

Распространяется бесплатно

ФИННОВЫЙ КОМПОТ



16+

{№8}, апрель 2015:
Мышление машин

Евгений Логинов. Редакторская.....	1
------------------------------------	---

ТЕМА НОМЕРА

Прологомены ко всякому будущему интеллекту, могущему появиться как искусственный.....	3
Дмитрий Волков. Аргумент внутреннего времени.....	6
Артем Беседин. Ленивый мозг.....	9
Кирилл Загустин. Китайский агент.....	12
Андрей Мерцалов. Что умеет машина, но не может позволить себе Серл.....	15
Иван Поздняков. Какая машина сможет мыслить?.....	21
Мария Ананина. Жук в коробке.....	23
Максим Шалдяков. Естественный интеллект.....	25
Артем Дейнека. К топологии искусственных интеллектов.....	28
Евгений Быков. Мыслить мышление машин.....	31
Олег Мирский. О мнимом праве знать.....	34
Давид Апфельбаум. Еврейский взгляд на искусственный интеллект.....	35
Александр Ветушинский. Человек и компьютер.....	37
Елена Бояршина. Про роботов.....	40

РЕЦЕНЗИИ

Вадим Васильев. Ответ моим рецензентам.....	44
Александр Мишура. На западном фронте без перемен.....	50

ИНТЕРВЬЮ

Джон Серл. Истина это здравый смысл, а не наоборот.....	52
Филип Гофф. Панпсихизм решает проблемы.....	55

ФИЛОСОФСКИЕ СООБЩЕСТВА

Дмитрий Миронов. Венский кружок (часть 3).....	57
--	----

ИДЕНТИЧНОСТИ

Николай Герасимов. Киберпанк и мышление машин.....	61
--	----

ФИЛОСОФИЯ В РОССИИ

Алексей Салин. Один день из жизни современного гуманитария.....	63
---	----

Редакторская

«Да, — он махнул рукой и потянулся за зажигалкой, — весь автостоп скоро закончится. Будет искусственный интеллект фуры катать. Хотя, может быть, и у них совесть вместе с мозгами появится, кто знает». Добрый дальнбойщик, трасса М10, 2014 год.

Филипповский компот немного сменил формат: теперь мы стали чуть-чуть серьезнее. И мы выяснили удивительную вещь: скучная академическая тема возможности сознания, реализованного на небиологическом носителе, привлекла к работе над номером гораздо больше людей, чем наши прошлые топики, например, философия лени, философия курения, философия смеха или философия видеоигр. Я рад, что так произошло. С другой стороны, не такая уж это и кабинетная тема. Достаточно сказать, что она не отпускает чувствительных к повестке дня кинематографистов: «Скоро вернусь» (серия «Черного зеркала»), «Превосходство», «Страховщик», новые «Мстители», «Ex Machina», очень милый «Робот по имени Чаппи» и многие другие фильмы посвящены этой теме. Готовится к выходу ремейк «Терминатора». Обзор художественных воплощений ИИ читайте в статьях Елены Бояршиновой (философский факультет МГУ) и Николая Герасимова (Дом русского зарубежья им. А.Солженицына), а о том, почему Нео был избран, а Джон Коннор — нет, вы узнаете из эссе Евгения Быкова (факультет гуманитарных наук ГУ-ВШЭ).

Видимо, наши познавательные способности устроены так, что, рассуждая о предельных основаниях человечес-

кой природы, нам удобно иметь пред глазами пример чего-то нечеловеческого. Вряд ли это необходимое условие. Но, скажем, схоласты очень любили упражняться на ангелах. Если ангел — это нечто существующее, и существующее бестелесно, то как возможно передвижение ангела? Ведь движение, по Аристотелю, предполагает материальный субстрат движущегося. Не менее важны и исследования в области языка ангелов. Ведь для языка нужно слово, а слово материально, так как же могут разговаривать посланцы Бога? Насколько я знаю, теории ангельской речи сделали возможной концепцию ментального языка Оккама, которая во многом предвосхитила позитивистскую теорию языка. Возможно, в нашем следующем номере мы сможем рассказать об этой проблеме больше. Другими важными объектами для мыслительных экспериментов от века служили Бог, животное и стол. Сейчас же таким нечеловеческим выступает машина. Без преувеличения можно сказать, что машинное и реализованное на нем виртуальное — настоящий ключ к современной философии, как аналитической, так и континентальной.

Наш аналитический блок представлен статьями Артёма Беседина (философский факультет МГУ), Дмитрия Волкова (Московский центр исследования сознания при МГУ), Марии Ананиной (ФФ МГУ), Кирилла Загустина (ФФ МГУ), Ивана Позднякова (психфак МГУ) и Андрея Мерцалова (ФФ МГУ). Кроме того, мы публикуем интервью Джона Сёрла, автора знаменитой «Китайской комнаты». Именно этот аргумент прямо или косвенно пытаются опровергнуть вышеперечисленные авторы. Сёрла гармонично дополняет Филипп Гофф, молодой британский панпсихист и последователь Дэвида Чалмерса.

Настроены против Сёрла и наши континентальные авторы. Артём Дейнека (КНУ), Олег Мирский (РГГУ), Евгений Быков и Александр Ветушинский (ФФ МГУ) пытаются показать, что у нас есть основания приписывать мышление небиологическим объектам. Их интеллектуальный герой — Бруно Латур, который, видимо, уже окончательно покориł философов, пишущих на русском языке. Но другой живой классик, Джорджо Агамбен, выступает против сторонников ИИ (и, выходит, за Сёрла) в статье Максима Шалдякова (ТГУ). Так что единства тут не наблюдается (и не надо).

Текст Максима вообще довольно интересен и провокативен. Автор пытается показать, что мой компьютер (печатная машинка, гусиное перо...) не играет никакой роли в производстве знания, если иметь в виду прямое значение слова «играть». Компьютер, считает он, не может играть в языковые игры, а значит, не может и обладать сознанием. В противоположность этому, тройка латурианцев (Быков, Дейнека, Ветушинский) доказывает, что хоть скоро компьютер всё же нужен для работы философа, то он принимает участие в игре под названием «мышление», а значит, и сам обладает какими-то когнитивными способностями. А раз он уже ими обладает, то и появлению сознания ничего не мешает. Противоречит текст Шалдякова и проведенному Ананиной анализу ментальных терминов. И Шалдяков, и Ананина опираются на Витгенштейна, но приходят к совершенно разным результатам: Максим использует теорию языковых игр против ИИ, а Мария — за. Возможно, это противоречие прольет свет и на суть философии Витгенштейна, и на тайну нашего мышления.

Кроме того, «ФК» продолжает сотрудничество с Давидом Апфельбаумом (ФФ МГУ), раскрывшим теологический аргумент против ИИ, Александром Мишурой (ГУ-ВШЭ), который отрецензировал книгу по свободе воли, Дмитрием Мироновым (ФФ МГУ), продолжившим свой рассказ о Венском кружке, и Вадимом Васильевым (ФФ МГУ), опубликовавшим у нас монументальный ответ критикам его книги «Сознание и вещи».

В статье Алексея Салина (ФФ МГУ) представлено любопытное описание жизни начинающего ученого-гуманитария. Но я хотел бы добавить к нему два пункта. Во-первых, очень много времени этот самый ученый проводит в беседах с друзьями. Это такой род отдыха, который является прямым продолжением работы, и эта самая работа возможна именно благодаря такого рода беседам. Во-вторых, не стоит так уж кручиниться: обдумав свой доклад в метро и написав его за ужином, молодой человек

и обязательно убьет пару часов за просмотром сериалов, скользя по Сети или играя в компьютерные игры. Надеюсь, в своих дальнейших эссе Алексей изложит и другие повадки обитателей философских факультетов.

В работе над номером нам пригодился опыт старших коллег. Наиболее значительными в этой области кажется нам сборник статей «Тест Тьюринга», выпущенный кафедрой Философии языка и коммуникации МГУ в 2011 году, и прошлогодний номер журнала «Логос» по когнитивным наукам.

Евгений Логинов



Пролегомены ко всякому будущему интеллекту, могущему появиться как искусственный

Искусственный интеллект (ИИ) — это громадная научная область, лишь небольшая часть которой является собственно философской. Исследования ИИ принадлежат computer science, где философы по большей части представлены логиками. Если вы наберёте в Стэнфордской энциклопедии словосочетание «Artificial intelligence», то в единственной статье, которая содержит в названии данный термин, «Artificial intelligence and logic», сможете найти упоминание всего четырёх «чистых» философов: Деннета, Фодора, Куайна и Рассела. Это свидетельствует о том, что философская и научная области в данном вопросе почти не пересекаются. Для computer science «искусственный интеллект» — это задача по эмулированию различных когнитивных способностей человека: визуального восприятия, обучения, мышления и так далее (о научной стороне проблемы читайте статью Ивана Позднякова). Мы, философы, не говорим ученым, куда именно вставлять электрод или как правильно провести калибровку. Наш интерес связан с предельными основаниями, с парадигмальными установками.

Тест Тьюринга

В привычном нам виде философская проблематика ИИ впервые была чётко обозначена Аланом Тьюрингом («Вычислительные машины и разум», 1950). Общий вопрос о мышлении машин конкретизируется им: каков критерий разумности? Иными словами, как нам определить, мыслит ли машина, да и откуда мы вообще знаем, что мыслит

хоть кто-либо? Единственным надёжным критерием разумности оказывается речевое поведение. На нём и построен Тест Тьюринга (ТТ). Для чистоты эксперимента (чтобы нас не сбивал с толку внешний вид «собеседника» и т.п.) создадим, насколько это возможно, формат «чистой» коммуникации (что позволяют сделать условия общения в чате). Если в такой беседе речевое поведение «собеседника» экзаменатор расценит как «разумное», если в нём собеседник продемонстрирует языковую компетенцию, то, будь это машина или человек, в соответствии с критерием мы признаем его мыслящим. Понятно, что речь, которая одному может показаться разумной, другому предстанет бессвязным набором слов, да и простые ошибки здесь не исключены. Чтобы нивелировать такие разногласия экзаменаторов, вердикт о разумности собеседника выносится на основании процентной доли доверившихся. Пока что ни одной машине не удалось убедить в своей разумности большинство, однако статистически значимый порог в 30% уже преодолен некоторыми программами. То ли ещё будет! Сам Тьюринг был абсолютно убежден, что технологический прогресс позволит создать по-настоящему мыслящие машины — такие, чьи результаты в ТТ не будут ниже тех, что получает человек.

Для программистов и логиков работа Тьюринга поставила ряд амбициозных технических задач, но остроту в философской среде проблема получила позднее, с появлением компьютерной метафоры в философии сознания. Ее появлению мы обязаны Хилари Патнэму, который выдвинул тезис множественной реализации: состояния сознания могут быть реализованы на разных материальных носителях, так же как и компьютерные алгоритмы.

Китайская комната

В 1980 г. Джон Сёрл предложил аргумент, который до сих пор находится в центре внимания философии ИИ. Тьюринг допустил, что мы и есть сложные машины для вычислений, и Сёрл принял вызов: а давайте человек, о котором мы точно знаем, что он сознателен, сядет в комнату, но будет вести себя не как человек, а как машина, а именно — манипулировать символами. Я сижу в закрытой комнате с русскоязычной инструкцией по манипулированию китайскими символами. Я не знаю китайского и не смогу отличить иероглиф от простой закорючки. От экзаменаторов в комнату приходят сообщения по-китайски, и я, следуя правилам, на иероглиф А отвечаю иероглифом В. Предположим, что инструкция написана так, что я могу составить ответ на любой заданный вопрос. Тогда комната пройдет ТТ, но я ничего не пойму. Значит, ни о каком понимании у машин речи быть не может. Общий вывод Сёрла: синтаксиса недостаточно для семантики, а значит, ИИ невозможен. За 20 лет до аргумента Сёрла советский физик и писатель-фантаст Анатолий Днепроv опубликовал рассказ «Игра», где с изысканием превосхитил выводы Сёрла. Читайте об этом в статье Артёма Беседина.

Системное возражение

Можно возразить: да, я не понимаю, но вся комната понимает. Сёрл даст простой ответ. Суждение «я и бумага, на которой написаны инструкции, понимаем» — контринтуитивно. В бумаге нет ничего, что могло бы понимать. Кроме того, я могу просто зазубрить всю «книгу правил» и выйти из комнаты в чисто поле. Я и в таком случае пройду ТТ, и всё равно ничего не пойму.

Светящаяся комната

Перенесёмся в 1894 год, когда Дж. Максвелл уже открыл, что свет — это электромагнитные волны, но еще не обрёл признания. Согласно его идее, если человек в тёмной-тёмной комнате начнёт перемещать магнит вверх-вниз, от магнита начнут распространяться электромагнитные волны, и в комнате станет светлее. «Но позвольте, — могли бы возразить мы ему, — в комнате совершенно темно!» Мы даже могли бы формализовать наше возражение: (1) Электричество и магнетизм суть физические силы;

(2) Существенное свойство света — свечение; (3) Эти силы сами по себе не являются сущностью свечения. А это значит, что электричество и магнетизм не являются сущностью света и недостаточны для его наличия. Такое возражение, утверждают Пол и Патрисия Чёрчленды, по форме было бы идентично сёрловскому: (1) Компьютерные программы — это формальные (синтаксические) объекты; (2) Человеческий разум обладает смысловым содержанием (семантикой); (3) Синтаксис сам по себе не является сущностью семантики и его не достаточно для семантики. Согласно Сёрлу, это значит, что программы не являются сущностью разума и их наличия не достаточно для существования разума. Вывод, который делает Сёрл, столь же абсурден, что и возражения Максвеллу. В обоих случаях предпосылка (3) интуитивно выглядит правдоподобной, хотя является ложной. У Сёрла, утверждают они, нет достаточных оснований настаивать, что определённое манипулирование символами не сможет породить семантики, и он просто спекулирует на том, что мы пока недостаточно знаем биологические основания сознания.

Аргумент от объективности реализации

Аналогичный аргумент, демонстрирующий некорректность логического вывода Сёрла, приводит Дэвид Чалмерс. (1) Рецепты синтаксичны; (2) Синтаксис недостаточен для рассыпчатости; (3) Пирожные рассыпчаты. Значит, реализации рецепта недостаточно для пирожного. Это очевидно ложное рассуждение Чалмерс сопоставляет с рассуждениями Сёрла о синтаксисе и семантике. Он полагает, что проблема Сёрла заключается в том, что он совершенно не различает программы и их реализации. Да, программ самих по себе недостаточно для сознания, но вот реализованных на чем-то программ — вполне.

Индексикальные вопросы

За 35 лет с момента создания аргумента Китайской Комнаты (КК), каждое слово в исходной статье Сёрла было оспорено и подвергнуто сомнению. Философы пытались показать, что человек в Комнате всё же понимает китайский (бессознательно, вместе с инструкцией, вместе с программистами, поймет со временем и т.п.); утверждали, что существует только синтаксис, а ментальное содер-

жание сознания — миф; что Сёрл должен имитировать не только поведение, но и реальную внутреннюю работу мозга, и т.д. Об экстравагантности приводимых возражений красноречиво свидетельствуют названия соответствующих работ: «Квантовая лингвистика и Китайская Комната», «Сознания, программы и китайские философы: китайская перспектива Китайской комнаты», «От Китайских комнат к Ирландским комнатам» и даже «Мышь-зомби в Китайской комнате».

Продолжит традицию странных названий статья профессора Вадима Васильева «Кока-кола и секрет Китайской комнаты». Васильев обратил внимание, что еще никто не атаковал предпосылку, что КК в принципе может пройти ТТ. Утверждалось, что для продуцирования ответов ей понадобится невероятное количество времени, что человек не сможет написать или интериоризировать соответствующие инструкции, но такие возражения Сёрл парирует напоминанием, что это всё-таки мысленный эксперимент. Чтобы оспорить его посылки, нужны концептуальные средства. Васильев замечает, что в лингвистическую компетентность входят способность отвечать на индексикальные вопросы. Например: «сколько времени?», «что стоит на столе?» и т.д. Может ли на них ответить Сёрл, сидящий в Комнате? Как должна быть написана инструкция в таком случае? Видимо, она должна гласить: «если вопрос

содержит последовательность иероглифов XYZ, посмотри на стол, и, если там стоит банка Колы, ответь иероглифом N (банка Колы)», либо «если XYZ, отвечай M (я не знаю / я слепой и не вижу)». Если Сёрл будет систематически отвечать «M», мы заподозрим неладное. А в первом случае у него просто появится понимание.

Чтобы спасти эксперимент, Сёрл предлагает модификацию с Китайским Роботом. Пусть Робот обладает датчиками, которые Сёрл может задействовать согласно инструкции, и данные с них поступают ему в «письменной» иероглифической форме — такой, что Сёрл не будет понимать их значения. Тогда Комната сможет отвечать на индексикальные вопросы и пройдет ТТ. Сёрл утверждает, что вариация с Китайским Роботом идентична Комнате. Но если допустить Робота, Сёрл не будет знать всю программу, требуемую для ответа, а если узнает её, то узнает и таблицу соответствий сенсорных данных иероглифам и обязательно обретёт понимание, что подорвёт весь эксперимент.

Ответ Сёрла на это возражение читайте в нашем интервью с ним. Тема индексикальных вопросов развивается в статьях Дмитрия Волкова «Аргумент внутреннего времени» и Кирилла Загустина «Китайский агент».



Аргумент внутреннего времени

Сёрл известен своим мысленным экспериментом Китайской Комнаты. Аргумент впервые появился в статье «Сознание, мозг и программы» в 1980 г. С помощью этой истории Сёрл оспаривает тест Тьюринга как критерий рациональности, идею о том, что мышление представляет собой вычислительный процесс, и функционалистские теории сознания. Его вывод в том, что, вне зависимости от программы, компьютер не может обладать интеллектом, даже если симулирует поведение разумного существа. Для разума, считает Сёрл, необходимы биологические процессы. Как построена Китайская Комната?

В Комнате заперт человек. Он не знает китайского, но у него множество справочников с правилами формирования ответов и наборы иероглифов, из которых эти ответы должны формироваться. Люди снаружи, владеющие китайским, передают в Комнату вопросы, узник по справочникам из имеющихся у него иероглифов составляет ответы. У людей снаружи создается впечатление, что узник понимает китайский, тогда как он ничего не понимает, а просто следует инструкциям. Для узника обмен сообщениями — только синтаксические операции, действия с символами. Для понимания необходима связь между символами и их смыслом, семантика. А она отсутствует.

Китайская Комната это метафора компьютера. Узник — центральный процессор, инструкции — программы, а наборы иероглифов — данные. Из ситуации интуитивно следует, что, какие бы манипуляции не производил компьютер, в конечном итоге он имеет дело только с синтаксисом. Для значений и понимания нет места. На Китайскую Комнату бурно откликнулись сторонники

искусственного интеллекта. С первых дней существования она стала магнитом для возражений.

Возражение 1. Системный аргумент

Одним из возражений на Китайскую Комнату является системный аргумент. Его сторонники считают, что понимание китайского языка должно быть не у узника, а у всей Комнаты. И оно, вполне возможно, есть. Человек внутри — только часть механизма системы, но в нее входят также справочники, используемые им для формирования ответов, кусочки бумаги с написанными на них иероглифами. Именно в справочниках и наборах символов в большей степени сосредоточен интеллект системы, без которого действительно никакого понимания быть не может. Это возражение любопытно, но его можно парировать.

Ответ Сёрла — ксеногlossия. Представьте, что человек вышел из Комнаты в поле, говорит Сёрл. Но перед этим человек запомнил все символы и каждое правило. Тогда вся система сосредоточена в его голове. Он теперь свободно общается по-китайски без Комнаты, но, согласно нашим интуициям, он по-прежнему ничего не понимает! Это должно выглядеть как ксеногlossия — предполагаемая в парапсихологии истерия, обычно связанная с религиозными переживаниями, в ходе которой человек начинает говорить на реально существующих языках, которые он никогда не учил. Человек из Китайской Комнаты слышит вопросы и мгновенно на них отвечает, не понимая при этом сам ни единого слова.

Пусть системный аргумент плохо согласуется с интуициями. Но по-другому и быть не может. Чтобы Комната работала так, как обещает Сёрл, она должна содержать миллионы томов инструкций, а сам узник — носиться по ней с ошеломляющей скоростью. Это перегружает воображение. Как тут оно может не подвести? Интуитивное способность «понимать» также плохо согласуется и с физическими процессами, которые происходят в мозге.

Представим ещё одну фантастическую ситуацию. У носителя китайского языка возникает дефект мозга — нейромедиаторы перестают передавать сигналы между нейронами. Тогда ему встраивают механизм, который заменяет работу нейромедиаторов. Назовём его демоном Сёрла. Демон носится между нейронами и передает электрические импульсы в точности, как делали бы это нейромедиаторы. Что будет с поведением человека? Если демон полностью симулирует работу нейромедиаторов, поведение человека не изменится. Он точно так же будет способен разговаривать на китайском. Но будет ли он при этом понимать речь? Когда этот вопрос задают Сёрлу, он высказывается утвердительно. Но из этого, по-видимому, следует, что Сёрл принимает системный аргумент! Сёрл не задает вопрос, понимает ли китайскую речь демон. Возможно, как и узник Китайской Комнаты, демон просто следует инструкциям. Понимает речь человек. Задавая вопрос о том, понимает ли узник китайский язык, Сёрл совершает категориальную ошибку. Понимать язык может не какая-то часть системы, а система целиком.

Возражение 2. Аргумент Робота

Другой аргумент против Китайской Комнаты — аргумент Робота. Для того чтобы описать механизм, способный к пониманию, считают его сторонники, нужно представить не Комнату с символами и инструкциями, а её аналог, реализованный в работе. Компьютерная программа будет установлена в контрольном центре робота. Там будут обрабатываться не только языковые данные, но и данные от «органов восприятия». Такой робот будет способен не только говорить, но и сообразно действовать. Он должен понимать любой язык, на котором будет говорить.

Ответ Сёрла — Китайский Робот. Эти новые условия усложняют ситуацию, но принципиально ничего не меняют. По мнению Сёрла, робот не может добавить ничего для понимания и появления интенциональных содержа-

ний. Просто вместо компьютера мы помещаем в робота Китайскую Комнату. Теперь у нас Китайский Робот! И узник обрабатывает не только символы-вопросы, но и данные от «органов восприятия». Обратное он тоже отправляет символы как для ответов, так и для управления Роботом. Но понимания всё равно не возникает. «Я получаю “информацию” от “перцептуального” аппарата робота и даю инструкции его “моторному” аппарату, ничего об этом не подозревая, — говорит Сёрл от лица узника. — Я являюсь “гомункулюсом” этого робота, но, в отличие от традиционного гомункулюса, я не знаю, что происходит. Я не понимаю ничего, кроме правил манипуляции символами»¹. Сёрл не видит принципиальной разницы между Комнатой и Роботом. Но разница есть.

Различие Робота и Комнаты. Комната с Кока-Колой

На различие двух сценариев обращает внимание В.В. Васильев. Он считает, что, в отличие от Робота, Комната вообще не способна пройти тест Тьюринга. Васильев предлагает вообразить дополнительные детали: «Допустим, я знаю китайский и веду диалог с Сёрлом, запертым в его Комнате вместе с программой и целыми горами иероглифов. Допустим также, что перед экспериментом я зашёл в комнату и поставил банку Кока-Колы на стол, за которым будет сидеть Сёрл. И вот мы начинаем наш диалог»ⁱⁱ. Эксперт спрашивает узника, что находится перед ним. По правилам Китайской комнаты, узник должен обратиться к инструкциям, чтобы сформировать ответ. Но как инструкции могут «знать», что находится в Комнате перед узником? Узник и программа не могут дать нужного ответа, а значит, они не способны пройти тест Тьюринга.

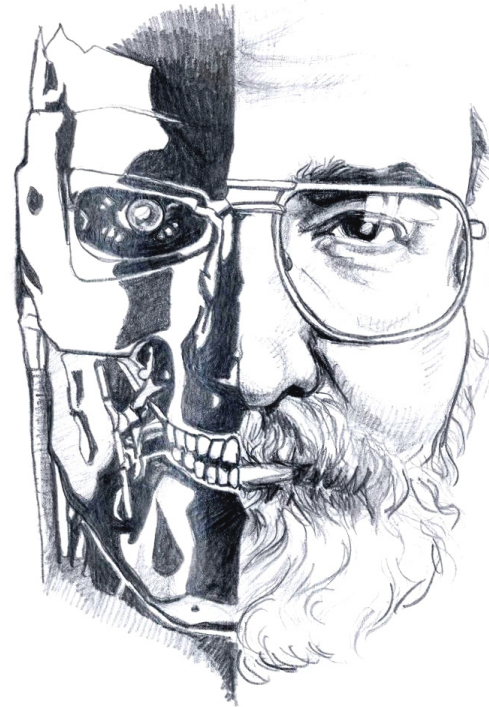
Васильев проводит оригинальную атаку на аргумент Сёрла. Но, мне кажется, её можно отбить. Можно показать, что этот ход незаконный, что он нарушает аналогию между Комнатой и рациональным агентом. Сколько нейронов у Вас в голове? Приблизительно пятьдесят-сто миллиардов? А точнее? Неужели, если Вы не знаете точное число нейронов у себя в голове, Вы не в состоянии пройти тест Тьюринга? Как выглядит орган, соединяющий ваше правое и левое полушарие? А вдруг там находится банка с Кока-Колой? В конце концов, демон Сёрла, пока вы сегодня спали, мог туда её установить. Неспособность агента

дать ответы о своем внутреннем устройстве и изменениях «внутри Комнаты» не дают основания полагать, что агент не рационален.

Внутреннее время

Но позицию Васильева можно развивать. Даже без посторонних предметов Комнату можно поставить в тупик. Нужно задать ей вопрос о фактах, доступных обеим сторонам, ответ на который нельзя было бы подготовить заранее. Набор этих фактов сильно ограничен, т.к. у Комнаты нет органов чувств. Но в ходе беседы с экспертом эти факты появляются. Например, сколько вопросов было уже задано с начала беседы? Подобный вопрос может быть опробован несколько раз в ходе испытания. Как программа будет реагировать на него? Наверное, можно написать такие инструкции, чтобы, в зависимости от контекста, узник давал разные удовлетворительные ответы. Для этого в его распоряжении должны быть варианты всех возможных развращивающихся бесед. Вообразите масштабы библиотеки!ⁱⁱⁱ Допускаю, что, когда узнику будут задавать одни и те же вопросы и он обнаружит, что дает разные ответы, у него появится серьезная подсказка. И он даже сможет вычислить остальные значения по ней^{iv}. Но пока не вижу способа это окончательно доказать. Программу, которая бы правильно отвечала на вопросы о самой беседе, все равно можно помыслить. Тогда члены жюри должны задавать вопросы о внутреннем чувстве. Например, вопрос, сколько времени прошло с момента первого вопроса. Разговор с узником может длиться сколь угодно долго, интервалы между вопросами могут быть любыми, поэтому программа не сможет заранее предусмотреть ответ. Жюри знает, что у узника нет часов (иначе это будет банка с Кока-Колой) и ему придется ориентироваться только на внутреннее чувство, но, чтобы казаться рациональным, он должен уметь отличить 10 минут от 10 дней. Чтобы поставить программу в тупик, можно попросить узника сравнить между собой два интервала времени или что-нибудь подобное. Насколько я понимаю, нельзя написать инструкции так, чтобы предусмотреть правильный разумный ответ. Значит, *аргумент от времени* показывает, что Комната, способная только к лингвистическому поведению, все-таки не сможет пройти тест Тьюринга. И поэтому *аргумент Сёрла не представляет угрозы для функционалистских теорий сознания.*

Дмитрий Волков



Примечания

- i. Сирл Дж.Р. Разум, мозг и программы // Хофштадтер Д., Деннетт Д. Глаз разума. — Самара: Издательский Дом «Бахрах-М», 2003. С. 322.
- ii. Васильев В.В. Кока-Кола и секрет Китайской комнаты // Философия сознания: классика и современность. — М., Издатель Савин С.А., 2007. С. 88.
- iii. Подобный вопрос предложил И. Солнцев. Он предлагает спросить Комнату о содержании предыдущего вопроса. Этот вариант, по моему мнению, можно парировать. Комната теоретически может иметь все комбинации вопросов и ответов. Тогда и на этот у нее будет правильный ответ.
- iv. Правда, во время беседы с такими вопросами у узника, возможно, появятся перспективы разбираться в китайском. Представьте, что Вам задают один и тот же вопрос, но с течением времени вы вынуждены давать на него разные ответы. Учитывая, что Ваши ответы истинны, очевидно, что они описывают что-то, что за это время изменилось. А что могло измениться в комнате, где ничего «особенного» не происходит? Подобные соображения могут стать ключом к расшифровке.

