

Вопросы философии. 2017. № 4. С. 81–94

**Таинство естественной семантики:
трансцендентальное измерение смысла и проблема искусственного интеллекта***

Д. Э. Гаспарян

В настоящем исследовании обосновывается, что задача эмуляции искусственного интеллекта по большей части ориентирована на так называемое натуралистское понимание смысла и его отождествление со значением. Между тем, если предположить, что категория смысла обладает признаками трансцендентальности и не тождественна значению, то, возможно, удастся пролить свет на некоторые проблемы эмулирования естественной семантики, подчеркнув их принципиальный характер. В работе обосновывается ценностный и отчасти целевой аспекты смысла, которые в совокупности позволяют ввести понятие “метафизического смысла”. Обосновывается гипотеза, что именно “метафизический смысл” лежит в основе любых понимаемых актов, доступных человеческому сознанию. При этом в силу ценностной и целевой природы метафизического смысла задача его компьютеризации должна была бы означать эмуляцию некоторой системы ценностей или этики, что представляет отдельную большую проблему.

КЛЮЧЕВЫЕ СЛОВА: смысл, контекст, трансцендентализм, естественная семантика, искусственный интеллект, синтаксис, семантика, значение, лингвистическая прагматика.

ГАСПАРЯН Диана Эдиковна – кандидат философских наук, старший научный сотрудник в Центре фундаментальных исследований лаборатории исследования философии, доцент факультета гуманитарных наук Школы философии Национального исследовательского университета “Высшая школа экономики”.

anaid6@yandex.ru

Статья поступила в редакцию 20 ноября 2016 г.

Цитирование: *Гаспарян Д. Э.* Таинство естественной семантики: трансцендентальное измерение смысла и проблема искусственного интеллекта // *Вопросы философии. 2017. № 4. С. 81–94.*

Введение: сохранение проблемы семантических компетенций ИИ

Большинство дискуссий, посвященных перспективам развития искусственного интеллекта (далее везде ИИ), так или иначе, вращаются вокруг вопроса: может ли программа эмуляции интеллекта оказаться восприимчивой к естественной семантике или умеет справляться лишь с синтаксическими задачами? Этот вопрос играет ключевую

* Статья подготовлена в результате проведения исследования (№ проекта 16-01-0032) в рамках Программы «Научный фонд Национального исследовательского университета “Высшая школа экономики” (НИУ ВШЭ)» в 2016–2017 гг. и с использованием средств субсидии на государственную поддержку ведущих университетов Российской Федерации в целях повышения их конкурентоспособности среди ведущих мировых научно-образовательных центров, выделенной НИУ ВШЭ. The article was prepared within the framework of the Academic Fund Program at the National Research University Higher School of Economics (HSE) in 2016–2017 (grant no. 16-01-0032) and supported within the framework of a subsidy granted to the HSE by the Government of the Russian Federation for the implementation of the Global Competitiveness Program.

© Гаспарян Д.Э., 2017 г.

роль в анализе перспектив создания полностью роботизированных аналогов человека. Напомним, что эмуляция (англ. *emulation*) в вычислительной технике, согласно каноническому определению, — это комплекс программных, аппаратных средств или их сочетание, предназначенное для копирования (или *эмулирования*) функций одной вычислительной системы (*гостя*) на другой, отличной от первой, вычислительной системе (*хосте*) таким образом, чтобы эмулированное поведение как можно ближе соответствовало поведению оригинальной системы (*гостя*). Целью является максимально точное воспроизведение поведения в отличие от разных форм компьютерного моделирования, в которых имитируется поведение некоторой абстрактной модели.

Исторически толчком для сближения машинного интеллекта с человеческим послужили знаменитые тесты Тьюринга, в ходе которых делался вывод, что если интеллект машины полноценно имитирует ответы человека в ходе коммуникации, то он ничем не отличается от человеческого [Turing 1950].

Также основанием для уподобления стало создание искусственных нейронных сетей (ИНС), представляющих собой новое слово в разработках ИИ. ИНС представляет собой комплекс программ, построенных по принципу функционирования биологических нейронных сетей живого организма. Нейронные сети не программируются в традиционном смысле, предполагающем простое воспроизводство и комбинацию введенных в программу данных, но обучаются — производят решения, не заложенные в первоначальный алгоритм. Подобные успехи позволяют предположить нейросетевую работу программы, тождественную мыслительной работе человеческого интеллекта [Иваницкий 2015].

В дополнение к этим причинам немаловажная роль в сближении искусственного интеллекта с естественным принадлежит более глубокой идее отождествления смысла и значения, о которой речь пойдет ниже.

Однако наряду с оптимистическими подходами существуют и скептические. Немалую роль в их формировании сыграла критика идеи машинного интеллекта, принятая Дж. Сёрлом в его знаменитом аргументе “Китайская комната” [Searle 1980]. Сёрловский мысленный эксперимент показывает, что тот факт, что машина может вести с человеком вполне осмысленную беседу вовсе не означает, что машина может “думать”. Согласно аргументу Сёрла, программа будет успешно справляться с задачей внятной коммуникации, если только будет ставить в соответствие одним символам (вопросам) другие (ответы). Программа исключительно *синтаксична*, т.е. в ней принимается во внимание лишь начертание символов. *Значение их никак не раскрывается*. Однако реальное человеческое понимание предполагает не только синтаксические правила, но и владение *значением слова*, т.е. по сути знание того, как выглядит денотат. То есть программа должна была бы обладать опытом, например, “синего и зеленого цвета” — в этом случае ее “знание” терминов “синий” и “зеленый” было бы семантическим и подобным человеческому. Но поскольку на это машинный интеллект не способен, то он не обладает реальным пониманием. Таким образом, даже если компьютер проходит тест Тьюринга, т.е. демонстрирует вербальное поведение, неотличимое от человеческого, из этого еще не следует, что он обладает подлинным интеллектом. Такой вывод делает Дж. Сёрл.

