает; оно является собранием соревнующихся поведений. Из локального хаоса их взаимодействий в глазах наблюдателя возникает согласованный паттерн

поведения. <...> Даже на локальном уровне – писал Брукс – *мы не обладаем репрезентациями традиционного ИИ*. Мы никогда не используем знаки, обладающие семантикой, которая может быть приписана им» [Ibid., 90].

Таким образом, как видно из приведенных пассажей, в рамках своего подхода Брукс всячески отрицал использование централизованных репрезентации в смысле классического «старого доброго искусственного интеллекта». Однако наиболее известный тезис Брукса, кажется, предполагает нечто большее: «Когда мы исследуем интеллект очень простого уровня, мы обнаруживаем, что эксплицитные репрезентации и модели мира просто мешают. *Лучше использовать мир как его собственную модель*» [Ibid., 81].

Несколько замечаний может быть высказано в связи с предложенной Бруксом программой исследований. Во-первых, все, что было сказано им в отношении репрезентаций, очевидно, являлось частью его собственного частного инженерного подхода. В какой степени по этим максимам можно судить о работе действительных процессов познания является отдельным, независимым вопросом (об этом речь пойдет чуть ниже). Во-вторых, несмотря на стремление строить роботов, которые были бы в максимальной степени ориентированы на использование мира как его собственной модели, пожалуй, наиболее интересное из созданных на основе предложенных Бруксом принципов созданий, робот Тото, строило распределенные и динамически обновляемые (как бы добавили психологи, когнитивные) карты местности (офисных коридоров и комнат), которые затем могло использовать для планирования кратчайшего маршрута с целью выполнения данного задания. Какие бы архитектурные программные особенности не воплощало в себе это устройство, ясно, что способность строить карты местности и использовать их для оптимизации маршрутов выходит за пределы принципа использования мира как его собственной модели. Наконец, в-третьих, возвращаясь к вопросу об отношении концепции Брукса к анализу действительных процессов познания, более общие ее недостатки и ограничения были указаны в обстоятельной работе философа и когнитивиста из Калифорнийского университета в Сан-Диего Рика Граша [Grush 2003].

Граш, развивший собственную репрезентационную программу, именуемую им теорией эмуляции, указал на то, что как раз таки необходимость действия в реальном мире в условиях ограниченного времени ввиду организации биологических агентов свидетельствует в пользу использования ими определенных типов репрезентаций. Дело в том, что, как прекрасно известно, скорость распространения импульсов (потенциалов действия) по нервным волокнам значительно уступает скорости движения сигналов по электрическим проводам. Соответственно, мозг (в особенности крупных организмов) с целью обеспечения быстрых и аккуратных действий, полагает Граш, ссылаясь на родственные его теории работы в области моторного контроля, мог бы использовать некие виртуальные копии, эмуляции, посланных им эффекторам команд, корректируя таким образом действия еще до прибытия относительно медленных действительных сигналов обратной связи (здесь Граш ссылается на данные о том, что мозг начинает корректировать инициированные

им действия еще до прибытия реальных сигналов об их исполнении от эффекторов). Тогда как «роботы Брукса, – замечает Граш, – не нуждаются в том, чтобы быть невероятно быстрыми и получают свою сенсорную информацию посредством электрических сигналов по очень быстрым проводам, а не медленным биологическим аксонам» [Ibid., 84].

Помимо всего прочего, по замечанию Граша, недостатком стратегии использования мира как его собственной модели была бы ограниченность агентов только тем, что непосредственно находится перед органами чувств (при условии их безупречной работы). Тогда как не секрет, что существенной характеристикой продвинутого интеллектуального поведения является способность «оценки контрфактических ситуаций и предсказания будущего» [Ibid., 85]. Вкратце, для организмов, которые не пытаются предвидеть возможные варианты развития событий и не позволяют «своим гипотезам умирать вместо себя» [Деннет 2004, 94], любое действие может привести к фатальной ошибке.

