

ДИСКУССИИ

А.Н. Фатенков

НАТУРНЫЙ УМ В КОНТРАТАКЕ НА ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ (ПОЛЕМИЧЕСКИЙ ОТКЛИК НА «ЭВОЛЮЦИЮ РАЗУМА» РЭЯ КУРЦВЕЙЛА)

Фатенков Алексей Николаевич – доктор философских наук, профессор кафедры отраслевой и прикладной социологии. Национальный исследовательский Нижегородский государственный университет им. Н.И. Лобачевского. Российская Федерация, 603022, г. Нижний Новгород, пр. Гагарина, д. 23; профессор кафедры социально-гуманитарных наук. Приволжский исследовательский медицинский университет. Российская Федерация, 603005, г. Нижний Новгород, пл. Минина и Пожарского, д. 10/1; e-mail: fatenkov@fnsn.unn.ru

Обращаясь к геометрическим метафорам, автор защищает природно фундированный ум и – в контексте анализа концепции Рэя Курцвейла – обсуждает возможность решения человеком и машиной так называемых неразрешимых задач. Обсуждение данной проблемы ведется в сопряжении с проблемами иерархии, однородности и неоднородности, противоречия и аналогии, смысла и информации. Приводятся аргументы в подтверждение следующих идей: 1) цифровые технологии не преодолевают принципиальной ограниченности аналоговых технологий, что влияет на возможности искусственного интеллекта; 2) натуральный ум мыслит самобытно (не по аналогии), для него нет неразрешимых теоретических задач; 3) в здоровом человеческом мозгу нет ни избытка, ни недостатка, он не нуждается в искусственном усовершенствовании; 4) неполнота актуализации потенциала мозга не слабость, а сильная особенность человеческого ума; 5) теория искусственного интеллекта характеризуется избыточным редукционизмом и фетишизацией скорости.

Ключевые слова: натуральный ум, искусственный интеллект, Р. Курцвейл, однородность, неоднородность, аналогия, противоречие, иерархия, смысл, информация

Для цитирования: *Фатенков А.Н.* Натурный ум в контракте на искусственный интеллект (полемический отклик на «Эволюцию разума» Рэя Курцвейла) // Философский журнал / Philosophy Journal. 2022. Т. 15. № 3. С. 172–183.

В неприятии, отвержении чего-то или кого-то всегда важную роль берет на себя человеческое иррациональное. Вместе с тем этот ментальный, а вернее, психосоматический акт и соответствующее состояние человека никогда не являются сугубо иррациональными.

Некоторым образом разъяснит ситуацию и вдобавок высветит сущностную черту натурального ума обращение к Ф.М. Достоевскому. Его герои что-то искренне отвергают, чем-то и кем-то искренне дорожат. Они принимают

во внимание все возможные аргументы, будучи в готовности сходу отвергнуть их. Как в случае с Иваном Карамазовым: «Пусть даже параллельные линии сойдутся и я это сам увижу: увижу и скажу, что сошлись, а все-таки не приму»¹. Знаменитый и концептуально емкий, многогранный тезис. Рассмотрим его.

Действительно, можно ли по поведению линий (фигур одномерных) однозначно судить о характере пространства (реалии трехмерной, соразмерной человеку)? В однородном (гомогенном) пространстве есть и пересечения, и параллели; причем параллельными могут быть и прямые линии, и ломаные. В неоднородном (гетерогенном) пространстве параллелизм, строго говоря, невозможен, если не редуцировать неоднородность к комплексу сегментов, каждый из которых, отличаясь по плотности от других, внутри себя однороден. В гетерогенной среде линии только сходятся и расходятся. Однако изначальная фиксация непараллелизма никак не гарантирует нам встречу именно с гетерогенной средой.

Брат Иван не принимает не возможность схождения одномерных фигур – он отвергает гомогенность пространства, в которой все фигуры, даже проективно объемные, невольно становятся одномерно-плоскостными. И тогда, чтобы хоть как-то воспрепятствовать фатальной геометрической редукции, низведению плоскости к точке, не остается ничего другого, как табуировать схождение параллелей.

Что скрывается за жадной точечной гегемонии? Во-первых, нигилистическая по своей сути стратегия замещения минимизированной реальной (чувственно воспринимаемой) фигуры виртуальным (лишь мыслимым) означающим без означаемого, когда гомогенность отождествляется с пустотой, а точка, следовательно, с условно дробным отсутствующим в континуально отсутствующем. Во-вторых, обнаруживается и радикально прогрессистский замысел о допустимости проведения через идеальную точку сколь угодно числа линий, ведь с графической точкой на плоскости это допущение невыполнимо: какими бы тонкими ни были фигуры, неизбежно наступит момент, когда для начертания очередной линии места уже не останется.

