

Философское «я», время и когнитивные науки

Михайлов И. Ф., ИФ РАН, Москва

Аннотация: В недавней истории когнитивных наук имела место дискуссия между классическим или традиционным направлением, рассматривающим свой предмет как формальный процессинг символьных значений, и направлением, основанным на теории сетей и нейрокомпьютиге, и пытавшемся в своих моделях приблизиться к реальному функционированию человеческого мозга. В литературе она обозначается как дискуссия между компьютеризацией и коннекционизмом. В статье предпринята попытка показать, что коннекционизм демонстрирует свои сильные стороны в этой конкуренции, если его адепты понимают каузально-динамическую природу их модели и отказываются от символьно-репрезентационистских объяснений в духе когнитивного классицизма. Такая версия коннекционизма совместима с сетевыми теориями общества, а вместе они образуют методологию, позволяющие элиминировать из философии и когнитивной науки некоторые лишние сущности с помощью анализа объективного и субъективного временных рядов.

Ключевые слова: когнитивная наука, коннекционизм, сети, «я».

Главными концептами когнитивной науки в её классическую эпоху были «репрезентация» и «вычисление (компьютация, процессинг)» — т.е. сознанию приписывалась архитектура, аналогичная фон-неймановской машине: процессор, память, ввод/вывод и т.п. В сознании, согласно этому взгляду, есть структуры, представляющие (репрезентирующие) данные о внешних объектах, и есть процессоры, осуществляющие вычислительные процедуры над этими данными. Когнитивисты различных школ дискутируют относительно того, что представляют собой репрезентации и вычисления, однако данный подход можно считать достаточно общим, чтобы он мог охватить текущий мейнстрим когнитивной науки, включая некоторые разновидности коннекционизма — направления в когнитивных исследованиях, моделирующего ментальные процессы с использованием нейросетевых моделей.

Итак, когнитивный мейнстрим основывается на гипотезе, что ментальная жизнь основана на репрезентациях, аналогичных структурам компьютерных данных, и вычислительных процедурах, аналогичных вычислительным алгоритмам. Однако доминирующая аналогия «сознание—компьютер» оказалась перед лицом конкурирующего подхода, использующего другую аналогию — головной мозг.

Коннекционизм предложил модель, в которой нейроны и их связи играют роль репрезентирующих структур данных, а возбуждения и распространение активации нейронных ансамблей играют роль обрабатывающих алгоритмов. Коннекционистский вариант когнитивной науки, таким образом, основан на своего рода трёхместной аналогии между сознанием, мозгом и компьютером, где каждый из элементов модели может пролить новый свет на представления о других. Таким образом, в когнитивной науке нет единой вычислительной модели когнитивных способностей, поскольку различные школы в компьютерной науке и различные подходы к программированию

предлагают различные модели ментальных процессов. Большинство из нас работает на компьютерах, использующих серийные процессоры, которые в каждый момент времени обрабатывают одну инструкцию (хотя и очень быстро), тогда как мозг и некоторые недавно появившиеся компьютеры основаны на идее параллельного процессинга, который позволяет делать множество операций одновременно.

В последнее время происходит интеграция комплекса когнитивных наук с нейрофизиологией, которая порождает то, что можно было бы назвать комплексом нейронных наук. Эта интеграция происходит отчасти в экспериментальной области, что вызвано появлением новых инструментов изучения мозга, таких как функциональная магнитно-резонансная томография, транскраниальная магнитная стимуляция и оптогенетика. Отчасти эта интеграция происходит на уровне теоретизирования, из-за прогресса в понимании того, как большие популяции нейронов могут выполнять задачи, которые до сих пор объяснялись когнитивными теориями, истолковывающими ментальную жизнь в терминах правил и понятий.

Коннекционистские сети, состоящие из простых узлов и соединений между ними, очень полезны для понимания таких психологических процессов, как видение аспектов, принятие решений, выбор объяснения, и порождение значений в процессе освоения языка. Моделирование различных психологических экспериментов показало психологическую актуальность коннекционистских моделей, которые, однако, лишь очень приблизительно соответствуют реальным нейронным сетям.

Теоретическая нейронаука представляет собой попытку разработать математические и вычислительные теории, а также модели структур и процессов в головном мозге человека и других животных. Она отличается от коннекционизма большей биологической точностью и реалистичностью в моделировании поведения большого числа нейронов, организованных в функционально значимые церебральные области. С развитием науки вычислительные модели мозга приближаются как к более реалистичным имитациям нейронов, отражающим электрический и химический аспекты их деятельности, так и к моделированию взаимодействия между различными областями мозга, таких как гиппокамп и кора головного мозга. Эти модели не являются строгой альтернативой компьютеризированных моделей, пытающихся имитировать логический вывод, следование правилу, концептуальные структуры, аналогии и образы, но должны в конце концов соответствовать им и продемонстрировать нейрофункциональный базис когнитивных процессов.

Когнитивная наука как междисциплинарная область исследований относится к философии тройко. Во-первых, психологические, вычислительные и другие результаты когнитивных научных исследований как бы освежают взгляд на традиционные философские проблемы в гносеологии, метафизике и этике, иногда довольно существенно меняя угол зрения. Во-вторых, сама когнитивная наука может служить объектом философской критики, особенно в отношении её главного предположения, что мышление основано на репрезентациях и вычислениях. В-третьих, когнитивная наука может быть объектом философии науки, где осуществляется анализ её методологии и предпосылок.

Гиперсетевая теория когниции

В (Михайлов 2015a, Михайлов 2015b, Михайлов 2015c, Михайлов 2015d) мною была разработана теоретико-методологическая основа междисциплинарной интеграции, получившая название «гиперсетевой подход» или, может быть, гиперсетевая теория (ГСТ). Эта рабочая рамка представляет собой синтез коннекционизма — нейросетевого подхода в когнитивных науках — и сетевой концепции общества. Она предполагает, что в ходе эволюции обе сети — церебральная и социальная — объединяются в единую суперсистему, давая нашему виду решающее эволюционное преимущество.

Если мы примем метафору нейрокомпьютера — а я её принимаю, — мы должны признать, что у нейрокомпьютера нет операционной системы в привычном смысле. Скорее, он функционирует на основе встроенного ПО узлов своей сети. Это встроенное ПО состоит из примитивных правил порога, после превышения которого возбуждение, полученное от соседних узлов, передаётся дальше, другим соседним узлам. Говоря об «операционной системе» сознания или общества, мы должны понимать, что общество — это не примитивная сеть. Это сеть, состоящая из многих нейрокомпьютеров, т.е. сетей. Поэтому важно подчеркнуть, что метафора операционной системы человеческого мира предполагает, что такая система не программируется и не устанавливается никаким высшим существом. Скорее, она создаётся в режиме обучения в процессе эволюционного взаимодействия человеческого мира со своим окружением. Некоторые называют её культурой.