Наконец, наша дискуссия в данной части работы была бы неполной, если бы мы обошли вниманием предложенный почти в то же самое время О'Риганом с несколько иных позиций, но по своей сути родственной максиме Брукса принцип «мира как внешней памяти» [O'Regan 1992]. Эта идея возвращает нас к тому, что говорилось здесь о теории О'Ригана ранее. Основной мишенью критики и в этом аспекте его теории для него служит идея внутренних (подобных фотографиям) перцептов, существование которых, как уже говорилось, О'Риган неистово отрицает. Для восприятия, с его точки зрения, напомним, необходимо применение навыков обследования мира и способность направить внимание на то, что в данный момент доступно органам чувств. Мир, утверждает О'Риган, представляет собой внешнее хранилище, к деталям которого мы можем последовательно обращаться, делая их доступными для внимания и высокоуровневых когнитивных процессов. Благодаря этому, по О'Ригану, устраняется потребность во внутренних репрезентациях (перцептах).

Как пишут О'Риган и Ноэ, идея мира как внешней памяти послужила стимулом для развития области исследований феномена «слепоты к изменению». Этот феномен, заключающийся в неспособности субъектов обнаружить различие между двумя последовательно предъявляемыми изображениями в силу некоего отвлекающего фактора наподобие краткой вспышки по всей площади экрана в промежутке между ними, по их мнению, является прямым вызовом идее формирования внутренних репрезентаций зрительных сцен целиком. Однако свидетельствуют ли эти результаты о том, что мы не формируем репрезентаций воспринимаемых сцен вообще, в какой бы то ни было форме?

Здесь необходимо вспомнить то, что говорили в своем комментарии к статье О'Ригана и Ноэ Саймонс и Шолль, а именно что слепота к изменению не может служить доводом в пользу существования или несуществования внутренних репрезентаций. К примеру, мы можем обладать репрезентациями начальной и измененной сцен, однако просто быть не в состоянии сопоставить их прямо. Нам представляется более перспективной та мысль, что субъекты

могут формировать упомянутые выше ансамблевые репрезентации множества объектов на зрительной сцене, при нарушении общих статистических свойств которой не обнаружить изменение становится практически невозможно (см.: [Cohen, Dennett, Kanwisher 2016, 327]). Мы вернемся к обсуждению этой идеи позднее.

(3) *Возражение к аргументу об отсутствии альтернатив вычислительной программе.* Данный аргумент носит более глобальный характер, нежели предыдущие два. Его выразителем стал философ Тим ван Гелдер, представивший основные положения своей концепции в получившей известность статье «Чем может быть познание, если не вычислением?» [van Gelder 1995]. Примечательно, что в ней практически не содержится прямой критики идеи внутренних репрезентаций и их применения к анализу процессов познания. Вместо этого акцент был сделан на том, чтобы продемонстрировать возможность альтернативного вычислительной перспективе проекта изучения познания, в рамках которого понятие репрезентации оказалось бы лишено своей ведущей объяснительной роли.

Так, к примеру, в свое время в качестве довода в пользу классической вычислительной программы в когнитивной науке указывалось на отсутствие каких-либо сопоставимых альтернатив [Newell, Simon 1976]. ван Гелдер попытался оспорить это утверждение, полагая, что когнитивные системы должны пониматься как динамические системы, а как познание – как разворачивающаяся в реальном времени эволюция пространства состояний (т.е. ряда возможных состояний) этих систем. Это, как явствует из рассуждений ван Гелдера, может означать, что наиболее адекватной концептуальной рамкой для понимания таких систем оказывается теория динамических систем и связанный с ней математический аппарат дифференциальных уравнений, а не служившие долгое время эталоном для когнитивных ученых математическая теория вычислимости и классическая символическая логика.

Центральное место в работе ван Гелдера, вокруг которого строится большая часть дискуссии, занимает образ центробежного регулятора Уатта – значительного изобретения эпохи промышленной революции, благодаря которому была решена важнейшая проблема поддержания постоянной скорости парового двигателя, несмотря на постоянные колебания в нагрузке. Это устройство представляло собой вертикальный вал с подвешенными с каждой стороны на рычагах двумя тяжелыми шарами. Движение шаров регулировало положение механически связанного с рычагами дроссельного клапана в паропроводе, контролирующего подачу пара в цилиндр. Вал регулятора соединялся с валом двигателя, поэтому малейшие изменения в нагрузке на двигатель находили соответствующее отражение в скорости вращения самого регулятора. Если нагрузка на двигатель существенно снижалась, это приводило к увеличению скорости вращения регулятора. В результате шары на рычагах, соединенных с валом, под действием центробежной силы расходились в стороны, заставляя клапан перекрывать подачу пара в цилиндр в объеме, необходимом для стабилизации скорости двигателя в связи с произошедшим снижением нагрузки. (В случае увеличения нагрузки шары сходились ближе к