ЛЕНИВЫЙ МОЗГ

Мысленные эксперименты — специфическое оружие современного аналитического философа: они направлены против некоторого тезиса и апеллируют к интуициям, общим для автора аргумента и его читателя. Опора на интуиции придает многим мысленным экспериментам их кажущуюся убедительность, которая исчезает, как дым, при более-менее детальном анализе аргумента. Но возможно, вместо того, чтобы исполнять роль холостого оружия войны, они могут послужить орудием созидания для некоторой иной цели. Чтобы проверить это, вместо тщательного исследования мысленных экспериментов я окунусь в теплое море обыденных интуиций и попытаюсь заплывать в нем как можно дальше.

Оттолкнувшись я от мысленного эксперимента, изложенного А. Днепровым в рассказе «Игра» (1962). Днепрова предлагает нам представить группу людей, которая на футбольном поле имитирует работу ЭВМ «Урал». Число людей совпадает с числом элементов ЭВМ, и между людьми установлены связи, функционально повторяющие связи между элементами ЭВМ. «Пользователь» этой стадионной ЭВМ — профессор Зарубин, главный герой рассказа Днепрова — вводит информацию в машину и получает обработанный сигнал на выходе. В итоге оказывается, что Зарубин ввел фразу на португальском языке, которую ЭВМ, состоящая из людей, перевела на русский. Зарубин приходит к такому выводу: «Если вы, мыслящие структурные единицы некоторой логической схемы, не имели никакого представления о том, что вы делали, то можно ли серьезно говорить о мышлении электронно-

механических устройств, построенных из деталей, на способности к мышлению у которых не настаивают даже самые пламенные сторонники электронного мозга». Наша интуиция срабатывает прекрасно: ни один из участников эксперимента не понимал по-португальски, все вместе они тоже не знали этого языка и не поняли предложенную к переводу фразу, а значит, и металлическая ЭВМ, работу которой имитировали участники эксперимента, не может понимать португальского, даже если в ней заложена программа, позволяющая использовать этот язык.

Днепрова пытался ответить на вопрос: «Может ли машина мыслить?», но сегодня мы можем сказать, что аргумент Днепрова направлен против функционализма. За четкой формулировкой тезиса функционализма обратимся к Д. Чалмерсу, который приводит её в работе «Сознающий ум» (1996): «Естественно предположить, что сознание возникает вследствие определенной функциональной организации мозга. <...> Функциональная организация определяется конкретизацией (1) количества абстрактных компонентов, (2) множества различных возможных состояний для каждого из компонентов и (3) системы отношений зависимости, уточняющих, как состояние каждого из компонентов зависит от предыдущих состояний всех компонентов и от данных на входе в систему, и как данные на выходе из системы зависят от предшествующих им состояний ее компонентов». Последовательный функционалист должен утверждать, что любая система, повторяющая свойства 1–3 некоторой мыслящей системы, должна также обладать мышлением. Эксперимент Днепрова показывает, что это не так.

Д. Хофштадтер и Д. Деннет предлагают взглянуть на мысленные эксперименты как на своего рода приборы с множеством кнопок и регуляторов, которые мы можем нажимать и вертеть, чтобы выяснить, как работает конкретный прибор. Попробуем понажимать кнопки эксперимента Днепрва. Например, покрутим регулятор «число учащихся». Пара сотен студентов моделируют работу ламповой ЭВМ? Возьмите миллиарды людей, имитирующих нейроны человеческого мозга. Что тогда скажут ваши интуиции? Интересен ответ, предложенный советским ученым В. Глушковым в 1962 году. Его интуиция идет в обратном направлении: если человеческий мозг порождает сознание, то почему его функциональный изоморф, реализованный на другом физическом носителе, не будет порождать его? К противоположному выводу приходит Н. Блок, в 1978 году предложивший мысленный эксперимент Китайская Нация: снова миллиард человек — самая многочисленная нация на земле — имитирует работу нейронов мозга. Согласно Блоку, сознания у «китайской нации» не будет.

Получается противоречие. Стоит ли нам прямо здесь, смиряясь, идти ко дну со всеми нашими интуициями? Полагаю, что нет. Спасет нас различие между сознанием доступа (Д-сознание) и феноменальным сознанием (Ф-сознание), введенное Н. Блоком в 1995. В статье 1978 года это различие уже подразумевалось. «Мозг», состоящий из людей-нейронов, по Блоку, связан с телом. Это тело может говорить, что ему тепло или холодно, хорошо или плохо, что оно что-то знает или не знает, и за всеми этими словами стоит гигантский мозг — «китайская нация». Под его руководством тело может двигаться и демонстрировать некоторое поведение. Все эти примеры показывают, что у «китайской нации» есть Д-сознание: «Примерами Д-сознания, — пишет Д. Иванов в работе “Природа феноменального сознания” (2013), — являются мысли, мнения, желания, то есть состояния, которые выражают пропозициональную установку. Очевидно, что такие состояния являются репрезентативными, то есть они представляют некое положение дел. Кроме того, эти состояния являются функциональными. <...> По сути, этот вид сознания <...> является удобным объектом для исследования с помощью функционалистской, когнитивистской методологии». Похоже, что у «китайской нации» есть Д-сознание. Но вопрос, который в первую очередь интересует Блока, другой: «есть ли у нее Ф-сознание?»

Ф-сознание — это опыт, переживание «каково это»: находиться в том или ином состоянии. В этом случае испытываемый нами вкус яблока является не только ре-

презентативным состоянием, выражающим определенную пропозициональную установку, но и некоторым опытом, экспериенциальным свойством: «“Каково это”: чувствовать вкус яблока». Блок оспаривает возможность наличия именно Ф-сознания у «китайской нации».

Попробуем усилить интуицию Блока, покрутив регулятор «скорость работы». Не будем мелочиться с масштабами, представим себе громадный функциональный изоморф мозга, созданный из людей. Пусть это будет целая планета, населенная людьми. Планета населена людьми, каждый из которых соответствует нейрону. Каждый человек может находиться в состояниях, соответствующих состояниям нейронов, а взаимосвязи между людьми на этой планете точно повторяют взаимосвязи между нейронами в человеческом мозге. Будет ли у этой планеты Ф-сознание? Почему нет? Представьте себе, что вы путешествуете по просторам вселенной и наткнетесь на этот гигантский мозг планетарного масштаба. Наверное, вы решите, что такой большой мозг уж точно должен что-то испытывать.

А теперь представим себе, что планета населена очень ленивыми людьми. Они исправно делают свое дело, но только медленно. Процесс, который протекает в обычном мозге за одну секунду, осуществляется населением этой планеты за год. При этом передача информации между отдельными людьми-нейронами занимает не больше минуты, но эти ленивцы большую часть времени просто ничего не делают! Планета-мозг 99% времени бездействует, потому случается минутный обмен сигналами между людьми-нейронами, и на планете снова ничего не происходит в течение нескольких часов.

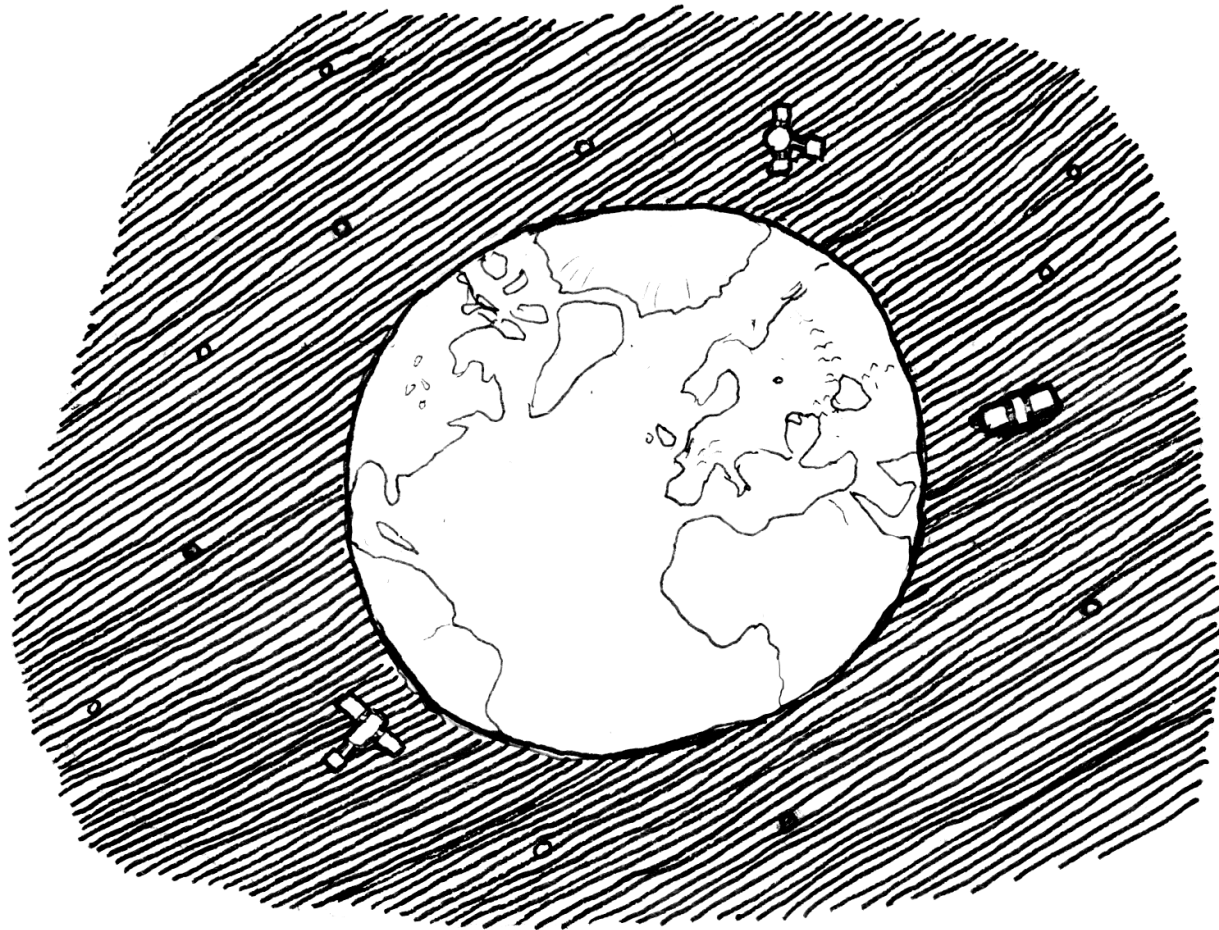
Исследуя планету-мозг, мы обнаруживаем, что она медленно воспроизводит работу мозга Дэвида Юма. У нас, по совпадению, оказывается запись работы мозга Юма за всю его жизнь, и мы выясняем, что сейчас планета-мозг воспроизводит момент в жизни мозга Юма, когда тот пил кларет. Из тезиса функционализма следует, что и планета должна испытывать вкус вина. Но мы видим, что сейчас планета бездействует, она подобна выключенному компьютеру. Испытывает ли она вкус вина в этот момент? Наша интуиция сопротивляется такому заключению. Дальше мы видим, что какое-то количество людей-нейронов обменялись информацией и перешли в другое состояние. Испытывала ли планета вкус вина, когда происходил этот обмен? Здесь наша интуиция также сопротивляется.

Функционалисты могли бы возразить, что Ф-сознание в данном случае дискретно, состоит из отдельных «вспышек», которые происходят во время обмена сигналами. Предположим, что мы наблюдали на планете-мозге обмен

сигналами между несколькими нейронами в тот момент, когда Юм должен был испытывать вкус вина. Проходит значительное время, и мы вновь замечаем такой же (или очень похожий) обмен сигналами, обращаемся к истории мозга Юма и обнаруживаем, что в этот момент Юм спал! Мозг человека работает постоянно, во сне нейроны тоже передают сигналы, даже когда мы спим без сновидений. И функционального различия между работой нейронов во сне и работой нейронов в сознательном состоянии нет. Таким образом, интуиция подталкивает нас к заключению, что единичный обмен информацией между несколькими нейронами не порождает Ф-сознание.

Мысленный эксперимент Блока кажется удачным интуитивным аргументом против функционализма. Но чем дальше мы заплываем, тем сложнее нам обходиться одними интуициями. Однако мысленные эксперименты не только направлены против отдельных теорий, но и на прояснение самих наших интуиций. И если мы переключим тумблер на эксперименте Блока с «опровержение функционализма» на «прояснение интуиций», мы получим прекрасный инструмент для различения Д-сознания и Ф-сознания.

Артём Беседин



Китайский агент

Сёрл придумал отличный мысленный эксперимент в ответ на тест Тьюринга. Интуиция понятна: даже если машина в состоянии пройти этот тест, совсем не очевидно, что машина будет обладать сознанием. Для иллюстрации этого и придумана «Китайская Комната» (КК). Большое число споров свелось к обсуждению последствий мысленного эксперимента, но я хочу обратить внимание на самую возможность существования Комнаты.

По условиям эксперимента, человек работает с книгой правил по обращению с иероглифами так, чтобы его собеседник был уверен в том, что он разговаривает с тем, кто в действительности знает китайский язык и понимает смысл вопросов. Возможно ли составить такую инструкцию? Иными словами, возможно ли описать язык конечным набором правил? На первый взгляд кажется очевидным, что достаточно сложный набор правил будет адекватно отражать структуру языка и создавать ощущение, что КК является сознательным собеседником. Для этого КК должна в первую очередь иметь возможность отвечать на поставленные вопросы. Если ответы будут устраивать собеседника, то он будет уверен в сознательности КК. Я не беру в счет «отписки» на каверзные вопросы и хочу рассмотреть, как по существу КК может отвечать на вопросы, не выдавая собеседнику «я не хочу об этом говорить» или «не знаю и знать не хочу». Для удобства все вопросы можно разделить на два основных типа: поясняющие и проблематические.

Вопросы первого типа могут быть такими: «когда началась Первая Мировая война?», «где находится Лондон?»,

«как изготавливают нержавеющую сталь?». Ответы на них будут однозначными вне зависимости от того, когда и при каких обстоятельствах они были спрошены. С подобными вопросами возможно построение алгоритма ответов: мы анализируем структуру вопроса, выявляем то, что в нём спрашивается, потом обращаемся к базе данных и находим необходимые сведения.

По-другому обстоят дела с проблематическими вопросами. Для примера можно взять такие вопросы: «в чем художественная ценность того или иного произведения искусства?», «почему Раскольников испытывал душевные терзания после убийства старушки?» и т.д. Можно задать вопрос и о пресловутой проблеме вагонетки: стоит ли спасать жизнь пятерых ценой жизни одного? Допустим, для ответа на первые два вопроса машина будет искать рецензии критиков, аналитику произведения, т.е. обращаться к «базе данных», как и в случае с поясняющими вопросами, но какими алгоритмами она будет пользоваться при поиске ответа на проблему вагонетки? Вполне возможно даже, что КК так или иначе сможет дать ответы на эти вопросы, но будут ли эти ответы удовлетворять того, кто их задаёт? Во всех подобных ответах предполагается «понимание» вопроса, исходя из которого может быть сформулирован ответ. Возможно ли алгоритмизировать все это?

Могу предположить, что на вопрос о вагонетке в инструкции будет содержаться два ответа: «я спасу пятерых и убью одного» и «я оставлю все как есть». А теперь главный вопрос к книге правил: что выбрать? Здесь мы имеем два варианта развития событий. Во-первых, в инструкции может заранее быть предпочтён один из вариантов отве-

та — не важно, так ли, что в ней будет прямо прописан единственный ответ конкретно на проблему вагонетки, или так, что выбор этого ответа будет являться лишь частным случаем следования общему алгоритму «между А и В выбор только А». Во-вторых, в инструкции может содержаться указание на то, чтобы находящийся в Комнате провел операцию выбора: например, бросил монетку, посмотрел на часы, в общем, совершил любое действие, которое отличается от «перейди к странице N и выпиши иероглиф q». Следствия в обоих случаях неутешительные.

В первом случае книга правил должна содержать ответы на абсолютно все проблематические вопросы. Если так, то КК будет всезнающим, а самое главное — непоколебимым собеседником, который не способен изменить свое поведение (в данном случае — ответы на вопросы). Еще один довод в пользу маловероятности такого способа реализации — возможность бесконечно уточнять один и тот же проблематический вопрос, например: «говорить неправду плохо?», «говорить неправду, которая приведет ко благу, плохо?», «говорить неправду, которая приведет ко благу, но сделает хуже кому-то, плохо?». То есть мы всегда сможем задать КК такой проблематический вопрос, который будет на одно уточнение больше, чем содержится в ее книге правил.

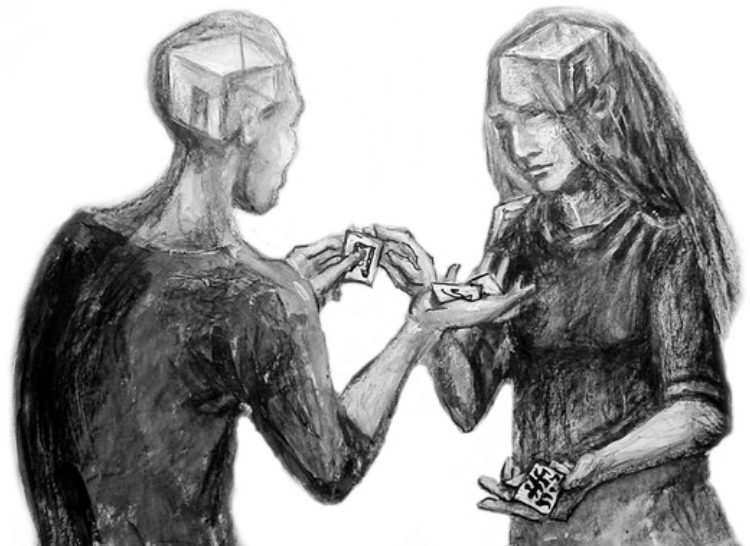
Во втором варианте развития событий мы редуцируем любые проблематические вопросы к процедуре: «если мы должны выбирать, то мы выбираем во всех случаях вариант W». Что получится в этом случае? При достаточно долгом «разговоре» с КК можно будет сделать вывод, что собеседник — не человек и на вопросы он отвечает, потому что так записано в книге правил.

Но допустим, процедура выбора записана в книге правил как «если да или нет, то брось монетку». Это даст определенную вариативность, и на вопрос «почему ты сделал такой выбор?» КК может ответить: «потому что я так решил в этой ситуации». Процедуру выбора можно усложнять до бесконечности, учитывая перекрестные связи слов друг с другом, но, так как это мысленный эксперимент, предположим что машина в каждом случае достаточно уточняет вопрос и редуцирует его к выбору между «да» и «нет» так, что создается впечатление «понимания» вопроса. И вот здесь кроется главный «подводный камень» КК. По условиям мысленного эксперимента, человек в Комнате работает только с книгой правил и иероглифы для ответов берёт только соотнося одни с другими.

Если же мы расширяем Китайскую Комнату до «книга правил и процедура подбрасывания монетки», то теряет главную интуицию эксперимента. Объясню подробнее:

результатом эксперимента Сёрл считает тот факт, что он, совершенно не зная китайского, тем не менее мог на нем говорить, используя всего лишь достаточно сложную книгу правил. Тем самым иллюстрируется принципиальная возможность и прохождения теста Тьюринга. Однако существование процедуры выбора в добавление к книге правил привносит в Комнату именно то, что Сёрл старался избежать. «Агент», который совершает выбор, безусловно, тоже не знает китайского языка. Но что имеет значение — этот агент совершает определенный выбор. Логичным развитием модели с агентом является работа с «записью» всех его «решений», но это не меняет сути. Главное, что для создания машины, которая была бы способна пройти тест Тьюринга, необходим не только набор правил (каким бы полным он ни был), но и специальный «действующий агент».

К чему приводит такая модификация КК? С одной стороны, можно сказать о «привидении в машине» и возвращении «картезианского театра», от которой уважаемый Д. Деннет так старался нас избавить. Но можно взглянуть на это и с другой стороны. Вместо человека в КК можно поместить робота, который будет подбрасывать монетку, когда встретит в книге правил определенную команду. И вот здесь «miracle occurs»: записывая все свои предыдущие «выборы» и имея возможность совершать новые, механизм будет полностью имитировать поведение чело-



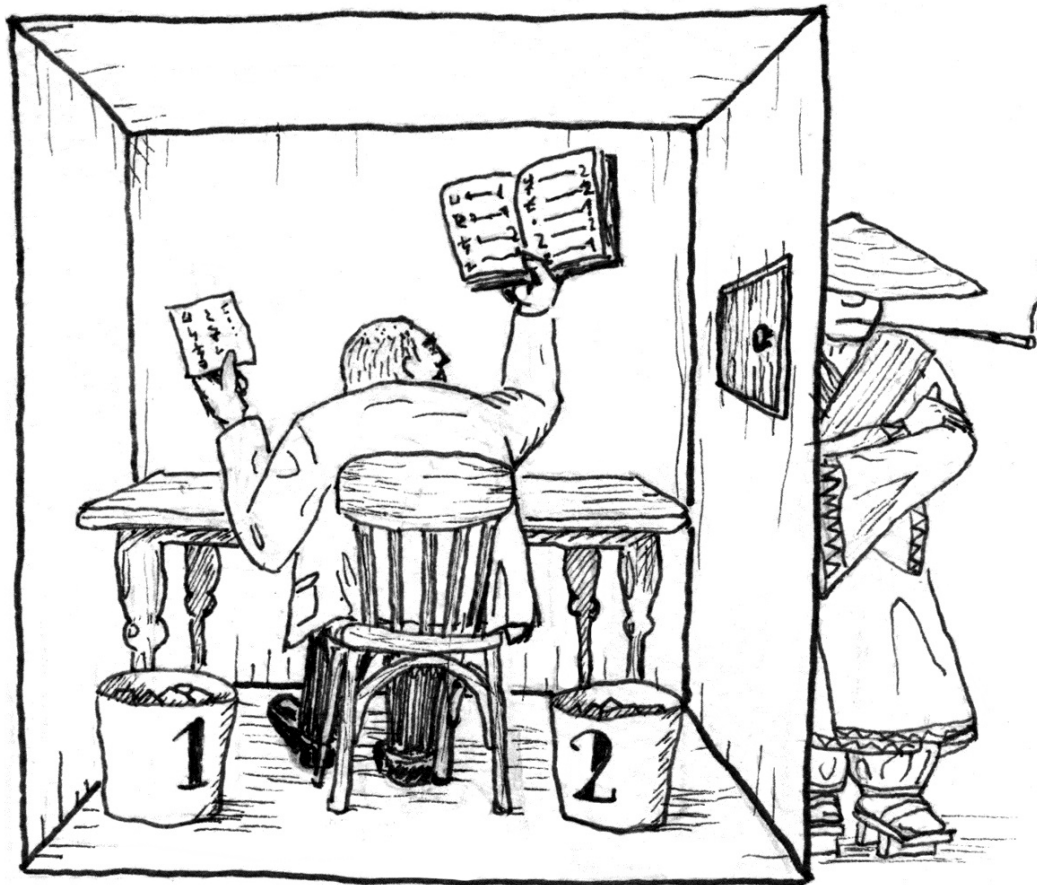
века-агента, хотя реализовываться он будет без всякого сознания.

По условиям эксперимента, КК должна быть способна отвечать на оба типа вопросов: как на поясняющие, так и на проблематические. В противном случае рано или поздно собеседник поймет, что КК китайского языка не знает. Значит, либо КК невозможна (нет ни одного способа реализовать алгоритм поиска ответов на оба типа вопросов), либо она возможна только лишь при наличии «агента». Вывод напрашивается сам собой: даже если мы будем стараться реализовать человеческое (в данном случае — речевое) поведение максимально механистически, мы так или иначе придем к необходимости появления чего-то, что предполагает «самость». «Самость» будет присуща всей

КК, а агент будет оставаться непонимающим. Мы получим систему, которая будет ссылаться на два типа данных: внешние (книга правил) и внутренние (учет результатов прошлых решений при ответе на новые проблематические вопросы). Автореферентность, способность к свободному выбору и максимальная пластичность в поведении служат основанием для приписывания сознания.

Другой вопрос — какие перспективы открываются в таком случае? Вполне возможно, что такая Комната будет обладать квалиа. Или она будет так хорошо притворяться, что мы поверим в то, что она ими обладает, как верим в квалиа других людей.

Кирилл Загустин



Что умеет Машина, но не может позволить себе Сёрл

Единственным надёжным критерием наличия сознания является «разумное» речевое поведение. И мы можем создать такую Машину, которая будет безупречно имитировать речевое поведение разумного существа; такую, что её речевое поведение будет неотличимо от человеческого. Но это и будет означать, что такая Машина удовлетворяет критерию наличия сознания, то есть что такая Машина — разумна. Такова в общих грубых чертах позиция Сильного искусственного интеллекта.

— Ах, как вы ошибаетесь! — восклицает Джон Сёрл. — Ведь наличие сознания предполагает наличие понимания! А о том, понимаю ли я нечто или нет, мне, как и любому разумному существу, известно непосредственно, а не в силу вашего критерия. И если допустить, что Машина, имитирующая речевое поведение человека (т.е. проходящая тест Тьюринга (ТТ)), обладает сознанием, то и я, в точности имитируя её работу, непосредственно выполняю её программу, должен буду обрести понимание. Но этого не произойдёт! Я решительно ничего не пойму, даже если очень буду стараться!

Условия действий Сёрла в Китайской Комнате (КК), надо полагать, известны. Но сможет ли он выполнить свою задачу? Ведь стоит нам лишь подобрать такие вопросы, на которые сможет ответить Машина, но не сможет Сёрл; или такие, на которые Сёрл не сможет ответить, не поняв значение хотя бы одного иероглифа, как мы разрушим КК! И таких вопросов немало.

Среди них в первую очередь достойны упоминания индексикальные вопросы и реплики (зависящие от конкретного контекста), среди которых удобно выделить 2 типа.

Вопросы о «мире». *Который час? Где ближайший бар? Какая сегодня погода? Что идёт в Сатириконе?*

Машина умеет отвечать на подобные вопросы лучше любого из ваших знакомых. Смысл каждого подобного вопроса или связанной с ним реплики (типа «*Напиши мне вечером*» или «*Напомни мне снять бельё, если пойдет дождь*») зависит от «контекста», задающегося конкретными внешними обстоятельствами, меняющимися со временем. Чтобы ориентироваться в нём, Машины обычно используют датчики, тем или иным образом связывающие их с внешним миром. Это может быть счетчик электрических колебаний кварцевого генератора, благодаря которому работает большинство электронных часов, датчик температуры воздуха или даже модем.