Достоинство моего подхода, вкратце, состоит в том, что он предлагает единую «сетевую» онтологию для когнитивных и социальных наук, что принципиально позволяет использовать один и тот же формальный аппарат, осуществляя «сквозные» социально-когнитивные исследования.

Как с этой точки зрения мы можем определить, что есть «когниция»? Когнитологический мейнстрим связывает этот концепт с тем, что мы понимаем под словами «помнить», «думать», «рассуждать» и т.п. Однако этот словарь имеет отношение скорее к «народной психологии», чем к науке (см., напр., Stich 1985). Высказывания "народной психологии" имеют тот же эпистемологический статус, что и «превращённые формы» у Маркса: не будучи истинными *sub specie aeternis*, они верны и необходимы в рамках определённой функциональности системы. Однако же, точка зрения науки — это именно точка зрения *sub specie aeternis*, пусть даже только в нормативном плане.

С моей точки зрения, когнитивные акты осуществляются как подстройка гиперсети в соответствии с поведением рационального агента и его последствиями. Паттерн возбуждения определённых нейронов головного мозга, которое имеет место в ходе когнитивных актов, мейнстримом рассматривается как «репрезентация», в то время как, с точки зрения ГСТ, это аналог программной библиотеки, которая хранит типичные процедуры для типичных обстоятельств и активируется в соответствующих случаях. Способность к запоминанию — это самая примитивная когнитивная способность. В то же время, мышление и рассуждения — благодаря системе социальных связей и приобретённым социально-когнитивным функциям, к которым относится, например, язык, — надстраиваются над нею в качестве более общей способности видеть свою собственную и чужую память в качестве части социальной задачи или проблемной ситуации. Если некто встречает существ, которые также могут хранить нечто в памяти,

он или она склонны описывать такую ситуацию примерно так: «некий факт *P* имеет место, и наш собеседник *A* знает, что это так». Тогда, «знать» означает «хранить в памяти тот факт, который сохраняется в памяти широкого круга собеседников».

Такая позиция, безусловно, является функционалистской. Более того, она совмещает ментальный функционализм — понимание интенциональных актов и состояний как функций, независимых от природы носителя, — с лингвистическим: в соответствии с нею, такие слова, как «знать» не только не относятся к ничему, кроме функциональных отношений в мире, но их собственная функция в языке связана не с референциальным значением, но исключительно инструментальна. Такое слово не называет внешний предмет или сущность, не указывает на них, а создаёт то, что можно было бы назвать *коммуникативной модальностью*: «*P* имеет место, но *A* об этом не знает».

Мы также можем говорить о *коммуникативной интенциональности*: это означает, что ментальное состояние или высказывание, порождающее значение, не только направлено на объект, но также с необходимостью предполагает собеседника, который может быть даже несуществующим, но, тем не менее, эпистемически отличаться от говорящего — например, не знать, то, что знает говорящий.

Коннекционизм и антирепрезентационизм

К настоящему моменту накоплен большой массив литературы по проблемам репрезентаций и значений в коннекционизме. Семантический потенциал коннекционизма был предметом дискуссий, по крайней мере, с момента выхода знаменитой работы (*Fodor and Pylyshyn 1988*), выводы которой оспаривались в (*Smolensky 1988*) и многих других работах различных авторов. Фодор и Пилишин, будучи «классическими» когнитивистами, утверждали, что у коннекционизма недостаточно вычислительных средств для объяснения семантической композициональности языка — аргумент, который побуждает исследователей продолжать попытки предложить коннекционистские модели, демонстрирующие композициональность и систематичность (см., напр., *Frank Haselager and van Rooij 2009*).

Но большинство из них неявно понимает значение как пропозициональное значение. Я же предполагаю, что теория значения должна быть построена таким образом, чтобы она могла охватить ещё и «значение» в рамках некоторых неязыковых практик, таких как, скажем, мимика, интонации и паузы в театральной игре, или в квази-языках искусства. Существенной поддержкой такому исследованию будет переформулирование проблемы *qualia* в рамках коннекционистской позиции.

Как нейронные сети могут содержать репрезентации?

По словам Кристиансена и Чейтера, есть две основные опасности для коннекционистской теории значения: её несоизмеримость с классической лингвистической теорией значения на основе символического подхода и, наоборот, непривлекательность символической парадигмы с коннекционистской точки зрения, что влечёт риск потери достижений первой в рамках второй. «В частности, не

представляется возможным обсудить отношения между коннекционизмом и теорией значения без заигрывания с этими опасностями, так как традиционная теория значения связана с символическими, лингвистическими репрезентациями» [Christiansen and Chater 1992, 4].

Существует определенная теоретическая рамка, которая, насколько я могу судить, наиболее распространена и упоминаема в качестве «коннекционистской теории значения». Устойчивый результат обработки входных данных с помощью коннекционистской модели — паттерн активации — это уникальное сочетание значений активации нейронов скрытого слоя, которое математически выражается как вектор. Паттерны располагаются на абстрактном, конструктивно определяемом пространстве состояний сети: «Величины, определяющие пространство состояний данной коннекционистской системы, представляют собой мгновенные уровни активации каждого из узлов в сети. Таким образом, пространство состояний сети часто называют её "активационным пространством"» [Horgan 1997, 20].

Как нейронные сети могут представлять квалиа?

Репрезентационалистская версия коннекционизма работает в отношении строго выделенного фрагмента феноменального опыта (скажем, некоторой квалии) и стабильного паттерна активации нейронной сети. Но с этой позиции, на мой взгляд, не удастся сохранить различие типа и экземпляра при переходе от репрезентации к её нейронному носителю: впечатление от одного и того же объекта не обязательно есть одно и то же впечатление, тогда как нейронный носитель отдельных случаев данной репрезентации может быть одним и тем же физически. Есть два аспекта этой проблемы идентичности типа и экземпляра. Во-первых, существует ли прямая корреляция между определенным качеством феноменального впечатления и определенным сочетанием узлов нейронной сети, связей между ними и их параметров. Эта проблема, на мой взгляд, решается, главным образом, эмпирически. Во-вторых, был ли материальный носитель впечатления идентифицирован математически или физически.