валу, тем самым приоткрывая клапан, чтобы больше пара подавалось к поршню в цилиндре.) Таким образом, достаточно простое устройство почти совершенно справлялось с работой, которую человек мог делать весьма неаккуратно.

При этом, настаивает ван Гелдер, в процессе работы регулятора не возникает ничего, что можно было бы назвать репрезентацией – объектом, замещающим или свидетельствующим о положении дел в другом объекте. Так, угол наклона рычагов с закрепленными на них шарами, по его мнению, является не репрезентацией скорости двигателя, а одним из прямых каузальных факторов, определяющих ее (точнее, стоящих с двигателем в отношении взаимоопределения). А поскольку стандартное понимание вычисления в когнитивной и компьютерной науке предполагает, что вычислительный процесс должен происходить на основе репрезентаций, то центробежный регулятор, заявляет ван Гелдер, не может считаться вычислительным устройством в этом смысле. В отличие от вычислительных устройств, оперирующих дискретным, пошаговым образом, указывает он, процесс работы центробежного регулятора является непрерывным, плавно разворачивающимся в каждый момент времени. Для описания этой работы, замечает ван Гелдер, безукоризненно подходят дифференциальные уравнения, определяющие скорость изменения угла наклона рычагов как функцию их текущего угла наклона, скорости изменения их угла в данный момент времени, а также текущей скорости двигателя [van Gelder 1995, 356].

Что же дает ван Гелдеру право полагать, что на основе этой аналогии можно судить и о работе процессов познания в общем или, как он говорит, что центробежный регулятор Уатта оказывается «предпочтительнее машины Тьюринга как ориентир для моделей познания» [Ibid., 381]? Главное и фактически единственное свидетельство, к которому он апеллирует, заключается в том, что естественные когнитивные процессы «всегда происходят *во времени* (здесь и далее курсив автора – М.С.), что не просто означает, что они охватывают определенный промежуток времени подобно любому физическому процессу, включая обычное цифровое вычисление, а то, что детали *времени* (продолжительность, темп, ритмы и т.д.) всегда критичны для системы, которая оперирует в реальном теле и среде» [Ibid., 379]. Эта мысль разъясняется на примере ЭВМ, которой была бы делегирована та же самая задача контроля скорости парового двигателя. Помимо ограничений, связанных с выполнением задачи к определенному сроку и организацией операций машины в определенной последовательности, считает ван Гелдер, не было бы ничего, что необходимым образом бы диктовало время совершения конкретной операции, ее длительность и время до начала следующей операции (Ibid., 354). Тогда как ни одно из действий центробежного регулятора не обладает таким же произвольным характером во временном плане в его отношении к работе двигателя.

В качестве ответа на данный аргумент о темпоральном характере естественных процессов познания необходимо вспомнить, что существует и продолжает совершенствоваться целый класс устройств, таких, как системы управления беспилотных автомобилей и летательных аппаратов, которые

вполне успешно справляются с даже более сложными задачами контроля в реальном времени (и реальном мире!), будучи при этом вычислительными и репрезентационными в смысле, который оказывается неприемлемым для ван Гелдера. Поэтому такого рода аргумент единственно не может использоваться как решающее свидетельство в пользу вывода о том, что процессы познания в общем должны быть организованы на манер центробежного регулятора Уатта, а не, скажем, машины Тьюринга.

Завершая обсуждение проекта ван Гелдера, мы можем отметить, что ему действительно удалось показать возможность альтернативного понимания феноменов, бывших некогда в безраздельном ведении классической когнитивной науки. По его признанию, эта задача не предполагала установления истинности предлагаемой альтернативы – динамической концепции познания [van Gelder 1995, 346]. В этом отношении ван Гелдеру вполне удалось достичь заявленных целей. Но чтобы составить реальную конкуренцию вычислительной и репрезентационной программе в исследованиях познания (понимаемой не только в классическом смысле) группе динамических подходов необходимо сделать шаг далее и попытаться показать свое превосходство в объяснении того же спектра феноменов и процессов, для понимания которых предлагаются репрезентационные модели.