Теперь – о карамазовском упрямстве (как выясняется, не сугубо иррациональным), в продолжение его. Признаюсь, в сегодняшней дискуссии вокруг искусственного интеллекта (ИИ) нахожусь среди тех, кто противится его безоглядному продвижению. У привносимого им комфорта, коммуникативных и медицинских вспоможений непомерно высокая цена (понимаю, что кто-то со мной и не согласится). В социальном плане диагностируются зачатки хронической тоталитарности: повышенный комфорт оборачивается повышенным контролем, никак не препятствующим, кстати, росту числа мошенничеств. В экзистенциальном плане отчетливо распознаются следы бесчеловечности, включая фиксируемое современной психиатрией цифровое слабоумие. И кем бы и какие бы аргументы ни приводились в защиту ИИ (даже не сугубо апологетические²), они не кажутся убедительными. Поэтому детально проработанные рациональные контраргументы тоже представлю. Цель – показать колеблющимся, взвешивающим вероятности, немалые возможности натурального ума, нисколько не нуждающегося для

¹ *Достоевский Ф.М.* Собрание сочинений: в 15 т. Т. 9: Братья Карамазовы. Л., 1991. С. 265.

² См., в частности: *Бостром Н.* Искусственный интеллект. Этапы. Угрозы. Стратегии. М., 2016.

достижения жизненных смыслов в подключении к электронному информационному облаку. Иначе говоря, под защиту берется стратегия, согласно которой когнитивные способности человека фундаментальны не только социально-культурно, но и биологически, и не одной лишь нервной системой, но всей нашей телесной организацией³. Умным может быть исключительно живое. Машину не оживить, и незакавыченно умной ей никогда не стать. Хотя соблазны, конечно, останутся. У тех, прежде всего, кто предъявляет претензии к природе, будто она им что-то недодала. Вот они и грезят о точке сингулярности, в которой естественный интеллект должен слиться с искусственным. Сингулярианское мировоззрение с его стремлением избавиться от тела, как справедливо отмечает Кори Пайн, насквозь мизантропично⁴.

Недавно уже доводилось говорить и писать о том, что отказ от тех, казалось бы, неоспоримых преимуществ, которые дает гуманитарии компьютер (пишущая машинка + библиотека + почта в одном компактном флаконе), и возвращение к доинновационному формату работы совершенно точно не отразятся негативным образом на ее качестве – скорее наоборот, с устранением спешки и суеты, детерминируемой скоростными возможностями «умной машины», качество текстов наверняка повысится⁵. Полтора пандемических года наглядно продемонстрировали паллиативность и образовательных онлайн-процедур: уровень знаний, умений и навыков студентов-первокурсников, обучавшихся в старших классах средней школы фактически заочно, крайне низок. Чем отвечать и восполнять пробелы? Где только позволялось, вел занятия очно, уважительно вспоминая Майкла Полани и его книгу о личностном знании, о необходимости и желательности непосредственного, без всякого экранирования, общения учителя и ученика⁶. (Не удивлюсь, если цифровое общество отправит труд Полани в реестр не рекомендованных к изучению.) Словом, привычный встречный выпад прагматического плана со стороны инноваторов – дескать, ругает то, чем сам с выгодой пользуется, – не страшен.

Сосредоточимся далее на преимущественно теоретических аспектах полемики с приверженцами ИИ, выбрав в качестве главного оппонента Рэя Курцвейла, открыто заявляющего, что без подключения к компьютерным ресурсам, расширяющим мозг, он чувствует себя неполноценным⁷. Заранее ожидаема, в общем-то, типичная для ИТ-шника экстраполяция компьютерной матрицы на структуру человеческого мозга и функционирующее человеческое сознание, как и встреча с характерным для сциентистского умонастроения жестким редукционизмом. Отдельные проявления его, впрочем, могут быть до некоторой степени любопытны.

Сомнительных упрощений в концепции Курцвейла предостаточно. Жизнь и эволюция сводятся им к процессу обучения, а далее к операциям с информацией. Мышление уравнивается с памятью, причем такой, применительно

³ См., в частности: Матурана У., Варела Ф. Древо познания. М., 2011.

⁴ См.: Пайн К. Живи, вкальвай, сдохни. Репортаж с темной стороны Кремниевой долины. М., 2019. С. 277.

⁵ См.: Фатенков А.Н. С цифровыми технологиями – в бесчеловечное будущее // Юридическая наука и практика: Вестник Нижегородской академии МВД России. 2020. № 1 (49). С. 217–219.

⁶ См.: Полани М. Личностное знание. На пути к посткритической философии. М., 1985.

⁷ См.: Курцвейл Р. Эволюция разума: как развитие искусственного интеллекта изменит будущее цивилизации. М., 2020. С. 414.

к которой забывание трактуется однозначно негативно; уравнивание педалируется, когда ставится задача подчеркнуть интеллигентный гомоморфизм человека и компьютера, и маскируется, когда приоритетно обсуждаются проблемы прогнозирования и предвидения будущего. Понимание отождествляется со статистическим анализом, которому приписывается иерархический (ранжирующий) характер. Мозг животного и человека подстраивается под машинный схематизм, трансформирующуюся систему нейронных модулей (блоков). В целом, обновленная вариация сциентизированного просвещенческого механицизма.