Представим себе следующий научный эксперимент: человек, живущий нормальной жизнью, подсоединяется к устройству, которое сканирует всю его мозговую биоэлектрическую активность и транслирует её, скажем, с помощью Wi-Fi. В соседней комнате находится ещё один человек с имплантированным Wi-Fi приемником и устройством, которое расшифровывает принятые сигналы, превращает их в паттерны нейронных возбуждений и заменяет ими активность собственного мозга этого человека. Если правы сторонники коннекционистского репрезентационализма, и квалиа в буквальном смысле являются паттернами нейронных возбуждений, тогда второй человек должен начать переживать субъективный опыт первого. Но на данном этапе эксперимента есть конкурирующая интерпретация, говорящая, что второй человек видит, слышит и чувствует то, что первый человек видит, слышит и чувствует, но делает это в рамках собственного качественного опыта с его невыразимой природой. И до тех пор, пока не предложено решающего эксперимента для этого вопроса, обе концептуальные интерпретации логически эквивалентны, и, следовательно, коннекционистский репрезентационализм ничего не добавляет к решению философской проблемы квалиа.

Но, на самом деле, идея такого эксперимента есть. Поместим обоих испытуемых в одно и то же окружение. И попробуем изменить приёмник второго человека таким образом, чтобы он мог как бы переключать потоки опыта, то есть подавать на второй в мозг попеременно характеристики своих и чужих паттернов нейронных возбуждений. Такой А/Б-тест или выявит качественные различия в переключаемых потоках опыта, или оставит эти переключения незамеченными. Первый из возможных исходов следует рассматривать как свидетельство в пользу отождествления феноменального опыта и его физического носителя в рамках репрезентационалистской версии коннекционизма.

Один знакомый пример. В компьютерной практике цвета обычно определяются с помощью некоторой традиционной условной математической модели: например, RGB. Но опытные пользователи знают, что различные мониторы, или даже различные веб-браузеры, по-разному воспроизводят цвета с одними и теми же RGB-идентификаторами.

Я хотел бы предложить гипотезу, которая может показаться немного нестрогой на данный момент, но всё же с некоторой степенью вероятности даст надежду на позитивный результат в исследовании квалиа. Нужно подчеркнуть важный момент, а именно то, что нейросеть, которую пытается имитировать коннекционистская машина, функционирует в принципиально аналоговом режиме, в то время как компьютерационалистские (классические когнитивные) модели представляют когницию как цифровой процессинг. Идея исходит из представления, согласно которому нейронная сеть функционирует скорее как причинно-следственный двигатель, чем как вычислительный процессор — что неявно подразумевают компьютерационалистские критики. Или, по словам Голдштейна и Слейтера, «это сенсорно-моторный процессор, который обучен вести себя определенными конкретными способами только через взаимодействие с случайными событиями этого меняющегося мира» (*Goldstein and Slater 1998, 313*). Существует очевидная аналогия с различными техниками записи и передачи звука: в то время как музыка на компакт-дисках или в компьютерных файлах кодируется потоком дискретных цифровых значений и впоследствии декодируется с помощью цифро-аналоговых преобразователей, звук от виниловых пластинок непосредственно вызван движением иглы по мелким изгибам краёв канавки. Многие аудиофилы предпочитают виниловые пластинки, так как считают, что аналоговые записи производят сигнал наивысшего разрешения и могут сохранить «теплоту» и «телесность» — т.е., качественную уникальность исходного звука.

В этом смысле, любые модели, представляющие когнитивные акты как операции со значениями ментальных символов — а именно таковы компьютерационалистские модели, основанные на гипотезе «языка мысли» — это цифровые машины по своей природе, поскольку они представляют значения как дискретные данные, подлежащие конечной алгоритмической обработке. Отсюда систематичность и композициональность, рассматриваемые как главные достоинства любой соответствующей семантической теории. Коннекционистские сети, напротив, имитируют причинно-следственные устройства, индуцирующие волны церебральных активаций, и они, таким образом, способны сохранять предположительно континуальную форму реальной психической деятельности. Когда речь идет о репрезентации языковых единиц, их лексические и пропозициональные значения расположены в виде дискретных точек в пространстве активации сети, имитируя вычисляемые цифровые репрезентации. Но в реальной когнитивной машине мозга всегда есть некий аналоговый «остаток», который воспринимается как своего рода «теплота»

ментальных образов, т.е. их качественная «каково-это-бытность» («what-is-it-likeness»), которая остаётся принципиально невыразимой дискретными средствами языка.

Понятое таким образом, противопоставление пропозициональных значений и квалиа с большей ясностью показывает, что происходит в «комнате Мэри» Фрэнка Джексона (Jackson 1982). Вплоть до момента, когда она покинула свою черно-белую комнату, Мэри знала только дискретные кодированные значения, почерпнутые из существующих теорий цвета. Дальше происходит то, что она обретает полноценное континуальное восприятие цветного мира. Если мы понимаем «знание» как исключительно пропорциональное отношение, то теперь она не *знает* ничего нового. Но есть многое, что было *показано* ей. Если это знание, то относящееся не к типу «знания, что», а, скорее, к типу «видения, как».

Витгенштейн показал, что в языковых играх квалиа иррелевантны — конструирование значения происходит без их участия. Можно ли вывести qualia из коммуникации? Рассмотрим две альтернативы: (а) "Чувствовать красное" и (б) "понимать, что такое "красное"". Возможно ли (а) без (б)?

Что видит начитанная Мэри, выйдя из чёрно-белой комнаты и впервые увидев пожарный гидрант или спелый помидор, если она знала заранее, что люди наделяют их одним и тем же свойством, которое ей до сих пор было "субъективно" недоступно? Воспользуемся игровым методом Витгенштейна и представим себе, что мы учим ребёнка: сложи сюда всё красное, а сюда — всё круглое. Почему мы считаем, что круглость, в отличие от красности — объективное качество? У нас есть язык (концепции) для его экспликации в диалоге: мы можем сказать, что круглый - это равноудалённый от центра, квадрат - это равносторонний прямоугольник и т. п. Для "красного" просто нет языка экспликации, но есть практики различения. Отсюда - представление о личном, закрытом для других контейнере, где содержатся квалиа. Наш собеседник скажет: «Но я же непосредственно и очевидно воспринимаю красный цвет!» Ответим: «Но вы непосредственно и очевидно воспринимаете и круглую форму и можете отличить её от квадратной. Только в этом случае у вас ещё есть и словесные инструменты для обоснования такого различения — язык экспликации. А для цветов его нет...».

Действительно ли компьютеризация настолько близка функционализму?

Нам необходимо уточнить логические отношения между компьютеризацией, функционализмом и репрезентационализмом. Компьютеризация утверждает, что сознание выполняет формальные вычисления над символическими репрезентациями. Таким образом, она предполагает репрезентационализм. Но предполагает ли она также и функционализм? Альтернативы функционализму представлены или бихевиоризмом, или натурализмом. Компьютеризация не может предполагать бихевиоризм, поскольку она возникла как попытка преодолеть ограничения последнего. Она также не может предполагать натурализм, поскольку, рассматривая психические состояния как сущности, для объяснения их она должна была бы стремиться скорее к причинно-следственной теории, а не к исчислению. Таким образом, компьютеризация, действительно, предполагает функционализм.