Примеры современных репрезентационных проектов исследования познания

В этой части работы мы постараемся продемонстрировать, что понятие репрезентации вовсе не утратило своего значения и представляет собой важный элемент ряда современных проектов исследования познания и разума. Вопрос об отношении между этими проектами является слишком обширным, чтобы получить необходимое освещение на страницах этой работы. Не предполагает настоящее обсуждение и возможности детального рассмотрения этих проектов. Вместо этого мы, обозначив их общие принципы, попытаемся показать, в каких именно отношениях эти проекты могут считаться репрезентационными.

(I) *Байесовская программа в современных исследованиях познания.* Первая программа, а точнее, группа моделей, подходов и концепций в психологии, нейронауке и области их философского осмысления, объединяемая рассмотрением процессов познания сквозь призму понятийного аппарата байесовской теории вероятностей, уже получала освещение на страницах этого журнала усилиями автора настоящей статьи [Сушин 2017]. Поэтому сейчас мы постараемся сосредоточиться в первую очередь на репрезентационной составляющей байесовской программы.

Отправным пунктом для нее, как нетрудно догадаться, служит теорема, первоначально открытая Т. Байесом, а затем независимо от него П.С. Лапласом, используемая для нахождения условной вероятности какого-либо события (т.е. вероятности события A при условии события B – например, вероятности заболевания при определенных симптомах) с учетом имеющихся у нас данных (скажем, вероятности симптомов при конкретном заболевании) и, самое

главное, нашего предшествующего знания о вероятности интересующего нас события (вероятности заболевания как такового). Последний компонент, в формальном отношении известный как априорная вероятность, позволяет приходиться к наиболее вероятной оценке события, отбраковывая при этом заведомо неправдоподобные гипотезы и интерпретации (допустим, гипотезу недомогания участника полярной экспедиции вследствие теплового удара).

Применительно к рассмотрению процессов познания сторонниками байесовской теории предполагается, что мозг аналогичным образом способен использовать имеющееся у него предшествующее знание для интерпретации и предвосхищения подчас недостаточно определенного сенсорного входа (например, для оценки удаленности воспринимаемого объекта, его размера, формы – скажем, является ли он выпуклым или вогнутым – и многого другого). Даже более того, как подчеркивается недавними довольно радикальными интерпретациями [Clark 2013; Hohwy 2013], организуемые мозгом перцептивные процессы являются по преимуществу предсказывающими и нисходящими: направленными от производящей предсказания и перцептивные ожидания внутренней модели мира (называемой также генеративной) «вниз» в сторону фиксирующих сенсорный вход органов чувств. Органы чувств в свою очередь могут подтвердить или не подтвердить дошедшие до них «сверху» предсказания. В последнем случае возникает так называемая ошибка в предсказании, которая влечет за собой внесение соответствующих корректив во внутреннюю модель мира, чтобы ее предсказания в лучшей степени соответствовали тому, что регистрируется органами чувств.

Важным элементом недавних научных и философских вариаций на эту тему является совмещение идеи байесовского перцептивного вывода с учением Гельгольца о бессознательных умозаключениях: считается, что описанные выше процессы производятся мозгом по большей мере без участия сознания. Тогда как представления Гельгольца, изначально связывавшиеся им с индуктивной логикой Дж.С. Милля [Helmholtz 1925], наполняются конкретным математическим смыслом.

Таким образом, мы можем отчетливо видеть решительную приверженность сторонников байесовской программы репрезентационному пониманию природы разума и процессов познания: мозг, с этой точки зрения, обладает внутренней моделью мира, которая пытается производить все более совершенные предсказывающие репрезентации того, что может быть воспринято им в следующий момент времени.