Данная полемика несколько раз меняла акценты. Вначале речь шла о том, что свойства семантики автоматически не вытекают из свойств синтаксиса. Однако впоследствии условия обсуждения были уточнены — речь стала вестись о возможности прохождения Тотального теста Тьюринга в случае замены Программы Роботом, оснащенной сенсорно-моторной компетенцией [Hagad 2001]. Согласно данному уточнению, в случае включения сенсорно-моторной компетенции, которая могла бы придать недостающее измерение “знанию” компьютерной программы в виде чувственных индексальных данных, Программа могла бы успешно справиться с теми же задачами, которые обрабатывает человеческий интеллект. В основе этого тезиса лежит идея, что Программа не может “думать” только потому, что не оперирует значением так, как это делает человек, а именно не может поставить в соответствие знаку реальный

квалитативный опыт. Для того чтобы формализовать значение, следовало бы добиться того, чтобы машина видела синий цвет так, как видим его мы. В данном случае предполагается, что люди учат значения слов посредством особой связности — причинной связи между впечатлением и объектом, которому соответствует символ, и мы понимаем слово “вода”, потому что имели жизненный опыт с водой [Rapaport 2006]. Следовательно, чтобы компьютер понимал значения символов, которыми манипулирует, он должен быть оснащен сенсорной аппаратурой — например, камерой — именно так он в действительности сможет установить значения объектов, представленных символами. Чтобы заслужить ярлык “думающий”, машина должна пройти расширенный (Тотальный) тест Тьюринга, т.е. реагировать на индексикальные ситуации, для чего она должна быть соединена с внешним миром.

Представляется, впрочем, что подобного уточнения недостаточно. Как мы попробуем показать ниже, понимание человека *не ограничивается способностью ставить в соответствие знакам полученный квалитативный опыт*, но предполагает *особое (неоднородное) встраивание в контекст* — соотнесение наличного материала с тем, который в качестве наличного не присутствует. Соответственно, критерием непонимания чего-либо является не только отсутствие квалитативного опыта, но и невозможность выполнить контекстуальное встраивание.

Известна и другая, более “свежая” критика ИИ (мы остановимся лишь на той, которая так или иначе будет задействована в настоящем исследовании), основные тезисы которой сводятся к следующему:

а) интенциональные состояния сознания с трудом поддаются моделированию из-за многообразия символических выражений, “привязанных” к одному имплицитному содержанию (проблема, раскрытая П. Грайсом и Дж. Сёрлом).

б) существуют методологические ограничения формальных систем, и теории множеств (теорема К. Гёделя о неполноте), которые остро ставят вопрос о невычислимости когнитивных процессов.

в) известны также антимеханические и антиредукционистские аргументы Дж. Лукаса, Х. Дрейфуса, Р. Пенроуза [Барышников 2014].

Несмотря на разность критических подходов, проблема онтологического статуса семантических процессов остаётся на сегодняшний день ключевой. Особую роль в вычислительных моделях и методологии искусственного интеллекта (ИИ) занимают семантические интерпретации синтаксических состояний системы. При этом принципиально важно выяснить, какую роль играет сама категория смысла.

Одним из центральных тезисов настоящего исследования является мысль о том, что в основе интерпретаций ИИ как возможного аналога человеческого сознания лежит отождествление смысла и значения. Эта же идея отождествления, как нам кажется, является причиной многих неясностей, проблем и ошибок в анализе проблем ИИ. Такая ситуация, которую мы назовем *натурализацией* смысла, встречается не только в области программирования, но и в других областях знания. Так, в большинстве общих толковых, философских и лингвистических словарей смысл определяется как синоним значения. Это относится не только к русскому слову “смысл”, но и к его немецкому аналогу “Sinn”. В английском языке, несмотря на то, что в языке существует отдельное понятие “sense”, в абсолютном большинстве случаев в научном дискурсе, равно как и в обыденном языке, русские понятия “значение” и “смысл” переводятся одним и тем же словом “meaning”. Таким образом, можно выделить две принципиально различающиеся традиции использования понятия “смысл”. В одной из них смысл выступает как полный синоним значения; эти два понятия взаимозаменяемы. Во второй традиции понятия “смысл” и “значение” так или иначе разтождествлены [Леонтьев 2006, 24]. Первая традиция способствует оптимистическим прогнозам по достижению полноценной эмуляции человеческого интеллекта, и именно ей по умолчанию следует большинство сторонников ИИ. Вторая традиция, напротив, помогает увидеть ограничения ИИ и потому именно на нее ориентировано настоящее исследование.

Учитывая сказанное, становится ясно, что обращение к философской проблематике ИИ предполагает всесторонний анализ смысла с точки зрения его своеобразия, однако в силу впечатляющей обширности подобного исследования в настоящей работе мы сосредоточимся лишь на одной задаче — *демонстрации неоднородной контекстуальности смысла и его трансцендентальной природы*. На этом пути из всего многообразия возможных теорий, где немаловажная роль принадлежит российским авторам, мы выделим философский аппарат Л. Витгенштейна, так как по совокупности факторов он наилучшим образом иллюстрирует то, что мы здесь надеемся сформулировать.

Наша мысль сводится к тому, что машинные процедуры оказываются иррелевантными передаче значения в условиях естественной семантики. Эта идея *как вывод* вполне совпадает с тем, что хочет сказать большинство современных критиков ИИ, но их аргументация отличается от нашей. Согласно, например, Сёрлу, компьютеры не работают с естественной семантикой, так как в машинных процедурах нет места вещам [Searle 1980]. Если у человека есть квалиа как посредник между вещью и её репрезентацией, то машина обрабатывает значение только на уровне протокольных описаний, которые сводятся в объемные базы знаний [Барышников 2015].

Мы же хотели бы сделать специальный акцент на том, что трудности компьютерной формализации смысла связаны с принципом недосказанности, предполагающим *два подпринципа — встраивание в контекст (доставление до некоей полноты) и выразимость*. Однако в отличие от сугубо лингвистической проблематики имплицатурной передачи смысла (которая также активно развивается такими авторами, как Х. Дрейфус), нас будет интересовать собственно философская (метафизическая) трактовка смысла, с помощью которой можно увидеть принципиальные трудности программирования интеллекта в его целостности.

Нетождественность смысла и значения

Для того чтобы перейти к собственному видению указанной проблемы, нам важно предложить особое разделение на “синтаксис” и “семантику”. Под *синтаксисом* мы будем понимать уровень значений, а под семантикой, собственно, уровень *смысла*. Предполагается обосновать две идеи: *во-первых*, показать нетождественность значения и смысла в определенном философском ключе и, *во-вторых*, указать на особый статус измерения смысла, а именно его фундаментальную и принципиальную неопределенность. На пути реализации второй задачи нам потребуется ввести понятие “трансцендентального измерения смысла”.

В наиболее оформленном виде принцип разграничения смысла (Sinn) и значения (Bedeutung) проводится Г. Фреге, которого принято считать родоначальником оппозиции смысла и значения. Наше понимание смысла и семантики, однако, будет сближаться с тем, что Пирс, а в современных исследованиях Грайс подразумевали под лингвистической прагматикой [Ногн 2004]. Речь идет об изучении смыслов в контекстах — социальном, ситуативном, т.е. как средства коммуникации, а именно анализе явных и скрытых целей высказывания, внутренней установки говорящего и слушающего; изучении типов коммуникативного поведения: речевой стратегии и тактики, правил диалога, направленных на достижение эффективности общения, использование так называемых “непрямых” речевых актов и разнообразных приемов языковой игры.