Более или менее ясно, что функционализм и репрезентационализм, каждый сам по себе, не подразумевают компьютеризацию. Но подразумевают ли они это в

сочетании? Мой ответ — да, если предполагается, что репрезентации обладают семантикой в виде некоей референции к чему-то внешнему, функционально являясь, таким образом, символами. В противном случае, предположим, что они таковыми не являются. Что бы заставило нас считать их репрезентациями? Как Пирсова теория семиозиса, так и наш здравый смысл говорят нам, что для того, чтобы нечто могло быть распознаваемой репрезентацией, оно должно быть, по крайней мере в возможности, читаемо каким-либо интерпретатором, вооружённым формальными правилами понимания. Другими словами, предикат «быть репрезентацией чего-то», на самом деле, трёхместный: он должен читаться как «быть репрезентацией чего-то для кого-то». Этот кто-то есть тот, кто декодирует значение данной репрезентации. Этот кто-то не обязательно должен быть человеком: на этом месте может быть абстрактный процессор, способный интерпретировать представления в соответствии с некоторыми правилами. Таким образом, без некоторой подсистемы означивания и извлечения значений никакие отпечатки одного объекта в другом не должны пониматься как репрезентации.

Отсюда можно сделать вывод, что компьютеризация вытекает из конъюнкции функционализма и репрезентационализма. Можно также заключить, таким образом, что намерение расстаться с компьютеризацией ради какой-либо другой парадигмы предполагает опровержение одного из конъюнктов. В нашем случае, мы оставляем позади репрезентационализм. В то время как компьютеризация полагается на репрезентационалистское видение, коннекционизм позволяет как сохранить репрезентации в качестве объяснительного принципа, так и обойтись без них. Эти версии различны в ответе на вопрос, могут ли паттерны активации рассматриваться как репрезентации внешних сущностей, от которых получают входные данные, или их роль чисто инструментальна, тогда как значение возникает скорее как функция деятельности актора в той или иной среде.

Может коннекционизм обойтись без репрезентационализма?

В некоторых публикациях анти-репрезентационализм Витгенштейна сочетается с анти-репрезентационалистской версией коннекционизма. К тому же, так называемые «новые коннекционисты» нашли какие-то пути адаптации Выготского (см., например, *Michaelian and Sutton 2013, 4*). В своем проекте «радикального коннекционизма», О'Брайен и Опи (*O'Brien and Opie 2002*) используют теорию интериоризации речи Выготского с одной важной новацией: не важно, в обычной ли внешней, или в интериоризированной форме, но язык понимается не как инструмент репрезентаций, но только как набор коммуникативных сигналов, направленных на манипулирование ментальным содержанием.

Точкой соприкосновения новых веяний в коннекционизме с Витгенштейном является теория значения. Витгенштейн решительно вывел значение из скрытой области «метальных состояний» в очевидный мир человеческих практик. В то же время, критики философских претензий коннекционизма отмечают, что семантические интерпретации нейронных репрезентаций основаны исключительно на структурных сходствах и причинно-следственных корреляциях внутренних паттернов активации и определённых входных данных, и при этом отсутствует какая-либо однозначная связь с репрезентируемыми объективными сущностями или категориями. Это, по их мнению, делает семантическую теорию, которую возможно связать с коннекционизмом, слабее, чем таковая компьютеризация.

Я же утверждаю, что, напротив, это и есть то, что делает философские претензии коннекционизма сравнительно сильнее, если мы только покинем

репрезентационалистскую парадигму. Теоретические сходства между коннекционизмом и Витгенштейном подробно обсуждаются в (*Mills 1993*). В частности, Миллз подчеркивает, что оба подхода подразумевают отказ от какой-либо формальной теории исчисления: почерпнутой из компьютеризации в первом случае и происходящей из собственных прежних взглядов Витгенштейна во втором. Далее он добавляет, что оба подхода во многом полагаются на обучение, существенную часть которого составляет тренировка на конкретных примерах. И, наконец, он указывает на то, что оба подхода рассматривают обобщения как перекрещивающиеся значения в стиле «семейных сходств» (*Mills 1993, 141-142*).

В то же время, Миллз выявляет существенное различие между этими двумя подходами: «Общее различие в исследовательских подходах, очень грубо говоря, состоит в том, что, если коннекционистское понимание владения языком основано на описании его в терминах внутренних состояний, подход Витгенштейна — это описание социального употребления слов» (*Mills 1993, 144*). В основе разницы подходов — различные теории значения: если более традиционный коннекционизм придерживается понимания значения как отношения репрезентации к тому, что ею репрезентируется, Витгенштейн идентифицирует его с моделями поведения в языковых играх, которые схватываются и обозначаются термином «употребление». Миллз предлагает определенный принцип комплементарности: «описательный подход Витгенштейна понимается как общая характеристика задачи, которую которую должна выполнить объяснительная теория владения языком, в то время как коннекционизм решает эту задачу через моделирование вычислительных процессов, которые обеспечивают это владение. Понятые таким образом, оба подхода взаимно дополняют друг друга, и то, как они это делают, представляет существенный теоретический интерес» (*Mills 1993, 145*).

Но эта роль вычислительных процессов в обеспечении владения языком не столь проста, как это может показаться с прямолинейно репрезентационалистской точки зрения. Говоря о когнитивных механизмах, с помощью которых мы имеем дело с лексическими и пропозициональными значениями, не следует полагать, что отношение между паттерном активации, представляющим слово, и тем, который представляет референт этого слова, является прямой репрезентацией значения. Скорее, коннекционистские состояния активации представляют собой средства обучения и запоминания, вместе будучи платформой для осмысленного участия человека в языковой игре. Или, как выразил это Миллз, «коннекционизм следует понимать как позицию, согласно которой понимание слова не идентично определенным нейрофизиологическим состояниям, а представляет собой способность правильно использовать слово, и что связанное с ней нейрофизиологическое состояние есть основание этой способности» (*Mills 1993, 150*).

Рассуждая о значении идей Витгенштейна для осмысленной коммуникации в интернете, Дэвид Ауэрбах пишет: «Слово есть развивающаяся сущность, которая несет свою историю с собой во времени» (*Auerbach 2015*). Вряд ли он именно это имел в виду, но это звучит так, как будто слово обладает контейнером для своей собственной истории. Между тем оно, очевидно, не больше, чем просто звук или надпись без каких-либо дополнительных технических устройств. С помощью чего, говоря онтологически, эта «своя история» переносится сквозь время? Кто-то может сказать, что таковой средой является память определенных людей, которые фактически используют это слово в данный момент. Тогда слово становится как бы лучом фонарика, в котором кружат мотыльки, обладающие памятью. Но, имея изменчивую память приходящих и уходящих индивидов, трудно обосновать устойчивость словесных значений. Ответ заключается в том, что слово является управляющей функцией в более или менее устойчивой структуре человеческих сетевых соединений.