Ситуация становится чуть занятнее, когда выясняется, что каждый из нейронных блоков проще и мозга как такового, и отдельного нейрона. Курцвейлом постулируется, что нейроны плотно прижаты друг к другу в модуле, между ними нет выстроенных отношений. Отсутствие такой реляционной структуры и делает модуль простым. Сложность же, прежде всего сложность мозга, обусловлена как раз наличием реляционного каркаса, образующегося и разрастающегося в процессе совместного параллельного функционирования нейронных блоков. Введенный постулат позволяет, во-первых, обнулить, по сути, биологическую специфику нейрона и не обращать внимания на его сложность; во-вторых, подготовить возможность легкой замены и дополнения естественного мозгового модуля искусственным; в-третьих, подчеркнуть приоритет сети отношений перед нереляционными элементами (блоками) этой сети; в-четвертых, эксплуатировать идею ранга, без которой не обойтись при легитимации прогрессивной эволюции разума.

Американский интеллектуал и бизнесмен много рассуждает об иерархии: и в глобальном масштабе, касательно мироздания, и в масштабе локальном, касательно мозга *Homo sapiens*. Более того, обсуждается пластичная иерархия (идея и вовсе мне близкая⁸). Но как же ее истолковывает Курцвейл?

Истолковывает с явными логическими прорехами. Когда восторгается эволюцией и прогрессом, иерархия трактуется им как процесс и результат роста сложности: человеческий мозг, благодаря развитию новой коры, сложнее мозга любого другого млекопитающего (с эволюционно-прогрессивной точки зрения так оно и есть), но одновременно заявляет, что «один нейрон устроен сложнее, чем вся новая кора»⁹. И в чем же тогда прогресс? Допустим (другого варианта нет), в том, что появляется упрощенная (соразмерная возможностям компьютера) сложность, каждый уровень которой отличается от предыдущего лишь числом и величиной повторяющихся структурных блоков и, соответственно, числом связей между ними, иными словами, отличается фиксируемыми количественными параметрами. Однако, даже если некоторые из этих отличий описываются экспоненциально (как полагает Курцвейл), у нас никогда не будет гарантий ни того, что они носят качественный характер, ни того, что любая из складывающихся комбинаторных композиций умнее отдельного нейрона, а тем более умнее мозга, не редуцируемого к ассоциации модулей.

⁸ См., в частности: Фатенков А.Н. К онтологическим истокам идеи подвижной иерархии // Вопросы философии. 2008. № 5. С. 47–57.

⁹ Курцвейл Р. Эволюция разума. С. 19.

По Курцвейлу, «важнейшая особенность новой коры – удивительная однородность ее основных структур»¹⁰. Он намеревается совместить ранжированность с гомогенностью (без которой априори невозможно полное сопряжение от природы полученного разума и технически смонтированного интеллекта)... и моментально реальная иерархия замещается у него симулятивной. Есть смысл настаивать: ранжированность аутентична только в гетерогенном пространстве, она и делает его собственно гетерогенным; живое существо характеризуется (в корректном приближении) органичной неоднородностью, не доходящей в здоровом состоянии до дуалистической и плюралистической расколотости. Всамделишная неоднородность, включая человеческую, становится серьезным препятствием для бесконечного прогресса: какой-то очередной рубеж восходящее движение может и не преодолеть. В частности, может не состояться (и, надеюсь, не состоится) слияние естественного и искусственного интеллекта.

Апологет слияния, ему деваться некуда, пытается одновременно сидеть на двух стульях: «иерархическом» (иначе прогресс недоказуем) и «однородном» (иначе недоказуем бесконечный прогресс, вернее, бесконечное движение в определенном направлении). В итоге нам предъявляют имитацию иерархии и заведомо содержащий погрешности гомогенный аналог негомогенного.

Солидаризируясь с Д. фон Нейманом, Курцвейл ставит компьютер с его якобы более пластичной структурой выше человеческого мозга. Вот, собственно, и результат спекулятивно-иерархического подспорья. Естественное не сопрягается, а поглощается искусственным. Агрессивный технократический «прогресс». Его агент, апеллируя уже к последним *IT*-разработкам, поясняет: современные компьютеры используют исключительно цифровую алгоритмику, тогда как человек в своих мыслительных процедурах – и цифровую, и аналоговую (якобы гораздо менее совершенную). Однако находим и такой примечательный тезис: «Аналоговые механизмы действия мозга можно моделировать с помощью цифровых механизмов, поскольку цифровые вычисления способны моделировать аналоговые значения с любой степенью точности (а точность передачи аналоговой информации в мозге достаточно низкая)»¹¹.