Необходимо признать, что к настоящему моменту байесовская программа требует дальнейшей серьезной разработки как в теоретическом, так и в эмпирическом отношении. В частности, байесовские модели в психологии были подвергнуты серьезной критике из-за их чрезмерной общности и, как следствие, проблем с фальсифицируемостью [Bowers, Davis 2012]. К тому же пока что имеются лишь чрезвычайно скудные и отрывочные нейрофизиологические данные, которые могли бы быть использованы в качестве поддержки приведенной выше схемы [Ibid.]. И все же несомненной заслугой байесовской программы в когнитивных исследованиях является то,

что она в большей степени, чем какая-либо иная теория в современной психологии и когнитивной науке, способствует разрушению представлений о восприятии как об однонаправленном восходящем, идущем от органов чувств к мозгу процессе (как если бы имеющееся у нас знание совершенно не влияло на то, что и как мы можем воспринимать – см.: [Palmer 1999, 85]).

(II) *Идея статистически суммированных зрительных репрезентаций.* Данная идея, восходя еще, возможно, к представлениям гештальтпсихологов о «принципах группировки», получила развитие в рамках современной науки о зрении в последние 10-15 лет [Haberman, Whitney 2012]. В ее основании лежит тот посыл, что естественные зрительные сцены обладают избыточной и повторяемой структурой на разных уровнях организации – к примеру, на уровне составных объектов лужайка включает в себе похожие друг на друга травинки, дом – одинаковые кирпичи и т.п. Соответственно, вместо того, чтобы воспринимать каждый схожий объект индивидуально, зрительная система, как полагают сторонники этого направления, могла бы формировать некоторые усредненные, или статистически суммированные (ансамблевые), репрезентации целых комплексов таких объектов.

Располагая такого рода идеями, едва ли не естественным представляется ход мысли попытаться использовать ее для решения известной проблемы несоответствия между нашим субъективным впечатлением богатства зрительного опыта и упоминавшимися выше объективными свидетельствами существенной ограниченности механизмов зрения [O'Regan 1992; O'Regan 2011; Ноэ 2014]. В частности, именно эту возможность вслед за другими исследователями в своей недавней работе попытались актуализировать М. Коэн, Д. Деннет и Н. Кэнвишер [Cohen, Dennett, Kanwisher 2016]. Как они отмечают, несмотря на то, что наблюдатели способны видеть лишь несколько объектов в высоком разрешении (благодаря фовеальной области глаза), та значительная область зрительной сцены, которая находится на периферии, может тем не менее восприниматься как «часть зрительного ансамбля [47]. Наблюдатели, – пишут Коэн, Деннет и Кэнвишер, – обладают сознанием не только нескольких объектов, но также всей сцены...» [Ibid., 327].

Такая способность формирования статистически суммированных репрезентаций, по их предположению, могла бы находиться в основании восприятия зрительных сцен в широком смысле, позволяя воспринимающим субъектам чрезвычайно быстро распознавать их суть и основные аспекты. Например, даже не совершая саккад, за время единичной зрительной фиксации (т.е. порядка 275–300 мс.) наблюдатели способны распознать суть зрительной сцены и несколько основных объектов на ней. Однако и за меньшее время (около 50–100 мс.), замечают Коэн, Кэнвишер и Деннет, воспринимающие все еще способны отчитаться о сути этой сцены и ряде ее ключевых параметров, включая глубину, возможность ориентации в ней, открытость и т.д.

Таким образом, по мнению авторов исследования, все эти данные говорят о том, что даже «в пределах единичного взгляда наблюдатели не просто обладают доступом к малому количеству изолированных объектов в море небытия; они обладают доступом к огромному количеству информации,

охватывающему все зрительное поле» [Ibid., 329]. Важно здесь отметить еще и то, что, по их предположению, эта способность быстро извлекать ключевую статистику зрительной сцены может играть существенную роль в руководстве дальнейшими действиями – и в первую очередь, конечно, саккадами. Как известно, вместо того чтобы инициировать движения глаз случайным образом, субъекты «систематически обращаются к тем частям зрительной сцены, которые оказываются наиболее информативными для текущей задачи [72]. Эта способность интеллектуально выбирать цели совершения саккад является возможной, – полагают Коэн, Деннет и Кэнвишер, – потому что наблюдатели способны воспользоваться полученным ими знанием о сцене, исходя из ее общей графической (image) статистики [73–75]» [Ibid.]