Возвращаясь к известному спору о композициональности, мы можем найти основания утверждать, что если мы обнаружили, что коннекционистские паттерны активации являются не прямыми репрезентациями, а скорее физиологическими предпосылками осмысленного поведения, то требование систематичности, отстаиваемое Фодором и Пилишиным, здесь не релевантно. Я склонен согласиться с Голдштайном и Слейтером в том, что коннекционистская «семантика не композициональна: [в ней] нет отображения высказывания 'а есть игра' на [объект] а и на некоторое свойство 'быть игрой' — или на а и на некоторый набор всех игр, — так как прежде нашего согласованного использования [этих слов] нет ничего, что устанавливает, как мы должны использовать 'быть игрой', и, таким образом, что это значит» (*Goldstein and Slater 1998, 302*).

Может ли коннекционизм быть объединён с непропозициональной и нементалистской теорией значения?

Таким образом, компьютерно-инструментальный подход известен своими хорошо разработанными инструментами анализа владения языком, и его научные заслуги многочисленны. Его слабым местом является то, что, объясняя, как люди развили эту способность, он постулирует некоторые лингвоподобные структуры на когнитивном этапе головного мозга, как будто язык всегда там находился. Отсюда начинаются некоторые проблемы этой теории, связанные с объяснением врожденности языковых структур и грамматического разнообразия реальных языков. Напротив, коннекционизм разработал концептуальную основу, которая особенно хорошо подходит для объяснения гибкости обучения, а также освоения грамматических форм и категорий. Но эта система не очень хорошо приспособлена для объяснения некоторых аналитических свойств семантики, приписываемых человеческим языкам. Исходя из этих соображений, выбор кажется довольно сложной задачей для философа. Но, как это часто бывает, правильный выбор заключается в том, чтобы просто изменить нашу точку зрения. Описанное выше противопоставление имеет смысл, если только мы принимаем значение за свойство некоторых внутренних репрезентаций. Но эта точка зрения не является ни логически необходимой, ни эмпирически доказанной. К тому же, согласно Витгенштейну, полагая так, мы становимся жертвами навязанной нам грамматики. Как показали некоторые концептуальные исследования, существует определенная взаимодополняемость между его взглядами и коннекционизмом.

Вся машинерия нейронной сети, как она представлена в PDP-моделях, существует не для процессинга значений. Она, как широко известно, не имеет конкурентов в эффективности распознавания речи и образов, обучения, спонтанной категоризации и обобщении. Как видно внимательному наблюдателю, это те самые способности, которые делают возможным наиболее известные языковые игры, описанные Витгенштейном, такие продолжение ряда натуральных чисел. Кроме того, это те способности, которые могут помочь правильно истолковать то, что имел в виду «ранний» Витгенштейн под «показыванием», противопоставляя его высказыванию, намечая, таким образом, перспективы для теории непропозиционального значения.

Это мощное сочетание приводит нас к такому заметному прогрессу, как обретение удовлетворительной теории квалиа, а также к возможности отказаться от пропозициональных рамок в теории значения. Но на этом пути еще многое предстоит сделать. В частности, коннекционистская онтология может быть дополнена некоторой некомпозициональной семантикой. Основным результатом этого исследования, как представляется, является понимание языковых значений, которые входят в наши аналоговые бортовые компьютеры, в качестве точечных воздействий социальной среды,

которые в принципе готовы к своего рода цифровой обработке в компьютерном стиле. Они надстраиваются над волнами мозговой активации как их супервентная и социально контролируемая функция.

Это именно тот самый момент, когда, как однажды сказал Витгенштейн, «объяснения должны заканчиваться». В этом пункте начинается коннекционистская когнитивная наука, не оставляя философам никаких сомнений в том, что на их законной территории никаких загадок не осталось.

Проблема «я»

Однако часть философов скажет, что на их законной территории осталась неизбывная загадка и любимая игрушка многих философских систем — человеческое «я», известное также как «самость» («self») или апперцепция. В соответствии с классической установкой в философии, «я» играет важную роль в структурировании сознания и в познании, осуществляя некий синтез впечатлений или значений.

На мой взгляд, в основе такой позиции лежит эссенциалистский подход к языку: убеждение в том, что всякое слово называет некоторую сущность. Каждый из нас называет себя «я», а поэтому — «у каждого есть своё Я». Философский «я»-дискурс, таким образом, представляет собой теорию, описывающую в третьем грамматическом лице значения терминов первого лица.

В то же время, есть подлинная загадка человеческого «я». Это тот вопрос, который задаёт себе (или окружающим) каждый ребёнок в какой-то момент: почему я — это я? Что будет, когда я умру? Однако Витгенштейн времён «Трактата» отнёс бы предмет и основания подобных вопросов к сфере мистического и невыразимого и потому не имеющего научного решения.

Вместе с тем, достойным рассмотрения выглядит то, что я назвал бы *парадоксом философской эгологии*:

ПФЭ: Я не могу быть бессмертным как биологическое существо, как вещь этого мира, но я не могу быть смертным как системообразующий принцип этого мира.

Парадокс основан на (1) допущении сущностного единства «я» и (2) понимании его как «контейнера» субъективных переживаний, которые (3) упорядочены во времени.

Может ли «я» умереть? Для выявления возможности ответа на этот вопрос выделим два аспекта, в которых может употребляться «я»: коммуникативный и качественный. В первом из них ПФЭ решается тривиально и нефилософски: «я» — это прагматический оператор языка, означающий "тот, кто это говорит". Поэтому никаким «системообразующим принципом» этот оператор не является, а биологическое существо в должное время умирает, как ему и положено.

Во втором аспекте ПФЭ не имеет места: представления о границах жизни — о дате рождения и дате смерти — целиком принадлежат коммуникативному аспекту. Но "коммуникативное Я" умереть не может, так же как не может перестать существовать «здесь» и «теперь».

Рассмотрим два ряда событий и фактов - субъективный (С) и объективный (О). Первый соотнесён с моими переживаниями и памятью, второй имеет языковую, коммуникативную, теоретическую природу. К первому относится, например, внешний облик моих родителей в моей памяти, ко второму - дата моего рождения.

Событие моего рождения (СМР) - событие О-ряда, которое не имеет места в С-ряду. Близкие к нему по времени события С-ряда, содержащиеся в моей памяти, по мере

приближения к СМР становятся всё менее чётко воспринимаемыми и всё более зависимыми от событий О-ряда: сообщений других людей, документальных записей и т. п.