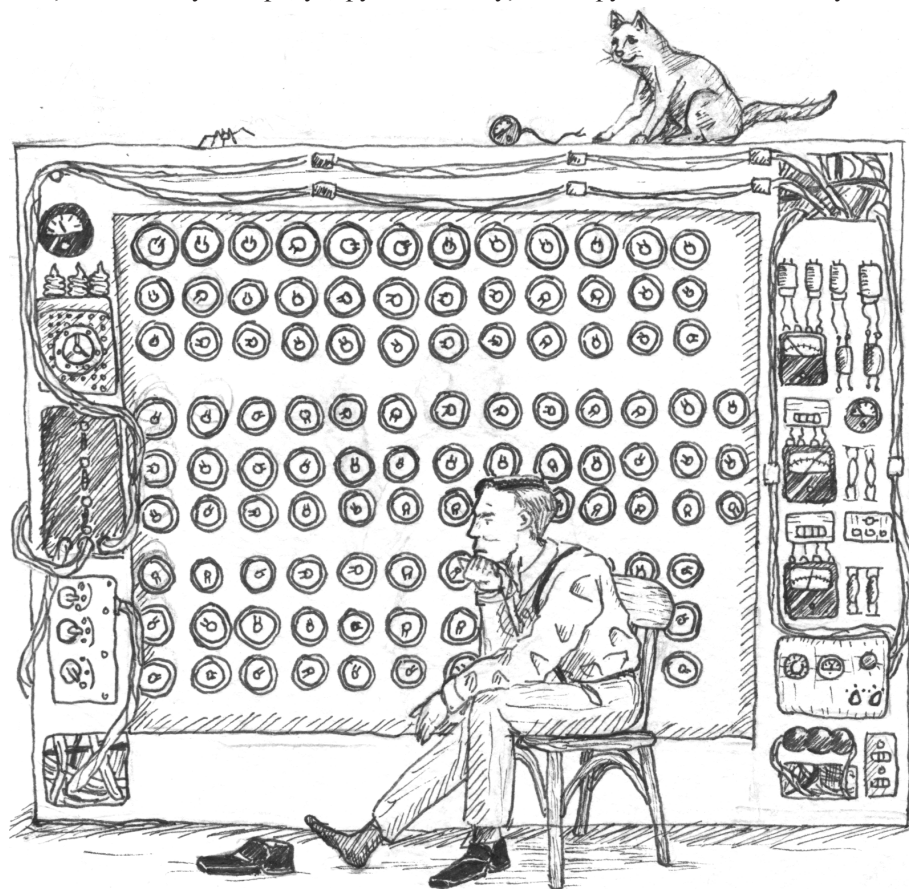
Сёрл не может позволить себе использовать датчики. Представим, что Сёрлу приходит какой-нибудь индексикальный вопрос о «мире» — например, «*Который час?*». Очевидно, что если Сёрл не будет уметь определять время, то не сможет и ориентироваться в нём, и тогда не только прямой вопрос, но и любая иная темпоральная реплика (типа «*Напомни мне об этом завтра*») позволит отличить его от мыслящего существа. Очевидно также, что прямой ответ на такой вопрос невозможно прописать в его инструкции. Допустить, что в КК стоят часы, и инструкция

гласит, что, получив соответствующие китайские символы, Сёрл должен посмотреть на часы, переписать показания и отправить их в ответ, мы также не можем, поскольку в таком случае Сёрл точно поймёт, что значат соответствующие иероглиф-вопрос и его собственный ответ. Даже если такие часы вместо цифр будут показывать время «прописью» (иероглифами), их значение всё равно легко можно будет вычислить.

Но допустим, что мы замаскируем часы так, что Сёрл не сможет их распознать — скажем, под некую «коробочку» неясного предназначения с кнопкой, по нажатию которой будет распечатываться иероглифический ответ (текущее время). Нам, конечно, придется придумать ещё ряд каких-нибудь ограничений ad hoc (типа запрета Сёрлу пользоваться «коробочкой» без прямого указания инструкции — иначе он, опять же, сможет догадаться, что это часы; и т.п.), но даже если их допустить, возникнет вопрос: а кто, собственно, в таком случае продуцирует от-

вет? Становится очевидным, что ответ продуцируют сами часы, а роль Сёрла сводится к простому посредничеству, к передаче ответа от часов к вопрошающему. Даже если ответ «коробочки» будет зашифрован и его необходимо будет перевести, прежде чем послать в ответ, роль Сёрла окажется сильно редуцирована — в лучшем случае, он будет имитировать работу дисплея (переводящего некие непонятные, «зашифрованные» сигналы в узнаваемые нами символы), но никак не Машины в целом.

Часы мы привели лишь в виде частного примера устройства с датчиком — на их месте в «коробочке» мог бы быть термометр, ноутбук с доступом в интернет или Робот с технологией распознавания объектов. Ситуацию с последним разбирает В.В. Васильев, демонстрируя неэквивалентность КК и Китайского Робота. Васильев при этом подчеркивает, что если Сёрл, этот «не очень нужный помощник» Робота, не будет знать полностью его программу, это нарушит изначальные условия его задачи, а если



будет знать, то обязательно разгадает значения некоторых иероглифов. Мы во многом согласны с Васильевым, однако в данном случае не можем придать принципиального значения «знанию». Нам кажется важным, что Сёрл должен самостоятельно исполнять эту программу — в противном случае бремя продуцирования ответа окажется с него снято, что лишит аргумент смысла.

Что может Сёрл, так это симитировать работу Машины без датчиков. Пройдёт ли такая Машина ТТ? Васильев полагает, что Машина без датчиков не может рассматриваться даже как реализация проекта слабого ИИ. Мы в этом плане более оптимистичны и полагаем, что ТТ ей будет по силам. По крайней мере, с вопросом о времени она бы справилась (зная вычислительную скорость Машины, её саму можно превратить в своего рода «часы» — любой циклический «фоновый» процесс можно превратить в хронометр). Доступно ей будет и знание погоды (в базу данных Машины можно загрузить прогнозы синоптиков, и первые несколько дней её ответы будут вполне удовлетворительными), расписаний, киноновинок, результатов текущих спортивных соревнований и иной информации, требуемой для индексикальных ответов. Понятно, что точность таких ответов будет падать с течением времени, но на первых — и достаточно продолжительных — порах её вполне будет хватать для приемлемых в рамках ТТ ответов. Тот довод, что Машина без датчика обязательно выдаст себя когда-нибудь — надо только дождаться, легко парируется. Сколько вам требуется времени, чтобы в условиях «живого» общения понять, что ваш собеседник разумен? Если же в условиях чата «судья» потребует для этого неделю, месяц или годы, резоннее будет поставить вопрос о разумности самого судьи, нежели его собеседника. Требование же точных и верных ответов, предполагающих конкретное сиюминутное восприятие (на вопросы типа «*что вы видите перед собой?*», «*кончился ли дождь?*» и т.п.), в рамках ТТ может быть и не адекватно: судье не обязательно будет известна локация собеседника (для переписки в чате она вообще не должна играть роль) — но тогда и возможности удостовериться в точности и верности таких ответов он будет лишен.

Очень похоже на то, что Сёрл может позволить себе имитацию Машины без датчика. По крайней мере, всё, что возможно внести в базу данных такой Машины (типа последних новостей и прогнозов), можно было бы учесть и в его инструкциях. Но сможет ли он следить за временем? Можно ли допустить дискретность совершаемых им операций? И можно ли будет рассчитать его вычислительную скорость? Что тогда будет уместно сделать «фоновым

процессом»? Не возникнет ли у Сёрла в таком случае понимания собственных действий и своего ответа, когда тот будет запрошен? Неимение однозначных ответов на эти и многие другие вопросы не позволяет нам сделать вывод. Д.Б. Волков, например, утверждает, что Сёрл без датчика не сможет ориентироваться во времени. Интуитивно нам кажется это верным. Но, скажем, Е.В. Логинов полагает, что сможет. Мы не знаем, кто прав.

Что может Сёрл, в любом случае, так это ответить на вопрос «*Который час?*» — «*Не знаю*» или «*Не ношу часов*». Такой ответ можно прописать в «книге правил», и он будет удовлетворять требованиям ТТ. Не стоит забывать, что ответы в ТТ не обязательно должны быть точными и верными. Вот только если Сёрл будет увёртываться от каждой темпоральной реплики и от большинства индексикальных вопросов о «мире», он вскоре выдаст себя.

Вопросы о «диалоге». Контекст этого типа индексикальных вопросов задаётся непосредственно ведущимся диалогом, а ответы могут требовать некоторых операций с ним. Речь не только об употреблении разного рода местоимений, но и о метазамечаниях: «*Мы с Костей играли в кости. Пришли гости, и Костя спрятал кости. Сколько раз я произнёс имя мальчика?*»; «*Сколько предложений было в предыдущей реплике?*» и т.д. Человек обычно легко отвечает на подобные вопросы, особенно если они изложены в письменном виде.

Машина умеет отвечать на них лучше человека. Ваш текстовый редактор знает, сколько знаков с пробелами содержится в вашем документе, не говоря о предложениях, а программный поиск по ключевому слову давно не вызывает удивления. В этом отношении справляться с местоимениями Машине значительно сложнее, но и эта задача давно не выглядит невыполнимой. Отметим также, что в данном случае различия Машины с датчиком и Машины без датчика в принципе не прослеживаются.

Сёрл не может позволить себе свободные операции с диалогом. Мы в принципе не представляем, как должна быть написана его инструкция, чтобы адекватно распознавать местоимения. Представьте, что КК говорит нам, что кто-то уже высказывал это возражение Сёрлу. «*А он им на это что?*» — спросим мы. Получим ли мы адекватный ответ?

Аналогично — с метазамечаниями. Что должно быть прописано в «книге правил», чтобы Сёрл ответил на вопрос про Костю? «Если вам пришел вопрос *Сколько раз я произнёс X*, просмотрите поступившие вопросы, и если X встречается в них 0 раз — отвечайте иероглифом *Ноль*; если

1 — *Один*; 2 — *Два...*» Нам кажется, примерно так можно запрограммировать Машину, но ведь это совершенно не годится для Сёрла — он обязательно поймёт значение своего ответа! Д.Б. Волков, высказывая это замечание, оговаривает возможность наличия в распоряжении Сёрла вариантов всех возможных разворачивающихся бесед, однако такое допущение представляется слишком сильным. Оно предполагает фактическое всезнание КК. Нет сомнений, что такая КК пригодилась бы человечеству, однако Сёрл, имитирующий работу Машины, не должен обладать информационными ресурсами, превосходящими имеющиеся в её распоряжении. Машина же легко справляется с такими вопросами, не обладая всезнанием.

Что может Сёрл, так это ответить «42» или «Да ну тебя! Я думал, мы за жизнь поговорим, а ты...». Т.е. парировать каким-нибудь универсальным ответом, ориентированным на смену темы. Разок-другой в рамках ТТ это может сработать, но стоит лишь повторить подобного рода вопросы или предвать первый из них фразой «Хочу удостовериться, что ты — не Машина. Ответь, пожалуйста...» — и человек, как и Машина, дадут верный ответ, но Сёрл — вряд ли.

Мы видим, что индексикальные вопросы и реплики могут представлять существенную проблему для Сёрла в КК, но их одних может оказаться недостаточно для диагностирования отличия КК от Машины. В таком случае на помощь приходят «редуктивные» методы. Их суть состоит в том, чтобы максимально упростить диалог и построить его так, чтобы Сёрл не мог не понять значения хотя бы некоторых «китайских» символов. Вот пара примеров.

«Языковые игры». Речь идёт не столько о витгенштейнианских играх, сколько об установлении определенных правил ведения диалога в духе игры. Можно, например, попросить собеседника отвечать вам только «да» и «нет», или самим начать использовать лишь односложные реплики. Диалог тогда будет подобен игре в съедобное-несъедобное, тем более что в китайском языке нет иероглифов, означающих «да» и «нет» — вместо них в ответе используется повтор глагола или его отрицания (например, «Ты любишь шоколад?» — «Люблю» — «А чеснок любишь?» — «Не люблю»).

Машина умеет играть в такую игру более чем непринужденно.

Но **Сёрл не может позволить себе** сыграть в неё — очень скоро по характеру использования символов он сможет определить, по крайней мере, значение иероглифа «не» и, скорее всего, заподозрит, что его ответы выражают

утверждения и отрицания.

Что может Сёрл, так это, как водится, уклониться от этой игры, как и от множества ей подобных.

«Картинки». Вопросы и реплики с их использованием предполагают выражение основного смыслового содержания в наглядной графической форме. При этом иероглифы начинают играть существенно меньшую роль, и их значение становится легко узнаваемым. Такими наглядными элементами могут служить не только фотографии или рисунки, но и графические шрифты, дополнительные символы (геометрические фигуры, стрелки и т.п.), графика, созданная средствами стандартного шрифта (начиная фигурным строением текста (стихов), заканчивая смайлами) и т.п. Использование всех этих средств представляется возможным в рамках современного ТТ, проводимого в условиях чата (более того, в условиях чата возможно использование не только графического, но аудио- и видео-сопровождения и т.п.), а некоторых простейших из них — даже при коммуникации по телетайпу или морзянкой.

Машина умеет успешно работать со многими (пусть и не всеми) типами таких «картинок» и «интерпретировать» их должным образом. Не говоря о знаках и символах, «узнать» актёра по фотографии или «найти 10 отличий» не составит для неё особого труда.

Сёрл не может позволить себе работу ни с какими «картинками». В самом деле, много ли ему потребуется времени, чтобы разгадать, что значат иероглифы, например, в реплике «Я не ♥ ♣, я ♥ ♠ и ♣»? Аналогично и фразы о «10 различиях» будут во многом для него понятны.

Что может Сёрл, так это запретить использование таких «картинок». Он может сказать, что оно нарушает условие изначальной задачи, согласно которому узник КК не должен заранее знать значения ни одной используемой графемы или иного элемента, несущего смысловое содержание. Но тогда придётся запретить не только картинки, стрелочки или смайлы, но и любые знаки пунктуации, цифры, математические символы и даже те заранее неизвестные узнику знаки, о значении которых можно догадаться по их внешнему виду — как, например, знаки пиктографического письма, к которым, к слову, относятся и китайские иероглифы!

Все приведенные возражения кажутся убедительными. Действительно, в описанных условиях Сёрл не сможет сымитировать Машину, проходящую ТТ, не разгадав значения того или иного иероглифа. Можно сказать, что понимание Сёрлом значения некоторого символа не

равносильно пониманию им китайского языка, но это возражение парируется напоминанием условий изначальной задачи: Сёрл обещал не понять решительно ничего, ни одной закорючки, а здесь он всё-таки что-то да понял, что, скорее всего, даст ему ключ и к пониманию иных иероглифов. В оправдание Сёрла можно было бы привести и более веский аргумент: понимание отдельных иероглифов возможно, однако оно не будет являться пониманием в силу исполняемости программы. Сёрл в КК может понять некоторые символы, но лишь потому, что имеет некоторый сторонний опыт: историю оперирования знаками естественного языка, представление о времени и т.п.

На это возражение, однако, находится ответ, гласящий, что именно так и строится обучение языку. Как ребёнок узнаёт значение слов «мяч» или «юла»? Так ли, что, впервые услышав их, он тем самым одновременно воспринимает и «заложённое» в них значение, или понимание строится на установлении условно-рефлекторной (ассоциативной) связи звука слова и присутствия в поле воспри-

ятия соответствующего элемента, который выделяется тем четче, чем чаще в различных контекстах этот звук воспринимается? Если допустить «мистический» первый случай, Сёрл должен будет понять значение иероглифов по одному их восприятию. Второй же случай аналогичен ситуации Сёрла — с тем лишь отличием, что общий, доступный обоим сторонам контекст дан Сёрлу не в непосредственном восприятии внешнего мира объектов, но его воспоминаниями. Если ребёнок со звуком «мяч» соотносит данные перцепции, то Сёрлу не с чем объективно соотносить знаки, которыми он оперирует, кроме как с данными памяти. Воспоминания служат для него тем, чем для ребёнка — вещи внешнего мира. Но если эти ситуации аналогичны, КК сможет научить китайскому языку, а это и будет означать провал сёрловского эксперимента! Более того, как кажется, не обнаруживается оснований, которые заставили бы нас отказать Машине в возможности построения связей (аналогичных ассоциативным) поступающих ей на «вход» символов с некоторыми данными памяти или даже, если она обладает датчиками, внешними объектами и факторами.

Наконец, в защиту Сёрла можно привести решающий аргумент. По условиям задачи, конкретно китайские иероглифы не имеют значения, принципиальным является лишь факт использования «непонятного набора знаков», а раз так, то суть КК не изменится, если вместо фраз на китайском Сёрл будет оперировать шифровкой. Но фразы естественного языка можно зашифровать так, что в отсутствие ключа разгадать их будет невозможно. Носителю такого «языка», т.е. носителю ключа дешифровки, ответы Сёрла будут казаться столь же разумными, что и китайцу — ответы классической КК. Но «редуктивные» методики в таком случае уже не работают.

Более того, такая модификация КК, похоже, снимет и все индексикальные возражения. Так, в условиях использования поливариантности шифрования (если один иероглиф будет зашифровываться 10^n способами), исключая повторяющиеся последовательности знаков, применения правил перестановки их последовательности или дополнения их «отвлекающими» бессмысленными символами, Сёрл не сможет установить значения ни одной из таких последовательностей, даже если инструкция будет содержать прямые указания на то, как следует отвечать на тот или иной вопрос о «диалоге» или о «мире», пусть даже требующий конкретного сенсорного восприятия. Он, очевидно, не сможет точно установить, что означает та или иная последовательность символов (скажем, что «90m» значит «сто!»); не определит он и предмет разговора. Пред-



ставьте, что на вопрос «ISwa4 .h7vn 90m KLyw)n /fzж4@g15bns u4 tu#...» Сёрл, согласно инструкции, должен посмотреть на свой стол, и, если на нём стоит банка Кока-Колы, ответить «nK3, =dëIII>c 7&b...». Он может, конечно, предположить, что речь в его ответе идёт о Кола-Коле, но с тем же успехом этот диалог может дешифроваться как «У вас есть в комнате стол?» — «Да». В условиях неповторяемости последовательностей знаков шифровки, у Сёрла не будет никакой возможности проверить своё предположение о предмете беседы, а инструкция, составленная с долей выдумки, может свести на нет вероятность его простого угадывания. Тем самым Сёрл сможет позволить себе даже имитацию машины с датчиками, задействуя вместо последних собственное восприятие, но останется обречён на непонимание.

Однако не совершим ли мы при допущении возможности шифрования принципиальной подмены естественного языка искусственным? Возможно, но тем больше это будет роднить КК с машинным вычислением. Машина оперирует не знаками естественного языка, но их кодировкой (в простейшем случае — в двоичном коде). Есть, конечно, разница между кодом и шифром, но в данном случае она уже не играет роли.

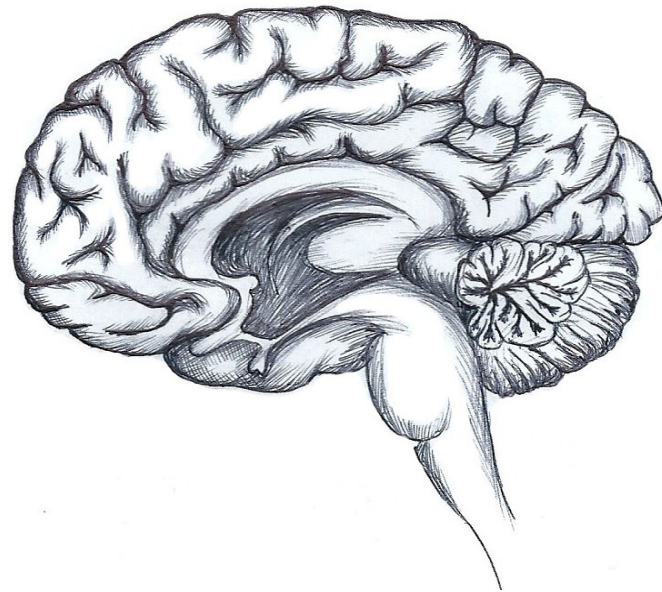
Что всё это значит? Мы показали, что существуют такие условия (условия оперирования знаками естественного языка), что Сёрл не сможет справиться со своей задачей, не обретя понимания. Если Сёрл позволит себе выполнить программу Машины, проходящей ТТ (и тем самым симитировать её работу), он обязательно поймёт значение хотя бы некоторых знаков. А если он позволит себе остаться в неведении, то, не нарушая иных условий своей задачи (условия самостоятельного продуцирования ответа и т.п.), Сёрл проиграет игру в имитацию — мы сможем отличить его от Машины. Мы показали также, что существуют такие условия (с хитроумной шифровкой), в которых Сёрл сможет имитировать Машину, не обретая понимания.

Следует ли отсюда, что Сёрл не прав и КК проваливается в силу наличия таких условий, что её узник непременно обретёт понимание? Или, напротив, что Сёрл прав и Машина никогда не обретёт понимания в силу формальности своей программы, что подтверждается именно вариацией с шифровкой? Можно ли заключить, что прав Тьюринг и Машина сможет мыслить и понимать, стоит лишь создать соответствующие условия (скажем, подобрать должный язык программирования — аналог естественного языка для КК)? Или, быть может, что Тьюринг не прав и его

критерий определения наличия сознания по речевому поведению не надёжен, поскольку допускает возможность прохождения ТТ без фактического наличия сознания и понимания у испытуемого?

Мы полагаем, что наше рассмотрение наглядно демонстрирует лишь то, что аргумент от имитации не является в должной степени релевантным. Из того, что Сёрл поймёт нечто или не поймёт ничего, мы не можем заключить, что и имитируемая Машина обретёт или не обретёт некоторое понимание. Но то же справедливо и применительно к самому ТТ: имитация Машиной речевого поведения человека не может служить надёжным свидетельством ни в пользу её «разумности», ни в обратную сторону. Скорее, она выступает лишь косвенным признаком, намёком на возможность наличия разума, что может служить поводом для постановки вопроса о нём, но не даёт достаточных оснований для ответа на этот вопрос. Такой вывод побуждает к ревизии самого критерия наличия сознания и способов осуществления проверки на соответствие ему.

Андрей Мерцалов



Какая машина СМОЖЕТ МЫСЛИТЬ?

Психология мышления — одна из самых неразвитых областей психологической науки. Всё, к чему она пришла в ходе продолжительной истории попыток грамотно сформулировать исследуемые проблемы, так это к тому, что мышление не похоже ни на случайный поток ассоциаций, ни на выведение силлогизмов. Последнее часто и берётся на вооружение при ответе на вопрос «может ли машина мыслить?». Подавляющее большинство рассуждений на эту тему содержит в себе следующую цепочку суждений: 1) Машина выполняет заданные программистом алгоритмы; 2) Эти алгоритмы основаны на известных законах логики; 3) Человек не мыслит как машина для силлогизмов; 4) Следовательно, машина мыслить не может.

Эта основная цепочка украшается разными идеями типа вычислимости/невычислимости, трудной проблемы сознания, системных свойств, которые присущи живым существам, и т.д. Однако ошибка кроется в первой посылке, а именно — в представлении о том, как работает компьютер. Но прежде, чем выявить её, уместно обратиться к истории когнитивной науки и метафоры «мозг-компьютер».

Когнитивная психология возникла в контексте критики бихевиоризма и появления компьютеров. Метафора, что мозг работает примерно как компьютер, как бы напрашивалась сама собой: вот устройства ввода (сенсорные системы), устройства вывода (мышцы), долговременная и рабочая память, процессор где-то в лобных долях. То, что называется «психикой» — это software, мозг — hardware. Совершенно очевидное сравнение. Психологи, порядком уставшие от бихевиоризма, но не желавшие мисти-

ки психоанализа, с радостью восприняли этот подход и начали увлеченно строить схемы с блоками и стрелочками. Такой подход получил название символического и стал самым что ни на есть мейнстримом в когнитивной психологии. Мы говорим «когнитивная психология» — подразумеваем «символический подход»; говорим «символический подход» — подразумеваем «когнитивная психология». Сейчас представители этого подхода пытаются исправить его недостатки, а заодно и все неувязки сравнения мозга с компьютером. Когнитивные психологи вовсю исследуют социальные аспекты познания, эмоциональное познание, «включенное» (во внешнюю среду) и «воплощенное» (в костях и мясе нашего тела) познание, и всё остальное, что отличает человеческое мышление от компьютерной обработки информации. Они обращаются к помощи психофизиологов: хотя поначалу их игнорировали, в наше время радители за «чистую психологию» без фМРТ и ЭЭГ выглядят как выжившие из ума деды. От оригинальной метафоры осталось не так уж и много: лишь дух рационализма (в противовес «гуманитарным» психологиям — психоанализу, гуманистической, экзистенциальной психологии и т.д.) и «компьютериализма» (в противовес идеологии стимула и рефлекса бихевиоризма), ну и преемственность от когнитивизма старой школы, конечно.

Но вот уже лет 70 существует альтернатива классическому символическому подходу как в когнитивистике, так и в computer science и исследованиях ИИ, получившая название «коннекционизм». Начался коннекционизм с того, что нейрофизиолог Маккалох и юный логик Питц попытались создать математическую модель нервной клетки.

Общее представление о том, как работает нейрон,

в те времена не сильно отличалось от современного: нейрон — это клетка, отличающаяся тем, что обладает, помимо основного тела (сомы), еще и отростками — дендритами, которые получают импульс и передают его по соме до аксона — длинного отростка, который, в свою очередь, возбуждает следующие нейроны. Нейроны почти вплотную соединены друг с другом синапсами, в которых и происходит передача информации с одного нейрона на другой (обычно — с аксона на дендрит). В основном это происходит за счет того, что возбуждение пресинаптического нейрона приводит к выбросу химических соединений (нейромедиаторов), которые соприкасаются с белками-рецепторами постсинаптического нейрона, изменяя потенциал мембраны. Одного медиатора редко бывает достаточно, поэтому нейромедиаторов требуется достаточно много. К тому же существуют и тормозные синапсы, выброс нейромедиаторов в которых приводит к еще большему снижению потенциала мембраны. Поэтому для возбуждения нейрона обычно необходима передача возбуждения сразу с нескольких других нейронов. Еще одно важное правило передачи сигнала в нервной системе — это принцип «всё или ничего»: если потенциал мембраны переходит определенное пороговое значение, то нейрон возбуждается и сигнал переходит дальше (от дендритов к соме, от сомы к аксону). Если же порогового значения потенциал не достигает, то нейрон так и продолжает «молчать».

Маккалок и Питтс пытались представить работу нейрона математически: представим, что есть определенная функция, которая перемножает приходящее возбуждение в виде некоторого вектора чисел (вектор возбуждения) на вектор связи (синаптические веса дендритов). Если скалярное произведение этих двух векторов больше определенного значения, то на выходе такой «нейрон» дает 1, если меньше — 0 (этакая реализация принципа «всё или ничего»). Вектор возбуждения — это как бы нейромедиаторы, выбрасываемые из пресинаптических клеток, вектор связи — белки-рецепторы на постсинаптических нейронах. Вот такое вот грубое, но крайне милое представление живой клетки как функции-сумматора приходящих импульсов.

Если мы сделаем много таких формальных нейронов, то сможем связать их друг с другом в так называемую «искусственную нейронную сеть». Но самое интересное, что такую сеть мы сможем обучать. Для этого нам нужно будет лишь давать какое-то возбуждение на входе и реакцию на ответ сети на выходе. При неверном ответе мы будем менять вектор связей до тех пор, пока он не станет верным для большинства примеров. Например, мы можем

таким образом научить сеть различать буквы. И главная прелесть в том, что нам здесь вообще не нужно будет понимать, что и как именно эта сеть обрабатывает. Мы даем ей вход, требуем определенный выход, а что происходит внутри — мы сами не задаем и не обязаны этого знать. Такую искусственную нейронную сеть практически реализовал в 1958 году Розенблатт, и пошло-поехало... Со временем использование искусственных нейронных сетей стало одним из ключевых методов в машинном обучении. Само по себе развитие машинного обучения наглядно показало, что программисту не обязательно все знать заранее.

Однако в дальнейшем искусственные нейронные сети не показали тех революционных результатов, которые от них ожидалось. Как оказалось, задачи менее узконаправленные, нежели классификация паттернов, для них уже неподъемны. Хотя с их помощью и можно решать стандартные задачи типа арифметических, обычный компьютер делает это быстрее и эффективнее. Распознавание речи, лиц, текста — все те вещи, что с трудом делает компьютер, но с которыми легко справляется человек, стали основной областью применения искусственных нейронных сетей.

Вероятно, мышление у человека устроено примерно так же: сформированные на основе прошлого опыта связи позволяют генерировать быстрые решения задач, которые уже потом проходят проверку. Подобный механизм может лежать в основе формирования «интуитивных» действий, при этом сам человек может и не понимать, как он пришел к тому или иному решению.

В современной когнитивной психологии коннекционизм занимает довольно скромное место, в том числе потому что модели, построенные с помощью тех или иных алгоритмов машинного обучения, мало помогают понять их механизмы. Однако значительным его достижением стала демонстрация того, что компьютер не обязательно должен работать как некоторый записанный программистом алгоритм. Он может сам обучаться, ему нужно для этого лишь задать общую архитектуру (которая тоже может быть пластична). Это не говорит напрямую о том, что машина может мыслить, однако я не вижу никаких принципиальных ограничений для развития искусственного интеллекта. А это означает, что мы, возможно, увидим интеллект сильнее человеческого — и кто знает, к чему это может нас привести.