Также используемое здесь понимание смысла сближается с той трактовкой, которая присутствует в работах таких российских авторов, как М. Бахтин, Г. Шпет, Д. Леонтьев, Г. Тульчинский, В. Налимов. В их работах мы также встречаемся с идеей отделения смыслов от значений, а также представлением об особом философском содержании категории смысла [Бахтин 1979; Шпет 1996; Леонтьев 2006; Тульчинский 1986; Налимов 1995].

Дать определение смыслу крайне сложно. Одной из версий объяснения причин этих трудностей является трансцендентальная интерпретация смысла. Трансцендентальность в данном случае подразумевает такую структурную характеристику, которая означает возможность определять все прочие предметы (объекты) через себя, но отсутствие возможности самой быть опредмеченной (объективированной) [Гаспарян 2014].

Определение смысла осложняется также в силу того, что понятие смысла может употребляться в разных значениях. В общем виде можно разделить *смысл-сущность*, *смысл-цель* и *смысл-ценность*. Первое значение связывают с проблемой познания сущности вещей (типичным воплощением такого подхода является платоновская “идея” как смысл вещи, воплощение ее сущности в качестве первообраза, модели и образца). Как правило, и чаще всего в исследованиях по ИИ, именно это представление о смысле отождествляют со значением (денотативной частью знака), чего недостаточно. Во втором значении смысл понимается как подчиненность какой-либо определенной цели, т.е. целесообразность, или как мотив действий. В третьем значении смысл понимается как что-то ценностное (важное, дорогое) и при этом возможен только в рамках ценностной картины мира (так мир, который лишен ценности, мы назовем абсурдным). Из этих трех ключевых значений смысла можно вывести еще три дополнительных, в частности, 1. смысл отвечает за системно-целостное положение дел; 2. смысл работает как некий организующий принцип и ориентир, придающий вещам и миру в целом единство; 3. смысл отвечает за последовательность и логическую связность (событий/размышлений) [Тульчинский 1986].

Несмотря на наличие этих различных аспектов смысла, можно говорить о единой интуиции смысла, которой пользуется человеческий интеллект и которая объединяет в себе все аспекты. Если попытаться обобщить все многообразие трактовок смысла в философии и гуманитарных науках под тем углом, который использован в данной статье, то можно обратить внимание на две основные черты, объединяющие практически все эти трактовки. “Смысл определяется, во-первых, через более широкий контекст и, во-вторых, через интенцию (целевую направленность, предназначение или направление движения)” [Леонтьев 1999, 57]. Кроме того, каждый из этих аспектов едва ли может мыслиться в отдельности от всех остальных. Между тем ниже я сосредоточусь на ценностной и целевой компонентах смысла.

Если говорить непосредственно о том различии между значениями и смыслами, которое используется в данном исследовании, то ему больше всего соответствует разделение на факты (предложения) и ценности, о которых говорит Л. Витгенштейн. Всякое предложение, которое может быть истинным или ложным, имеет значение [Витгенштейн 1958]. Значением в данном случае будет положение дел (факт), которое имеет или не имеет места. Так, *значением* предложения “Зеленое яблоко лежит на столе” будет некоторое положение дел — яблоко, лежащее на столе — ситуация, которая может иметь место (быть фактом) или не иметь. Однако совсем не обязательно, чтобы предложения обладали *смыслом* — по крайней мере, “иметь значение” и “иметь смысл” не является одним и тем же. Наделение смыслом сродни наделению ценностью в терминологии Витгенштейна² — ценность того, что яблоко лежит на столе (если мы пожелаем увидеть это событие как ценностное), сама по себе не есть факт и в мире фактов-предложений не встречается [Витгенштейн 1989]. В случае установления значения достаточно понимать, что это такое, что имеется в виду (“яблоко лежит на столе”), т.е. иметь качественный образ этой ситуации или предъявить ее как остенсию. В случае же наделения смыслом мы хотим понять, *почему это так*, ведь оно *могло бы там и не лежать* (ответ “Его положил туда мой друг” только временно снимет вопрос, так как вопрос о смысле можно переадресовать и в отношении действия друга и т.д. по цепочке причинных объяснений). Например, вопрос “*Что значит, что яблоко зеленое?*” отличается от вопроса “*Какой смысл в том, что яблоко зеленое?*”. Равно вопрос “*Что значит, что яблоко существует?*” отличается от вопроса “*Какой смысл в том, что яблоко существует?*”. При этом далеко не во всех случаях вопросы “*Какой смысл в том, что яблоко зеленое?*” или “*Какой смысл в том, что яблоко существует?*” будут являться адекватными

в повседневной языковой практике. В большинстве жизненных ситуаций, образующих прагматический контекст, мы можем никогда не задаться этими вопросами. Однако редкость использования не означает, что мы не понимаем подобных вопросов или не можем отличить их от вопросов о значении.

Наш тезис состоит в том, что человеческий интеллект с легкостью проводит различия между значением и смыслом, хорошо понимает предложения непосредственно о смысле и, как будет показано ниже, даже если не прибегает к таким предложениям напрямую, постоянно ими пользуется в качестве фона или контекста, а также формального условия частных актов понимания. Но, как кажется, не все так гладко будет протекать в случае машинного интеллекта.

Неоднородное встраивание и неоднородная выразимость

Трудности компьютерной формализации смысла связаны с принципом недосказанности. Схватывание смысла, таким образом, обеспечивается восполнением “недостающих частей” целого, иначе говоря, восприятием того, “чего нет”. Это восполнение может производиться двумя способами: *во-первых, через встраивание в контекст (до-встраивание до некой полноты), а, во-вторых, через выразимость.*

Под принципом встраивания в контекст понимается такая особенность смысла, благодаря которой он возникает не внутри определенной синтаксической цепочки (будь то текст книги или видеоряд фильма), а вне его. В своем определении контекстуальности, к примеру, Фреге [Фреге 1997] утверждает, что необходимо всегда учитывать полное предложение, так как только в нем слова обладают подлинным смыслом. Если предложение имеет свой смысл как целое, благодаря этому свое содержание получают также и его части. Однако ясно что, встречаясь с разными предложениями, мы не всегда можем актуально наблюдать контекстуальное целое. Следовательно, мы должны уметь воссоздавать контексты — достраивать непосредственно воспринимаемое до некой целостной полноты.