Помимо этого, мы сталкиваемся с временной однонаправленностью сознания: память хранит события прошлого, но не будущего. Онтология прошлого предстаёт как соотносённость событий О-ряда с содержанием памяти. Онтология будущего, в свою очередь, — как проекция возможных событий О-ряда на предполагаемые факты С-ряда. Прошлое дано в том или ином виде, будущее только предполагается, у него нет актуального С-ряда.

Т. о., СМР не сопоставлено ни одно впечатление С-ряда. Но то же можно сказать о событии моей смерти (СМС). Я предполагаю, что в О-ряду СМС будет соответствовать определённая дата. И я могу предположить, что в этот момент времени ему будут сопоставлены какие-то переживания С-ряда. Но в каждый настоящий момент СМС является таким же О-событием без сопоставленных ему С-впечатлений, как и СМР. Причём первое — исключительно концептуальная конструкция: события будущего предполагаются, но не переживаются. Чистая же "теоретичность" СМР — эмпирическая загадка: почему я ничего не помню об этом?

Однако несомненно то, что ни одной из О-границ моей жизни (ни СМР, ни СМС) не соответствует ни одно событие С-ряда.

Итак, я имею два ряда содержаний моего сознания: С-ряд, состоящий из "здесь и теперь" и памяти, но не содержащий никаких представлений о собственных границах; и О-ряд, заставляющий меня полагать, что я в определённый момент родился и в определённый момент умру, но всецело зависящий от культуры и обучения. На какой из них нужно положиться?

В «Загадке человеческого Я» (Михайлов 1976) Феликс Михайлов пишет, что наглядная разгадка — это слепоглухонемые дети, успешно социализированные и получившие хорошее образование. В каком-то смысле это верно, поскольку показывает, что подлинным значением термина "я" не является ничто из относящегося к С-ряду: ни его события, ни сам он в целом. Людей, не обладающих "нормальным" С-рядом, можно научить правильно пользоваться словом "я". Думаю, что этому же можно научить и компьютер в исключительно коммуникационном аспекте.

А в качественном? Очевидно, что компьютер можно научить хранить в памяти изображения и совершать какие-то действия, если эти изображения узнаваемо похожи на актуально воспринимаемые образы, чтобы, например, смоделировать поведение ребёнка, который тянется к матери, узнав её. Но можно ли научить компьютер сохранять в памяти "образы" его включения и отключения?

Опираясь на свой С-ряд, компьютер никогда не сможет определить, что значит для него "быть выключенным". Но из этого не следует, что он будет включён вечно.

Итак, можно утверждать, что в ПФЭ смешиваются два концепта: (я1) "я" как особая позиция говорящего в коммуникативной системе и (я2) приватные ощущения и переживания (проприоцепция и т.п.). Я1 не может умереть, поскольку не является определённым существом или вещью, а только способом языковой автореференции говорящего. А носитель приватного содержания (я2) умереть вполне может и, скорее всего, это сделает, но не может в каждый данный момент отразить этот факт в своих впечатлениях С-ряда.

Кроме того, высказывание "я умру" по смыслу эквивалентно высказыванию "моё тело умрёт". Чего нельзя сказать, например, о высказывании "я знаю": всё же это не моё тело знает, что я родился в определённый день определённого месяца и года. Физикалист, однако, скажет, что именно тело это знает, поскольку мозг — часть тела. Контраргумент может быть следующим: "я знаю", "он знает" — не внутренние состояния

тел, а их позиции в системе коммуникации как акторов. Моё тело, включая мозг, хранит процедуры, благодаря которым я могу осмысленно включиться в межличностную коммуникацию, где другие могут сказать обо мне: «Он знает, что *p*».

Ещё раз: "я" (в смысле я1) не может умереть, как не может пройти "сейчас" и нельзя покинуть "здесь". Смерть становится проблематичной, когда "я1" смешивают с "я2": психологическим "контейнером" событий С-ряда. Первое не может умереть по чисто концептуальным основаниям, а второе не способно на качественную репрезентацию события собственной смерти.

Рассмотрим теперь трансценденталистские притязания на философскую исключительность «я». Введём для этого различие трансцендентального (Ят) и психологического (Яп) «я». Когда говорят, что Ят есть конститутивный принцип мира, условие единства опыта и т. п., смешивают по ложной аналогии я2 (в предыдущей части анализа) — "только я могу иметь мои ощущения" — с я1, придавая последнему свойства исключительного собственника семантики осмысленных высказываний о мире, каковым оно не является.

Само время является одной из опорных конструкций О-ряда: если С-переживаниям не сопоставлены О-события, локализованные на временной шкале, то я, скорее всего, не смогу определить, какие из впечатлений, хранимых моей памятью, получены "раньше", а какие -- "позже". Я могу только отличить С-впечатления "здесь и теперь" от более или менее отчётливых содержаний памяти.

Можно было бы сказать, что ПФЭ не имеет места, поскольку С-ряд, переживаемый я2, не образует мир (мир упорядочивается в коммуникации), а я1 не является никаким системообразующим принципом.

Но на чём же основана неистребимость концепции "бессмертия души"? В её основе — представление о привилегированном положении моего тела как единственного наблюдательного пункта моего я2: если тело умирает, то прекращается киносеанс, а кто же тогда встанет и выйдет из зала?

Ответ может состоять в следующем. Концепция "жизни" принадлежит я1, поскольку, как уже было сказано, только на временной шкале О-ряда возможны "до" и "после". С-ряд, принадлежащий я2, локализован на этой шкале только во взаимодействии с культурой, а сам по себе не имеет временных характеристик, и поэтому не имеет смысла говорить о его "прекращении".

Мой интенционально предполагаемый собеседник говорит: «Я не могу представить себе свою смерть и то, что будет после неё.» — А могу ли я представить себе своё рождение и то, что было до него? Стоит лишь абстрагироваться от представления об однонаправленности времени, и мистерии становятся обыденностью. Если душа должна пережить тело, потому что невозможно представить себе полное отсутствие представлений, то почему нас не смущает полное отсутствие представлений о её дотелесном существовании, да и о моменте воплощения?

"До" и "после" — навигационные инструменты, которыми "душа" (я2) изначально и не владеет. Только в системах координат культуры оказывается, что всё то в С-ряду, что не локализовано в "здесь и сейчас", представляет собой или содержание снов, или память о событиях "прошлого". Поэтому, "что будет после моей смерти?" -- это не тот вопрос, ответ на который стоит искать в тайниках моей души. Качественное «я» умереть может, но не может иметь этого опыта, также как не имеет опыта собственного рождения. Я знаю, что я родился, но не помню этого. Я знаю, что умру, но... Если "душа" может иметь посмертные впечатления, почему у неё нет пренатальных? Загадка коммуникативного «я» тривиальна и решается эмпирически. Загадка качественного «я» относится к сфере мистического, невыразимого, и средствами известных нам науки и философии не решается.