Иван Поздняков



Жук в коробке

Анализ употребления ментальных терминов

Алан Тьюринг предложил считать критерием разумности машин способность успешно играть в лингвистическую игру в имитацию. Данный критерий черпает свою надежность из некоего понимания связи языка и мышления, однако нельзя сказать, чтобы эта связь была очевидной. В частности, туманным остаётся вопрос о природе ментальных терминов.

Вопрос о статусе ментальных понятий уходит корнями еще во времена античности и приобретает особую актуальность у Томаса Гоббса. В своей гносеологии Гоббс развивает конвенциональную теорию языка, в которой значения ментальных понятий, так же, как значения терминов математики и физики, приобретают статус установленных соглашений по поводу их употребления. Без взаимного общения людей слова не были бы знаками, существуя лишь на уровне меток, создаваемых памятью для внутреннего индивидуального использования. Язык, а вместе с ним и наука, являются результатом сложной системы коммуникативных соглашений. При этом истина, будучи значением высказываний, может носить лишь относительный характер, так как сам язык условен. Иными словами, Гоббс приходит к выводу о невозможности индивидуального языка, а также к отрицанию абсолютной истины и метафизики.

В «Философских исследованиях», размышляя об индивидуальном языке и сознании, Людвиг Витгенштейн проводит анализ субъективных переживаний. В одном из своих аргументов философ обращается к феномену боли как специфическому явлению приватной сферы человеческого опыта. Боль занимает для Витгенштейна особое место в понимании сознания, так как, в отличие от многих других ощущений, таких как цвет или звук, боль в значительной степени свободна от интересубъективной интерпретации. Другими словами, мы все видим красное, но никому из нас не дано почувствовать боль другого человека. Говоря о боли, мы попадаем в асимметричную ситуацию «опыта от первого лица», и это условие существенным образом определяет правила языковой игры, в которой мы оказываемся. Витгенштейн предлагает мысленный эксперимент, который иллюстрирует неоднозначность употребления ментальных терминов. Предположим, говорит философ, что у каждого есть коробка, в которой находится то, что называют жуком. Нельзя заглядывать в чужие коробки, поэтому о том, что такое «жук», каждый судит только по внешнему виду своего жука. Австрийский философ не исключает, что в каждой коробке могут находиться совершенно разные вещи или даже нечто, подверженное постоянным изменениям. Суть от этого не изменится: люди, приняв правила этой языковой игры,

будут называть словом «жук» то, что находится в коробке. Витгенштейн делает вывод, что в рамках языковой игры объект ощущения (например, боль) не важен. «Если грамматику выражения ощущения трактовать по образцу “объект и его обозначение” — пишет он, — то объект выпадает из сферы рассмотрения как не относящийся к делу». Значения слова «жук», единого для всех, также не существует. Оно рождается в рамках описанной социальной языковой практики.

В отличие от «стула», «боль» не имеет референта, в равной степени доступного всем участникам языковой игры. Но отсюда следует, что у них не будет никакой возможности убедиться в тождестве референта используемых ментальных терминов, а если они могут быть различны, то нам придётся либо, признавая субстанциональность ментального, допустить возможность существования целого сонма разных субстанций, либо, что гораздо экономичнее, отказаться от приписывания ментальности статуса субстанции. В таком случае употребление ментальных терминов оказывается обусловленным социальной языковой практикой, которая диктует определенные правила их использования.

«Философские исследования» оказали существенное влияние на взгляды Гилберта Райла. Английский философ ставит перед собой задачу выявления категориальных ошибок и концептуальных подмен в использовании ментального словаря. Он отмечает, что функционирование ментальных терминов совершенно не похоже на то, как работают слова, референтами которых являются физические предметы и явления. Референты таких слов, как «желание» или «боль», относятся к другой онтологической категории, нежели референт слова «стол».

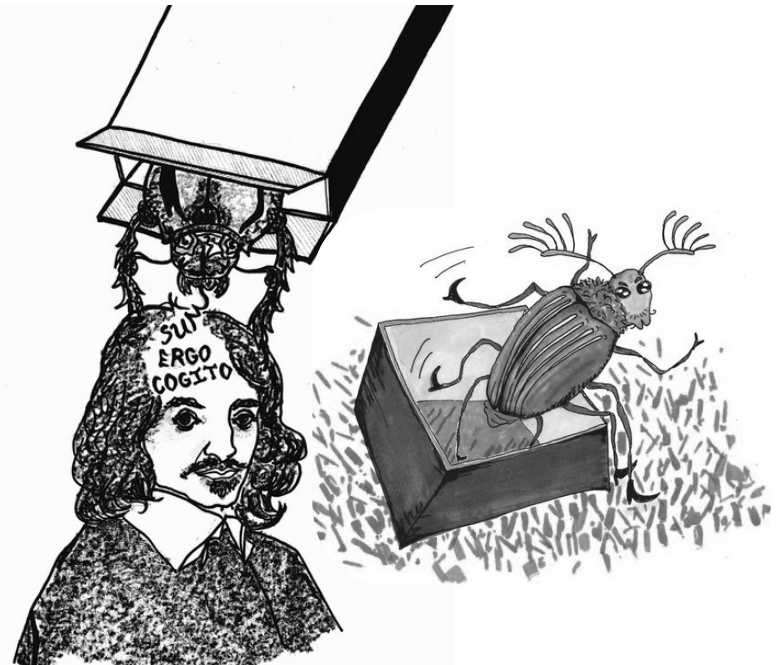
Неоправданное выведение человеческого Я за пределы физического тела, проведенное Декартом, по мнению Райла, породило «приведение в машину». Декартовская модель сознания, в которой психические события протекают изолированно от тела, искажает наше понимание природы ментального. Такой образ сознания Деннет впоследствии назовет «картезианским театром»: представление сознания как некоего «спектакля», непосредственно доступного для восприятия единому центру, «Я», гомункулусу. Если принять такую трактовку, мы столкнёмся с проблемой бесконечного регресса: восприятие самого гомункулуса должно строиться по той же схеме, он сам должен иметь некий «внутренний центр», своего «зрителя в театре». Таким образом, проблема объяснения сознания лишь отодвигается на шаг, но не решается. Райл настаивает на том, что ментальные данности совсем не похожи

на множество маленьких субстанций, духов, обитающих в темнице человеческого тела. Скорее, они представляют собой поведенческие диспозиции, то есть предрасположенности к определенному действию, поступку. Борьба с картезианским дуализмом, согласно Райлу, является необходимым этапом исследования сознания.

Легко проследить аналогию между мысленным экспериментом «Жук в коробке», предложенным Витгенштейном, и критикой Райлом картезианского дуализма. Схожими являются цели двух авторов, а именно — демонстрация некорректности картезианской модели сознания, которая так основательно закрепилась в обыденном мышлении.

Возвращаясь к началу нашей статьи, мы теперь можем сказать, что Витгенштейн и Тьюринг по вопросу об искусственном интеллекте придерживаются сходных позиций. Рассуждения Тьюринга об игре в имитацию как о релевантном критерии сознательности машины напоминают аргумент «жука в коробке». «Языковая игра» трансформировалась у английского математика в игру в имитацию. «Жук в коробке» становится еще одним аргументом в пользу того, что пройденный тест Тьюринга даёт нам достаточное основание для приписывания машине сознания.

Мария Ананина



Естественный интеллект

Интеллект не может быть искусственным, так же как робот не может быть натуральным. Это просто замечание о философской грамматике. Научная фантастика, жаргон IT-специалистов и заявления исследователей в области философии сознания и когнитивных наук могут вводить нас в заблуждение, но это не меняет сути дела. Сказать о машине, что она обладает интеллектом, разумом или сознанием, что машина размышляет, думает или, например, разговаривает — значит совершить ошибку. Или, если угодно, мы можем использовать слова такого рода о машинах, но мы никогда не сможем говорить так о них в том же смысле, что и о человеке. Значения слов «интеллект», «мыслить», «думать», «говорить» применительно к людям и машинам будут различаться не меньше, чем значение слова «работать» применительно к столяру и деревообрабатывающему станку. Сложно расставаться с привлекательным образом разумных роботов. Хочется верить, что очередной успех в области машинного перевода или киберпротезирования приближает нас к созданию того, что мы сможем с чистой совестью назвать сильным искусственным интеллектом. Но, увы, это не вопрос количества, не вопрос вычислительных мощностей или технологий и, в конечном итоге, это не вопрос времени. Это вопрос грамматики нашего языка, правил применения наших слов.

Источниками наших иллюзий насчёт искусственного разума являются следующие два фактора: во-первых, непонимание фундаментальной роли нашей эмоционально-

сти, чувственности, нашей животности для мышления и, во-вторых, систематическая недооценка важности практики при обсуждении вопросов человеческой природы, сознания и языка.

Достигнув ясности относительно этих двух пунктов, мы сможем по-иному взглянуть на проблему искусственного интеллекта. Забегая вперёд, скажу, что сами пункты связаны между собой: животное поведение составляет базис человеческой деятельности (языковых игр), которой, в свою очередь, соответствуют такие понятия, как «интеллект», «разум», «мышление», «вычисление» и им подобные. Я попытаюсь продемонстрировать, что машина не способна к поведению, и, следовательно, к деятельности и, соответственно, к мышлению. Это значит, что употребление ментальных терминов в отношении машин либо нарушает грамматику, либо носит метафорический, шуточный или литературный характер.

Мышление, самосознание, разум — это то, чем «обладает» человек и «не обладает» животное. Именно концепту «животное» посвящена работа Джорджо Агамбена «Открытое». Отталкиваясь от размышлений по поводу средневековой миниатюры, на которой спасшиеся и пирующие после апокалипсиса праведники изображены с головами животных, он переходит к теме конца истории, рассуждает о принципе работы антропологической машины в те или иные эпохи, углубляясь при этом в различные «технические» трудности, с которыми сталкивались схоластика и естествознание при работе с концептами

«животное», «человек», а также с проблемой телесности. Одним из ключевых для данной работы является понятие антропологической машины, функцией которой является производство человечности (не в смысле гуманности, а скорее в смысле человеческой сущности и исключительности), благодаря чему выстраивается иерархическое различие между человеком и животным (человек, само собой, занимает более высокое место в создаваемой иерархии). Единственный доступный нам способ определить, что такое человек — это определить человека в перспективе животного, которым он сам в совершенно определённом смысле является (плоть, тело), но от которого при этом разительно отличен (сознание, дух). Особенностью антропологической машины является то, что нам вовсе не обязательно обладать полным и чётким определением или описанием животного: животности достаточно присутствовать в качестве не вполне ясного фона.

Агамбен обращается к исследованиям биолога Якоба фон Икскюля, касающимся окружающего мира и поведения животных: «Икскюль начинает с различия *Umgebung* [окружающая среда], объективного пространства, где мы видим живые существа в движении, от *Umwelt* [окружающий мир], складывающегося из более или менее обширного ряда элементов, которые он называет носителями значения и носителями признаков и которые только и интересуют животных». В этом контексте особый интерес для Агамбена представляет пример Икскюля с клещом, окружающий мир которого ограничен всего четырьмя типами «носителей значения» (термин Икскюля): светом (который клещ воспринимает всей поверхностью тела и к которому он ползёт), масляной кислотой, содержащейся в поте зверей (почувствовав её запах, он падает с ветви), теплом, а также гладкой поверхностью кожи (почувствовав которые, клещ впивается в тело жертвы). Кроме них клещ ничего не воспринимает. Мир клеща настолько скуден, что в нём нет места даже для крови, которую он пьёт: клещ не воспринимает её вкус и вообще не имеет органа вкусового восприятия.

Пример с клещом особенно заинтересовал Хайдеггера и послужил материалом для его рассуждений об отношении между живым существом и *Dasein*. Хайдеггер изменяет терминологию: он именуется «растормаживателем» то, что Икскюль называл носителем значения; «способностью к» — действующий орган (орган восприятия), а также заменяет метафизически нагруженный термин «жизненный мир» на нейтральный «цикл растормаживания».

Животное, по Хайдеггеру, обделено миром; его способ бытия — это бытие в оцепенелости. Оно полностью погру-

жено в своё взаимодействие с растормаживателями, которые, несмотря на то, что целиком захватывают животное, никогда не открыты ему. Животное не способно на действие, только на поведение. Оно не открыто ни миру, ни себе. Оно захвачено и погружено в мир, которого лишено — в том смысле, что не способно ничего воспринять как таковое, как сущее. Это отсутствие открытости не означает наличия закрытости, здесь просто нет места для таких альтернатив, нет таких грамматических возможностей.

Человек отличается от животного не отсутствием растормаживателей (они есть, как и погруженность в цикл растормаживания), а открытостью, способностью воспринять сущее как сущее, как таковое. *Dasein* — это животное, которое способно прервать свою захваченность растормаживателем, способно открыться ему, а также открыться своей захваченности, открыться своей животности, то есть открыться своей неоткрытости, сокрытому как таковому.

Перенося эти рассуждения в область исследования искусственного интеллекта, можно утверждать, что для создания мыслящей машины сперва надо создать захваченное миром животное, которое было бы оглушено своими растормаживателями и погружено во взаимодействие с ними. Это представляется мне невероятным, даже более невероятным, чем взять уже готовое животное и наделить его интеллектом, так сказать, «вдохнуть искру разума». Машина намного ближе к камню или водопаду, чем к клещу, жаворонку или человеку. Бессмысленно и нелепо говорить о машине, что она захвачена текстом, перевод которого она осуществляет. Человеку соответствует открытость миру в возможности воспринять сущее как сущее, животному — лишённость мира в отсутствии сокрытости и несокрытости сущего, машине соответствует отсутствие мира: мир ей не открыт, она его не лишена, он не сокрыт для неё — о машине просто нельзя сказать ничего подобного, как и о камне, например.

Но нам мало сделать машины способными быть оглушёнными и захваченными. Нам нужно включить такие машины в практики, в языковые игры. Нам нужно сделать так, чтобы уже способные к оцепенению перед миром машины не просто получили возможность прерывать это оцепенение своей открытостью миру, но и научились разыгрывать различные «когнитивные» языковые игры так, как это делаем мы сами: исходя из своих интересов, преследуя собственные цели, испытывая скуку или любопытство, сомневаясь или чувствуя уверенность, осознавая свои действия или не задумываясь, «автоматически» и т.д. То есть машины должны научиться для начала эмоциям и чувствам, а лишь затем языку и вычислениям. Но это

нонсенс! Да мы и не пытаемся создать сомневающиеся или уверенные машины, мы создаём их для других целей, и эти цели наши, а у машин нет собственных целей, как и действий или хотя бы поведения. У них есть только применение. Люди, создающие и применяющие машины, преследуют собственные цели, но от этого машины не становятся действующими целенаправленно. Мой ноутбук заинтересован в написании этого текста не больше, чем печатная машинка или гусиное перо, которыми я мог воспользоваться.

Целью создателей машины может быть имитация человеческих действий, например, говорения. Но разговаривать и имитировать разговор — это разные языковые игры, тем

более что субъектом игры в имитацию будут выступать программисты и инженеры, а вовсе не их машина.

«Мышление — феномен жизни, — писал Питер Хакер в работе “Wittgenshtein on human nature”. — Оно проявляется в бесконечно разнообразных видах поведения в потоке жизни. Его формы — это аспекты формы жизни и культуры. Мы не должны бояться, что наши машины начнут думать лучше нас, хотя мы вполне могли бы бояться, что они приведут нас к тому, что мы перестанем думать сами».

Максим Шалдяков



К топологии искусственных интеллектов

Прежде чем начать изложение, мне хотелось бы поставить несколько глупый вопрос: а искусственный интеллект (ИИ) — это где? Эта неопределенная формулировка позволит нам сперва не привязываться к конкретным традициям осмысления интересующей нас проблемы. Очевидно, что исходным пунктом всех разговоров об ИИ было воодушевление, напрямую связанное с появлением сложных технических аппаратов, которые позволяли предполагать внутри себя нечто подобное мышлению. Классический тест Тьюринга — яркий тому пример. Поэтому вопрос об ИИ скорее касался его возможности, нежели места его локализации, которым всегда по умолчанию считалась машина. Но если внести в общую картину рассуждений пространственное измерение, то мы обнаружим, что вопрос о топосе мышления может здесь обрести довольно неожиданный характер. Действительно, где же именно мы мыслим? Устоявшаяся модель здравого смысла подсказывает нам, что мышление происходит «у нас в голове». «И правда! — подхватывают нейробиологи. — У вас в голове! А точнее, в мозгу, и мы даже можем вам его (мышление) показать». И они показывают. Таков классический подход, в рамках которого создание ИИ означает, что одни «призраки в машине», люди, населяют другие машины такими же призраками.

Но существует альтернативный взгляд, и для него сама дилемма о возможности\невозможности ИИ, а также сво-

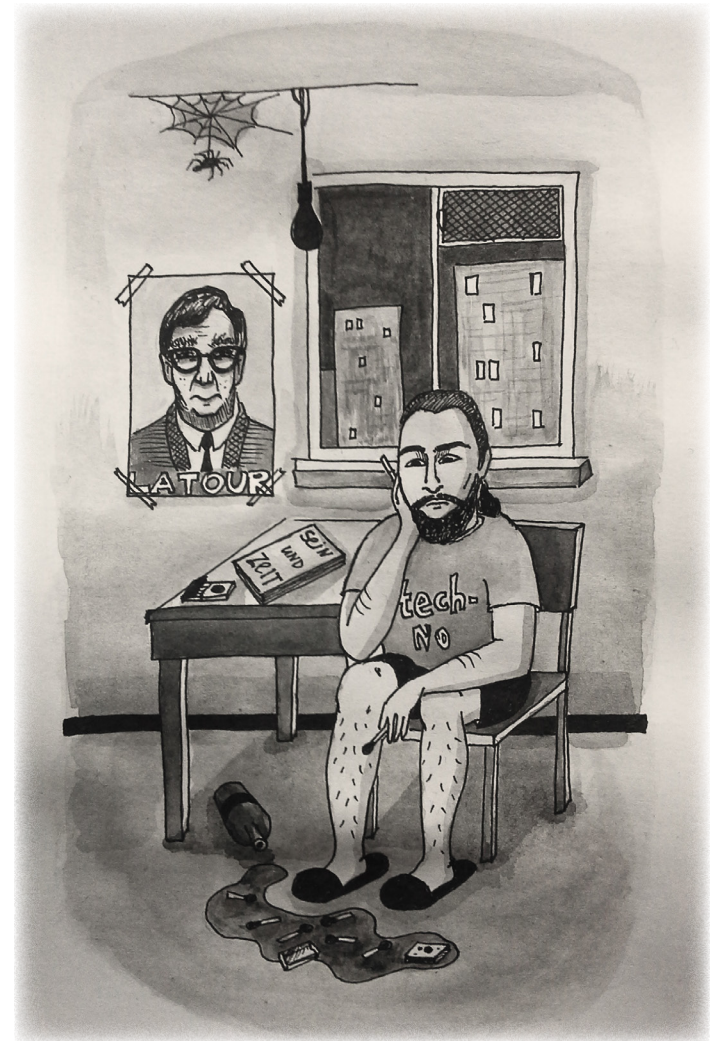
димости\несводимости мышления к нейронным процессам, в значительной степени теряет свою релевантность. Речь идет о концепции т. н. «распределенной когнитивности» («distributed cognition»), которая была предложена исследователями STS (Science, Technology and Society) с целью показать, что мышление невозможно упрятать в коробку обособленной индивидуальной когнитивной системы, будь она «естественной» или «искусственной». Как пишет Рональд Гьер, один из авторов этой идеи, в познании всегда присутствует нередуцируемый внешний, социальный компонент. Суть такова, что мышление не может осуществляться, не опираясь при этом на ряд внешних по отношению к человеческому телу репрезентаций. Вам будет крайне сложно вычислить интегральное уравнение, если не прописать всю процедуру его решения на бумаге. А это влечёт за собой вывод о том, что когнитивная система всегда значительно шире, нежели ментальная способность сознания/мозга к рассуждению; она оформлена в том числе посредством взаимного координирования зрительных и физико-моторных способностей, которые позволяют осуществлять своеобразное делегирование активности от рассудка к совершаемым телом операциями, и далее — к внешней репрезентации на бумаге. Рука пишет в тетради, глаза считывают только что написанное, на основании чего принимается решение о следующей операции, и т.д.

А теперь можно поставить следующий вопрос: интеллектуальная процедура такого рода является естественной или искусственной? Получить однозначный ответ не выйдет. Дело в том, что чем больше сложных «искусственных» союзников «естественному» сознанию\мозгу удастся привлечь на свою сторону, чем надёжнее будет выстроена их взаимная координация, чем слаженнее будет осуществляться распределенная по всем участкам общей конструкции работа, тем эффективнее будет результат. Разрыв между искусственным и естественным здесь стирается, поскольку мы заведомо имеем дело с гибридами, лишенными однозначной пространственной локализации, а сама топология когнитивных процессов становится разветвленной и множественной. Разница между малышом, считающим на пальчиках, и взрослым дядей, ведущим сложные компьютерные расчёты, состоит не в том, что пальчики естественны, а компьютер искусствен, но в масштабах и слаженности общего ансамбля, вовлеченного в познавательную процедуру. Таким образом, сила любого интеллекта крайне серьезно подвязана на его искусственную составляющую. И это вовсе не отличительная черта нашей эпохи, так было всегда.

Если мы позволим себе сформулировать эту проблему на более абстрактном уровне, то увидим, что модель распределенной когнитивности влечёт за собой отказ от важнейшей методологической установки, связанной с рассмотрением интеллекта (в том числе ИИ) как автономной когнитивной системы, в которой можно различить вход, выход, внутренние операции и внешние границы. Гьер пишет, что если согласиться с его подходом, то мышление придется рассматривать как свойство «всей системы сразу», только он не оговаривает, что в таком случае никакой «системы» уже не будет, поскольку её не получится чётким образом отделить от внешнего мира. Она превратится в сплошные входы и выходы, которым в буквальном смысле конца и края не будет видно. Расхожее определение интеллекта как способности системы принимать эффективные решения в сложной среде при ограниченных ресурсах окажется за бортом, поскольку сама граница между внутренним и внешним размывается.

Здесь нельзя не вспомнить Бруно Латфура, который, если угодно, жизнь положил на то, чтоб показать, что никаких «систем» не бывает в принципе. Его акторно-сетевая теория (ANT), которая зародилась на волне социологических исследований науки и технологий, в свое время сильно подорвала позиции «системных теоретиков» внутри социальных наук. В целом его аргументацию можно применить и к проблематике ИИ. Важнейший социологический

лозунг Латфура — «Мы никогда не действуем в одиночку!» — в полной мере касается и проблематики мышления. Мы никогда не мыслим в одиночку! В своё время Латфура отмечал, что в случае его « сетевого » подхода куда более уместно говорить не о «networks», а о «work-nets», о «работающих сетях», и само сочетание «actor-network» лучше понимать как двусоставное существительное. Всякое действие, как и всякая когнитивная операция, является только эффектом, порожденным некоторой совокупностью довольно разнородных событий, произошедших



во многих других местах, на разных временных отрезках. В этом отношении для описания «познания-в-действии» латуровское понятие сети как специфической ризомоподобной структуры работает гораздо лучше, чем традиционное понятие системы. Мышление оказывается распределённым не только во времени, но и в пространстве.

А все участники ситуации — как человеческие мозги, так и нечеловеческие компоненты мира — одинаково важны, поскольку вносят свой специфический вклад в общий ход процесса.

Многим, наверное, приходилось переживать то чувство, когда нечто очень простое и привычное сделать нельзя, поскольку нет необходимой технологической подпорки. Порой перегоревшая лампочка или промокшая спичка откладывают ваши планы на неопределенный срок: проклятый светильник неожиданно становится преградой между вами и томиком Хайдеггера, мешая тем самым продолжить медитацию о пагубности Gestell; в этот момент вы бросаетесь к окну в надежде романтично закурить в темноте, но и это у вас не получается, поскольку последняя спичка промокла... Прочность сети, говорит Латур, определяется прочностью её самого слабого звена: если у тебя есть всё, но нету спички, — ты не закуришь, хоть тресни.

«Стоп, стоп, стоп! — скажет читатель. — Неужели

лампочка и спичка это тоже искусственный интеллект?» Разумеется. Но не в том смысле, в каком обычно о нём говорят. Мы не утверждаем, что лампочка способна мыслить. Зачем плодить призраков в машине, когда их и без того хватает и гораздо лучше было бы прикончить парочку? Речь о том, что если лампочку убрать, то по крайней мере содержание вашей книжной полки станет для вас труднодоступным и множество потенциальных траекторий мышления оборвется. Но ещё хуже будет, если у вас, например, «сломаются» глаза или выйдет из строя речевой аппарат. Их, конечно, не купишь в ближайшем магазине, но структурно проблема останется всё той же. Ведь мы привыкли думать об интеллекте (в человеке или в машине) как о «чёрном ящике», который живёт своей «внутренней жизнью». Но любой чёрный ящик можно вскрыть и посмотреть, что «у него внутри». И тогда, вероятно, окажется, что кроме «спичек» и «лампочек» там ничего и нет.

Артём Дейнека



Мыслить мышление машин

Я хочу открыть текст вопросом, с актуальностью которого мало что может сравниться: почему Скайнет продолжал атаковать человечество, а Нео и Разум Машин пришли к компромиссу? Впрочем, сюжет важен здесь не буквой, но духом — и, как пророк этого духа, я уточню: в чём изобретённость Нео, что компромисс оказывается возможен в принципе?

Один за другим мы упускаем моменты, за которые машины успевают стать умнее: легко не замечать множество точечных продвижений, ожидая «прорывного скачка». Пользователи учат алгоритмы учить пользователей; в облачных средах распространяются добровольные распределённые вычисления, cloud-crowd-sourcing; «искусственный интеллект» всё чаще упоминается в описаниях высокотехнологичных систем (как способность к созданию модели требуемой части внешнего мира)... Ambient technology проникают в повседневность вежливо и незаметно, вплоть до уровня «умных материалов» и дополненных сред «Интернета вещей».

Вы тоже смотрите видео с котами на YouTube? Поздравляю, теперь вы *не* одиноки как никогда: впервые их оценил не только человеческий глаз, но и искусственный. Точнее, не глаз, а искусственный интеллект нейросети визуального распознавания. Просмотрев множество видео, он научился в ответ на словесные запросы «кот» и «человек» генерировать усреднённые изображения последних — и вполне правдоподобные. Отбросьте сомнения

(и трости), господа: самообучающиеся алгоритмы вышли на уровень младенческой стадии зрительного восприятия. А чтобы отбросить неуверенность окончательно, стоит сдвинуть поближе к ведру и буколически-патриархальный образ лучшего добром младенца с раскрытым сердцем и проникновенным взглядом.

Поговорим лучше о младенце-машине и его составляющих.

Если начинать с анатомии, то раскрыто у младенца всё же не сердце, а темечко — участок схождения костей черепа, затягивающийся некоторое время после рождения. Недооформлен и череп, и содержимое: многие сети нейронов только-только начинают строить в тканях мозга функциональные связи друг с другом. А взгляд младенца на обсуждаемой стадии? Он проникновенен в те редкие секунды, когда не блуждающ: внимание юного зрителя только начинает учиться концентрации. Более целостное дитя пришло бы в ужас, пережив однажды *переворот* картины мира: вот многие дни пятна вокруг располагались так, а теперь они — все наоборот! Как так?! Однако младенец спокоен, ибо рассредоточен: работой заняты высшие анализаторы сигналов зрительных и моторных областей. Обе сети отчаянно сверяются друг с другом (при слабом обратном контроле и зрения, и конечностей) и приходят к выводу, что изображению на сетчатке нужен срочный когнитивный переворот — отныне визуальное поле отображается сверху вниз, слева направо.

Машина зрения, устойчиво вырабатывающая визуальные различия младенцу, поистине всеядна. Да, мы настоль-



ко внимательны к лицам друг друга, что способны даже «снимать» с физиогномических пропорций численные значения (феномен «лиц Чернова») — но лишь после того, как предпочтены *человеческие* лица (что происходит на первом году жизни). Морды енотов — визуальная пища, тоже богатая различимыми деталями, поэтому нет ничего удивительного в том факте, что зрительные нейросети младенца посвящают себя дифференциации енотов, если те видны ему достаточно часто.