Курцвейл будто (а может, и вправду) не замечает, что проговаривается, что цифровой метод оказывается опять же аналогом, пусть и усовершенствованным, – аналогом аналога. Декларируемое превосходство компьютера сразу становится шатким. Оно исчезает вовсе, если иметь в виду еще три момента: 1) плавность аналогии зачастую ценнее дробности цифровых операций (с чем вынужден согласиться отчасти и автор «Эволюции разума» в ответе критикам¹²); 2) цифра сама по себе есть универсально-абстрактный аналог конкретно сущего (семь цветов радуги и столько же самураев в культовом фильме, к примеру); 3) рассуждение по аналогии с аристотелевских времен оценивается как ненадежное (какой цвет какому персонажу Куросавы приписать?).

Иначе говоря (усилим натиск), цифровые и аналоговые схемы, пусть они даже количественно преобладают в человеческом мышлении, не схватывают

¹⁰ Курцвейл Р. Эволюция разума. С. 54.

¹¹ Там же. С. 285.

¹² См.: Там же. С. 405.

и не выражают его атрибутивного качества. Пространство натурального ума неоднородно, в силу чего он и самобытен. В нем возникают порой такие смысловые композиции, для которых принципиально невозможны аналогии и параллели. Отсюда неустранимое недопонимание между людьми (отнодь не однозначно негативное, впрочем: всецело понятое грозит сделаться безразличным нам). Но и отсюда же способность натурального ума справляться с так называемыми неразрешимыми задачами.

Соотнося аналогию и оцифровку в контексте канторовской теории множеств, любопытный тезис формулирует Ник Ланд: «Перевод из аналога в цифру стирает информацию. Вкрадывается хаос»¹³. С замечанием британского философа-акселерациониста отчасти можно согласиться, уточнив, правда, несколько моментов. Во-первых, в результате очерченного трансфера теряется не столько информация, сколько смыслы, несводимые к информации. Во-вторых, упомянутый хаос не онтологический (тот порой может быть продуктивным), а гносеологический (этот без шансов на продуктивность). Онтологический хаос (как мощный, так и скудный) не познается адекватно ни упорядоченной, ни беспорядочной мыслью. Здесь требуется неординарная – уходящая от крайностей бинаризма, от механически перебегающих единиц и нулей – собственно человеческая мысль. В-третьих, гносеологический хаос обусловлен прежде всего разъединением рационального и иррационального в человеческой психике. Слепящий свет и непроницаемая тьма одинаково удручающи. Они не оставляют места здоровой иррациональности, без которой невозможно понимание – высший, не сугубо рациональный и не безумный уровень осмысления. Если я захочу кого-то или что-то понять, быть может, и пойму; не захочу – не пойму никогда и ни за что. Именно здоровой иррациональностью наполнено ментальное пространство, благоволящее к разрешению нетривиальных задач, для которых математики устанавливают единственно правильное решение и одновременно невозможность его отыскания.

Против решений чисто гипотетических, допустимых в каком-то одном (в некоторых) из множества возможных миров, активно выступал математик-интуитивист Курт Гёдель. В его философских воззрениях платонизм сопрягается с витализмом¹⁴; Яакко Хинтика усматривает в них также и концептуальные следы «вечного возвращения»¹⁵. Интуитивизм Гёделя – воззрения людей невыводимы полностью из формулируемых ими аксиом – прекрасный пример здоровой иррациональности, недвусмысленно заключающей, что «человеческий ум (даже в области чистой математики) бесконечно превосходит возможности любой конечной машины...»¹⁶.

Согласно Гёделю (в книге Курцвейла он упоминается, но гораздо большее внимание в ней уделяется, разумеется, Алану Тьюрингу и его «машине»), исчерпывающий ответ не может не быть противоречивым. Утверждение о непротиворечивости определенной системы аксиом и правил «не является выводимым из этих аксиом и правил при условии, что эти аксиомы и правила непротиворечивы»¹⁷. Натурный ум, не чуждый диалектике

¹³ Ланд Н. Сочинения: в 6 т. Т. 2: Киберготика. Пермь, 2018. С. 105.

¹⁴ См.: Гёдель К. Некоторые основные теоремы в основаниях математики и их следствия // Хинтика Я. О Гёделе. М., 2014. С. 166–200.

¹⁵ См.: Хинтика Я. О Гёделе. М., 2014. С. 80.

¹⁶ Гёдель К. Некоторые основные теоремы в основаниях математики и их следствия. С. 177.

¹⁷ Там же. С. 174.