Объектное, эссенциалистское понимание «я» может предполагать следующие варианты:

(1) «я» как термин, относящийся к физическому телу, что, по видимости, происходит в высказываниях типа «Я болею» или «Я потолстел»;

(2) «я» как термин, обозначающий эмпирическое содержание памяти и сознания, например, «В школе меня не любили», «Я плохо запоминаю цифры»; и, наконец,

(3) декартовское *ego cogito*, а также кантианское понимание «я» как априорного представления, необходимо сопровождающего любой акт мышления.

Рассмотрим первый случай. Предположим, случилось страшное, и я прибавил в весе. Тогда высказывание «Игорь Михайлов поправился» (в плохом смысле слова) будет истинно как в устах другого человека, так и в моих устах, в последнем случае лишь немного странным по способу выражения. Тогда как «Я поправился» будет истинно только в моих устах. Это ясно указывает на то, что условия истинности высказывания с «я» лежат в сфере прагматики, то есть, свойств конкретной коммуникационной ситуации, а не только в сфере семантики, как в случае и «именным» высказыванием. Если «я» и «Игорь Михайлов» не являются синонимами, а демонстрируют взаимозаменяемость только при определённых прагматических условиях, то логично было бы рассматривать роль термина «я» в языке в качестве своего рода прагматического оператора. Можно говорить о таких стилистически обусловленных синонимах этого термина, как «ваш покорный слуга», «автор этих строк», а также английское 'this man' (хотя под другим углом зрения и при другой прагматике их можно рассматривать как дескрипции).

Нетрудно заметить, что те же аргументы имеют силу и для варианта (2). То есть, вне зависимости от того, о каких эмпирических свойствах личности — физических или интенциональных — мы говорим, «я» может заменить другой термин, если и только если заменяемый термин является именем или дескрипцией говорящего.

Теперь о декартовском, кантианском, гуссерлианском — или вообще трансцендентальном, — понимании «я». Не вдаваясь в тонкости, все три подхода можно объединить на одном общем основании: имплицитное присутствие мыслящего «я», лишённого каких бы то ни было психологических определённостей, полагается в всех трёх традициях — картезианской, кантианской и феноменологической — совершенно необходимым для того, чтобы «чистое» мышление было возможным. Во всех трёх традициях трансцендентальное «я» получается путём освобождения эмпирического «я» от несущественных свойств, с тем чтобы оставить то единственное, которое выражает его сущность: мышление у Декарта, внутренняя связь суждений у Канта и направленность на предмет у Гуссерля. В свою очередь, освобождая все три концепции от несущественных различий между ними, думаю, было бы корректно сказать, что — Гуссерль выразил суть точнее всех — сущность чистого «я» состоит в его «о-чём-то-бытности»¹: я не могу придумать лучшего перевода слова 'aboutness', придуманного Сёрлем.

В этом пункте возникает законное подозрение: не используется ли здесь слово «я» вне его категориальной принадлежности, а именно — как существительное, как синоним слова «сознание»? Если это не так, то необходимо показать, что сознание раскрывает свою «о-чём-то-бытность» — то есть, интенциональность, — только в перспективе

¹ В этом пункте завзятые феноменологи могут обрушиться на меня с критикой, поскольку в переусложнённой, на мой взгляд, системе Гуссерля чистое «я» является не обладателем, а скорее условием интенциональных содержаний: оно есть субъект сознательных актов (*cogitationes*), которые конституируют интенциональные предметы. Но если его деятельность состоит в этом и более ни в чём, то логически всю конструкцию вполне можно упростить.

первого лица, только в моей собственной интроспекции. Этот подход в какой-то мере роднит Декарта с Гуссерлем, тогда как у Канта он не столь очевиден: именно поэтому концепция трансцендентального единства апперцепции выглядит у него немного *ad hoc*.

Попробуем использовать естественную асимметрию грамматических лиц: эффект, проявляющийся в том, что, если высказывание «Он не знает, что $2*2=4$ » имеет смысл, то высказывание «Я не знаю, что $2*2=4$ » его не имеет. Но «знание» всё же выражает некую эпистемическую — или, как я уверен, коммуникационную, — позицию, тесно связанную с истинностью подчинённого предложения. Мы же, вслед за тремя великими философами, пытаемся увидеть сущность чистого «я» в более простом и абстрактном свойстве — свойстве «быть о чём-то», что необязательно должно быть связано с истинностью или реальностью этого чего-то. Для этой цели лучше подходит интенциональный предикат «думать» («полагать») — что, по видимому, и имел в виду Декарт.

«Я не думаю, что расстояние от Земли до Луны составляет 384 400 км» — «Ты не думаешь, что расстояние от Земли до Луны составляет 384 400 км» — «Он (она) не думает...» Как, видим, есть вполне представимые ситуации, в которых это предложение в любой из личных грамматических форм окажется осмысленным. А следовательно, использование личного местоимения в качестве философской категории представляет собой некий атавистический след эпохи интроспекционизма и, в общем и целом, не имеет смысла.

Наш беглый грамматический анализ показывает, что перспектива первого лица имеет значение для некоторых, но не для всех интенциональных актов. А следовательно, «трансцендентальное» присутствие «я» во всех проявлениях сознания не требуется, что автоматически лишает его статуса трансцендентальности.

Иначе говоря, наша техника состоит в следующем. Трансцендентальность предполагает аподиктичность — то есть, всеобщность и необходимость. Но если присутствие чего бы то ни было в сознании всеобще и необходимо, то отрицание этого присутствия должно быть невысказуемо (как существование физического мира вне пространства и времени). И тогда достаточно показать мыслимость отрицания некоторых интенциональных предикатов в первом лице, чтобы лишить чистое «я» статуса всеобщности и необходимости, а следовательно — трансцендентальности. Но, поскольку мы это продемонстрировали, то теперь мы вправе утверждать, что трансцендентальное «я» — это философская фикция.

Со времён Витгенштейна известно, что — по крайней мере, некоторые — философские проблемы обусловлены неправильным пониманием категориальной структуры и функционирования нашего естественного языка. Символы, которые мы используем, ничего не говорят об их категориальной принадлежности. Отсюда возникает иллюзия, что слова «стол», «я», «человек», «мышление» функционируют одинаково, то есть обозначают некоторые объективно существующие сущности, и задача исследователя — их определить и классифицировать. На самом же деле — в некотором приближении — мы можем говорить о слова-константах, словах-переменных и словах-операторах. И, несмотря на наше интуитивное стремление отнести все имена нарицательные ко второй категории, некоторые из них, а также многие, если не все, местоимения, относятся к третьей. Так, слово «я» вовсе не обозначает никакой сущности или состояния (вроде трансцендентального единства апперцепции), а является оператором аутореференции — «тот, кто это сейчас говорит», подобно оператору 'this' в некоторых языках программирования². Аналогично, слово «человек» может быть

² см. об этом Михайлов, И. Ф. Субъектность как выбор: теория и практика // Человек вчера и сегодня.