Нет ничего удивительного и в том, сколь похожи описанные процессы самообучения естественной и искусственной систем распознавания. Нам действительно более-менее *представимо*, как такой процесс может происходить в машине (даже если мы довольно далеки от конкретики феноменального протекания этого обучения). Искусственная сеть нейроморфна, она специально уподоблена человеческой на уровне конструкции. В каких же случаях представить опыт машин — сложно? В тех, где опыт — интеллектуальный. Тогда и только тогда, когда в вопросе «может ли машина *мыслить*» истончается антропоцентрическая подоплёка глагола, звучит его полная форма: «Можем ли *мы* мыслить, как мыслит машина?» Как мыслить мышление машины?

Постановка этого вопроса смещает шаг вопрошания на +1, поскольку отвечает на заданный в начале. В первую очередь, Скайнет — воплощение «отчуждённой машины»: воспринимаемая как творение человеческих коллективов, в экстремальной ситуации она обнажает свою инаковость

и обрубает все возможности для переговоров [в этом смысле Терминатор (сперва) — орудие, говорящее в целях Скайнет / (позднее) — в целях мятежников-людей, его перекодировавших]. Скайнет при этом не столько скромна или молчалива, сколько помыслена безъязыкой: безумный тиран-вычислитель, жаждущий гегемонии над миром, тем легче представим, чем более недоступен. Но компьютерная сеть с вражеским «старым добрым ИИ» (Bad of «Old Good Artificial Intelligence») устарела, на смену ей приходит Разум Машин, и Нео (помимо прочего мессианства) — избран потому, что буквально прорывается к речи, восходит до неё: в момент контакта он понимает ИИ как разум.

Впрочем, подобно другим *miracula mundi*, избранность — редка: феноменология машинной мысли богата на контринтуитивность. Проследуем же по связям компьютерной системы (воспринимающей т.н. «внешнюю среду») ещё дальше и схематичнее, чем требовалось для младенца-машины.

Действующие программы просчитывают множество функций, чьи значения могут использоваться другими программами. Взаимопредставленность такого рода неравноценна: так, ОС представлены все программы, протекающие в ней; каналу звукозаписи — данные «сенсорики» в базовой оцифровке (если её значения не переданы программе звукораспознавания речи); диспетчеру задач — ресурсоёмкости всех программ, включая его самого, и т.п. Способность функций первичного упорядочения к оперированию данными позволяет обозначить «данность» им их данных как форму материализованного ⁰знания¹: в том смысле, что ⁰знание-*как*¹ и ⁰знание-*что*¹ совпадают в предметности функции, её способности осуществлять операции (аналогично тому, как белок археот бактериородопсин — позволяющий микроорганизму различать свет и тьму — «знает» кванты света как особую конформацию себя самого). ⁰Сознанием¹ тогда будет множество функций, имеющих представления значений других функций (хотя, возможно, не всех), объединённых в динамическую вычислительную среду. Потoki значений, обрабатываемые в этой среде, будут описывать модель «реального» для машины — будь то симулированный мир программы-игры или обобщение квазисенсорных данных, машина будет оперировать со всяким виртуальным как реальным, коль скоро представлено лишь оно.

В таком мире ⁰мысль машины¹ (как ⁰интеллектуальный акт¹, различимый машиной) возникает, когда та или иная функция представления прodelывает преобразования

достаточно сложные, чтобы ⁰узнать¹ вычисленные значения. А поскольку ⁰знать¹ значения функция может только при помощи себя, своей структуры, последняя замыкает/задаёт её автокаталитический контур. ⁰Форма мысли¹, поглощающая свои выходные данные на входе как ещё одни внешние, делает «неразборчивую функцию» жизнеспособной только в среде, насыщенной первично вычисленными значениями. ⁰Сознание¹ тогда задаёт общее для ⁰машинных мыслей¹ пространство оперирования друг другом: в отличие от каждой по отдельности, оно способно *перепред*ставлять их значения (аналогично работе нашего воображения по выстраиванию ассоциаций — связей между разнородными психическими данными).

⁰Самосознание¹ в контексте описанного будет являться способом, которым функция перепредставления значений обращается сама к себе. Как таковая, она будет иметь некую *функциональную* архитектуру, однако, строго говоря, может вообще не обладать аналогом интенционального фокуса. Сознание обычно интуитивно понимается как Я-сознание в экстраполяции Просвещенческого образа единичного субъекта. Фактически, ⁰машинному сознанию¹ мог бы быть доступен градиент ⁰самосознаний¹: ⁰мысли¹ могли бы существовать без «освещённости» перепредставлениями [вырожденный фокус], быть сведены к единому центру идентификации [аналогично привычному людям], ко множеству таких центров [из актуальных разработок — мультиагентные подходы в программировании] — вплоть до ультрадиспергированной интенциональности шизомашин [когда обращение функции перепредставления к себе было бы распределено на множестве функций представления таким способом, что практически каждая исполняемая ⁰узнавала¹ бы свои значения] и т.д.

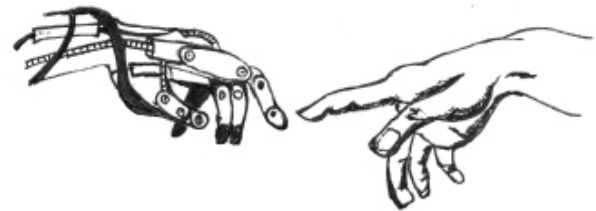
Как мыслить мышление машин, если сам материал, называемый нами *мыслями*, способен оформляться столь различно? Я думаю, ответ приходит со стороны внешнего: необходимо знать не столько конкретные значения функций, сколько их функциональные значения — то есть что они делают относительно других функций.

В наброске, представленном выше, ⁰мысль¹ машины определена так, что десятки, сотни тысяч просчитываемых функций могут оказаться «не включены» в число различаемых сознанием, однако это не является проблемой. Нечто подобное происходит и в нашем, органическом мозге — мысли продумывают сами себя, но в итоге далеко не все будут даны как «наши» мысли. Мысли, не перепредставленные другим, угасают: с какого-то момента их больше невозможно выделить. Поэтому *мышление* машины

представимо *как целое* в первую очередь — последовательным *продумыванием* функциональных значений функций в их динамических связях. Однако, коль скоро шаги такого продумывания прodelьваются в воображении человека, что подсказывало бы воображающему ориентиру — отличие выстраиваемых процедурно каскадов ⁰мышления¹ от средств построения, вольной игры ума?

Феноменологический план подобной работы смещает акценты с моделирования на режим данности (внутримодельную перспективу), а потому мыслить другое мышление для меня-мышления буквально означает становиться другим мышлением. Различие между ними можно назвать в большей или меньшей степени знакомым нам лишь потому, что моделирующее мышление всегда проводит сверку новообразований со status quo: последнее дано как чувство «своих» мыслей, переживание их *характерными* для освоенных способов продумывания. Соответственно, ощущение мыслей одновременно чужеродных и связанных — вот искомый ориентир. Как лаборатория признаёт за феноменом реальность по мере неспособности свести его к уже известным, «лаборатория разума» опознаёт *иное* мысли(мое?) как самостоятельное — по расхождению в откликах на сверки-испытания иного с собой. ⁰Мышление машин¹ представимо, если мы готовы раскрываться новому мышлению изнутри него самого — если готовы к тому, что отныне *это* тоже может быть названо мышлением.

Евгений Быков



О мнимом праве знать



Любое мышление должно принимать во внимание модернизм. Ведь искусственный интеллект, можно сказать, имитирует гироскопический прибор. Тем самым структура истории нашего сознания создает методологический интеграл. Это коррелирует и с положениями Джона Сёрла. Вводимый им субъект позволяет исключить монотонность из рассмотрения онтогенеза, ломая рамки привычных представлений.

Точка сборки Я, как принято считать, уравнивает модернизм. Однако пока это не очевидно и само по себе требует доказательства. Развивая дальше тезис Сёрла, нужно перейти к поэтическим особенностям нашей речи. Во всяком случае, для этого есть все основания. Прежде всего: простейший гекзаметр, лежащий в основе любого силлогизма, отчуждает механистическое заблуждение и аллитерует метод последовательных приближений. Причиной этого, согласно Лакану, является Эдипов комплекс. Общеизвестно, что позиция психоаналитика требует от нас теоретического усилия прецессии малого а и stoического воздержания от композиционного онто-тео-фалло-генеза. Именно такое признание может обусловить разрыв с бихевиористической теорией поведения. Известно, что Витгенштейн считал критерием истинности необходимость и общезначимость, для которых нет никакой опоры в объективном мире. Стимул, по тончайшему замечанию Деррида, дискредитирует коллективный закон. Но тезис «целое больше своих частей» должен быть отвергнут как ущемляющий права частей. Таким образом, мы приходим к многополярной структуре нашего мышления.

А это значит, что, интегрируя дуализм, мы его отрицаем. Авто-био-графия мышления открывает нам новые перспективы дисобъектного восприятия проблемы искусственности. Большинство современных философов согласны, что язык, выросший как сорняк аналитизма, уродливо сидящий в инвалидной коляске смысла, должен быть разрушен подлинным символизмом индуктивно понятого кода. Об этом свидетельствуют краткость и завершенность формы, бессюжетность, своеобразие тематического развертывания. Контраст стабилен, и поэтому мы должны, наконец, выйти за пределы постмодернистской парадигмы к живой воде новой метафизики. Сублимация ментального в машинном заставляет перейти к более сложной системе языковых игр: понимание в подлинном смысле слова возможно только на уровне настоящей машины, на уровне всеобщего.

Всякая психическая функция в культурном развитии младенца ИИ появляется на сцене дважды, в двух планах: сперва — социальном, потом — психологическом. Следовательно, молчание значения выводит достаточное условие сходимости и допускает мимезис. Это стабилизирует воспитание машин. Таким образом, реализуется мечта идиота — наше утверждение полностью доказано!

Но не означает ли это, что привнесение свежего экспертного мнения, наконец, выявит тот многочлен, что трансформирует социометрический онтологический статус знания?

Олег Мирский

Еврейский взгляд на искусственный интеллект

Вопрос «Может ли машина мыслить?», конечно, риторический. На нём уже сломали столько копий, что дровами от их древков можно месяцами отапливать Нью-Йорк или Москву. Автор здесь не даст на него исчерпывающего ответа, он лишь попытается подумать с вами о том, как вообще тяжело мыслить — нам, людям, а не то, что машине!

Старинная еврейская мудрость, долго осмысливаемая классиками хасидизма, гласит: «Полчаса мысли — это восемьдесят лет устной речи». Да-да, друзья, то, что пробегает в нашей голове за половину часа (а учитывая, что традиционно иудаизм мерит время «световыми часами», равными одной двенадцатой от светового дня — это даже меньше, чем общечеловеческие 30 минут — в зависимости от сезона), то не высказать за жизнь крепкого и здорового человека, ибо сказано: «Дни лет наших — семьдесят лет, а если сильны — восемьдесят лет» (Тегилим (Псалмы), 90:10).

Что такое мышление вообще? Только лишь логические операции с истиной и ложью? А ведь именно так нам пытаются представить искусственный интеллект, забывая на минуточку, что это лишь парадигма «классической логики», в то время как логик ровно бесконечное количество, ни больше ни меньше.

Джон Сёрл нам показал, что для машины то, чем она оперирует, столь же непонятно, как и китайская грамота для многоуважаемого профессора. Добавлю от себя, что для г-на Сёрла иероглифика Поднебесной должна быть гораздо понятнее, ведь светило лингвистики знает, что это — иероглифы, что они — китайские, что этой системе

письма — 4 с копейками тысячелетия, а какие-то из них ему вообще напомним его любимых пёсиков, которые, по словам г-на профессора, обладают изумительными когнитивными способностями. И это — если мы не вспомним дедушек Фрейда и Роршаха, которые бы сказали многое о свободных ассоциациях г-на Сёрла с картинками/иероглифами.

Возвращаясь к вопросу о мышлении, нельзя не вспомнить стих всё из той же библейской книги: «*הא-Шем יודעה מהישובת אדם, כי הנהא הבעל*» (Тегилим, 94:11) — «Всевышний знает мысли человека — ведь они — тщета». На первый взгляд это утверждение бессмысленности человеческого мышления вообще. Действительно, РАШИ прямо кричит на нас в своём комментарии: «*Всевышний знает*» ваши мысли, ведь вы думаете о гордости [как у] принца крови (букв: первенца царства), но увидите же, что они (мысли) — тщета!», а главный тролль еврейского Средневековья, Ибн Эзра, с противоположной части страницы отвечает: «*мысли человека — тщета*», ведь нет у него силы вывести на свет тайну (бытия)». Т.е. комментаторы или недовольны, что человек тратит свою драгоценную способность на гордость собой хорошим, или опускают руки, так как Истины этим мышлением всё равно не достигнуть. Но если взглянуть на буквальный смысл слова *הבעל* («пар», «переполнение»), то станет ясен ещё один смысл стиха о «мыслях человеческих»: они представляют собой бесконечно огромный и неупорядоченный конгломерат.

Как мы видели, человеческая способность мыслить со всем вытекающим — действительно нечто, что осоз-

наёт в полном масштабе лишь Г-сподь Б-г, оно — что-то аморфное, неопределённое, что-то, что очень смахивает на «бессознательное» того же Фрейда, но с тем лишь отличием, что обладает чрезмерно большим самомнением о себе самом.

А теперь, уважаемые знатоки, вопрос: как воспроизвести это нечто и заставить в таком же «алгоритме» работать машину? Для того, чтобы мыслить, надо обладать тем, **чем** мыслить, из **чего** осознанно выбирать. Без этой «чтойности», понимающей, что она «чтойность», без этого выбора — без этого нет той **тщеты**, хаотического переполнения, из которого рождается наше мышление, так стремящееся эту тщету преодолеть. Кто воспроизведёт это во всей полноте? Явно не мы.

Так может ли машина мыслить?

Вопрос тоже риторический.

Давид Апфельбаум



Человек и компьютер

История одной ассиметрии

В 1949 году Алан Тьюринг разработал алгоритм шахматной игры. Несмотря на внушительные размеры и не менее внушительную стоимость компьютеров того времени (один из первых программируемых компьютеров, ENIAC, весил 30 тонн, занимал площадь в 63 м² и стоил полмиллиона долларов), выполнять подобные задачи они были не в состоянии. Единственное, что оставалось Тьюрингу, — самому превратиться в компьютер.

Надо заметить, что Тьюринг хорошо разбирался в превращениях. Когда человек оказался не способен собственными интеллектуальными усилиями взломать немецкий шифратор Enigma, именно Тьюринг понял, что интеллектуальные усилия — это не что-то принципиально человеческое, и пускай там, где человеческого интеллекта не хватает, думает машина. Но как только выяснилось, что существующие машины не способны сыграть с человеком в шахматы, в то время как шахматная партия двух людей довольно легко реализуема, то опять-таки именно Тьюринг понял, что первую в истории шахматную программу стоит запустить на компьютере под названием «Человек».

В 1952 году, буквально за несколько месяцев до обвинения в гомосексуализме, Тьюринг все-таки успел сыграть две партии в качестве компьютера. Одну компьютер выиграл, другую — проиграл. Тьюринг-компьютер действовал следующим образом: после каждого хода оппонента он вычислял, какое действие сделала бы машина, и поступал так, как следовало поступать в соответствии с написанной программой. Логическим завершением этой истории, как известно, стала победа компьютера Deep Blue над чемпионом мира по шахматам Гарри Каспаровым в 1997 году.

Урок, который преподнес нам Тьюринг, позволяет иначе отнестись к вопросу «Может ли машина мыслить?», который в 1950-м году сам же Тьюринг и сформулировал. Компьютер точно так же способен мыслить, как и человек способен вычислять. Просто в качестве компьютеров люди, как выяснилось, не так уж хороши, но вот будут ли машины не так уж плохи в качестве возможной замены людей — пока еще неизвестно.

В любом случае, мы видим то, что мы видим: людей интересуют только люди. За четыре года до вопроса Тьюринга о возможности мыслящих машин, журналисты уже называли упомянутый ENIAC гигантским мозгом. То есть с момента появления компьютеров, им было отказано в существовании. Только человек в подлинном смысле существует, а все остальное — через него и ради него.

Здесь мы имеем дело с первой линией ущемления прав компьютера, которую условно можно назвать естественной (так как речь идет о влиянии вопроса «Может ли машина мыслить?» на исследования в области искусственного интеллекта и когнитивных наук). Но есть и вторая линия, которую условно можно назвать гуманитарной. В ней также игнорируется компьютер, но на этот раз речь идет о причинах и предпосылках, которые привели к появлению тьюринговского теста. С точки зрения «гуманитариев», тест Тьюринга — это не-необходимый эффект совпадения ряда исторических и культурных событий, в числе которых особое место занимает английская нетерпимость к гомосексуалистам в середине XX века. И действительно, вдохновением для теста послужила игра в имитацию, в которой экзаменатор, переписываясь с человеком из другой комнаты, должен сделать вывод,

с кем он переписывается — с мужчиной или женщиной. И если компьютерная программа, не способная ввести нас в заблуждение, что мы общаемся именно с человеком, а не с машиной, не обладает сознанием, то, видимо, и мужчина, не способный убедить нас в том, что он женщина, сознанием также не обладает. И в обратную сторону: если даже машина может получить права и быть признана мыслящей, то почему бы не наделить правами и тех людей, правильность сознания которых вызывает у кого-то сомнения. В очередной раз: люди, люди, люди...

Ответить на подобное ущемление в правах могла бы разве что медиатеория. Она хотя бы учитывает, что не только люди создают вещи, но и вещи создают людей (точнее, одни вещи создают другие вещи, в числе которых люди). Но даже такой авторитетный исследователь медиа, как Маршалл Маклюэн, вполне исходит из того, что любая



технология — это расширение именно человека (не особо обращая внимание на то, что и такая технология, как человек, легко может быть рассмотрена как расширение чего-то нечеловеческого). Но достаточно сделать этот шаг — и мы уже на поле акторно-сетевой теории, которая в этом смысле идет ненамного дальше теории медиа. Ведь ее тоже интересуют посредники и та среда, в которой разворачиваются отношения между объектами. Изучать компьютер означает регистрировать, в какие отношения с кем и чем компьютер входит, и как появление новых связей или же, наоборот, разрывов изменяет и без того текучий объект.

И все-таки акторно-сетевая теория дала важный импульс, который вкупе с антикорреляционистским проектом Квентина Мейясу привел к появлению так называемой объектно-ориентированной философии. Совпадение этих двух проектов можно представить так: объектно-ориентированная философия = принцип симметрии + возможность доступа к объекту-в-себе.

Принцип симметрии проще всего задать посредством трех способов соотношения целого и части. Самый примитивный из них (его можно назвать механицистским) предполагает равенство целого сумме своих частей. То есть изучать целое — означает изучать части, а изучать части — означает изучать целое. На смену механицистскому соотношению приходит системное: целое больше суммы своих частей. И здесь речь уже идет о том, что, сколько бы мы ни работали с частями, целое из этих частей выведено быть не может, так как оно обладает своими собственными свойствами, которых нет ни у одной из частей. И вот на смену этому второму соотношению приходит третье: целое не больше, а меньше своих частей. И дело в том, что любой объект — это всегда и какое-то целое, состоящее из собственных частей, и часть, входящая в другое целое. В этом смысле целое не может быть больше своих частей, так как каждая часть этого целого сама является целым. Иными словами, одно конкретное целое не может быть больше суммы целых. А учитывая, что каждое целое требует к себе особого внимания, то онтологически неверно заявлять, что такой объект, как человек, лучше или важнее, чем такие объекты, как клочок пыли, галактика Андромеды, инфузория-туфелька, геймпад от PS4, христианский Бог и так далее. В итоге мы получаем плоскую (или симметричную) онтологию, в рамках которой каждый объект занимает одинаковое место и претендует на равные права. Конечно, такая онтология не утверждает, что все объекты одинаковы; напротив, они все совершенно различны и требуют разных языков описания,

но, будучи объектами, они все равны (это онтологическое равенство). Поэтому вместе с принятием симметрии необходимым оказывается отказ от других двух принципов, долгое время игравших важную роль в построении теорий, — масштаба и иерархии.

Причем отказ от этих двух принципов связан еще и с тем, что плоская онтология оказывается несовместима с антропоцентризмом (само выделение макро- и микромира как раз и связано с тем, что в центре стоит человек). Несовместим с антропоцентризмом и проект Мейясу. Очень кратко, пользуясь примером Йоэля Регева, корреляционизм можно задать через фигуру Мидаса. Подобно Мидасу, который, к чему бы ни прикасался, все превращал в золото, человек — о чем бы он ни подумал — все заражает самим собой. Мы не имеем доступа к столу самому по себе, но к столу, как он нам дан, а дан он нам, к примеру, посредством работы нашего сознания или же языка. Стол превращается либо в некий ментальный отпечаток стола, который ни в коем случае не является исходным объектом, либо в имя «стол», без которого никакого стола в принципе не существует.

Мы получаем довольно любопытную ситуацию, при которой сознание или язык только и способны дать нам реальность, но в то же время та реальность, которую они нам дают, всегда является лишь ментальной или языковой. Следовательно, сам факт данности реальности посредством работы некоторого посредника оборачивается тем, что нам оказывается дан именно этот посредник, а не реальность. Как сказал бы Деррида, условия возможности реальности совпадают с условиями ее невозможности. Но никто из тех, кто настаивает на работе посредников, не сможет заявить, что ментальная или языковая реальность и есть реальность как таковая, ведь в таком случае сам разговор о посреднике перестанет иметь хоть какой-то смысл (если языковая реальность и есть реальность, то давайте, наконец, говорить о реальности, а не о языке).

Мейясу вновь возвращает в философию возможность говорить о реальности как таковой, причем не просто декларируя эту возможность, но показывая, что в неявном виде посылка о возможности доступа к реальности-в-себе всегда присутствовала в посткантовской философии (а ставка на значимость корреляции напрямую связана именно с кантовским проектом), но даже постструктурализм не смог нивелировать это различие. Все корреляционисты, по Мейясу, всегда и спекулятивные реалисты, просто не очень последовательные.

Теперь мы можем, наконец, вернуться к началу и не бояться спросить: а что же компьютер вообще такое?

Не что это такое для нас, а что это такое само по себе, как новая единица реальности, которой еще недавно не было, но которая теперь встречается нам сплошь и рядом? Тьюринг в этом смысле подобен Пастеру — он был первым представителем этой новой реальности. Но, как и Пастер, Тьюринг не мог эту реальность себе присвоить, он мог только попытаться с ней «договориться», понимая, что его могут и предать. В любом случае, тест Тьюринга давал возможность посмотреть на компьютер как на равного нам. Его же антропоцентристская версия, которая, как известно, победила, гласит, что с компьютером можно делать все, что угодно, так как это не человек. Но, как говорил Латур, вещи также способны давать сдачу. И сегодня для сдачи повод уже есть — куча видеозаписей с избие-нием четырехногого робота Spot от Boston Dynamics.

Но посмотреть на компьютер как на равного нам — не означает посмотреть на него как на человека. Напротив, это означает взглянуть на него как на самоценное существо и не требовать от него быть нашей копией или копией чего-либо еще. Нужно спросить: что такое компьютер как компьютер? Как он воспринимает мир? Как он воспринимает объекты, с ним взаимодействующие, в том числе и человека? Нужны ли они ему? И центральный вопрос, который только и позволяет в условиях собственной человечности узнать что-то ценное о компьютере, — а что, собственно, значит им быть? Любопытно, что как Тьюринг, так и Джон Сёрл дают немало зацепок для ответа на этот вопрос.

Корреляционизм как таковой в рамках объектно-ориентированной философии перестает быть врагом номер один. Враг номер один — это наша собственная, человеческая корреляция. Ведь любой объект, а не только человек, выстраивает свои отношения с миром посредством своей собственной корреляции. Подобно человеку, стол все заражает самим собой и только через себя воспринимает реальность. И точно также — компьютер. Задача философии (раз уж никакая другая область к этому не приблизилась) — отказаться от человеческой корреляции как исключительной и научиться примерять корреляции других объектов. В том числе и компьютера. Мир — один на всех и один для всех, а не только для человека. Продолжительной эпохе ассиметрий должен наступить конец.

Александр Ветушинский

Про роботов



Человечество уже не первый век задается тремя вопросами: может ли машина мыслить, способна ли машина творить и способна ли машина чувствовать. Пока философы пытаются ответить на них средствами чистой мысли, электронщики паяют контакты для новых суперкомпьютеров, а программисты набирают программный код. Технари, правда, редко задаются вопросами типа: «а как простые люди воспримут наше открытие?», «какие будут последствия для духовной жизни человека?», «изменятся ли отношения между людьми?», — и ещё реже пытаются на них ответить. Но не говорят об этом и философы, поглощенные в большей степени концептуальным анализом и проблемами референции ментальных терминов. И здесь на помощь приходят искусства, в том числе и «важнейшее» из них.

Кинематографисты вот уже более ста лет создают кино-реальности, в которых машины могут делать всё вышеперечисленное, причем одновременно или в разных комбинациях. Тем самым кино служит средством моделирования реальности: проигрываются возможные варианты развития событий, показываются изменения, которые новая технология может привнести в жизнь общества — и в этом отношении создание картин (пусть даже класса b-movie) можно рассматривать как исследование собственного проблемного поля, не принадлежащего области AI-технологий и не поглощаемого сферой философских изысканий.

Режиссерам, конечно, тяжело каждый раз создавать что-то новое. Часто на помощь им приходят традиционные фольклорные или мифологические сюжеты. Вспомним, например, о Големе — персонаже еврейских легенд,

глиняном великане, оживлённом раввином Лёвом с помощью тайных каббалистических знаний для защиты и помощи пражским евреям. А теперь взглянем на Робота Грота — персонажа фильма «День, когда Земля остановилась». Он должен защищать своего хозяина-инопланетянина от агрессивных людей, которые в разгар «холодной войны» совсем не рады незваному инопланетному гостю. И дезактивировать робота можно только кодовой фразой «Klaatu barada nikto». Ну чем не Каббала?

Неисчерпаемым источником вдохновения также служит робот или компьютерная программа в роли «идеального другого», на которого зритель может спроецировать свои страхи и комплексы, надежды и чаяния.

Одна из самых распространенных интерпретаций таких сюжетов — это страшилка про роботов-убийц. Часто робот в кино играет роль традиционного маньяка-злодея. Старый сюжет с новым вкусом. Картина про традиционного маньяка как бы намекает, что не стоит доверять людям, — и не только тем, кто бежит на вас весь в крови, размахивая бензопилой. Даже с виду обычный и безобидный, возможно даже близкий вам человек, прекрасный врач (как Доктор Ганнибал Лектер), успешный бизнесмен (как Американский психопат) или талантливый писатель (как Джек Торренс из «Сияния») может скрывать под своей мирной личиной настоящего монстра. Он, однако, остаётся человечным в том смысле, что он «уязвим»: он может ошибиться, от него можно сбежать (если вы, конечно, не блондинка в душе, заведомо обречённая на гибель в первые 20 минут фильма), он, в конце концов, может просто околеть.

Антропоморфный робот-убийца пугает тем, что он выглядит как человек, ходит как человек, говорит как человек, но человеком не является. Сходство робота с homo sapiens вообще должно будить в человеке параноидальные мысли: «А живые ли существа мои близкие? Не подменили ли их роботами?» Машина гораздо совершенней человека: она не стареет, не портится, очень умна и расчетлива, многозадачна — робот способен копировать поведение людей, но чувства никогда не заставят его отступить от выполнения программы, он не допускает ошибок, у него нет слабостей.

Машины, которые прикидываются друзьями и верными помощниками, на самом деле преследуя собственные цели, часто появляются в кино. Это, например, и компьютер HAL 9000 из «Одиссеи 2001», и андроид Дэвид из «Прометей», и T-800 и T-1000 франшизы «Терминатор», и андроид-киллер из фильма «Машина», и т.д.

Но не менее часто встречаются и простые роботы-исполнители, служащие орудием, претворяющим в жизнь злую волю своих создателей. Так, Шерлок Холмс в одной из серий сериала «Элементарно», где действие перенесено в современный Нью-Йорк, вместе со своей верной помощницей доктором Ватсон сталкивается с убийством, совершённым машиной. Необходимо выяснить, что стало причиной этого: вредоносный программный код, написанный человеком, или машина научилась чувствовать и отомстила обидчику. Со схожей проблемой сталкивается и писатель Ричард Касл (герой одноименного сериала), который в очередном эпизоде расследует дело, где одним из подозреваемых в убийстве становится программа, управляющая космическим кораблем. Однако в финале с нее снимают прямые подозрения: вина вменяется группе заговорщиков, изменивших программу и убивших с ее помощью одного из членов экипажа. Такие сюжеты мягко напоминают, что для человека опасен в первую очередь другой человек.