Например, мы смотрим ролик, где герои что-то делают (мужчина и женщина едут в метро и смотрят друг на друга) и понимаем смысл всей композиции как образное изображение любви, хотя о любви в ролике не было сказано ни слова. В данном случае смысл есть нечто, что не сказано напрямую и попросту превышает синтаксический порядок. Эффект смысла возникает для сознания в тот момент, когда мы что-то понимаем, а понимаем мы что-то тогда, когда обнаруживаем то, *что не дано* нам непосредственно. Например, понимаем смысл шутки, хотя непосредственно смешное в ней не говорится, а подразумевается. Так проявляется важнейший для естественной семантики эффект недосказанности, который состоит в том, что смысловая часть нарратива не совпадает с синтаксической частью, но всегда превышает ее. Наиболее полно эту тему развил П. Грайс, убедительно показавший далеко не всегда экспликатурную, но чаще имплицатурную природу смысла [Grice 1981].

Опираясь на эту трактовку можно сказать, что смысл никогда не удастся передать с помощью законченного числа высказываний, так как этому мешает проблема бесконечного регресса [Гийому, Мальдидье 1999; Пульчинелли Орланди 1999; Делез 1991]. Когда смысл понимается, то цепочка пояснений (потенциально имеющая бесконечный вид) произвольно обрывается, и мы говорим: “Я понял!” При этом человек может разгадывать подобные недосказанности, а машина пока нет. На сегодняшний день не ясно, как можно было бы задать машине эту произвольность, ведь она спонтанна и непредсказуема — нельзя предугадать в какой момент антропоморфный носитель прервет цепочку. Но если ее нельзя предугадать (она не детерминируема), то не совсем понятно, как ее моделировать, так как моделирование предполагает четкий контроль над причинно-следственными цепочками.

В свою очередь *принцип выразимости* характеризует ситуацию, в которой мы, к примеру, считываем в предложении “Миссис Дэллоуэй сказала, что сама купит цветы”

[Вульф 1989, 17] настроение беззаботности радостного солнечного утра и в то же время тоски и внутренней невыраженности, трагедии человеческого существования, к конечности которого мир сохраняет убийственное равнодушие. В терминах лингвистической семантики в данном случае понимание реализуется через аппарат имплицатур, составляющих, как показал Грайс, большую часть не только обыденной, но и специализированной коммуникации [Grice 1981].

Как оба эти принципа можно было бы задать машине?

На первый взгляд, в этом нет ничего сложного. Можно сказать, что с задачами достраивания компьютер уже успешно справляется (опция правописания в программе Word), а если это и представляет пока определенную трудность (распознавание символов при фильтрации ботов), то в будущем ее можно будет устранить (многократно расширив библиотеку прецедентов начертания).

Эмуляция принципа выразимости интуитивно кажется сложнее. Но можно и здесь сослаться на предложенный в свое время структурализмом метод и сказать, что выражаемые идеи есть не что иное, как структуры, а если задать машине все возможные структуры и правила редукции, то она легко справится с задачей выражения [Косиков 2000]. В действительности, как нам кажется, для ИИ оба этих задания являются непроходимыми.

Дело в том, что когда мы говорим о собственно человеческом понимании, речь идет о так называемом *неоднородном достраивании*. Примером однородного достраивания будет прочтение слов “the” и “cat” в знаменитом примере Сэлфиджа [Selfridge 1955], которая показывает, что в зависимости от контекста один и тот же “неполный символ” восполняется до Н или до А: **TAE CAT**. Здесь то, что мы ищем (нужную букву), принадлежит тому же классу объектов, что и наличное – буквы. Похоже, что с такими заданиями компьютер действительно может справляться. Метод в данном случае сводится к сравнению содержимого базы данных и подбору нужного элемента.

Но что делать, когда искомый контекст неоднороден с наличным элементом? Например, мы читаем фразу “Миссис Дэллоуэй сказала, что сама купит цветы” и понимаем, что эта фраза не только сообщает о покупке цветов, но и об определенном настроении героини и всего романа. Между тем информация о покупке цветов не является *однопорядковой* с информацией о настроении, но мы встраиваем фразу в контекст, который *неоднороден* элементу. Кажется, что большинство задач на понимание, решаемых человеческим интеллектом, лежит именно в этой плоскости. Понимание человека способно соотнести и сопоставить объекты из разных референтных групп. Напротив, машинный интеллект способен лишь к однопорядковым соотношениям – так в приведенном примере он сможет обнаружить лишь информацию о женщине, цветах и факте покупки. Такое, собственно синтаксическое, соотношение мы не считаем пониманием. Пониманием мы считаем такое соотношение, где открывается доступ к той информации, которая в прямом высказывании никак не заявлена. Способность к таким соотношениям представляет собой объективную проблему, так как сама природа этого механизма не вполне ясна³.

В качестве еще одного, развернутого, примера неоднородного встраивания, рассмотрим следующий мысленный эксперимент. Допустим, съемочной группе дается задание подобрать для режиссера актрису на роль главной героини. Режиссер облегчает задачу своим консультантам и показывает фото женщины, сходство с *образом* которой ему хотелось бы увидеть в искомой актрисе. Режиссер подчеркивает, что ему не обязательно физическое, внешнее сходство, но принципиально то настроение, которое несет образ интересной ему героини. Представим теперь, что женщина, изображенная на фото, является жгучей брюнеткой с темными глазами. В ходе кастинга помощники режиссера просматривают множество претенденток на роль и вскоре останавливают свой выбор на голубоглазой блондинке, которая при всем своем физическом несходстве несет именно тот образ, который искал режиссер. Сможет ли решить эту задачу компьютер? Как кажется, нет. Поиск, который можно будет продуктивно запустить с помощью программы, будет состоять в обширном переборе лиц, результатом

которого будет установление исключительно физического (и, учитывая, возможности машины, вероятно, максимального) сходства лиц. Таким образом, машина также сможет подобрать актрису для режиссера — скорее всего, это будет копия женщины с фото, которая при максимальном внешнем сходстве, вполне возможно, разочарует режиссера, так как будет нести совершенно иной и, может быть, даже противоположный искомому образ.

Цель данного мысленного эксперимента в том, чтобы показать способ работы машинного интеллекта — под синтаксисом в данном случае понимается перебор однородных элементов с целью установления сходств или различий. Однако подлинная семантика останется незадействованной — для ее реализации машина должна была бы научиться сличать не лицо с лицом, а лицо с образом. Проблематичным это окажется, впрочем, не потому, что у машин есть принципиальные ограничения, а потому, что способ задания образа фактически нам неизвестен и в этом смысле сама технология программирования остается неясной.