(Гёдель с ориентацией на континентально-европейский реализм много ближе к диалектике, чем Тьюринг с характерным для британской интеллектуальной традиции номинализмом), не станет убегать от противоречий. Он постарается их осмыслить. И хотя нам никак не удастся воспользоваться здесь универсальным алгоритмом (его просто не существует в неоднородной среде), мы нередко достигаем желаемого. Когда, скажем, безусловно мыслим речку, одновременно движущуюся и недвижущуюся. Да, эта безупречность ограничена мгновением... но сколько будет длиться мгновение, никто не знает. Мгновенно исчерпывающий ответ мы обретаем тогда, когда не отделяем идею от восприятия и представления и не разбиваем ни интеллигибельное, ни чувственное на отдельные, тяготеющие к разрозненности штрихи. Противное означало бы экстраполяцию машинной специфики на человеческое сознание, к чему поступательно склоняет нас Курцвейл, связывающий нашу способность к распознаванию образов, в частности, с опорой на инвариантные фрагменты рисунка, не изменяющиеся от случая к случаю¹⁸. Выбирая такой путь, мы неизбежно, утратив образ, упрямся в конце концов в точки и пробелы. Смысл тем и отличается во многом от информации, что он не раскладывается на единицы. Он не делится без значимых потерь и ни на что другое. «Разделяя» чьи-то взгляды, мы ставим и себя, и того другого в двусмысленное положение. Фактически мы расписываемся в неполноте постигнутого нами и в нежелании мыслить глубже, обходясь заведомо упрощенными интерпретациями. «Не разделять» взгляды, идеи, чувства означает в одном случае не принимать их (на основании того, что в них нет ничего, с чем соглашаешься безоговорочно), однако, хотя и неприятные, они могут быть поняты нами достаточно глубоко; в другом случае, – наоборот, принимать (если в них есть хоть что-то безоговорочно значимое для тебя, пусть с чем-то в них заведомо и не согласен), устремляясь тем самым к предельной глубине понимания. Неразделенная любовь, в отличие от неразделенного, отринутного знания, способна остаться в нас во всей своей полноте. Ее бессмысленно крепить биохимическими инъекциями. А любовь разделенная, разложенная Курцвейлом на нейромедиаторы и их активность¹⁹, от подобного допинга, наверное, не откажется... превращаясь в вариацию борьбы или бегства.

Редукция высших психических, экзистенциальных способностей и состояний к электрохимическим реакциям служит, в логике Курцвейла, наглядной иллюстрацией «избыточности» мозга. Указание на его «сверхмерность» удобно, конечно, для стратегии сочетания человеческого разума с «оптимальным» искусственным интеллектом (оптимальность и копирует избыточность, и делает ее эффективней). Но действительно ли наш мозг сверхмерен? Нет! В нем нет ничего лишнего, как нет при отсутствии серьезных травм и недостатков. Да, эволюция человеческого мозга практически завершена, что, однако, не является веским основанием для его искусственного усовершенствования. Из того, что мы актуализируем только часть наличествующего потенциала, вытекает не избыточность нашего мозга, а его... человечность. Полнота актуализации стала бы катастрофой.

Знать всё, или хотя бы помнить всё – непосильно, тягостно. Правда подобное состояние и недостижимо, к счастью, для человеческого сознания

¹⁸ См.: Курцвейл Р. Эволюция разума. С. 46.

¹⁹ См.: Там же. С. 172–176.

ни на индивидуальном уровне, ни на коллективном. Помнить всё означало бы держать в уме не только все прежние смыслы и их оттенки, но и все возможные смысловые конфигурации. Каким бы вместительным ни был резервуар памяти, он в случае всепамятования моментально заполняется до краев и никаким новым смыслам в сознание человека уже не вместится. Мышление тем самым обрывается не начавшись. Если всё же мы удовлетворяемся ситуацией с неприбывающими смыслами, т.е. оказываемся в формате вечности или ее аналога, то сталкиваемся либо с непоколебимо-однозначной концептуальной конфигурацией (и тут мыслить незачем – надо просто внимать), либо с вереницей индифферентно-равнозначных концептуальных конфигураций (и тут мышление избыточно: достаточно элементарного чередования готовых матриц). Так или иначе, всеобъемлющая память обрекает нас на примитивное в интеллектуальном плане существование: не допуская интенциональности сознания, лишая его явных и неявных предпочтений, она делает невозможным ранжирование смысловых композиций, а стало быть, невозможным становится и понимание, ведь ему как минимум никак не обойтись без иерархически выстраиваемой демаркации противоречиво-истинного и противоречиво-ложного. Понимание, да и собственно мышление (по преимуществу рациональный пласт понимания), коренится в небеспамятном, разумеется, сознании, но в таком, которому присуще и забывание. Память с лакунами поддерживает или даже запускает процесс мышления и вместе с тем предопределяет невозможность его абсолютистского итога, полного знания обо всем. И замечательно! Ценность, и немалая, человеческого ума – в способности приближаться к полноте знания о чем-то, фиксируя это приближение в истинно-противоречивой мысли. Величие человеческого сознания – в способности понимать то, что до конца не познано.