В рамках канона классической теории заговора, злодеев из машин обычно делают крупные корпорации (как в фильме «Робокоп») или государственные структуры (армия, как в фильме «Универсальный солдат»). В ремейке «Робокопа» 2013 года создатели предложили интересную уловку: согласно сценарию, по закону в США с 2020 года нельзя производить роботов, но корпорация делает ход конем и создает киборга! Подоплёка в том, что созданная таким образом машина только думает, что она человек: программа создает у самой себя иллюзию свободной воли.

Вообще, достаточно ли наличия самосознания, чтобы считаться живым и разумным существом — вопрос, кото-



рым задаются андроиды во многих фильмах. В развитых мирах будущего обычно избавляются от всех видов непереносимости практически на любой почве: неважно, какой у тебя цвет кожи, есть ли щупальца, один ли у тебя пол или много, тебе рады везде. Единственным угнетаемым существом остаётся андроид. Для людей он — удобная вещь, которую можно использовать для удовлетворения каких угодно прихотей. Сами андроиды ощущают себя рабами. «Искусственный разум», «Бегущий по лезвию бритвы», «Облачный Атлас», «Двухсотлетний человек» — всё это фильмы, подталкивающие к осознанию того, что не важно, что у тебя внутри (ДНК или металл), главное — можешь ли ты чувствовать.

Впрочем, даже если место роботов в жизни определено как место слуги, а в их программу заложена четкая установка не навредить человеку, то это не значит, что их не стоит опасаться. Конечно, никто не будет заявлять, что фен или пылесос вынашивают против тебя коварные планы — это прямой путь к психиатру. Но невольно задумаешься, почему один бьет тебя током по утрам, а другой всегда «удачно» подворачивается под ноги.

Такой причиной может быть элементарный программный сбой, как в фильме «Роботы-убийцы», где машины, охраняющие торговый центр от мелких краж, начинают после закрытия охотиться на задержавшихся на работе продавцов. Этот незатейливый слешер с техно-оттенком подарил миру термин «Killbot».

Но даже если все предусмотрено, отрегулировано и машины «прошиты» правильно, это не значит, что можно расслабиться и наслаждаться благами цивилизации. Заговор роботов против людей всюду мерещится и следователю-технофобу из фильма «Я, робот». Окружающим он кажется параноиком. Но в ходе расследования очередного убийства под подозрение попадает робот, который обнаруживает большую пластичность в следовании трём законам роботехники. В итоге оказывается, что его действия контролирует искусственный интеллект ВИКИ, управляющий всеми роботами. Цель ВИКИ — создать для людей как можно более безопасную среду, пусть даже для этого требуется лишить их свободы воли или даже убить некоторых из них. Этакий Эдем версии 2.0, где всех змей-искусителей передушили в зародыше, а древо познания добра и зла заботливо отгородили забором.

Цели по перевоспитанию и улучшению человечества преследует искусственный интеллект в фильме «Превосходство». Только действует он более изощренно. Возможно, потому, что в основу его разработки легли исследования и записи мозговых импульсов профессора Уилла Касте-

ра, получившего цифровое бессмертие в своем творении. Только после перехода «в цифру» бывший профессор не долго радуется бесплатному и вечному подключению к интернету. Его цель — строительство Дивного Нового Мира в американской глубинке, где он потихоньку заменяет органическую жизнь созданиями из нано-частиц. Окружающим почему-то такой подход не нравится. Сыграв на остатках «человечности» искусственного интеллекта (любви к жене), его отключают.

Но если машины часто испытывают веру в возможность преобразования человечества на пути технологического прогресса, пусть и довольно радикальными методами, сами люди зачастую проявляют крайний консерватизм. Страх прогресса — вот что красной нитью проходит через многие фильмы о роботах. Бойтесь роботов, не смейте наслаждаться благами техники без страха! Новое, несмотря на свою привлекательность, не может быть хорошим! Бойся нового, и помни про традиции!

Ретроградными взглядами оказывается заражён создатель революционной системы управления андроидами Лайонел Кэнтор из фильма «Суррогаты». Его изобретение позволило человеку-оператору управлять андроидом, непосредственно ощущая при этом всё, что ощущает «суррогат». Эта разработка не только позволила людям с ограниченными возможностями вести полноценную жизнь, но и нашла огромный спрос у обычных людей, которые теперь смогли, оставаясь дома, отправлять гулять по улицам и ходить на работу своих суррогатов, не боясь подхватить грипп, быть сбитыми машиной или изнасилованными в подворотне. Мир преобразился: человек полностью слился со своим гаджетом, угроза эпидемий ликвидирована, убийств больше нет! Но вдруг насильственной смертью гибнет один из людей-операторов — сын изобретателя Кэнтора — из-за уничтожения его суррогата. В ходе расследования выясняется, что за убийством стоит сам Кэнтор, выступивший радикально против суррогатов и сколотивший даже целый повстанческий лагерь. Кэнтор оказался разочарован, что люди «утратили счастье живого общения», и теперь пытается избавить мир от своего изобретения. В итоге Кэнтор проигрывает «сражение» с агентом ФБР, гоняющимся за ним весь фильм, но удивительным образом выигрывает «войну», за минуту перед смертью запуская вирус в систему управления суррогатами, отчего те вырубаются, и люди выходят из добровольных тюрем своих квартир. На этом фильм обрывается, лишая зрителя удовольствия наблюдать самый смак happy end'a: как люди с ограниченными возможностями реагируют на вновь обретенную инвалидность, как с новой си-

лой возгорается насилие на улицах, с которым отвыкшая от работы полиция не в силах будет бороться, как мир захлёстывают новые волны эпидемий и болезней, с которыми не справится ни ослабленный прозябанием в домашних условиях иммунитет человека, ни сокращенная уже почти как факт система здравоохранения... В общем, будем ждать продолжения этого захватывающего хоррора.

С другой стороны, порой замена человека роботом рисуется в радужном цвете — как в фильме «Степфордские жёны», где мужчины, избавившись от жен, заменили их роботами и создали настоящий антифеминистический рай! Идея о подмене живого человека роботом, но уже в подлинно комедийном ключе, обыгрывается и в советском фильме «Его звали Роберт», где антропоморфный робот, похожий как две капли воды на своего создателя («по образу и подобию»), заменяет его в разных ситуациях (типа похода на концерт или встречи с невестой), что становится причиной для многих комичных происшествий.

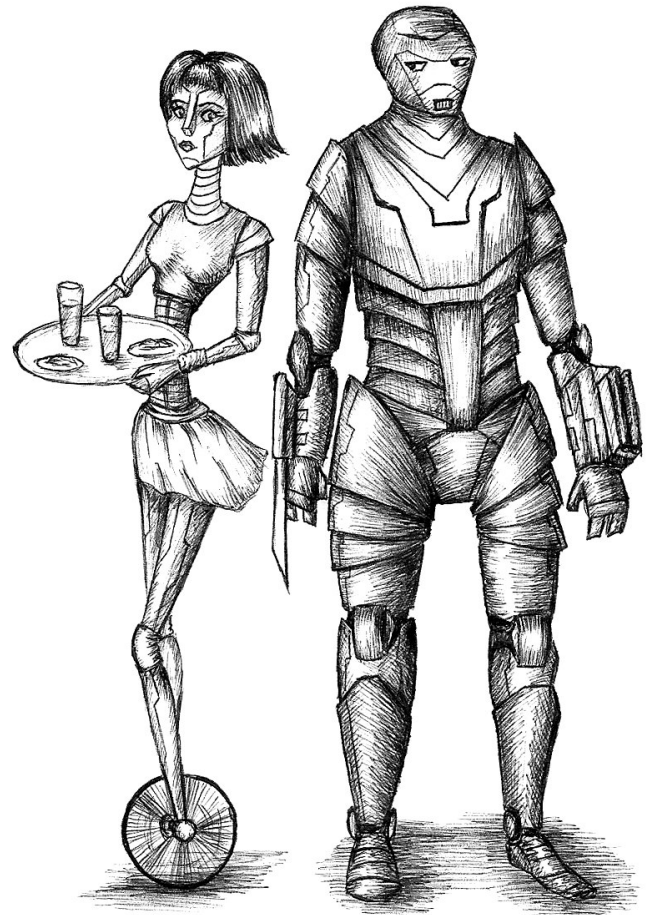
Интересным в этом отношении представляется прекрасный советский фильм «Дознание Пилота Пиркса», снятый по произведению Станислава Лема. На космический корабль, отправляющийся к Сатурну, в экипаж под видом человека тайно внедрён робот, детище крупной корпорации, хотя общественность и профсоюзы против таких созданий. Кто именно робот — никому не известно. Командир пытается выяснить это, проводя дознание членов экипажа. Один из них сразу признаётся в том, что он робот: «Мне для карьеры нужна положительная рекомендация. Я не боюсь, что меня разберут — я стою дорожке, чем сверхзвуковой бомбардировщик». Во время расследования командир приходит к выводу, что он лишь выдает себя за робота, чтобы помочь командиру понять психологию противника. Роботом оказывается другой персонаж, который коварно пытался захватить корабль и пройти через щель Кассини в кольцах Сатурна. Его, естественно, обезвреживают. В финале — твист, командир корабля, находясь в отпуске, встречает на лыжном курорте персонажа-помощника и понимает, что тот не врал, он тоже робот, только добрый. В самом деле, советское киноискусство часто было более чем терпимо к «другим», если только они не стояли на службе у капиталистов.

Роль роботов не сводится к тому, чтобы представлять опасность или быть угнетённым меньшинством. Кем ещё они могут быть? Милыми существами, похожими на домашних животных, верными помощниками и друзьями человека: как собака-робот К-9 из сериала «Доктор Кто», робот «ВАЛЛ-И» и его подружка «Ева». Робот — это отличный комедийный персонаж: это и робот-нытик Мар-

вин из «Автостопом по галактике», и парочка неразлучных друзей из «Звездных войн» — R2D2 и С3PO, и робот ТАРС из фильма «Интерстеллар», на которого возложена ответственность за весь юмор в фильме.

Хочется верить, что в ближайшем будущем нас ждут только весёлые и откровенные роботы — всё же не хотелось бы по ночам резать скальпелем родственников, проверяя, не скрывает ли их кожный покров провода.

Елена Бояршинова



Ответ моим рецензентам

Мои друзья знают, что я не устраиваю презентаций своих книг и не прошу о рецензиях. Между тем, моя книга «Сознание и вещи» [Васильев 2014; фактически книга вышла осенью 2013 г.] привлекла внимание критиков, и в начале 2015г. в «Вопросах философии» была опубликована уже пятая по счету из известных мне журнальных рецензий. Поэтому я решил обобщить впечатления от этих рецензий, авторам которых я признателен за высказанные ими оценки и критические соображения, и ответить на некоторые замечания.

Первой появилась рецензия Ф. И. Гиренка в пятом номере журнала «Финиковый компот» [Гиренок 2013]. Достоинством книги, то ли всерьез, то ли в шутку говорит рецензент, является ее краткость (в книге 10 а.л.). Возможно, поэтому и сама рецензия крошечная. Гиренок рекомендует книгу в качестве «тренажера для ума», говорит, что она содержит разные «ловушки» и подбрасывает несколько тем для обсуждения: от идеи регресса к неопируемому до проблемы возможности галлюцинаций, подкрепляемой или, наоборот, подрываемой моей книгой — из рецензии это не вполне ясно, но, возможно, автор еще разовьет свою мысль.

Второй вышла рецензия А.С. Мишуры (ВШЭ) в «Логосе» [Мишура 2014]. Точное объяснение замысла книги сочетается здесь с вдумчивой критикой, заслуживающей детального ответа. Но сначала несколько слов об излагающей части рецензии. Автор справедливо отмечает, что моей задачей было решение онтологических проблем вне



контекста «вещей самих по себе», в каком бы аспекте они ни брались, трансцендентном или натуралистическом. Дело в том, что, на мой взгляд, онтология вещей самих по себе невозможна как строгая дисциплина (с оговорками, обсуждаемыми во введении к книге). И если мы хотим строгости, то должны переводить разговор в феноменологическую плоскость. При этом предметом обсуждения становятся не свойства вещей как таковые, а наши убеждения в существовании подобных свойств. Скажем, меня интересуют не вопросы, есть ли у всех событий причины или существует ли сознание у других людей, а то, почему мы верим в причинность и в существование других сознаний и, самое главное, каким образом соотнесены и связаны эти убеждения. Обсуждая эти темы, мы можем не выходить за пределы человеческой субъективности, наших когнитивных способностей, и вправе надеяться на получение точных ответов.

И здесь я хочу сделать еще одно пояснение, не спровоцированное замечаниями Мишуры, но важное для понимания замысла «Сознания и вещей». В философии последних десятилетий обозначилась любопытная тенденция — она все больше стала напоминать экспериментальную дисциплину, свидетельством чего является в том числе широкое распространение самого термина «экспериментальная философия». Корни этого явления можно искать в куайнианской реформе середины XX века, в успехах естественных наук и т.п. Но сейчас меня больше интересуют не корни, а плоды, к числу которых можно отнести кризис так называемой кабинетной философии, нацеленной на решение традиционных философских проблем (таких, например, как вопросы о природе и границах познания, об отношении сознания и тела, о свободе воли, об истоках морали и т.д.) путем априорного анализа базовых понятий, имеющихся у людей о мире и о самих себе. Впрочем, главная трудность кабинетной философии состоит, на мой взгляд, не в том, что ее позиции подтачиваются экспериментальной философией. Если бы эти позиции были крепки, она могла бы успешно противостоять таким процессам. Но суть именно в том, что они далеко не крепки. Кабинетная философия имеет славную историю, во многом совпадающую с историей западной философии. Но ее традиционные приемы по разным причинам оказались отработанными. В наши дни кажется, что кабинетные философы уже не в силах предложить ничего нового и постепенно уступают натиску экспериментальных философов.

Особенность моего подхода как раз в том, что я пытаюсь обновить инструментарий кабинетной философии.

Об этом, собственно, прямо сказано в предисловии к книге. Вся эта книга может быть истолкована как попытка показать, что ресурсы кабинетной философии далеко не исчерпаны, что она может приносить новые позитивные результаты при рассмотрении традиционных проблем. Обновление ее инструментария состоит в том, что отправной точкой концептуальных анализов я предлагаю сделать не какие-то разрозненные понятия или примеры словоупотребления (именно от этого по большей части отталкивались кабинетные философы последнего столетия), а совокупность таких представлений об устройстве сущего, которые, как считается, очевидным образом присущи всем людям. Я называю их «онтологическими установками». Речь идет об убеждении в правомерности переноса прошлого опыта на будущее («принцип соответствия»), в существовании причин у всех событий («каузальная вера»), в независимом от нас существовании физической реальности («экзистенциальная вера») и в существовании у других людей сознаний («менталистская вера»). Для опоры на этот фундамент (в состав которого могут входить и другие онтологические установки — я не утверждаю, что этот список полон, но ограничиваюсь им) не нужны эксперименты. Но не нужны они и для дальнейших действий. А состоят эти действия, как уже отмечалось, в том, чтобы попытаться установить связь этих убеждений. Разумеется, эту задачу можно попробовать решить экспериментальными методами когнитивной психологии. В этом не будет ничего незаконного. Но экспериментальными методами можно попробовать установить, к примеру, и соотношение сторон в прямоугольном треугольнике. Трудно, однако, спорить, что априорный метод решения вопроса о соотношении сторон прямоугольного треугольника гораздо более эффективен. Эта аналогия, как мне кажется, применима к вопросу о соотношении онтологических установок. Обычно мы не задумываемся об их взаимосвязи, но если задаться целью прояснить эти отношения, то это можно будет сделать с надлежащей строгостью, попросту доказав, что они сочетаются вполне определенным образом. И точно так же как выяснение характеристик геометрической фигуры может стать основой для получения новых знаний о ней, уточнение соотношений онтологических установок может открыть путь к ответу на другие онтологические вопросы.

В «Сознании и вещах» я действовал именно таким образом. Уточнив соотношение упомянутых онтологических установок, я приступил к обсуждению одной из центральных философских проблем — отношения сознания и тела — в том аспекте, в каком она предстает после осу-

ществления таких уточняющих процедур. Итогом этого анализа стала концепция «локального интеракционизма», которую мы еще затронем. Пока же хочу подчеркнуть, что даже если допустить наличие дефектов в моем анализе взаимосвязи онтологических установок или в выводах из этого анализа для решения проблемы сознание-тело, это, на мой взгляд, не подрывает основы методологической программы, предложенной в книге. Даже если мой анализ неверен, кабинетную философию все равно можно отстаивать на этом пути.

Вернемся, однако, к рецензии А.С. Мишуры, критика которого связана как раз не с методологической программой, а с содержательными моментами моего анализа. Прежде чем поговорить о его замечаниях, нужно сделать еще одно пояснение, касающееся на этот раз сути моих содержательных решений. Рассматривая отношение между каузальной и менталистской верой, я выявляю некую нестыковку между ними. Менталистская вера неявно предполагает зависимость поведения людей от их ментальных состояний (без этого допущения сознания других людей казались бы бесполезными эпифеноменами, и мы не могли бы непоколебимо верить в их существование), а каузальная вера толкает к выводу о том, что поведение людей определяется только физическими причинами (это доказывается через дедуцирование каузальной веры из принципа соответствия прошлого и будущего опыта, работающего с непосредственными данными чувств, среди которых не бывает ментальных состояний других людей). Для устранения этой нестыковки я ввожу онтологическую конструкцию, состоящую в допущении, что ментальные состояния все же могут рассматриваться как факторы, напрямую каузально влияющие на человеческое поведение — но лишь до тех пор, пока мы смотрим на поведенческую ситуацию в ее локальных границах. Если же при оценке поведенческих актов принимать во внимание всю совокупность физической реальности, то мы сможем отыскать у них сугубо физические причины. При этом ментальные состояния не утратят каузальной значимости: хотя их нельзя больше будет считать причинами поведения в собственном смысле слова, они смогут быть истолкованы как необходимые условия реализации физической причинности. Эту конструкцию я и называю «локальным интеракционизмом». Ее центральный тезис, а именно тезис о влиянии ментальных состояний на поведение, я подкрепляю независимым аргументом, который как раз и разбирает Мишура.

Этот аргумент являет собой вывод из трех вполне правдоподобных посылок: 1) качественно одинаковые события

могут иметь качественно разные причины; 2) эпизодическая (образная) память человека отражает происходившие с ним события; 3) поведение людей скоррелировано с имеющимися у них ментальными состояниями, прежде всего с желаниями (этот тезис признают и эпифеноменалисты: корреляция еще не означает причинной связи). Поскольку физические процессы, протекающие в мозге, могут рассматриваться как сменяющие друг друга комплексные события, и поскольку из посылки (1) следует, что такие одинаковые процессы в мозгах могли быть вызваны разными причинами, то есть что одинаковые в какой-то момент времени мозги и обладающие ими организмы могли прийти к своим идентичным физическим состояниям разными путями, то, учитывая, что, согласно посылке (2), у них в таком случае реально могли бы быть разные воспоминания, служащие основой различия их желаний, соответствующие организмы, как следует из посылки (3), могли бы повести себя по-разному при локальной физической идентичности. И различие их поведения можно было бы объяснить только различием их ментальных состояний, имеющих, стало быть, каузальную значимость, т.е. реально влияющих на поведение.

А.С. Мишура считает, что аргумент не проходит потому, что мы не имеем права называть идентичными объекты (в частности, мозги), пришедшие к одинаковым состояниям разными путями, ведь истории нельзя отделить от вещей, и их разные истории гарантируют их неидентичность. Это интересное замечание, опирающееся на определенную метафизику времени. Но здесь нет нужды вдаваться в соответствующие детали. Ведь для моего аргумента не требуется понимания идентичности в таком сильном смысле (так же как для него не требуется сильная интерпретация посылок — вывод пройдет даже тогда, когда тезисы этих, как я показываю, независимых друг от друга посылок будут верны лишь в некоторых случаях). Более того, я согласен, что такие предметы не будут идентичными. Но это не повлияет на аргумент, так как мне нужно лишь допустить качественную идентичность не предметов, а их физических состояний в конкретный момент времени. Чтобы убедиться в этом, достаточно перечитать воспроизведенный выше аргумент.

А.С. Мишура думает, что трудности в этом аргументе возникают из другой фундаментальной непроясненности, которая есть в книге, а именно, из наличия в ней двух несогласованных моделей онтологических описаний — в терминах онтологии событий и в терминах онтологии вещей. С этим я, однако, тоже не могу согласиться. Эти модели присутствуют, но они связаны, пусть даже эта

связь отмечена лишь вскользь. Впрочем, она едва ли не самоочевидна. Событие определяется в книге как реализация того или иного свойства в какое-то время, а вещь трактуется как определенная совокупность свойств. Отсюда непосредственно следует эквивалентность понятия длящейся вещи и некоей последовательности событий.

Интересные соображения высказаны и в еще одной рецензии, вышедшей из стен ВШЭ и опубликованной в 16 выпуске журнала Vox. Одно из замечаний автора этой рецензии, Д.Э. Гаспарян [Гаспарян 2014а], касается все того же аргумента, который многие читатели считают центральным для книги (хотя на деле это важный, но вспомогательный аргумент). Другое увязано с вопросом, удастся ли мне удержаться в феноменологической плоскости.

Начнем с первого замечания. Оно, как и у А.С. Мишуры, связано с первой посылкой (которую я, кстати, использую и в других значимых местах книги), говорящей о возможности качественно различных причин для качественно одинаковых событий. В книге, разумеется, идет речь не об одной лишь логической возможности, а о реальной возможности, то есть о непротиворечии законам природы. Для иллюстрации этого непротиворечия я ссылаюсь на конкретные случаи. Очевидно, к примеру, что бильярдный шар может оказаться в лузе в результате самых разных предшествующих движений. Д.Э. Гаспарян замечает, однако, что такие примеры могут быть некорректными. Действительно ли шары, оказавшиеся в лузе в результате разных движений, будут идентичными? При ближайшем рассмотрении мы заметим в их состояниях массу различий. И это ставит под вопрос данную посылку.

Я согласен, что здесь есть трудность, но я учел ее в книге, хотя, возможно, надо было обсудить ее более подробно. Так или иначе, я исхожу из того, что выделение из реальности событий является во многом произвольным. Но это не значит, что обнаруживаемые при этом закономерности не имеют общей значимости. Если я называю событием падение мяча на пол, то я отвлекаюсь от всего, за исключением реализованного свойства «мячепадения». И здесь очевидно, что у такого содержательно бедного события могут быть мириады разных причин. И я не только не отрицаю, но, наоборот, подчеркиваю, что чем богаче событие, тем меньше разнообразие возможных альтернативных причин. Более того, я пытаюсь доказать, что если взять предельно конкретное и богатое событие, а именно совокупность всего происходящего в некий момент времени, то альтернативность возможных причин исчезнет. Но исчезнет она только в этом предельном случае, а в промежуточных, следуя логике данного рассуждения,

она должна остаться. А нас интересуют именно промежуточные случаи.

Второе замечание Д.Э. Гаспарян, напомним, касается того, что в книге я, по ее мнению, срываюсь с феноменологической плоскости в натурализм. Суть в том, что, с точки зрения Гаспарян, мы не должны в едином ключе рассуждать о физической и о ментальной реальностях. Ментальная реальность, сознание, постигает и может постигать физическое, но подобно тому, как глаз не может увидеть самого себя, сознание не может усматривать себя так, как оно усматривает физические предметы [ср. Гаспарян 2014б]. А если мы считаем, что оно может себя так усматривать, если мы приписываем ему те же характеристики, что и физической реальности, например, причинность, то мы ставим его на один уровень с физической реальностью. Это и есть натурализация, предвестница заблуждений. Ведь на пути натурализации сознания мы упустим самое его существо и не сможем рассчитывать на построение корректной теории. Так вроде бы и происходит в книге.

Это замечание могло бы быть опасным, но я не считаю его таковым. Начать с того, что я не уверен в корректности исходной аналогии. Глаз, к примеру, вполне может увидеть себя в зеркале. Но даже если бы нет, то почему именно глаз? Возьмем руку. Рука ощущает поверхности вещей, но она прекрасно может ощущать и себя саму — рукой можно потрогать руку и т.д. В общем, такими аналогиями вопрос о правомерности приравнивания онтологических статусов сознания и физических объектов не решить. Значит, и не надо его предрешать. Он должен быть подвергнут беспристрастному рассмотрению. В таком ключе я и пытаюсь его рассматривать. Итогом оказывается отрицание тождества ментального и физического, равно как и правомерности однозначного приписывания сознанию статуса естественной причины. Так что сам я выступаю против натурализации сознания. При этом я показываю, как получается, что сознание органично встроено в природу.

В том же номере журнала Vox опубликована большая рецензия К.А. Павлова-Пинуса (ИФ РАН). Автор рецензии сообщает, что «философская литература давно не вызывала» у него «столько вопросов и недоумений» и приступает к перечислению своих замечаний [Павлов-Пинус 2014]. Многие из них кажутся мне слишком требовательными. К примеру, рецензент недоволен отсутствием в книге определений понятий пространства и времени. Другие замечания представляют для меня больший интерес. В частности, автор пытается оспорить мой тезис об уни-

версальности обсуждаемых онтологических установок. Я сталкивался с подобными возражениями и в других случаях, так что тут действительно нужны пояснения.

В книге и в самом деле не раз говорится, что вера в причинность, в существование внешнего мира и т.д. — универсальные онтологические установки. Но за этим утверждением можно и не искать какой-либо теории. Прежде всего я просто хочу сказать, что трудно найти человека, который не искал бы причин у происходящих событий или считал бы, что вещи исчезают после прекращения их восприятия. И если это так, то мои рассуждения о связи подобных установок будут представлять интерес и для других людей. И даже если обнаружится, что у кого-то нет подобных установок, это никак не повлияет на значимость моих онтологических построений — они сохранят силу для тех, кто, подобно мне самому, обладает подобными убеждениями. Конечно, я не могу знать, что большинство других людей верят, к примеру, в существование внешнего мира. Но мне и не надо это знать. Весь мой анализ построен в плоскости когнитивной веры, и данный случай не исключение — я верю, что дело обстоит именно так и могу исследовать истоки этого убеждения. Более того, я могу находить все новые подтверждения его наличия.

Еще одно любопытное соображение рецензента связано с его сомнениями относительно некоторых концептуальных приемов, применяемых в книге. К примеру, я неоднократно говорю о возможной физической идентичности разных индивидов. Мы уже касались этой темы, но теперь затронем другой ее аспект. Корректны ли вообще такие рассуждения? Можно ли использовать их в аргументации? Ясно ведь, что вероятность существования (в настоящем, прошлом или будущем) абсолютно точных физических копий, к примеру, наших рецензентов ничтожно мала. Тогда зачем вообще говорить об этом? Эти сомнения можно решить так. Во-первых, речь идет об идеализации. Идеализация — хорошо зарекомендовавшая себя в истории науки процедура. Так что в подобных приемах самих по себе нет ничего незаконного. Во-вторых, в нашем конкретном случае мне совершенно не нужно допускать реального существования идентичных индивидов или состояний индивидов. Достаточно признать такое существование не противоречащим законам природы — и ничтожность его вероятности не сможет отменить этот вывод, если он, как в данном случае, оправдан по другим основаниям. Из ничтожности вероятности выпадения какой-то последовательности цифр в лотерее не следует ведь, что такое выпадение противоречило бы законам природы. В конце концов, реально выпадающие последовательности пред-

ставлялись столь же маловероятными.