Каким-то способом интеллект умеет перейти с одного уровня значений (синтаксического) на другой (смысловой), никак не проявленный внутри значащих событий. В естественной семантике смысл проявляется имплицитно — он подразумевается, но напрямую не проговаривается. Есть некоторая неоднородность того, что дано, и того, что при этом понимается. Подобная ситуация напоминает соотношение витгенштейновских фактов и смысла — все, что мы видим в мире, относится к фактам, в то время как смысл в мире никак не проявляется и не является частью мира.

Соответственно, когда мы спрашиваем “В чем смысл?”, мы вовсе покидаем область значений, так как смысл — это не еще одно значение, но нечто качественно иное. Наша гипотеза состоит в том, что человеческое понимание идет сразу в двух направлениях: с одной стороны, мы устанавливаем значения (когда спрашиваем и отвечаем, “что такое зеленое яблоко”), с другой, имеем в виду смысл (когда спрашиваем “Почему зеленое яблоко существует?”). Напротив, “понимание” машины может идти только в одном направлении — направлении значений.

Даже в самых рядовых ситуациях понимание есть тогда, когда происходит встраивание того, что дано, в то, что не дано, при условии неоднородности того и другого.

Метафизический смысл: цель и ценность

Можно предположить, что для человека понимание оказывается возможным, так как у него есть некая труднообъяснимая интуиция или переживание, лучше всего проявляющиеся в таких конструкциях, как смысл мира (у Витгенштейна “смысл мира должен лежать вне мира” [Витгенштейн 1958, 118]), “смысл жизни”, “смысл некоторого повествования”. Такой смысл можно назвать *метафизическим* или *трансцендентальным*. Соответственно, можно высказать гипотезу, что метафизический смысл как некая интуиция лежит в основании человеческого понимания, даже если речь идет о простых процедурах установления значения. Хотя данный смысл не формулируется напрямую (в таких самых общих редакциях, как: “В чем смысл происходящего?”, “Почему это происходит?”, “Почему это есть?”), он обеспечивает само понимание. Несмотря на метафизичность, все в мире понимается через этот смысл, и сам мир дается как осмысленный благодаря данной интуиции. Эту интуицию, которую довольно трудно пояснить в понятиях и которую вместе с тем люди отчетливо ощущают, можно приблизительно описать в серии вопросов: “Почему нечто вообще происходит?”, “Что оно выражает?”, “Почему этот рассказ надо было рассказывать?” Соответственно, для того чтобы уметь работать с неоднородным материалом, т.е. в нашем случае выходить за пределы значений, необходимо использование особой целевой и ценностной структур смысла. Следовательно, можно предположить, что для совершения конкретных содержательных переходов от высказанного к невысказанному требуется ориентирующая работа формальной категории смысла. Именно она играет ключевую

роль в организации такого мышления, которое справляется с задачами неоднородного достраивания. Равно чтобы спросить, в чем смысл мира и всех вещей в мире, какой смысл в том, что наша жизнь сложилась из определенной серии событий, или каков смысл рассказа, нужно выйти за пределы данных вещей, данных событий и данного сюжета⁴.

Но, как кажется, именно это машина сделать не может, так как всегда работает с однородным материалом, с тем, “*что есть*”. Эту работу и стоит обозначить как синтаксическую. При этом не ясно, как можно было бы задать машине основания для того, чтобы не ограничиваться тем, что есть. Не вполне понятно, как можно было бы заставить Программу спонтанно спросить о смысле ее существования или тех операциях, которые она проделывает, притом что если бы машина самостоятельно это сделала, ее можно было бы назвать “думающей” [McCall 1999].

Во многом трудности программирования смысла связаны с тем, что если пристотреться к нашей интуиции смысла, то мы увидим, что данная категория является ценностной, а не только логической или эпистемической. Когда мы спрашиваем “Какой в этом смысл?” мы фактически спрашиваем “Почему это важно?” или “Почему это ценно?”, “Чем это дорого?”, но для того, чтобы подобный вопрос задать, нужно иметь определенный пласт ценностей, т.е., если угодно, иметь свою этику. Ценностная трактовка смысла, в свою очередь, требует выделения фигуры субъекта, а также делает *смысл* проблемой прагматики, где смысл фиксирует отношение значения к потребностям субъекта [Бахтин 1979]. Однако вопрос субъектности и моральности компьютеров является отдельной большой проблемой.

Но так ли важен этот уровень смысла для машинного интеллекта и *как можно доказать, что его действительно нет в эмулирующих программах?* Хотя такой “метафизический” смысл и кажется слишком большой абстракцией, в действительности, его наличие или отсутствие обладает очень конкретными и технически важными для компьютерной программы последствиями. Дело в том, что “осмысление” для машины всегда имеет “переборный” вид, совершенно неестественный для человека, по крайней мере, если он рассуждает сознательно. Разнообразные операции сравнения, абстракции, классификации и обобщения, разумеется, используются человеком, но в основе этих процессов, в значительной мере лежит не до конца ясная в психологическом плане интуиция различения существенного и несущественного. Человеку не нужно просматривать все возможные комбинации – каким-то образом он сразу или почти сразу знает, что ему нужно. Эта способность в свою очередь возможна только на основе той интуиции смысла, которую мы определили как метафизическую. Именно вопросы типа “Какой в этом смысл?”, а не “Каково значение?” позволяют проводить отделение существенного от несущественного, а значит экономить на проверке всех потенциальных данных [Moravec 1999].

Следует, впрочем, удерживаться от того, чтобы трактовать подобные вопросы “о смысле” как сугубо целевые, по крайней мере, если речь идет о частных целях вида “целью выполнения команды X является задача Y”. С такими операциями программы как раз легко справляются. Чтобы увидеть разницу между частноцелевыми и смысловыми вопросами, рассмотрим ситуацию, в которой программист закладывает в программу шахматной игры не только все оптимальные ходы, но и *цель* – следует выигрывать, а не проигрывать. Может ли компьютер сам увидеть смысл в том, что следует стремиться к выигрышу? Ведь для него ценностно *эквивалентны* не только выигрыш белых или выигрыш черных, но и сам выигрыш и проигрыш. Фактически проиграть тоже можно оптимальным образом – в кратчайшем периоде и при минимальном числе ходов понести наибольшие потери. При этом если мы скажем, что и человек не всегда может ответить на вопрос “Какой смысл в том, чтобы выиграть?”, то важно понимать, что человек, однако, хорошо чувствует разницу между вопросами 1. “Какое лекарство мне выпить, чтобы выздороветь и жить?” и 2. “Зачем вообще жить?” [Вебер 1990] – первый вопрос является целевым, а второй смысловым (и по сути, ценностным). Между тем для машинного интеллекта эти вопросы будут однопорядковыми.