осмыслено как оператор родовой аутореференции — «я и любой, кто в принципе может понять то, что я сейчас говорю»³. Отличие оператора от переменной состоит в том, что его значение относится не к семантике, а к прагматике — к свойствам конкретной коммуникационной ситуации. Поэтому, что такое «человек» определяется параметрами такой ситуации и объединяется отношениями «семейного сходства»: каждый контекст употребления слова «человек» чем-то похож на некоторый другой, но нет ничего такого, чем они все были бы похожи друг на друга. «Человек» прагматически определяет коммуникационный универсум: это тот, кто в принципе может участвовать в коммуникации с помощью данного языка.

* * *

Продемонстрированные здесь рассуждения представляют собой концептуальную попытку расчистить философские завалы с тем, чтобы представители когнитивной науки могли спокойно продолжить движение в направлении демистификации человеческой души, не отвлекаясь на ложные сущности. Мне, как я надеюсь, удалось представить сильные аргументы в пользу следующих положений:

1. Коннекционизм как методологическая установка в когнитивной науке, приближающая нас к пониманию реального функционирования когнитивного аппарата человека, имеет ещё и то преимущество перед конкурирующими теориями, что его концептуальный аппарат позволяет осуществить «бесшовный» междисциплинарный синтез когнитивных и социальных наук на основе общей теории сетей.

2. Коннекционизм демонстрирует свои настоящие сильные стороны в конкуренции с компьютеризмом, будучи очищен от репрезентационалистского понимания таких ментальных феноменов, как квалиа и языковые значения. Нерепрезентационалистский вариант коннекционизма предполагает, что интенциональность и понимание языка являются супервентными эффектами когнитивно-социальной гиперсети.

3. Коннекционизм демонстрирует свои настоящие сильные стороны в конкуренции с компьютеризмом, будучи очищен от репрезентационалистского понимания таких ментальных феноменов, как квалиа и языковые значения. Нерепрезентационалистский вариант коннекционизма предполагает, что интенциональность и понимание языка являются супервентными эффектами когнитивно-социальной гиперсети.

4. Общая антирепрезентационалистская позиция позволяет сократить набор онтологических сущностей когнитивной теории за счёт отказа от представления о «я» или «самости» как о неизбежном объекте когнитивного исследования.

Междисциплинарные исследования. Выпуск 8. М.: ИФРАН, 2013

³ Не случайно, например, в немецком языке безличное местоимение неотлично по произношению и почти неотлично по написанию от слова «человек».

Литература

Auerbach D. (2015) The Limits of Language. Wittgenstein explains why we always misunderstand one another on the Internet. URL: http://www.slate.com/articles/life/classes/2015/09/take_a_wittgenstein_class_he_explains_the_problems_of_translating_language.html

Christiansen, Morten H., Chater, Nick. (1992) Connectionism, Learning and Meaning. *Connection Science*, 4, 227-252. 1992. Cited as a PDF document downloaded from URL: <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.127.8951&rep=rep1&type=pdf>

Churchland, Paul M. (1996) *Engine of Reason, the Seat of the Soul*. Cambridge, Massachusetts — Bradford Books, 1996

Ellis, Nick C. (1998) Emergentism, Connectionism and Language Learning. *Language Learning* 48:4, December 1998, pp. 631–664

Fodor, J. A., & Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3–71

Stefan L. Frank *, Willem F.G. Haselager, Iris van Rooij (2009). Connectionist semantic systematicity. *Cognition* 110 (2009) 358–379

Garzon, Francisco Calvo, Rodriguez, Angel Garcia (2009). Where is Cognitive Science Heading? *Minds & Machines* 19:301–318 DOI 10.1007/s11023-009-9157-3

Goldstein, L., Slater H. (1998) Wittgenstein, Semantics and Connectionism. *Philosophical Investigations*, Volume 21, Issue 4, pages 293–314, October 1998

HORGAN, T. (1997) CONNECTIONISM and the PHILOSOPHICAL FOUNDATIONS of COGNITIVE SCIENCE. *Metaphilosophy*. Vol. 28, Nos. 1&2, January/April 1997 0026–1068. Cited as a PDF document downloaded from URL: <http://thorgan.faculty.arizona.edu/sites/thorgan.faculty.arizona.edu/files/Connectionism%20and%20the%20Philosophical%20Foundations%20of%20Cognitive%20Science.pdf>

Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Quarterly* 32, 127–36

Joanisse, Marc F., McClelland, James L. (2015) Connectionist perspectives on language learning, representation and processing. *WIREs Cogn Sci* 2015. doi: 10.1002/wcs.1340

Michaelian K., Sutton J. (2013). Distributed Cognition and Memory Research: History and Current Directions. *Rev.Phil.Psych.* 4:1–24 DOI 10.1007/s13164-013-0131-x

Mills, S. (1993) Wittgenstein and Connectionism: a Significant Complementarity? *Royal Institute of Philosophy Supplement* 02/1993; 34:137 - 157

O'Brien, Gerard, Opie, Jon (1999). Connectionist Theory of Phenomenal Experience. *Behavioral and Brain Sciences* 22:127-48. Cited as a PDF document downloaded from URL: <http://srsc.ulb.ac.be/AI/papers/BBS-OO.pdf>

O'Brien, Gerard, Opie, Jon (2002). Radical connectionism: thinking with (not in) language. *Language & Communication*, Volume 22, Issue 3, July 2002, Pages 313–329

Smolensky, P. (1988) On the Proper Treatment of Connectionism. *Behavioral and Brain Sciences*. 11, 1–74

Stich, Stephen P. (1985) *From Folk Psychology to Cognitive Science: The Case Against Belief*. Cambridge: The MIT Press, 1985

Михайлов Ф.Т. (1976) Загадка человеческого Я. М., 1976

Михайлов И. Ф. (2015а) *Человек, сознание, сети*. М.: ИФРАН, 2015

Михайлов И.Ф. (2015b) К гиперсетевой теории сознания, *Вопросы философии* Москва, 2015, 11, С. 87–98

Михайлов И.Ф. (2015с) *Человек в сетевом обществе, Новое в науках о человеке. К 85-летию со дня рождения академика И. Т. Фролова / Отв. Ред. Г. Л. Белкина, ЛЕНАНД, Москва, 2015 , 364-387*

Михайлов И.Ф. (2015d) *Коммуникация и онтология мышления // Человек. No6. 2015. С. 23–31*