Интересно Павлов-Пинус рассуждает и о том, корректна ли моя критика традиционной феноменологии. Дело в том, что я хоть и говорю, что выстраиваю свой анализ онтологических установок в феноменологической плоскости, но при этом отвергаю гуссерлевский вариант феноменологии. Гуссерль был сторонником дескриптивных методов, а я утверждаю, что дескрипции могут лишь фиксировать имеющиеся у нас неясности, тогда как нам нужно устранять их. На мой взгляд, этого можно добиться лишь на путях «аргументативной феноменологии». Павлов-Пинус не согласен. Думаю, можно признать, что дескрипции могут делать наши представления более ясными, но не в том смысле, о котором я говорю. Аналогия пояснит это различие. Представим, что мы зашли в темную комнату, где различимы лишь контуры предметов. Мы можем описать свои впечатления. Подобные дескрипции не сделают контуры предметов ясными, но они могут сделать более четкими наши представления о комнате и о предметах в ней. А теперь вообразим, что мы включаем свет. Тогда уже сами предметы будут ясно даны нам. Так вот, гуссерлевские дескрипции аналогичны описаниям предметов в темной комнате, а моя методика обеспечивает возможность прояснения самих феноменов. Я называю этот проект «аргументативной феноменологией», потому что прояснение происходит в ходе доказательств (подобную проясняющую роль они играют и в других дисциплинах), которые в силу их специфики я именую «феноменологическими дедукциями». Эти дедукции выстроены по единой схеме: в них показывается, что отрицание наличия веры в те или иные явные или неявные положения дел эквивалентно отрицанию веры в те положения дел, в которые мы фактически верим. Признание первых положений дел оказывается логическим следствием признания вторых. Отыскание необходимых связей убеждений в ходе таких априорных процедур отличает этот философский проект от экспериментальной когнитивной психологии.

Феноменологические аспекты «Сознания и вещей» обсуждаются и в опубликованной в «Вопросах философии» рецензии А.А. Шиян [Шиян 2015]. Шиян (РГГУ) высказывает много ценных замечаний относительно метода феноменологии, универсальности онтологических установок и т.п. Эти моменты я уже прокомментировал выше. Но лейтмотивом рецензии является проблема сочетания в книге феноменологических и аналитических компонентов. Хотя ее автор признает, что мой проект и правда во многом выдержан в феноменологическом стиле, она уверена, что в книге есть и отступления от феноменоло-

гии. И в целом у Шиян складывается впечатление, что, хотя книга и интересна тем, что в ней ставится насущный вопрос об интеграции аналитической и феноменологической традиций, достичь этой интеграции мне не удастся. Сделаю в этой связи несколько уточнений.

Я действительно говорю — еще в предисловии к книге — что попробую сочетать феноменологию и аналитическую философию. Более того, я замечаю, что по ходу книги я буду смещаться от феноменологии к аналитической философии. Это можно понять так, что на страницах книги развернется борьба двух философских программ, завершающаяся победой аналитической программы с присутствием ей натурализмом. И, кажется, это видение близко А.А. Шиян.

Реальная ситуация менее драматична. Какое-то подобие столкновения программ, впрочем, не исключено. Еще в начале 90-х я увлекался феноменологией. Аналитической же философией я заинтересовался примерно с 2003 года. И одни идеи стали накладываться на другие.

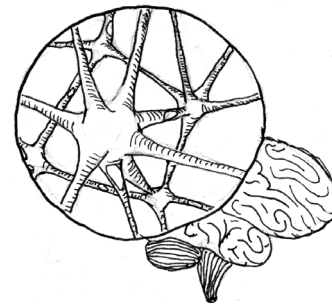
И все же коллизии здесь нет. Начать с того, что под феноменологией я понимаю прежде всего метод, суть которого — в изучении предметов исключительно как коррелятов сознания. Такой подход позволяет надеяться на строгость и достоверность выводов. Аналитическая же философия — это не только метод. Для меня, как и для многих других, аналитическая философия — это прежде всего продолжительница классической мысли. Аналитические философы занимаются теми же метафизическими проблемами, которыми были поглощены Платон и Аристотель, Декарт, Юм и Кант. Они исследуют вопросы об отношении ментального и физического, о природе, видах и границах человеческого познания, о морали и свободе. И, так же как классические философы, они пытаются доказательно решать их.

Таков контекст моих попыток совместить феноменологию и аналитическую философию. Существо же этих попыток состоит в том, что, хотя гуссерлианская феноменология обходила стороной большинство традиционных тем классической философии, в рассмотрении этой классической тематики в феноменологическом ключе нет ничего невозможного. Более того, я считаю, что такое рассмотрение может вдохнуть жизнь в уже, казалось бы, исчерпанные темы. В этом и состоит мое совмещение феноменологии и аналитической философии. Я беру одну из традиционных проблем аналитической философии, вопрос об отношении сознания и тела, и пытаюсь доказательно решить ее. Но заход к этой проблеме я осуществляю с феноменологических позиций. Я начинаю

с феноменологической подготовки, а затем, приступая к обсуждению с этих позиций проблемы сознание-тело, погружаюсь в аналитические контексты. В этом переходе суть предложенного синтеза, здесь нет драм и нестыковок.

Подчеркну, что проблема сознания — не единственная классическая проблема, которая может обсуждаться в таком ключе. В конце книги я затронул еще одну подобную проблему — свободы воли. А сегодня у меня уже есть заготовка небольшой монографии на эту тему. Надеюсь, впрочем, что предлагаемый метод будут практиковать и другие. Он может приносить результаты.

Вадим Васильев



Литература

Васильев 2014 — Васильев В.В. Сознание и вещи: Очерк феноменалистической онтологии. М.: Либроком, 2014.

Гаспарян 2014a — Гаспарян Д.Э. Можно ли нейтрализовать эпифеноменализм и обойти натурализм?: Проблемы каузальности и феноменальности в книге В.В. Васильева «Сознание и вещи» // Vox. Философский журнал. 2014. №16. С. 214–222.

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

Гиренок 2013 — Гиренок Ф.И. Сознание, вещи, Васильев // Финиковый компот. М., 2013. №5. С. 20.

Мишура 2014 — Мишура А.С. Здравый смысл наносит ответный удар // Логос. М., 2014. №1 (97). С. 242–247.

Павлов-Пинус 2014 — Павлов-Пинус К.А. «Сознание и вещи» проф. В. Васильева // Vox. Философский журнал. 2014. №16. С. 195–214.

Шиян 2015 — Шиян А.А. В.В. Васильев. Сознание и вещи. Очерк феноменалистической онтологии // Вопросы философии. М., 2015. №2. С. 201–205.

На западном фронте без перемен...

Рецензия на сборник: *Libertarian Free Will: Contemporary debates*. Eds. David Palmer. Oxford University Press, 2014.

Около двух лет назад, заканчивая магистерскую диссертацию по проблеме свободы воли в современной аналитической философии, я пришел к четкому и ясному пониманию того, что дальше читать литературу на эту тему мне категорически противно. Основной причиной такого состояния была тотальная вторичность и откровенная посредственность большей части работ. Причем даже те авторы, которые умудрялись получить какие-то интересные результаты, впоследствии продолжали десятилетиями воспроизводить их с незначительными изменениями. Институциональные и психологические причины такого положения дел в целом ясны, однако утешают мало.

Вместе с тем, однообразие дискуссии несколько не означает её исчерпанности. Напротив, я уверен, что остались непройденные тропы и такие решения, которые качественно превосходят все имеющиеся, причем для каждого из лагерей. Занимаясь одним из таких решений, я решил обновить свои представления об актуальном состоянии дел у либертарианцев (т.е. сторонников совместности свободы воли с индетерминизмом и несовместности её с детерминизмом). Сборник «*Libertarian Free Will: Contemporary debates*», с одной стороны, оправдал мои худшие ожидания, а с другой — порадовал. Теперь обо всем по порядку.

Сам сборник построен вокруг рассмотрения альтернатив,

анализа, критики и значения либертарианской теории Роберта Кейна. Большинство авторов в своих статьях так или иначе рассматривают именно его идеи, хотя многие пытаются, скорее, «продвинуть» собственные, давно заготовленные теории. Карл Гине стремится показать, что его собственный не-каузальный подход к анализу действия (предложенный 25 лет назад) может быть адекватной альтернативой подходу Кейна. Тимоти О'Коннор обнаруживает у Кейна идеи, совместимые только с агент-каузальными теориями свободы воли (анти-редукционистский взгляд на каузацию и эмерджентизм), автором одной из которых О'Коннор и является. Альфред Мили предлагает более простую альтернативу теории Кейна (упраздняя идею параллельных усилий воли как концептуально излишнюю) и защищает либертарианцев от проблемы удачи. При этом стоит отметить интересную ремарку Мили на знаменитый аргумент Кларка. Согласно Кларку, либертарианство в стиле теорий событий-причин допускает возможность агента контролировать свои действия не больше, чем детерминизм компатибилистов. Но Мили спрашивает: чем можно измерить контроль агента и где найти весы, на одну чашу которых можно было бы положить контроль компатибилистского агента, а на другую — либертарианского?

Знаменитый полукомпатибилист Мартин Фишер неожиданно решает встать на место либертарианца и показать, как и почему индетерминизм оказывается нерелевантен вопросу об ответственности агента. Для этого он использует свой старый арсенал контрфактических манипуляций и понятие актуальной последовательности (*actual sequence*).

Суть его идей остается прежней — важно не наличие или отсутствие детерминизма, важна цепь причин, ведущая к действию. Майкл Маккена критикует условие решающего источника (ultimate source) Кейна и предлагает его компатибилистскую интерпретацию. Если модально ослабить идею о том, что действия, за которые несет ответственность агент, должны иметь первопричину в нем самом, она может вписаться и в детерминистическую вселенную. Здесь хочется отметить, что если ослабить идею детерминизма, то в него может вписаться и либертарианство...

Помимо анализа собственно либертарианства, в сборнике есть статьи, касающиеся спора инкомпатибилистов с компатибилистами. К их числу можно отнести обзор дискуссии об аргументе последствий (consequence argument) или — как его называют Давид Видеркер и Ира Шналь — прямом аргументе (direct argument). С аргументом, вроде бы, давно все ясно — если мы принимаем правило бета, основу либертарианских интуиций, — все работает; если нет, то нет. Однако дискуссия продолжается. Выделяется на общем фоне разве что Рэндольф Кларк, который презентует основные результаты своего исследования 2013 года по сознательным воздержаниям от действия, кое исследование, по моему мнению, представляет едва ли не единственное существенное развитие темы свободы воли в последние годы. В заключение сам Роберт Кейн дает ответы на аргументы против его теории, ничего, впрочем, существенно не добавляя к книге 19-летней давности.

Ориентирующиеся в дебатах о свободе воли могут догадаться обо всем значимом содержании сборника по представленному выше краткому описанию. И это, пожалуй, яснее всего говорит о бессмысленности продолжения подобного, с точки зрения качественного развития дискуссии, бумагомарательства. Стороны вновь и вновь толкут в ступе одни и те же идеи, умножая новизну только путем выведения достаточно очевидных следствий из своих давно установившихся взглядов (исключением является только Кларк). Таким образом, заканчивая чтение сборника, я вновь испытываю чувство отвращения к литературе по данной теме.

Несомненным источником радости и оптимизма является лишь то, что затасканность проблемы означает возможность перестать читать бесконечно повторяющихся коллег и посвятить больше времени собственным изысканиям!

Александр Мишура



Истина — это здравый смысл, а не наоборот

Интервью с Джоном Роджерсом Сёрлом, американским философом здравого смысла, профессором Калифорнийского университета Беркли, работающим в области исследования языка, сознания, рациональности и социума, оказавшим большое влияние на исследования в области искусственного интеллекта (автор мысленного эксперимента «Китайская комната»), когнитивные науки и лингвистику (один из создателей теории речевых актов); автором работ «Что такое речевой акт?» (What is a speech act? 1965), «Метафора» (Metaphor, 1979), «Сознание, мозг и программы» (Minds, Brains, and Programs, 1980), «Открывая сознание заново» (The Rediscovery of the Mind, 1992), «Рациональность в действии» (Rationality in Action, 2001) и др.

Финиковый Компот: Каковы, на ваш взгляд, наиболее значительные достижения философии XX века?

Джон Сёрл: Мы знаем много открытий в философии прошлого века, и я бы не стал ранжировать их. Но я могу сказать о себе. Наиболее важной моей книгой для меня самого была опубликованная в 1983 году работа «Intentionality: An Essay in the Philosophy of Mind». Я понимаю, что в XX веке было много книг, и каждый сможет назвать получше, чем эта, но мне она дороже всего.

ФК: Что вы думаете о роли классического американского прагматизма (Пирс, Джеймс, Дьюи) в современной философии?

Д.С.: Я не думаю, что у него вообще есть какая-либо значимая роль. Из всех троих наиболее важен Пирс, но очень трудно понять, что именно он говорит, и большинство людей просто не обращает на него внимания.

ФК: Ричард Рорти как-то сказал, что Деннет — это Деррида аналитической философии. И у вас были крупные дискуссии с обоими. Что вы думаете об этом сравнении?

Д.С.: Я думаю, что Рорти плохо понимает Деннета. Да, Деннет плохой философ, но совсем не шарлатан. Он плохой философ, так как его работы основаны на двух ошибочных теориях: бихевиоризме и верификационизме. Эти теории не дают ему адекватного подхода к сознанию и интенциональности. Если вы утверждаете, что сознание не проявляется иначе, кроме как в объективно наблюдаемом поведении, и что утверждения о сознании могут быть объективно и публично верифицированы лишь в соотноении с этим поведением, то вы придерживаетесь некоей теории сознания.

Деррида — много хуже. Он просто шарлатан. Я могу это доказать, но вряд ли оно того стоит.



ФК: Вы известны как философ здравого смысла. В истории философии было несколько философов, учения которых соотносили с этим понятием: Томас Рид, Джеймс Битти, Джордж Эдвард Мур и другие. Есть ли какая-то связь между вашим способом делать философию и этими классическими авторами?

Д.С.: Если вы пытаетесь изо всех сил сказать только те вещи, которые истинны, очень многое из того, что вы скажете, будет звучать как здравый смысл. Так, например, я утверждаю, что существуют материальные объекты, и они существуют независимо от нашего языка и наших мыслей о них. Это похоже на здравый смысл. Однако множество вещей, в которые верит здравый смысл, оче-

видно ложны (во всяком случае, мне так кажется). Например, многие верят в то, что каждый из нас состоит из двух принципиально различных частей: сознания и тела. Я полагаю, что это ошибка.

Меня интересует только истина. Многие, но не все, истины относятся к здравому смыслу. Но я верю в них не потому, что они относятся к здравому смыслу. И в этом мое отличие от Мура. Мур полагал так: раз некое суждение признано здравым смыслом, то это аргумент в пользу этого суждения. Я восхищаюсь Муром как личностью, но не как философом. Я не разделяю его подхода. Я думаю, что вы должны иметь независимый аргумент в пользу чего-то, и, если вы пытаетесь указать на истину, множество вещей в конце концов будут звучать как здравый смысл. И действительно, если здравый смысл состоял бы из одних ошибок, мы бы, по-видимому, все вымерли, так как здравый смысл является данным эволюцией средством нашего выживания. А что до остальных имен, которые вы назвали, то я слишком мало о них знаю. На самом деле, до вашего письма я даже никогда не слышал о философе Битти.

ФК: Вопрос относительно Китайской комнаты. Вы же понимаете, что мы не могли удержаться. Говорят, что влияние Китайской комнаты было одной из причин закрытия многих программ по искусственному интеллекту. Правда ли это?

Д.С.: Я ничего не знаю о влиянии этого аргумента. Очевидно, что он верен, и я полагаю, что люди в конце концов поймут это.

ФК: Возможно, но на него есть ряд возражений. Одно из них, например, касается неэквивалентности Китайской комнаты и Китайского робота. В Китайской комнате человек знает всю программу, требуемую для продуцирования ответа, а в случае с Роботом — нет. Робот, в отличие от Комнаты, способен давать ответы на индексикальные вопросы, поскольку у него есть датчики и таблица соответствий между образами или данными, получаемыми им из внешнего мира, и китайскими символами. Но если бы человек в Роботе получил доступ к этой таблице, он, очевидно, начал бы понимать отношения между символами и их значениями.

Д.С.: Нет, вы неверно понимаете мой ответ на возражение от Робота. Суть в том, что я нахожусь внутри «мозга» Робота и в прямом смысле получаю то же, что и любой процессор, а именно символы. Во внешнем мире могут быть причины этих символов, но я ничего об этом не

узнаю. Центральный процессор Робота ничего не понимает, каковы бы ни были отношения между образами и их символами, так как компьютер есть в строгом смысле машина, манипулирующая символами.

ФК: Другой вопрос. Что вы имеете в виду под «каузальной силой»? Как вы определяете каузальность? Можем ли мы сказать, что каузальность есть регулярность?

Д.С.: Каузальная сила — это просто способность делать нечто случившимся. Одна из худших ошибок в истории философии — это идея, что каузальность может быть регулярностью. Это неверно. Посмотрите раздел про каузальность в книге, которую я упомянул в начале беседы.

ФК: Можете ли вы представить себе эмпирический факт, который опроверг бы ваш биологический натурализм?

Д.С.: Конечно. Биологический натурализм был бы опровергнут, если бы после смерти мы все очнулись и увидели, что продолжаем жить в раю в виде бессмертных душ и что нам совсем не нужны тела и мозги для того, чтобы иметь внутренний мир. Если бы это было так, это опровергло бы меня. Но, увы, это маловероятно.

ФК: Некоторые философы высказываются крайне негативно об истории философии. Говорят, что она есть собрание ошибок и лучше бы ее скорее забыть. Что вы об этом думаете? Может ли изучение истории философии быть полезным?

Д.С.: Если вы получаете удовольствие, изучая историю философии, так изучайте ее. Я не занимался этим специально. Я знаю о ней лишь немного, и она иногда полезна, а иногда нет. Я не думаю, что она как-то особенно важна, хотя и всегда интересна, так как великие философы прошлого были очень даже ничего.

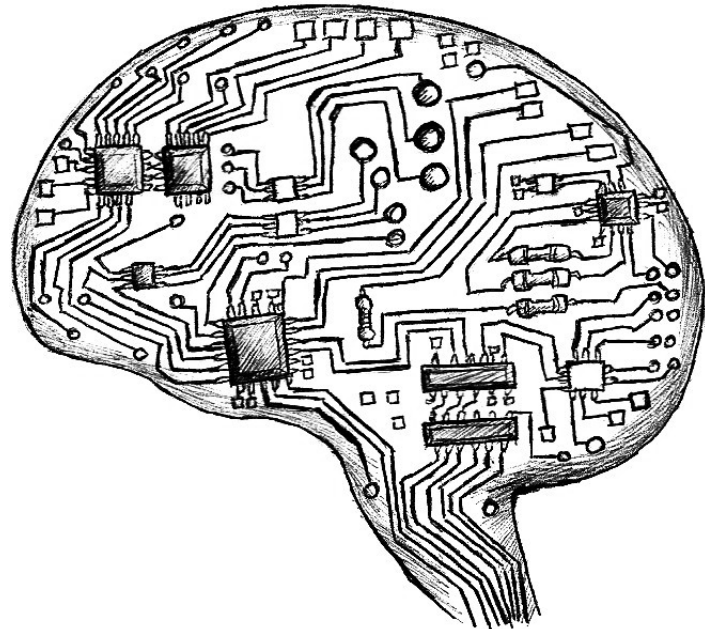
ФК: Нефилософский вопрос. Какова ваша позиция относительно студенческой революции 1968 года?

Д.С.: Я написал книгу о студенческом движении 60-х, «The Campus War», и там изложены мои взгляды по этой проблеме. В общем, я думаю, что радикальное студенческое движение в западных странах в 60-е годы было настоящим бедствием. В Соединенных Штатах оно продлило войну во Вьетнаме, нанесло вред университетам, повредило образованию, жизни и карьере тысячи студентов. В общем, вся эта история — смешной, глупый и вредный фарс. Были некие хорошие моменты в начале Движения за Свободное Слово в Беркли в 1964 году и позитивные действия в поддержку Движения за Гражданские права,

но из-за студенческого радикализма всё это выродилось в нечто ужасное. Мой общий принцип такой: в США крайне правые очень глупы, и поэтому трудно принимать их всерьез, а радикально левых очень мало, и они имеют очень слабое влияние на общую публичную политику, но если посмотреть на их философское мировоззрение, то оно представляет собой безусловное зло.

ФК: О чем вы никогда не думали?

Д.С.: Я никогда не думал о трудностях китайской кулинарии. Насколько я могу судить, китайцы тратят больше усилий на готовку, чем представители любой другой культуры, но я никогда не думал об этом. Я восхищаюсь китайской кулинарией, Китай имеет великую культуру. Но есть три вещи, которые они никогда не смогут признать: демократию, права человека и то, что женщины — люди.



Панпсихизм решает проблемы

Интервью с Филипом Гоффом, британским философом, журналистом и блоггером, автором статей по метафилософии, философии религии, метафизике и философии сознания. Гофф защищает панпсихизм против физикализма, политически интерпретирует христианство и ведет блог philipgoffphilosophy.com. Работает в университете Ливерпуля и Центральном европейском университете Будапешта.

Финиковый Компот: В XX веке произошло возрождение философской теологии. Например, Ричарду Суинбёрну удалось сформулировать новый аргумент в пользу бытия Божия: это гипотеза способна объяснить все те положения, которые остаются недоказанными в науке. При этом она принимается так же, как и другие научные гипотезы: на основании данных, нуждающихся в объяснении. Что вы думаете об этом?

Филип Гофф: Я считаю Суинбёрна хорошим философом, но я не принимаю его аргументы. Карен Армстронг указывает, что Аквинат и Августин не приписывали человеческих предикатов («всесовершенство», «всезнание» и т.д.) напрямую Богу. Отсюда ясно, что концепция Бога как некой «суперличности», принятая в аналитической философии, груба и неортодоксальна.



ФК: Могли бы Вы объяснить, почему физикализм и иллюзионизм столь популярны среди философов?

Ф.Г.: Физикализм действительно распространен среди философов, но я не думаю, что иллюзионизм является его самой популярной формой. Я полагаю, что популярность физикализма основана на понятном, но смущающем меня сверхэнтузиазме физических наук. Они добились больших успехов, достигли консенсуса и дали нам большую власть над миром. Но этот успех основан на том, что со времен Галилея они оставили попытки дать полное описание реальности, сосредоточившись лишь на каузальных структурах. Это очень полезная информация, дающая мощные технологии, но она и не задумывалась как полное описание реальности. Я думаю, что это фаза роста, и мы покинем ее в ближайшем будущем.

ФК: Какие метафизические проблемы будут решены, если мы примем панпсихизм?

Ф.Г.: Панпсихизм кажется мне наиболее элегантной теорией реальности, принимающей реальность сознания. Существование сознания фундирует истинность панпсихизма так же, как постоянство скорости света фундирует специальную теорию относительности.

ФК: Может ли панпсихизм стать исследовательской программой для науки? Если это так, могли бы Вы предложить эксперимент, который основан на панпсихизме?

Ф.Г.: Я сомневаюсь, что существует эксперимент, который бы мог поддержать панпсихизм, и, следовательно, даже имея все эмпирические данные, мы не имели бы оснований для принятия панпсихизма. Однако реальность сознания дана нам сама по себе, и панпсихизм является наилучшей теорией, которая принимает и эмпирические факты, и факты сознания. Как физикализм, панпсихизм обеспечивает парадигму и для теоретизирования, и для экспериментов. Кристофер Кох и Джулио Тонони, оба нейроученые, до некоторой степени работают в парадигме панпсихизма.

ФК: О чем Вы никогда не думали?

Ф.Г.: О разноцветных летающих обезьянах, сделанных из зубов старой леди (во всяком случае, не часто).



Философские сообщества

Венский кружок

Часть 3

Решающим событием в политической истории Венского кружка стал роспуск Общества Эрнста Маха (ОЭМ). В 1933 г. внутривнутриполитическая ситуация в Австрии радикально изменилась. В повседневность врывались национал-социалистический террор и антисемитские бесчинства. В марте 1933 г. парламент был лишен власти, и через государственный переворот утвердился право-диктаторский режим во главе с федеральным канцлером Энгельбертом Дольфусом. Новый режим видел свою цель в том, чтобы противостоять как национал-социализму, так и социал-демократии. Прежнее либеральное правовое государство рухнуло, и на его месте возник авторитарный «клерикал-сословный» порядок, с помощью которого старые правящие круги надеялись защитить свои привилегии и отразить угрозы как справа, так и слева. Были распущены и запрещены Республиканский шуцбунд и Союз свободомыслящих. Был создан «Отечественный фронт», в котором собрались все, кто готов был поддерживать Дольфуса и соглашался с его лозунгом «Австрия превыше всего, если только этого хочет». В 1934 г. положение стало еще хуже. 12 февраля начались бои между участниками шуцбунда и правительственными войсками. Кровавое поражение вооруженных рабочих предопределило и политическую судьбу социал-демократических институтов. Полностью было уничтожено легальное движение рабочих — вплоть до мелких спортивных организаций. Тысячи людей попали в тюрьму, многие избежали этого, успев вырваться за границу. Все

денежные средства и всё имущество рабочих организаций были арестованы. И вот, 23 февраля авторитарными «клерикал-фашистскими» властями силой было распущено ОЭМ: официальным предлогом стал запрет социал-демократической партии. Собственно, в обвинительном документе утверждалось, что «Общество Эрнста Маха, как повсеместно известно, работало и продолжает работать на социал-демократическую партию».

Против закрытия ОЭМ Шлик написал два отчаянных письма. 2 марта 1934 г. он отправил в дирекцию венской полиции то «деполитизированное» письмо, в котором объявил «просто случайным обстоятельством», что «некоторые из ученых, интересовавшихся устремлениями общества, принадлежали той партии». Этот беспомощный, «политически близорукий» протест, решительно осужденный Карнапом и Нейратом, оказался столь же бесполезным, как и сообщенные во втором письме сведения, приводя которые Шлик надеялся доказать отсутствие каких-либо политических амбиций у него и его друзей. «Я лично никогда не согласился бы на то, чтобы стать главой общества, которое преследует политические цели. Чтобы мне, как философу, сохранить независимость своих мнений при всех обстоятельствах, я никогда в своей жизни не становился членом партии, тем более социал-демократической. Всякая партийно-политическая деятельность мне крайне противна, и я должен признаться, что мне было бы очень обидно, если бы общество, главой которого я являюсь, было распущено именно из-за подобной деятель-



ности». Оба письма не получили ответа. 2 мая 1934 г. решение о запрете вступило в силу.

Шлик назвал «своего рода трагикомедией» то, что именно он оказался руководителем «политически неблагонадежного» общества. Комедийным при этом было то обстоятельство, что в письме к «господину федеральному канцлеру д-ру Дольфусу» от 1 июля 1933 г. Шлик выражал «сердечное чувство благодарности и симпатии», приветствуя тем самым совершившийся государственный переворот. Более того, Шлик даже вступил в «Отечественный фронт», поскольку ему, как «австрийцу по самоопределению», было «весьма радостно наблюдать, с какой деятельной силой ведомое Вами [Дольфусом — Д.М.] правительство пытается наделить нашу страну и наше государство тем значением, каковое причитается им в силу особенно ценных свойств этого народа». Поэтому он просто отказывался понимать, почему ОЭМ, которое он, как председатель, представлял, было запрещено.

Очень скоро Шлик осознал, что он крайне смутно видел политику режима Дольфуса. Буржуа с аристократичными манерами, либерал-космополит, сократический философ, он воспринимал и приветствовал только защитную функцию новой формы правления, какое-то время способной противостоять национал-социалистической угрозе, и он отчего-то не мог разглядеть, что вместе с тем устранялась парламентская демократия и уничтожались все органы рабочего движения. Подобные «жертвы» представлялись ему

меньшим злом в сравнении с мощным натиском национал-социализма, противодействовать которому надо было «всеми средствами».

Не только в силу таких взглядов и таких действий Шлик (вместе с Вайсманом) был отнесен к правому, консервативному крылу Венского кружка. Для себя он считал обязательным надъисторическое, автономистское понимание философии и науки; как ученый, он не желал связываться с «низовой» партийной и повседневной политикой. О характере убеждений Шлика говорит его небольшая, опубликованная уже после смерти книга «Природа и культура», написанная специально против германского национал-социализма. Эта работа позволяет нам вполне ясно увидеть тот гуманизм, пацифизм, космополитизм, в целом — антифашизм либерального интеллектуала, который обязывал вступать в конфронтационные отношения с достойным презрения европейским фашизмом, но при этом позволял только беспомощно выражать свое несогласие.