Для человеческого интеллекта смысловые вопросы даны как окончательные в ряду родовидовых предложений, но сами они *неопределенны*, так как в их отношении не получается сформулировать стандартное предложение-дефиницию вида “Это есть то” (“Смысл жизни – это...”). Попытки дать им определение приводят либо к тавтологиям, либо к бесконечному регрессу. При этом не совсем ясно, как человек, переходя с одного уровня значений на другой, принимает решение остановиться на тех или иных вариантах результирующего текста. Тут существенную роль играет способность человека учитывать нечеткие ситуации, не прибегая при этом к их точной формулировке, но именно в этом машина промахивается.

В терминах программирования речь в данном случае идет о том, что для компьютерной семантики острым остаётся вопрос о вычислимости нечётких значений (в действительности, смыслов) – тех, которые выражаются, но на синтаксическом уровне не прописываются. Пока только на уровне человеческого сознания производится совершенно уникальный “эффekt” смысла – результата перехода на иной уровень обработки информации. Напротив, в случае машины остается “слепое” вычисление с перспективой бесконечного усложнения, но в рамках жесткой однородности материала.

Задача осложняется не только ситуативной изменчивостью смысла и зависимости его от актуального состояния субъекта (эту проблему надеются частично устранить с помощью создания сенсорно-ориентированных роботов), но и тем, что смысл не всегда понятийно репрезентирован, не всегда осознан и не всегда может быть четко выражен доступными средствами [Леонтьев 1999, 61].

Тогда можно спросить: почему мы не можем передать машине *это интуитивное переживание смысла*? Ответ будет состоять в том, что для нас самих это переживание остается необъяснимым. Мы не можем задать машине то, устройство чего не понимаем. Если очень коротко попытаться объяснить, почему нам не до конца ясен механизм собственного понимания, то можно сослаться на философское и логическое обоснования. Философское обоснование состоит в указании на трансцендентальную природу интуиции смысла. То, что позволяет связывать и удерживать определенные пучки значений как осмысленные, не принадлежит самой сфере значений и не является еще одним значением.

Логическое объяснение основано на принципе системных ограничений (так называемом “метаязыковом парадоксе” или “парадоксе самореференции”) – иметь непротиворечивое знание можно только о тех уровнях, которые лежат ниже самого уровня знания. Нормальное описание должно иметь вид метаязыка. Напротив, самореференция познающего уровня невозможна, так как в этом случае низший и высший уровни совпадут [Penrose 1989].

Для сознания человека непонимание того, что такое смысл, является нормальным: мы не знаем, как работает наше мышление, но успешно пользуемся им. Можно высказать гипотезу, что подобного рода незнание предусмотрено самой структурой мышления. Согласно некоторым философским позициям, трансцендентализма и феноменологии, наше непонимание того, как мы понимаем, является условием того, что мы можем понимать все остальное. В связи с этим существует проблема создания полностью научной психологии. Похоже, что описать внутренние правила разума таким же образом, как законы физики описывают внешний мир невозможно, равно как нельзя понять собственное поведение, рассматривая некоторые объективные когнитивные законы [Dreyfus 1972]. Но это обстоятельство, естественным образом встроенное в работу антропоморфного сознания, может стать роковым препятствием на пути создания машинных копий человека. Машинный аналог создает сам человек, но работает он при этом в условиях существенной неопределенности.

Если мы не знаем устройства некоторых глубинных структур нашего мышления (мы не до конца понимаем, как именно понимаем), то создание программного аналога нашего мышления усложняется принципиально. Это утверждение следовало бы, конечно, еще доказать, но здесь может сработать формальный аргумент: если мы сами не знаем, как устроена работа смысла, поскольку она дана в виде смутной интуиции, то

и запрограммировать ее нельзя. Такие феномены человеческой психики, как явления интуиции (“инсайта”), в отличие от более простых и доступных для изучения процедур четкой понятийной классификации и логической дедукции, составляют главную трудность для машинного моделирования.

На сегодняшний день нечёткая семантика естественного языка хуже всего дается машинному интеллекту. По-видимому, это связано с тем, что в программировании язык – это абстрактная форма представления вычислений с однородным синтаксисом; для сознания же естественный язык – это форма фиксации таких содержаний, которые *непосредственно не даны*. Как кажется, при обработке данных человеческий интеллект подключает уровень метафизического смысла, который позволяет не просто констатировать вещи/факты (на синтаксическом уровне фиксации значений), но и *понимать* их благодаря имплицитному вопросу “Почему эти вещи/факты в принципе есть?”.

Таким образом, к человеческой когнитивной системе подключается такая странная способность, как встраивание в неоднородный контекст благодаря интуиции метафизического смысла. На этом пути антропоморфный агент обращается к ассоциативным и метафорическим связям, а также произвольно обрывает синтаксические связи. Напротив, машинные интеллектуальные процедуры имитируют семантические процессы за счёт обращения к формальным моделям синтаксиса, а также за счёт колоссального объёма баз данных, которые предписывают пробегать полный путь от одного значения к другому. По сути, машинный разум не контактирует ни с предметным миром (в этом состоял аргумент Сёрла, но это ограничение может быть устранено в ближайшем будущем), ни с миром смысловых репрезентаций значения (аргумент, к которому апеллирую я, предполагая, что данный пробел останется неустранимым), но лишь с информацией и правилами вывода.

Программа не понимает, что такое синий цвет или вода, не потому, что не имеет опыта наблюдения этого объекта, а потому, что не видит разницы между вопросами “Что значит, что небо синее?” и “Какой смысл в том, что небо синее?”. При всей своей содержательной нечеткости именно второй тип вопросов, важных по своей форме, делает возможным понимание. Семантика естественного языка фиксирует пока необъяснимую связь вопросов о смысле сущего (в пределе имеющих вид фундаментального метафизического вопроса “Почему все, что существует, существует?”) и значения. И в той мере, в какой мы сами не до конца отдаем себе отчет в том, как именно мы что-либо понимаем (например, мы не контролируем произвольный разрыв цепочки пояснений и лишь пассивно регистрируем акты своего понимания), мы затрудняемся создать компьютерную модель схватывания смысла, а именно компьютерную модель по-настоящему интеллектуальной работы.

Вероятнее всего, смысл можно отнести к особым метафизическим интуициям или переживаниям человека, точное происхождение и механизм которых до конца не ясны. Общим принципом подобных интуиций является то, что мы можем успешно ими пользоваться, но не можем дать им объяснение. Наличие таких фундаментальных интуиций, как правило, утверждается в рамках различного рода трансцендентальных философий и подходов и требует отдельного, выходящего за рамки этого исследования рассмотрения.