Безусловно, возможны и иные подходы к решению проблемы несоответствия субъективных впечатлений богатства зрительного опыта и объективных свидетельств серьезной ограниченности зрительных механизмов (одна такая возможность, к слову, предлагается обсуждавшимся нами выше сенсорномоторным подходом О’Ригана), и само понятие статистически суммированных репрезентаций в этом и других его аспектах является предметом активных споров среди исследователей. И тем не менее привлекательность этой идеи заключается в том, что зрительные ансамбли могут являть собой именно такой тип репрезентаций, который оказывается совместимым с требованиями действия в реальном мире (о чем мы писали во введении) и при этом позволяет не затратным образом кодировать значительные объемы информации, что дает возможность уйти от основных трудностей классического понимания репрезентаций (и делает затрудненной приведенную выше критику антирепрезентационистов).

(III) *Другие проекты.* Задача освещения всех репрезентационных концепций и проектов в современных исследованиях познания является столь масштабной, что превосходит компетенции какого-либо одного исследователя (и тем более не может быть реализована на страницах этой работы). Так, например, мы совершенно не имеем возможности хоть сколько-нибудь подробно обсуждать здесь понятие ментальных образов, дискуссия вокруг которого с разной степенью интенсивности продолжалась до самого последнего времени (и, возможно, еще не достигла предела). Впрочем, ясно, что вне зависимости от того, носят ли они пропозициональный характер (как утверждает философ и когнитивист З. Пылишин) или являются своеобразными квази-картинами (на чем настаивает главный оппонент Пылишина когнитивный психолог С. Косслин), их существование (что само по себе мало кто ставит под сомнение) являет собой вызов для противников идеи внутренних репрезентаций.

Другим примером сравнительно недавно возникших репрезентационных проектов является упоминавшаяся выше теория эмуляции Граша (к слову сказать, обладающая определенным сходством с недавними байесовскими когнитивистскими моделями). В ее основе находится идея, что мозг при помощи внутренней модели (называемой им эмулятором или прямой моделью) создает предвосхищающие сенсорный вход репрезентации, которые могут

использоваться либо для задач взаимодействия с миром в процессах восприятия и моторного контроля, либо в автономном режиме для зрительного/моторного воображения и моторного планирования [Grush 2003, 76–79]. И, как мы видели, согласно Грашу, задача действия в реальной среде, вопреки чаяниям антирепрезентационистов, ввиду специфической телесной организации агентов фактически ставит перед ними необходимость использования репрезентаций, а также, возможно, и эмуляционного механизма в общем.

Наконец, как утверждает в недавней работе философ Уильям Бектел [Bechtel 2016], репрезентационная установка была неотъемлемой частью недавних исследований системы пространственного ориентирования в мозге грызунов, за которые нейрочеловек Джон О’Киф и Эдвард и Мэй-Бритт Мозеры удостоились Нобелевской премии по физиологии и медицине в 2014 г. Как указывает Бектел, эти исследователи и их коллеги рассматривали гиппокамп и примыкающую к нему энторинальную кору в качестве нейронного субстрата создаваемых организмами карт местности и репрезентаций их текущего местоположения (что является развитием линии исследования, заложенной Толменом с выдвинутой им гипотезой когнитивных карт).

Заключение

В данной работе мы стремились продемонстрировать, что понятие репрезентации отнюдь не утратило своей значимости для современных проектов изучения познания и разума. Мы полагаем, что репрезентационная платформа в широком смысле по-прежнему сохраняет богатые объяснительные возможности, чтобы двигать будущие исследования в лучшей степени, чем какие-либо антирепрезентационные подходы. Вместе с тем замечания, подобные тем, что прозвучали со стороны О’Ригана и его соавторов, или которые могут быть сделаны на основе результатов исследований феноменов «слепоты к изменению» и «слепоты по невниманию», должны быть учтены. И хотя, как мы утверждали, на их основе нельзя заключать о несуществовании перцептов (как это делает тот же О’Риган), при рассмотрении зрительных репрезентаций нужно исходить из того, что они в первую очередь служат организмам для ориентации и действия в сложной динамической среде. Понятие статистически суммированных репрезентаций предоставляет перспективный путь для устранения противоречия между субъективно представляемым богатством опыта, которым располагают воспринимающие, и известными объективными ограничениями устройства и работы зрительной системы.