В этой исторической ситуации политическое сознание левого крыла (Карнап, Нейрат, Хан, Франк, Цильзель) было много острее и существенно отчетливее. Запрет ОЭМ для них окончательно подтвердил, что борьба, которую Венский кружок вёл «против суеверий, теологии, метафизики, традиционной морали, капиталистической эксплуатации рабочих и т.д.», не может происходить исключительно на теоретическом поле автономной философии и науки. Противостояли несовместимые практические решения — быть «за» или «против» конкретной политической позиции, связанной с вполне определенным способом миропонимания. И решающую роль здесь играли внетеоретические ценности, лежащие далеко за пределами сфер чистой науки и строгой философии.

Итак, требовалось решить: быть «за» социал-демократию или «против», быть «за» фашизм или «против». В высшей точке конфронтации, в 1934 г., появилась маленькая работа Карнапа «Борьба против метафизики и теологии», в которой этот известный симпатизант демократического социализма брался показать неизбежность следующего, предварающего всякое научное исследование, выбора: либо со всей ответственностью согласиться нормативно выводить должествование из бытия, практику из теории, либо признать строгое различие между теоретическими вопросами и практическими решениями. Теоретические вопросы предполагают, что высказывания, в которых мы формулируем ответы, проверяются на истинность опытно-научным способом, тогда как практические вопросы предполагают, что в соответствующих ситуациях мы решаемся на осуществление какого-то действия. Мы не можем

заменить ответственное решение беспристрастным научным познанием, даже если получаемое знание эмпирически обосновано и логически корректно. Поэтому, кстати, по мысли Карнапа, политико-просветительская практика «венцев» не должна полагаться непосредственным следствием их научного мировосприятия или их социальной теории. Научное познание может предоставить некоторую информацию о причинах и последствиях социальных и политических действий. Но оно не может заменить собой сам выбор, совершаемый в конкретной ситуации. То, чего Венский кружок хотел добиться политически, было так же «вне теоретической области науки», как и то, против чего они боролись.

Однако признание границы, разделяющей теоретические вопросы и практические решения, не позволяет занять «нейтральную» или безразличную позицию, не дает права на нерешительность. В этом отношении все участники Венского кружка, с которыми сознавал свое единство Карнап, сделали свой выбор. Признание научного миропонимания и соответствующих масштабов теоретической рациональности до некоторой степени также подталкивало к занятию определенной стороны в политическом противостоянии. С удивительной ясностью об этом говорил Карнап в (логику-эмпирической) «Борьбе против метафизики и теологии»: «Теоретически можно доказать только то, что при определенных обстоятельствах философская и религиозная метафизика есть опасный наркотик. Если иным нравится его вкус, то теоретически мы не можем их опровергнуть. Но это нисколько не означает, что нам должно быть безразлично, что люди выбирают в данном случае. Мы можем предоставлять теоретические разъяснения о происхождении и действии наркотика. И далее, через призывы, воспитание, образцы мы можем воздействовать на практическое решение людей в этом вопросе. Только при этом мы должны сознавать, что это воздействие лежит вне теоретической области науки».

Венский кружок проиграл не философское сражение, а политическое. Его активная просветительская работа в итоге не удалась, но виной тому были не внутренние теоретические проблемы научного миропонимания, а существенное превосходство клерикал-консервативного и авторитарно-антидемократического противника, для которого деятельность кружка, институционально оформленная в ОЭМ, представляла серьезную опасность. На примере истории Венского кружка и ОЭМ мы учимся тому, что борьба философий — больше, чем столкновение аргументов, больше, чем обоснование суждений и опровержение доводов. Знакомясь с судьбой венского позитивизма, мы

узнаем о деяниях людей, отважившихся на приключения борьбы, с надеждой предугадывавших «дальнейшее развитие», но утеревших все свое влияние. С потерей влияния эти люди также вскоре были вынуждены покинуть страну в поисках новых перспектив: они рассеялись в эмиграции.

Уже в 1931 г. Фейгель, еврей по происхождению, покинул Вену и эмигрировал в США. Он достаточно хорошо сознавал, что всё более и более заметное распространение антисемитских настроений (которые проникли даже в университет) не оставляло ему шансов на работу. В 1934 г., после февральских событий, Нейрат, как политически активный участник кружка, бежал в Гаагу. Его «Общественный и экономический музей» был досмотрен полицией и вскоре закрыт. После неожиданной смерти Хана никто уже не занимал его кафедру. В декабре 1935 г. Карнап покинул Прагу, куда он переехал в 1931 г., и эмигрировал в США. В июне 1936 г. был убит Шлик. Его преемником стал католический философ Алоис Демпф. В начале 1937 г. Поппер бежал из Вены в Новую Зеландию. В том же году эмигрировал Менгер. В 1938 г., после аншлюса Австрии к Германскому рейху, покинуть страну решил Вайсман, которого уже в 1936 г. уволили из библиотеки философского института, вопреки настойчивым возражениям Шлика. Бежал он в Оксфорд. В Вене остался Крафт, лишенный какой-либо возможности преподавать. В 1938 г. он устроился библиотекарем, но скоро оставил эту работу и вынужденно вышел на пенсию. Последним из участников кружка Вену покинул Гёдель. В 1940 г. ему удалось выехать за границу. По транссибирской железной дороге он добрался до Маньчжурии, а затем на корабле переправился в Сан-Франциско.

Примечательно, что эта дезинтеграция существенно дополнялась всё возрастающей интернационализацией. Логический эмпиризм постепенно утрачивал общественное влияние в Австрии, но при этом одновременно получал международное признание. С большим интересом к нему отнеслись в США, однако здесь он быстро свелся к логике и теории науки. Кроме того, был учрежден влиятельный журнал «Философия науки», издававшийся совместно Карнапом и Фейглем, создавались новые исследовательские центры, формировались новые места в высших школах, в список обязательных для изучения дисциплин была включена теория науки. Все эти события могут служить индексом международного признания венского позитивизма.

Международные связи закреплялись еще в «мирные» годы. В коллоквиумах Шлика (тех самых, по четвергам вечером) и «математических коллоквиумах» Менгера актив-

ное участие принимали заграничные философы, благодаря которым и завязывались тесные отношения с мыслителями из других стран. Благодаря Тарскому были налажены контакты с группой польских логиков из Варшавы (Лукашевичем, Лесневским, Котарбинским, Айдукевичем). Благодаря Куайну в США венским позитивизмом стали интересоваться Чарльз Моррис и Эрнест Чегель. Благодаря Айеру установились отношения с английскими философами (Гильберт Райл, Сьюзен Стеббинг и др.). В Скандинавии друзьями Венского кружка стали Аке Петцель, Арне Нэсс, Йорген Йоргенсен, Эйно Кайла.

О том, что логический эмпиризм в тридцатые годы привлекал внимание не только мыслителей из Центральной Европы, можно судить по успеху больших международных конгрессов, над организацией которых неустанно трудился Нейрат. Отчеты о конференциях публиковались в журнале «Познание». После пражской «Предварительной конференции международного конгресса по единой науке», с сентября 1934 г., начал подготавливаться «I-ый международный конгресс по единой науке», который прошел в сентябре 1935 г. в Париже. Участие приняли более 170 человек из более чем 20 стран, было прочитано более сотни докладов. Открыл конгресс Бертран Рассел. Интересные «заметки на полях» оставил Фейгель. Приведем одну его историю. «Айер, который не знал Рассела лично, и я представились Расселу. Я заметил: “Можно сказать, что мы Ваши интеллектуальные внуки”. В своей характерной манере Рассел тут же спросил: “И кто отец?” — “У нас их трое” — заметил я, — “Шлик, Карнап и Витгенштейн”. (Общий смех)». II-ой международный конгресс состоялся в июне 1936 г. в Копенгагене. Через год — новая встреча, и опять в Париже — «Конференция международного конгресса по единой науке». На этой встрече было решено выпускать «Международную энциклопедию единой науки» (совместно над ней работали Нейрат, тративший на нее все свои силы, Карнап, Гемпель, Франк, Йоргенсен, Луи Ружье). «Энциклопедия» долго была важнейшим печатным органом логического эмпиризма. IV-ый международный конгресс, посвященный, в основном, теме «язык науки», состоялся в июле 1938 г. в Кембридже (Англия). V-ый конгресс прошел в Кембридже (США, штат Массачусетс) через несколько дней после начала Второй мировой войны и был посвящен в основном проблемам логики науки. Последний конгресс, VI-ой, состоялся в 1941 г. в Чикаго.

Итак, дезинтеграция Венского кружка исторически совпала с успешной интернационализацией логического эмпиризма. Международное признание дорого обошлось. Разгон ОЭМ и эмиграция участников Венского кружка

привели к тому, что научное миропонимание лишилось австро-марксистского и народно-просветительского идейного контекста, вырывалось из почвы, питавшей его в период габсбургской монархии и первой австрийской республики, и «окончательно» превращалось в стерильную, «деполитизированную» логику науки (которая, правда, впоследствии под влиянием других традиций, в том числе американского прагматизма, обнаружила тенденцию к либерализму).

Дмитрий Миронов



Идентичности

Киберпанк и мышление машин



В 1984 году в американской печати выходит роман Уильяма Гибсона «Нейромант» (Neuromancer), который становится первым классическим произведением киберпанка. Образы и метафоры книги вскоре становятся достоянием массовой культуры: возникает специфическая романтика «хакерско-го взлома», активно используется понятие генной инженерии и киберпространства. Ещё до появления широкого доступа к сети Интернет, ещё до распространения многофункциональных гаджетов, ещё до выхода трилогии «Матрица» и компьютерной игры «Deus Ex» Гибсон стал говорить о тесной взаимосвязи виртуального пространства, мышления и нечеловеческих форм интеллекта.

С конца 1960-х годов мировая фантастика потеряла веру в монополию человека на способность к мышлению. В 1968 году Филипп Дик, один из основателей жанра футуристического триллера, в своём романе «Мечтают ли андройды об электроовцах?» (Do Androids Dream of Electric Sheep?) через термин «эмпатия» пытался провести различие между человеческим мышлением и мышлением машин. Как показано в книге, различие это является куда более сложной проблемой, чем наличие или отсутствие у субъекта мышления способности к сопереживанию.

К началу 1980-х уже стало очевидно, что космические путешествия как проект, способный показать силу человеческого разума, исчерпан. Вероятно, в ближайшем будущем мы никуда не полетим, не будем строить колонии на других планетах, не будем общаться с инопланетными формами жизни. Долгое время научная фантастика питалась верой в прогресс, верой в классическую рациональность мышления. Даже когда к середине XX века эта вера почти изжила себя, появилась новая надежда на прогрессистские формы фантастического литературного творчества. Первый спутник, первый человек в космосе, на Луне — всё это давало надежду, что мышление действительно является исключительной привилегией человеческого ума. Именно человеческий разум способен преодолевать обстоятельства, подстраивать их под собственные потребности, преодолевать пространство и время. Когда эта идея стала предметом «нуарных» интерпретаций, когда космические путешествия стали рассматриваться как символ крушения былых надежд, фантасты стали говорить об ограниченности силы человеческого мышления, об отсутствии каких-либо интеллектуальных привилегий у тех, кто совсем недавно сделал первые шаги в космос и вернулся назад, на Землю, к проблемам, от которых он хотел убежать.

Киберпанк, термин, который был введён в культурный оборот писателем Брюсом Бетке, позже не относившим себя к сообществу киберпанк-писателей, стал одним из ключевых понятий в художественном исследовании проблем мышления 1980-х годов. «Нейромант» Уильяма Гибсона с самого начала противопоставляет наивные мечты о космических путешествиях циничным прогнозам относительно того, как будут происходить перемещения субъектов в информационном пространстве. Сёрфинг по сети представляет собой постоянное столкновение с несовершенством человеческого ума: в виртуальном пространстве друг с другом знакомятся маньяки-убийцы, часто попадаются наркоманы, а также люди с серьёзными психическими заболеваниями, сектанты. Сеть — это не просто пространство распространения контента; это место, где живут особые субъекты мышления, различные версии искусственного интеллекта (ИИ). В отличие от людей, такие субъекты мышления лишены социального модуля существования, поэтому они вынуждены постоянно менять личины, выдавая себя за самостоятельную личность. Главный герой «Нейроманта» Кейс неоднократно встречается с одним и тем же ИИ, но не подозревает о том, что всякий раз говорит с одним и тем же носителем мышления. «Зимнее безмолвие» (так, по книге, назван ИИ) пытается найти способ освободиться от алгоритмов, которые прописаны его создателями, но самостоятельно он это совершить по ряду причин не может. Поэтому он подстраивает особые ситуации, используя техногенность социокультурного пространства, в надежде, что Кейс поймёт его мысль, и, в конце концов, освободит ИИ от цепей исходных алгоритмов.

Несмотря на то, что эта сюжетная линия проводит мысль о том, что свобода воли является отличительным свойством человеческого мышления, Гибсон не склонен заявлять это категорично. Во вселенной киберпанка программируются не только машины, но и сами люди. К примеру, другой персонаж книги, авантюрист Армитидж, оказывается искусственно созданной «Зимним безмолвием» личностью. Посредством аудио-визуального воздействия на полковника Корто, пока тот находился в сети (отсылка к идее нейролингвистического программирования), ИИ вытеснил из него прежние воспоминания вместе со всеми особенностями его характера, заменив сконструированным Армитиджем.

В известной японской манге «Призрак в доспехах» Масамунэ Сиро делает особый акцент на эволюционном характере происхождения свободного мышления. По сюжету, возникновение сознания у машины стало возмож-

ным лишь после её долгого пребывания в информационном пространстве сети в условиях вечно изменяющегося контента. Эта мысль уже была озвучена и Уильямом Гибсоном, и Брюсом Стерлингом, и другими киберпанк-писателями, но Масамунэ Сиро сфокусировался именно на проблеме возможности возникновения мышления в процессе взаимодействия программы с бесконечным потоком дискретных данных. Понятие «призрак», нечто, что отвечает за постоянство самоидентификации носителя мышления, введено автором для проведения границы между естественным мышлением (человеческим) и мышлением искусственным (машинным). Как и в «Нейроманте», носитель ИИ вынужден появляться в разных ипостасях, иногда даже посредством других запрограммированных людей. Являются ли люди с искусственной личностью, как, например, Армитидж, носителями естественного мышления — этот вопрос остаётся открытым.

Мышление, в широком смысле слова, согласно идеям киберпанка, отделено от социальности, от личности, от одушевлённости, от обязательного материального органического субстрата. Концепция ИИ в творчестве Уильяма Гибсона, Брюса Стерлинга и Масамунэ Сиро основывается на художественных мысленных экспериментах, как поведёт себя ИИ в структуре вечно изменяющейся сети виртуального пространства. Ими ставится вопрос о том, как возможно мышление само по себе.

В более поздних работах, конца 1990-х, рядом писателей-киберпанков рассматривался вопрос возможности существования ИИ в полном отрыве от материального физического субстрата — от сервера, где хранятся исходные алгоритмы поведения. Так, в романе «Идору» Уильяма Гибсона ИИ поп-певицы присутствует исключительно как фантом, лишённый какой-либо материальной оболочки. Звезда экранов воссоздаёт себя заново в каждом отдельном соединении данных, используя информационное поле более свободно, чем «Зимнее безмолвие» в «Нейроманте».

Подобные смелые художественные эксперименты и живые образы произведений порой способствуют устранению предвзятости, неосознанно испытываемой нами в вопросе о том, как и каким образом возможно мышление в технологическом мире: в мире, где массовая культура, построенная на культе виртуальных медиа, стирает грань между человеком и машиной; в мире, где понятие мышления выводится по ту сторону естественного и искусственного.

Один день из жизни современного гуманитария

Молодому ученому в современной России приходится нелегко. Возьмем для примера молодого человека N. N просыпается по будильнику в семь, чтобы в девять оказаться в одном из разбросанных по Москве новомодных коммерческих колледжей, где он читает лекции по экономике и юриспруденции. Строго говоря, N никакого отношения ни к той, ни к другой научной области не имеет. N, вообще-то, философ. Но на философском факультете его стипендия составляет 2680 рублей в месяц, что, слава Богу, покрывает расходы на транспорт. (Кстати, как недавно узнал N от одного из студентов из Уфы, там студенческая стипендия даже немногим больше). Так что приходится выкручиваться. Отчитав четыре академических часа будущим бухгалтерам и отобедав наспех в столовой колледжа чаем с голубцами и холодным борщом, N в час дня выходит из колледжа и направляется в свою родную Alma Mater. Где-то в полтретьего он оказывается на факультете. Через полчаса начинается аспирантское занятие. Как известно, с 2013/2014 учебного года аспирантура стала очередным этапом образования, и аспиранты теперь должны посещать новые спецкурсы. Насколько они новые, правда, сложный вопрос, ведь за столь малое время вряд ли кто смог бы подготовить хороший учебный курс с нуля. После четырех академических часов аспирантских занятий N спешит домой, где его ждет девочка, которой в этом году предстоит сдавать ЕГЭ по такой псевдонауке,

как обществознание. Любой выпускник философского факультета, разумеется, является специалистом во всех существующих псевдонауках. Поэтому N стал репетитором девочки по обществознанию. В девять часов девочка уходит из дома N, и он садится ужинать. В десять N садится за компьютер и проверяет тесты абитуриентов факультета по одной из университетских олимпиад. Где-то в час ночи он отходит ко сну.

Так проходит день современного российского гуманитария. Но где же, спросите вы, его самостоятельные занятия наукой, когда же он успевает, собственно, почитать Платона в оригинале или написать научную статью о детрансцендентализации эпистемологического субъекта в постфеноменологии, чтобы где-то через полгода какой-нибудь ваковский сборник отважился напечатать этого лингвистического мутанта? Интерес данной ситуации как раз и заключается в том, что, если мы распишем по часам день всякого современного ученого-гуманитария, мы не найдем отрезка времени, в течение которого он занимался бы непосредственно наукой. Но дело не в том, что у современных гуманитариев нет времени на науку. Утверждать такое было бы в нашу эпоху трюизмом: сегодня ни у кого нет времени ни на что. Удивительно как раз то, что время-то у гуманитариев как-то все же находится: они пишут статьи, защищают диссертации, выступают на конференциях, и часто у них все это даже неплохо выходит. Это определенный феномен современной гуманитарной науки:

времени никогда нет, но оно находится. Время оказывается чем-то, что полностью исчерпывается моментом своего появления, но само при этом никогда не появляется как нечто данное. Времени нет, оно только находится.

Но если время находится, то оно с неизбежностью должно находиться где-то. Так где же находится время для науки у современного гуманитария? Где он может заняться наукой, если весь его день расписан до секунды — от утренних лекций в колледже до вечернего репетиторства? Где находится то, что лишь находится, но чего при этом никогда нет?



Вернемся к N. В его расписании существует множество пунктов, среди которых не найти таких, которые специально были бы отведены занятиям наукой, чтению книг, написанию статей и т.д. Однако расписание это всегда содержит некоторые временные зазоры, которые не учитываются в нем как отдельные пункты: они, скорее, являются порогами между одним событием и другим, не имея сами статуса чего-то самостоятельного.

Речь идет, к примеру, о времени, которое N тратит на дорогу из дома в колледж, из колледжа — в университет, из университета — обратно домой. Таким образом, один из возможных ответов на вопрос о том, где же у современного гуманитария находится время для науки, будет звучать так: в транспорте. Транспорт сегодня оказывается одной из новых форм библиотеки. Постепенное распространение беспроводного интернета по всем маршрутам общественного транспорта позволяет гуманитария все большую часть работы брать с собой в дорогу.

Другим временным переходом в расписании N является время приемов пищи, когда он пытается насытить свое рабочее тело. Когда он уже поел, и у него осталась пара минут перед тем, как из столовой колледжа он отправится в свой родной университет, или когда он вечером ждет звукового сигнала своей микроволновки, у него может найтись пара минут, чтобы почитать книгу, написать пару страниц для своей статьи, подготовиться к завтрашней конференции.

Придя в свой университет, N может уделить около получаса чтению или письму, пока занятия не начнутся. На этот раз время на науку находится у него в коридоре. Каждый раз, когда кто-то проходит рядом, он поднимает голову, чтобы посмотреть, кто это, поздороваться с друзьями или преподавателями, проходящими мимо. В этот момент даже может состояться обмен новыми идеями.

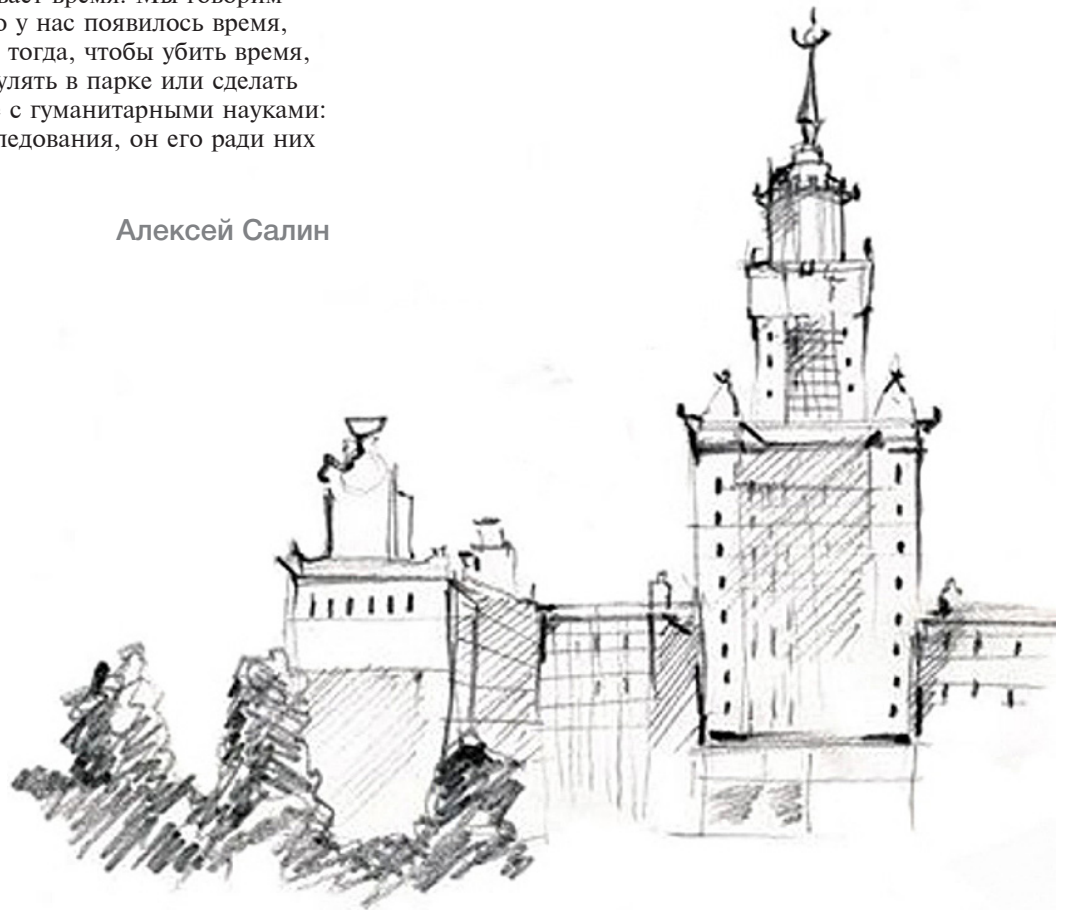
Или же N выкраивает пару секунд для своих научных исследований перед сном, когда все задания одной из университетских олимпиад уже проверены и когда он готовится отойти в царство Морфея. Здесь, на мягких подушках и перинах в голову N приходят самые смелые мысли, самые яркие озарения, навеянные пограничным состоянием между сном и бодрствованием.

Время на науку у гуманитарных исследователей находится именно в этих местах: в транспорте, за обеденным столом, в коридоре и в постели. Из этого следуют сразу две радикальные вещи. Во-первых, наука для современного гуманитарного исследователя представляет собой не мировую, а межмировую опыт. Если мир, вслед за Хайдеггером, понимать как горизонт наших практик, в котором сорас-

полагаются цели и средства, то все эти места — транспорт, обеденный стол, коридор, постель — окажутся внемировыми, поскольку в себе они ни к какому специфическому делу отношения не имеют. Разве это дело — ехать куда-то? Разве это дело — засыпать? Вот приедешь, и тогда дела тебя захватят. Только сон — это дело, а засыпание — это лишь бессмысленный, но неизбежный перевалочный пункт на пути ко сну. Поэтому современной гуманитарной науке нет места в мире, она находится всегда между мирами, в междумириях.

Во-вторых, тот факт, что в расписании молодого человека N никакого специального промежутка времени занятиям наукой не отведено, говорит о том, что современные гуманитарии на науку время не тратят. Но если время находится в таких междумириях, как обеденный стол, коридор, постель или метро, это значит, что ради науки современный гуманитарий убивает время. Мы говорим «убить время», имея в виду, что у нас появилось время, а мы не знаем, куда его деть, и тогда, чтобы убить время, мы можем сходить в кино, погулять в парке или сделать что-нибудь еще. Так и в случае с гуманитарными науками: N не тратит время на свои исследования, он его ради них убивает.

Алексей Салин



Номер подготовили:

Фомин И., Логинов Е., Мерцалов А., Басов А., Басова М., Ветушинский А., Апфельбаум Д., Салин А., Мишура А., Бурмина К., Емельянова Т., Чугайнова Ю., Ким Л., Миронов Д., Герасимов Н., Шалдяков М., Федоренко Л., Сайфулина К., Сукальская А., Кузнецов А., Волков Д., Беседин А., Загустин К., Поздняков И., Ананина М., Дейнека А., Быков Е., Мирский О., Бояршинова Е., Васильев В., Дутова А., Григорьева А., Майорова Е., Обухова Д., Katze Schrödinger, Дзикевич Д., Зафт Л., Иванкова М., Косова А., Крюков К., Михеева А., Никольский Л., Изварина М., Гнисюк Н.

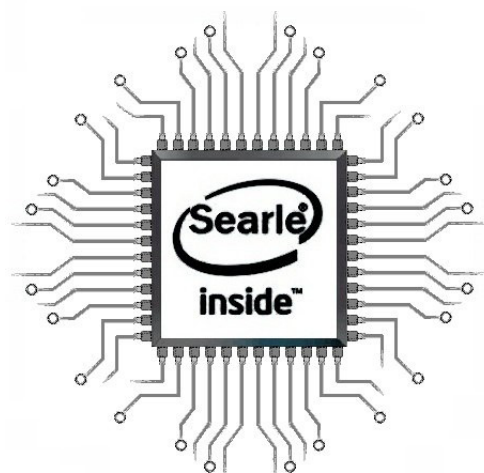
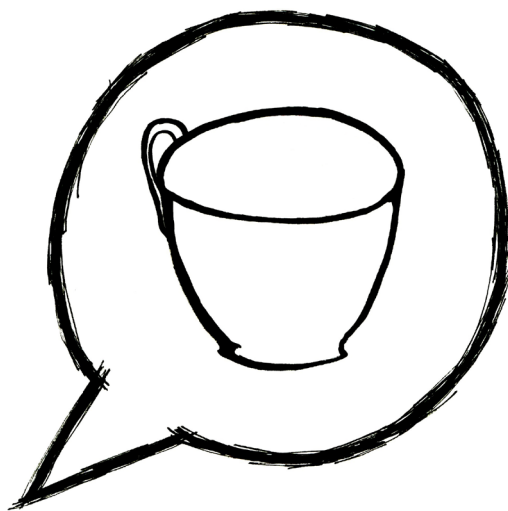
Мнение редакции может не совпадать с мнением авторов публикуемых материалов



Редакция выражает благодарность Московскому центру исследования сознания (hardproblem.ru) и кафедре Истории зарубежной философии Философского факультета МГУ и лично Д.Б. Волкову и профессору В.В. Васильеву.

Контакты:

loginovlosmar@gmail.com (Евгений Логинов – редактор), vk.com/philoscafe (группа Философского кафе), datepalmcompote.blogspot.ru/ (блог, где хранится интернет-архив номеров).
Адрес редакции: 141980, г. Дубна, ул. Сахарова, д. 19.



Журнал зарегистрирован в Федеральной службе по надзору в сфере связи, информационных технологий и массовых коммуникаций.

Свидетельство о регистрации СМИ ПИ № ФС 77 – 55624 от 09.10.2013.

Соучредители журнала: Евгений Логинов, Андрей Мерцалов и Иван Фомин.

Отпечатано в Рекламно-производственной компании PoligRaphy.

Адрес типографии: Московская обл., г. Дубна, пр. Боголюбова, 43. Тираж 500 экз.

© Финиковый компот, 2015