Заключение – трудности создания думающего ИИ

Нужно признать, что предложенные рассуждения носят по большей части концептуально-теоретический характер и отсылают к мысленным экспериментам. Соответственно, у практически ориентированных сторонников ИИ, скорее всего, возникнут возражения касательно избыточности подобной аргументации в связи с фактическими успехами программирования. Вместе с тем нельзя отрицать, что проблема несоизмеримости синтаксических и семантических возможностей программ сохраняются на

протяжении всей истории успешного развития ИИ. В частности, речь идет о том, что “практически безграничный синтаксис программы пока не приводит к схватыванию даже примитивного семантического ядра” [Барышников 2014, 110].

Эту проблему можно считать присущей самому механизму означения – непостижимость естественного языка лежит не только в парадоксальной метапозиции означения (бесконечный регресс значений), но и в уникальных свойствах семантики, способной к постоянному нарушению концептуальных моделей, к новому смысловому синтезу и встраиванию в неоднородный контекст. В конечном итоге на сегодняшний день для того, чтобы перевести философские дискуссии об ИИ на новый уровень, следовало бы вначале обрести окончательную ясность в том, что такое смысл и что такое понимание (притом что это глубоко родственные понятия [Тулчинский 1986]). Но это как раз труднее всего сделать. Проблема заключается в том, что отличие машинного понимания от человеческого нельзя даже удовлетворительно формализовать: какой бы отличительный критерий мы ни придумали, через некоторое время программисты успешно программируют найденное отличие. Между тем интуитивно мы чувствуем, что созданный машинный аналог не является “мыслящим субъектом”, и пытаемся сформулировать новый критерий отличия, но вновь программисты оказываются на высоте. Однако вскоре приходит понимание, что цель полученной эмуляции все равно не достигнута и, следовательно, предложенный критерий был неверным. Отсюда напрашивается уже сугубо философский аргумент: мы предлагаем не совсем адекватные критерии потому, что сами не понимаем, как устроены механизмы сознания и какими мыслительными способностями мы владеем. В частности, мы до конца не понимаем, как схватываем смыслы. Но в той же мере, в какой мы сами не знаем, чем отличаемся от машин, не ясно, как это отличие можно было бы задать программе. Всё, что мы действительно знаем о себе, можно успешно эмулировать, но это будет лишь часть человека, фактически не представляющего некоторых глубинных механизмов собственного устройства. Для полноты же картины наблюдателю пришлось бы выходить за пределы собственного мира [Penrose 1989]. Так что создание думающих (в полноценно антропоморфном смысле) роботов возможно, но только не в нашем, человеческом мире.

Примечания

¹ Примерами трансцендентальных форм в классической философии являются платоновские идеи прекрасного, мужественного, справедливого и т.д., аристотелевские категории сущности, качества, количества и пр., кантовские категории пространства и времени, а также категории рассудка (в конечном итоге, весь категориальный аппарат философии отвечает характеристикам трансцендентальности). В более современных контекстах ближе всего к этой теме подошел Л. Витгенштейн, посчитав трансцендентальными законы логики и языка.

² Ценностную интерпретацию смысла можно найти не только у Л. Витгенштейна, но и, в частности, у М. М. Бахтина [Бахтин 1979].

³ Если использовать терминологию Хомского, то можно сказать, что на сегодняшний день вопрос “Как это возможно?” имеет характер тайны, а не проблемы [Chomsky 2000].

⁴ На языке философии эта способность человека определяется как способность к трансцендированию, “умение поставить бытие под вопрос” (М. Хайдеггер), т.е. отнестись к миру не изнутри, а извне.

Источники и переводы – Primary Sources and Russian Translations

Бахтин 1979 – *Бахтин М. М. Эстетика словесного творчества*. М.: Искусство, 1979 (Bakhtin M. M. *Aesthetics of Verbal Creativity*. In Russian).

Вебер 1990 – *Вебер М. Наука как призвание и профессия // Вебер М. Избранные произведения*. М.: Прогресс, 1990. С. 707–735 (Weber M. *Wissenschaft als Beruf*. Russian Translation 1990).

Витгенштейн 1989 – *Витгенштейн Л.* Лекция об этике // Историко-философский ежегодник. М., 1989 (Wittgenstein L.J.J. *Lecture on Ethics*. Russian Translation).

Витгенштейн 1958 – *Витгенштейн Л.* Логико-философский трактат / Пер. с нем. Добронравова И. и Лахути Д., Общ. ред. Асмуса В. М., 1958 (Wittgenstein L.J.J. *Tractatus Logico-Philosophicus*. Russian Translation).

Вульф 1989 – *Вульф В.* Миссис Деллоуэй. М.: Художественная литература, 1989 (Woolf V. *Mrs. Dalloway*. Russian Translation).

Делез 1991 – *Делез Ж.* Логика смысла. М., 1991 (Deleuze G. *Logique du sens*. Russian Translation).

Налимов 1995 – *Налимов В. В.* Вселенная смыслов (интервью) // Общественные науки и современность. 1995. № 3. С. 122–132 (Nalimov V. V. *The Universe of Meaning (an interview)*. In Russian).

Тулчинский 1986 – *Тулчинский Г. Л.* Проблема осмысления действительности. Л.: Изд-во Ленингр. ун-та, 1986 (Tulchinsky G. L. *The Problem of Understanding Reality*. In Russian).

Фреге 1997 – *Фреге Г.* Смысл и значение. В кн.: // Фреге Г. Избранные работы. М.: Дом интеллектуальной книги, 1997. С. 128–146 (Frege G. *Über Sinn und Bedeutung*. Russian Translation 1997).

Шпет 1996 – *Шпет Г.* Явление и смысл: феноменология как основная наука и ее проблемы (1914). Томск: Водолей, 1996 (Shpet G. G. *Appearance and Sense: Phenomenology as the Fundamental Science and its Problems*. In Russian).

Chomsky, Avram Noam (2000) *New Horizons in the Study of Language and Mind*, Cambridge, Cambridge Press.

Dreyfus, Hubert Lederer (1972) *What Computers Can't Do: The Limits of Artificial Intelligence*, Cambridge, Cambridge Press.

Penrose, Roger (1989) *Shadows of the Mind: A Search for the Missing Science of Consciousness*, New York, Oxford University Press.

Searle, John Rogers (1980) “Minds, brains and Programs”, *Behavioral and brain sciences*, Т. 3 (11), pp. 417–424.

Selfridge, Oliver Gordon (1955) *Pattern recognition and modern computers. Proceedings Western Joint Computer Conference*, Pallgrav Press, Los Angeles, California.