Очерченная выше позиция позволяет, в частности, категорически не принимать и «не разделять» пропагандируемую Курцвейлом модель сознания, аргументированно выделяя при этом в ней главное и второстепенное. Оппонент уверяет, что человек мыслит линейно, а компьютер научит его мыслить экспоненциально – как будто экспоненциальная кривая не линия. Нет, человек (тот, кто не подстраивается под компьютерный шаблон) мыслит качественно иначе – объемно: так же, как и чувствует, и вместе с чувствованием. Восприятие музыкального произведения в концертном зале не равносильно для нас восприятию тех же мелодий посредством технических стереосистем, имитирующих трехмерность. И не в предвидении будущего, на что упирает Курцвейл, заключена важнейшая особенность нашего ума. Нет, любая мысль о будущем продумывается в настоящем. Именно осмысление настоящего (недаром оно синонимично подлинному) важнее и труднее всего, что фиксирует даже «линейная» логика Аврелия Августина. Логика Курцвейла линейна без оговорок, тенденциозно исчисляющего характера. Ее квинтэссенция – в утверждении о том, что «самые важные эволюционные достижения *Homo sapiens* были количественными...»²⁰. Налицо примитивная редукция качества к количеству, от которой не спасает и возведение исчисляемых величин в степень (согласно курцвейловскому «закону ускорения отдачи»). Вновь подчеркну: никакой, даже порядковый, количественный прирост не гарантирует изменения качества, а тем более качественно-го роста.

²⁰ Курцвейл Р. Эволюция разума. С. 415.

Наряду с «не разделяю и отвергаю» мы вольны, как уже говорилось, «не разделять, но принимать» – и не по причине внешнего насилия, не угодая ложной авторитарности, не вследствие, стало быть, своей бесхребетности, а, напротив, со всей личной ответственностью беря всё на себя, несмотря на то, что соглашаешься далеко не со всем вбираемым. «Умной машине» это не под силу. А у меня так, к примеру, и с Ф.М. Достоевским (хотя религии не симпатизирую), и с М. Хайдеггером (хотя никогда не поддерживаю тезис о том, что бытие не есть сущее). Чтобы обрести подлинное, понимающее мышление, «мы должны сначала прыгнуть на почву, на которой мы, собственно, стоим»²¹ – сногсшибательное противоречие. И далее еще одно: подлинное мышление вращается в стихии многозначности – притом что какая-то мысль заслуживает того, чтобы быть одной-единственной, которая по-настоящему обдумывается²². Не прогресс, а вечное возвращение. Не линия, как и сколь угодно гнущаяся, а объемный пульсар.

Неполнота актуализации возможностей мозга, открывая перед человеком творческие перспективы, обуславливает «лишь» мгновенность истины (без притязаний на вневременность) и насыщенность мышления противоречиями, причем отнюдь не все они истинным образом ориентированы. Осмысление противоречий включает в себя выстраивание надежной иерархии. В том числе соотносящей в психике интеллигибельное и волевое. То есть важно не столько то, что одна наша идея может противоречить другой, сколько то, что любую из них, как и возможный их мыслительный синтез, мы вольны как принять, так и не принять. Важен, к примеру, не сам ответ на вопрос, умирать в конце концов человеку или не умирать, а то, отвергается ответ или не отвергается. В аутентично человеческом мышлении атрибутивен зазор между получением и принятием/отклонением смыслов. И это не сугубо временной зазор. Отнюдь не всегда принятие или отклонение происходит после обнаружения смысла. До его обнаружения какие-то процедуры с ним, конечно, невозможны вовсе. Одновременность же всегда чревата вневременным. Никаким увеличением скорости операций, материально отечающих за процесс мышления, деятельность натурального ума кардинально не изменить.

Искусственный интеллект не терпит ни противоречивой дистанции между получением информации и ее принятием/непринятием, ни противоречий как таковых. Симптоматична ремарка Курцвейла: существенное ограничение новой коры головного мозга «заключается в том, что в ней нет механизма, устраняющего или хотя бы оценивающего противоречащие друг другу данные, отчасти это объясняет весьма распространенную нелогичность человеческих рассуждений. <...> В компьютерной новой коре можно предусмотреть процесс, выявляющий противоречивые данные для их последующего пересмотра»²³. На ликвидацию противоречий и направлены – безнадежно, есть уверенность – ускоренные машинные исчисления. Теоретическая сверхзадача ИИ – сверхскоростным методом обеспечить решение математически неразрешимых задач – заведомо невыполнима. Компьютерная отбраковка противоречий позволит в лучшем случае лишь асимптотически

²¹ Хайдеггер М. Что зовется мышлением? М., 2007. С. 53.

²² См.: Там же. С. 60, 126.

²³ Курцвейл Р. Эволюция разума. С. 293.

приближаться к исчерпывающему решению, но не достигать его. Параллельные линии вот-вот сойдутся, но каждый раз что-то идет не так...