Источники и переводы – Primary Sources and Russian Translations

Гибсон 1988 – Гибсон Дж. Экологический подход к зрительному восприятию. М.: Прогресс, 1988 (*Gibson J.J. The Ecological Approach to Visual Perception. Russian Translation 1988*).

Деннет 2004 – *Деннет Д.С.* Виды психики: на пути к пониманию сознания. М.: Идея-Пресс, 2004 (*Dennett D.C. Kinds of Minds: Towards an Understanding of Consciousness. Russian Translation 2004*).

Ноэ 2014 – *Ноэ А.* Является ли видимый мир великой иллюзией? // Логос. 2014. № 1. С. 61–78 (*Noë A. Is the Visual World a Grand Illusion? Russian Translation 2014*).

Helmholtz, Hermann von (1925) “Concerning the Perceptions in General”, *Treatise on Physiological Optics*, Vol. 3, Ch. 26, 3rd edn, Translated by J.P.C. Southall, Op. Soc. Am., Section 26, reprinted Dover, New York, pp. 1–37.

Ссылки – References in Russian

Сущин 2014 – *Сущин М.А.* Концепция ситуативного познания в когнитивной науке: критический анализ: дис. ... канд. филос. наук: 09.00.01 / Сущин Михаил Александрович: М., 2014.

Сущин 2016 – *Сущин М.А.* Возможно ли объединить сенсорно-моторную и предсказывающую программы в современных исследованиях познания? // Аршинов В.И. Асеева И.А., Буданов В.Г., Гребенщикова Е.Г., Гримов О.А., Каменский Е.Г., Майнцер К., Москалев И.Е., Пирожкова С.В., Сущин М.А., Чеклецов В.В. Социо-антропологические измерения конвергентных технологий. Онтологии и коммуникации / Отв. ред. И.А. Асеева, В.Г. Буданов. Курск: ЗАО «Университетская книга», 2016. С. 199–216.

Сущин 2017 – *Сущин М.А.* Байесовский разум: новая перспектива в когнитивной науке // Вопросы философии. 2017. № 3. С. 74–87.

In Defense of the Hypothesis of Internal Representations in Modern Research on Perception and Cognition

Mikhail A. Sushchin

This article considers three main challenges to the idea of internal representations, which was one of the theoretical foundations of cognitive science for a long time, stated in the recent literature. The author examines each critical argument in turn and concludes that the notion of representation retains its significance for the study of cognition and mind. In contrast to anti-representational approaches, the author refers to several projects in modern cognitive investigations, including the recent Bayesian models of perception and action, and the concept of visual ensembles, which in the fundamental sense remain committed to the notion of internal representations. According to the author, the study of visual representations must take into account that they are the result of functioning of imperfect visual system and serve primarily for orientation and action in complex dynamical environment.

KEY WORDS: representation; cognitive science; sensorimotor approach; cognition; perception; action.

SUSHCHIN Mikhail A. – CSc in Philosophy, Senior Research Fellow of the Institute of Scientific Information on Social Sciences of the Russian Academy of Sciences, Senior Research Fellow of Department of Philosophy and Sociology of Southwest State University.