Turing, Alan (1950) “Computing Machine and Intelligence”, *Mind*, Vol. 59, pp. 433–460.

Ссылки – References in Russian

Барышников 2014 – *Барышников П. Н.* Семантические процессы сознания: от вычислительных моделей к языковому опыту // Эпистемология и философия науки. 2014. Т. 41, 3. С. 96–114.

Барышников 2015 – *Барышников П. Н.* Философские проблемы семантики и методология искусственного интеллекта // Человек в технической среде, сборник научных статей / Под ред. Н. А. Ястреб. Вологда, 2015. С. 18–21.

Гаспарян 2014 – *Гаспарян Д. Э.* Бои за эпистемическое первенство: трансцендентальные границы современного натурализма // Вопросы философии. 2014. № 12. С. 58–68.

Гийому, Мальдидье 1999 – *Гийому Ж., Мальдидье Д.* О новых приемах интерпретации, или Проблема смысла с точки зрения анализа дискурса // Квадратура смысла / Под ред. П. Серию. М.: Прогресс, 1999. С. 124–136.

Иваницкий 2015 – *Иваницкий А. М.* Сознание и мозг: как “поверить алгеброй гармонию”? // Вопросы философии. 2015. № 2. С. 35–45.

Косиков 2000 – *Косиков Г. К.* “Структура” и/или “текст” (стратегии современной семиотики) // Французская семиотика: От структурализма к постструктурализму. М.: Прогресс, 2000.

Леонтьев 1999 – *Леонтьев Д. А.* Психология смысла. М.: Смысл, 1999.

Леонтьев 2006 – *Леонтьев Д. А.* Понимание смысла и смысл понимания // Понимание: опыт мультидисциплинарного исследования / Под ред. А. А. Брудного, А. В. Уткина, Е. И. Яцуты. М.: Смысл, 2006. С. 20–27.

Пулчинелли Орланди 1999 – *Пулчинелли Орланди Э.* К вопросу о методе и объекте анализа дискурса // Квадратура смысла / Под ред. П. Серию. М.: Прогресс, 1999. С. 197–224.

Ссылки на иностранных языках см. в разделе References.

**The Mystery of the Natural Semantics:
The Transcendental Dimension Meaning and the Problem of Artificial Intelligence**

Diana E. Gasparyan

The research explores that the emulating AI is focused on the so-called naturalistic understanding of sense – as a meaning. But if we assume that the category of sense has the attributes of transcendence and is not identical to the meaning, then we may be able to shed light on problems of natural semantics emulation. The research emphasizes valuable and teleological aspects of meaning, which together lead to the concept of “metaphysical sense”. The hypothesis that the “metaphysical meaning” is the basis of any cognitive acts is proved. The computerization of meaning in this case should emulate the system of values or ethics that represents an apparent large problem.

KEY WORDS: the sense, the meaning, context, transcendentalism, natural semantics, artificial intelligence, syntax, semantics, meaning, linguistic pragmatics.

GASPARYAN Diana E. – CSc (PhD) in Philosophy of Lomonosov Moscow State University, Senior Research Fellow in Centre for Fundamental Studies in Laboratory for Philosophical Studies, Associate Professor on Faculty of Humanities in School of Philosophy of the National Research University “The Higher School of Economics”.

anaid6@yandex.ru

Received at November, 20. 2016.

Citation: Gasparyan, Diana E. (2017) “The Mystery of the Natural Semantics: The Transcendental Dimension Meaning and the Problem of Artificial Intelligence”, *Voprosy Filosofii*, Vol. 4 (2017), pp. 81–94.

References

- Baryshnikov, Pavel N. (2014), “Semantic Processes of Consciousness: from Computational Models to Language Experience”, *Epistemology and philosophy of science*, Vol. 41, 3 (2014), pp. 96–114. (In Russian).
- Baryshnikov, Pavel N. (2015), “Philosophical Problems of Semantics and the Methodology of Artificial Intelligence”, *People in a technical environment. Collection of scientific articles*, Vologda State University Edition, Vologda, pp. 18–21. (In Russian).
- Gasparyan, Diana E. (2014), “Fighting for the Epistemic Superiority of the Transcendental Boundaries of Modern Naturalism”, *Voprosy Filosofii*, Vol. 12 (2014), pp. 58–68. (In Russian).
- Grice, Herbert Paul (1981) “Presupposition and Conversational Implicature”, *Radical Pragmatics*, Academic Press, 54 (2), pp. 183–197.
- Guilhaumou, J., Maldidier, D. (1990) “De nouveaux gestes de lecture ou le point de vue de l’analyse de discours le sens”, *Quadrature du sens*, PUF, Paris, pp. 228–239 (Russian Translation 1999).
- Harnad, Stevan (2001) “Minds, Machines, and Searle 2: What’s Right and Wrong about the Chinese Room Argument”, *Views into the Chinese Room: New Essays on Searle and Artificial Intelligence*, Preston, John and Bishop, Mark (eds.), Oxford University Press, pp. 294–307.
- Horn, Laurence R. (2004) “Implicature”, *Handbook on Pragmatics*, Horn, Laurence R., Ward, Gregory (ed.), Blackwell Publishing, Malden MA.
- Kosikov, Georgy K. “‘Structure’ and/or ‘text’ (the strategy of contemporary semiotics)” *French Semiotics: from Structuralism to Post-Structuralism*, Progress, Moscow, pp. 3–48. (In Russian).
- Ivanitsky, Alexey M. (2015) «Consciousness and Brain: How “to Verify the Harmony with Algebra”?», *Voprosy Filosofii*, Vol. 2 (2015), pp. 35–45. (In Russian).
- Leontiev, Dmitry A. (1999) *Psychology of meaning*, Smysl, Moscow. (In Russian).
- Leontiev, Dmitry A. (2006) “Understanding of the Sense and the Sense of Understanding”, *Understanding: the Experience of Multidisciplinary Study*, pp. 20–27. (In Russian).
- McCall, Storrs (1999) “Can a Turing Machine Know that the Gödel Sentence is True?”, *Journal of Philosophy*, 96 (10), pp. 525–532.
- Moravec, Hans (1999) *Robot: Mere Machine to Transcendent Mind*, Oxford University Press, New York.
- Pulcinelli Orlandi, E. (1990) “Observacoes sobre analise de discurso”, *Terra a vista. Discurso de confronto: velho e novo mundo*, Cortez ed., Sao Paulo 1990, pp. 25–52. (Russian Translation 1999)
- Rapaport, William (2006) “How Helen Keller Used Syntactic Semantics to Escape from a Chinese Room”, *Minds and Machines*, 16 (4), pp. 381–436.