Поклонение скорости – одна из самых отличительных черт теории искусственного интеллекта и ее эмпирической реализации. Такого рода культ напрямую связан с превознесением времени: чем выше скорость, тем большее расстояние покоряется за одно и то же время. С точки зрения Курцвейла, основа мироздания – информация. Мировой разум философов-панлогистов дробится им на элементарные составляющие, которые наделяются способностью вступать в многочисленные, но поддающиеся типологическим упрощениям комбинации. Чем быстрее происходит выстраивание комбинаторных структур, тем скорее несовершенная поначалу система «информация – материя» превратится в высокоорганизованный конвергентный компьютер. Превращение скорости в фетиш включает в себя редукцию всегда содержательного пути к бессодержательному расстоянию, что корректно лишь в простейших случаях и только в однородной среде. В неоднородной среде подобное сведение неминуемо обернется неулаживаемыми погрешностями.

Для натурального человеческого ума, не спешащего попусту и не озабоченного изгнанием противоречий, теоретически неразрешимых задач не существует. Контекстуально все они решаемы. Ни одна задача не является абсолютно внеконтекстуальной. Те, которые приближаются по характеру к внеконтекстуальным, имеют предрешение, никаким прогрессом не отменяемое (как в случае со смертностью человека). И тут, повторяю, важно не столько само предрешение, сколько остающееся неустранимо противоречивым его принятие или непринятие человеком.

Проект Курцвейла и K° – подойдя к точке сингулярности, обрести нескончаемое физико-техническое существование (жизнью его не назвать) – структурообразующий элемент сегодняшней прогрессистской конъюнктуры. Это эпицентр информационной экспансии, умаляющей человеческие смыслы. Всё равно, принимается ли проект за чистую монету или за PR-акцию ради продажи новых компьютерных разработок. Религиозный человек его отвергнет, думая вместе с И.И. Тхоржевским о совершенно ином: «...Легкой жизни я просил у Бога. / Легкой смерти надо бы просить». Натурный человек светской культуры, поняв просьбу, от себя скажет: «Одна из главных жизненных удач – умереть вовремя».

Список литературы

- Бостром Н. Искусственный интеллект. Этапы. Угрозы. Стратегии / Пер. с англ. С. Филина. М.: Манн, Иванов и Фарбер, 2016. 496 с.
- Гёдель К. Некоторые основные теоремы в основаниях математики и их следствия / Пер. с англ. В.В. Целищева // Хинтиikka Я. О Гёделе. М.: Канон+, 2014. С. 166–200.
- Достоевский Ф.М. Собрание сочинений: в 15 т. Т. 9: Братья Карамазовы. Л.: Наука, 1991. 699 с.
- Курцвейл Р. Эволюция разума: как развитие искусственного интеллекта изменит будущее цивилизации / Пер. с англ. Т.П. Мосоловой. М.: Эксмо, 2020. 448 с.
- Ланд Н. Сочинения: в 6 т. Т. 2: Киберготика / Пер. с англ. Д.Я. Хамис. Пермь: Гиле Пресс, 2018. 208 с.
- Матурана У., Варела Ф. Древо познания / Пер. с англ. Ю.А. Данилина. М.: Прогресс-Традиция, 2011. 224 с.

- Пайн К. Живи, вкальвай, сдохни. Репортаж с темной стороны Кремниевой долины / Пер. с англ. К. Казбек, И. Мельникова. М.: Индивидуум, 2019. 304 с.
- Полани М. Личностное знание. На пути к посткритической философии / Пер. с англ. М.Б. Гнедовского, Н.М. Смирновой, Б.А. Старостина; под общ. ред. В.А. Лекторского, В.И. Аршинова. М.: Прогресс, 1985. 344 с.
- Фатенков А.Н. К онтологическим истокам идеи подвижной иерархии // Вопросы философии. 2008. № 5. С. 47–57.
- Фатенков А.Н. С цифровыми технологиями – в бесчеловечное будущее // Юридическая наука и практика: Вестник Нижегородской академии МВД России. 2020. № 1 (49). С. 217–219.
- Хайдеггер М. Что зовется мышлением? / Пер. с нем. Э. Сагетдинова. М.: Академический Проект, 2007. 351 с.
- Хинтиikka Я. О Гёделе / Пер. с англ. В.В. Целищева и В.А. Суровцева. М.: Канон+, 2014. 256 с.