sushchin@bk.ru

References

- Attneave, Fred (1961) “In Defense of Homunculi”, *Sensory Communication*, ed. W. A. Rosenblith, MIT Press, pp. 777–782.
- Bechtel, William (2016) “Investigating neural representations: the tale of place cells”, *Synthese*, Vol. 193, No. 5, pp. 1287–1321.
- Bowers, Jeffrey S. & Davis, Colin J. (2012) “Bayesian Just-So Stories in Psychology and Neuroscience”, *Psychological Bulletin*, Vol. 138, No. 3, pp. 389–414.
- Brooks, Rodney (1999) *Cambrian Intelligence: The Early History of the New AI*, The MIT Press.
- Clark, Andy (2013) “Whatever next? Predictive brains, situated agents, and the future of cognitive science”, *Behavioral and Brain Sciences*, Vol. 36, No. 3, pp. 181–204.
- Cohen, Michael A., Dennett, Daniel C. & Kanwisher, Nancy (2016) “What is the Bandwidth of Perceptual Experience”, *Trends in Cognitive Sciences*, Vol. 20., No. 5, pp. 324–335.
- Craik, Kenneth (1967) *The Nature of Explanation*, Cambridge University Press.
- Crick, Francis & Koch, Christof (2000) “The Unconscious Homunculus”, *Neuropsychanalysis: An Interdisciplinary Journal for Psychoanalysis and the Neurosciences*, Vol. 2, No. 1, pp. 3–11.
- Dennett, Daniel C. (1991) *Consciousness Explained*, Little Brown.
- Grush, Rick (2003) “In Defense of Some ‘Cartesian’ Assumptions Concerning the Brain and Its Operation”, *Biology and Philosophy*, Vol. 18, No. 1, pp. 53–93.
- Haberman, Jason & Whitney, David (2012) “Ensemble Perception: Summarizing the scene and broadening the limits of visual processing”, *From Perception to Consciousness: Searching with Anne Treisman*, eds. J. M. Wolfe, L. C. Robertson, pp. 339–349.
- Haugeland, John (1998) *Having Thought: Essays in the metaphysics of mind*, Harvard University Press.
- Hohwy, Jakob (2013) *The Predictive Mind*, Oxford University Press.
- Johnson-Laird, Philip (1983) *Mental Models: Towards a Cognitive Science of Language, Inference, and Consciousness*, Cambridge University Press, Cambridge.
- Johnson-Laird, Philip (2013) “Mental Models Perspective”, *The Oxford Handbook of Cognitive Psychology*, Ed. D. Reisberg, Oxford University Press, pp. 650–667.

Newell, Allen & Simon, Herbert (1976) “Computer Science as Empirical Inquiry: Symbols and Search”, *Communications of the Associations for Computing Machinery*, Vol. 19, No. 3., pp. 113–126.

Palmer, Stephen A. (1999) *Vision Science: Photons to Phenomenology*, MIT Press.

O’Regan, J. Kevin (1992) “Solving the “Real” Mysteries of Visual Perception: The World as an Outside Memory”, *Canadian Journal of Psychology*, Vol. 46, No. 3, pp. 461–488.

O’Regan, J. Kevin (2011) *Why Red Doesn't Sound Like a Bell: Understanding the feel of consciousness*, Oxford University Press.

O’Regan, J. Kevin & Noë, Alva (2001) “A sensorimotor account of vision and visual consciousness”, *Behavioral and Brain Sciences*, Vol. 24, No. 5, pp. 939–973.

Scholl, Brian J. & Simons, Daniel J. (2001) “Change blindness, Gibson, and the sensorimotor theory of vision”, *Behavioral and Brain Sciences*, Vol. 24, No. 5, pp. 1004–1006.

Tolman, Edward (1948) “Cognitive Maps in Rats and Men”, *The Psychological Review*, Vol. 55, No. 4, pp. 189–208.

van Gelder, Tim (1995) “What Might Cognition Be, If Not Computation?”, *The Journal of Philosophy*, Vol. 92, No. 7, pp. 345–381.

Sushchin, Mikhail A. (2014) *The Theory of Situated Cognition in Cognitive Science: A Critical Analysis*. Candidate of Sciences in Philosophy Dissertation, Moscow (in Russian).

Sushchin, Mikhail A. (2016) “Is It Possible to Integrate Sensorimotor and Predictive Frameworks in Modern Cognitive Investigations?” // Arshinov V. I., Aseeva I.A., Budanov V.G., Grebenshchikova E.G., Grimov O.A., Kamenskii E.G., Mainzer K., Moskalev I.E., Pirozhkova S.V., Sushchin M.A., Chekletsov V.V. *Socio-Anthropological Aspects of Converging Technologies. Ontologies and Communications*, I. A. Aseeva, V. G. Budanov, Universitetskaya kniga, Kursk, pp. 199–216 (in Russian).

Sushchin, Mikhail A. (2017) “The Bayesian Mind: A New Theory in Cognitive Investigations”, *Voprosy Filosofii*, 2017, Vol. 3., pp. 74–87.