Natural intelligence in a counterattack against artificial intelligence (a polemical response to “How to Create a Mind” by Ray Kurzweil)

Aleksey N. Fatenkov

Lobachevsky State University of Nizhny Novgorod. 23 Gagarin Ave., Nizhny Novgorod, 603022, Russian Federation; Privolzhsky Research Medical University. 10/1 Minin Sq., Nizhny Novgorod, 603005, Russian Federation; e-mail: fatenkov@fsn.unn.ru

By turning to geometric metaphors, the author defends the naturally based intelligence and – in the context of analysing Ray Kurzweil’s conception – criticises the artificial intelligence. The focus is on the ability of man and machine to solve the so-called unsolvable problems. This issue is discussed in conjunction with the issues of hierarchy, homogeneity and heterogeneity, contradiction and analogy, meaning and information. Arguments are given to support the following ideas: 1) digital technologies do not overcome the fundamental limitations of analogue technologies, thus AI capabilities are also irreparably limited; 2) the natural intelligence works in an original way (not by analogy), so there can be no insoluble theoretical problems for it; 3) in a healthy human brain, there is neither excess nor deficiency, it does not need artificial improvement; 4) incomplete actualisation of the brain potential is not a weakness, but a strength and special feature of the human mind; 5) the theory of artificial intelligence is characterised by excessive reductionism and fetishization of speed.

Keywords: natural intelligence, artificial intelligence, R. Kurzweil, homogeneity, heterogeneity, analogy, contradiction, hierarchy, meaning, information

For citation: Fatenkov, A.N. “*Naturnyi um v kontratake na iskusstvennyi intellekt (polemicheskii otklik na ‘Evolutsiyu razuma’ Reya Kurtsveila)*” [Natural intelligence in a counterattack against artificial intelligence (a polemical response to “How to Create a Mind” by Ray Kurzweil)], *Filosofskii zhurnal / Philosophy Journal*, 2022, Vol. 15, No. 3, pp. 172–183. (In Russian)

References

- Bostrom, N. *Iskusstvennyi intellekt. Etapy. Ugrozy. Strategii* [Superintelligence. Paths, Dangers, Strategies], trans. by S. Filin. Moscow: Mann, Ivanov i Farber Publ., 2016. 496 pp. (In Russian)
- Dostoevsky, F.M. *Sobranie sochinenii, T. 9: Brat’ya Karamazovy* [Works, Vol. 9: The Brothers Karamazov]. Leningrad: Nauka Publ., 1991. 699 pp. (In Russian)

- Fatenkov, A.N. “K ontologicheskim istokam idei podvizhnoi ierarkhii” [To Ontologic Sources of Idea of Mobile Hierarchy], *Voprosy filosofii*, 2008, No. 5, pp. 47–57. (In Russian)
- Fatenkov, A.N. “S tsifrovymi tekhnologiyami – v beschelovechnoe budushchee” [With Digital Technologies – into an Inhuman Future], *Yuridicheskaya nauka i praktika: Vestnik Nizhegorodskoi akademii MVD Rossii*, 2020, No. 1 (49), pp. 217–219. (In Russian)
- Gödel, K. “Nekotorye osnovnye teoremy v osnovaniyakh matematiki i ikh sledstviya” [Some Basic: Theorems on the Foundation of Mathematics and their Implications], trans. by V.V. Tselishchev, in: J. Hintikka, *O Gedele* [On Gödel]. Moscow: Kanon+ Publ., 2014, pp. 166–200. (In Russian)
- Heidegger, M. *Chto zovetsya myshleniem?* [Was heisst Denken?], trans. by E. Sagetdinov. Moscow: Akademicheskii Proekt Publ., 2007. 351 pp. (In Russian)
- Hintikka, J. *O Gedele* [On Gödel], trans. by V.V. Tselishchev and V.A. Surovtsev. Moscow: Kanon+ Publ.; Reabilitatsiya Publ., 2014. 256 pp. (In Russian)
- Kurzweil, R. *Evolyutsiya razuma: kak razvitie iskusstvennogo intellekta izmenit budushchee tsivilizatsii* [How to Create a Mind: The Secret of Human Thought Revealed], trans. by T.P. Mosolova. Moscow: Eksmo Publ., 2020. 448 pp. (In Russian)
- Land, N. *Sochineniya, T. 2: Kibergotika* [Works, Vol. 2: CyberGothic], trans. by D.Ya. Khamis. Perm': Gile Press Publ., 2018. 208 pp. (In Russian)
- Maturana, H.R. & Varela, F.J. *Drevo poznaniya* [The Tree of Knowledge: The Biological Roots of Human Understanding], trans. by Yu.A. Danilin. Moscow: Progress-Traditsiya Publ., 2011. 224 pp. (In Russian)
- Pein, C. *Zhivi, vkalyvai, sdokhni. Reportazh s temnoi storony Kremnievoi doliny* [Live Work Work Work Die: A Journey into the Savage Heart of Silicon Valle], trans. by K. Kazbek and I. Melnikov. Moscow: Individuum Publ., 2019. 304 pp. (In Russian)
- Polanyi, M. *Lichnostnoe znanie. Na puti k postkriticheskoi filosofii* [Personal Knowledge: Towards a Post-Critical Philosophy], trans. by M.B. Gnedovsky, N.M. Smirnova and B.A. Starostin; ed. by V.A. Lektorsky and V.I. Arshinov. Moscow: Progress Publ., 1985. 344 pp. (In Russian)