

11. Иванов В.И., Барат В.А. Акустико-эмиссионная диагностика. М.: Изд-во «Спектр». - 2017. 368 с.

12. Фундаментальные и прикладные проблемы комплексной диагностики. Многотомная серия «Безопасность России». М.: МГОФ «Знание». - 2017. 902 с.

Сведения об авторах

Николай Андреевич Махутов - проф., д.т.н., главный научный сотрудник ООО «НИИ Транснефть», Россия, Москва, 117186, Севастопольский проспект, д. 47а. Тел. факс – 8 (495) 9308078, e-mail: kei51@mail.ru

Валерий Иванович Иванов – проф., главный научный сотрудник ЗАО «НИИИИ МНПО Спектр», Россия, Москва, 119048, ул. Усачева, д. 35, стр. 1 Тел. – 8 (499) 1355211, e-mail: ivi444@mail.ru

Игорь Евгеньевич Васильев – старший научный сотрудник Института машиноведения РАН, Россия, Москва, 101990, Малый Харитоньевский пер., д. 4. Тел. – 8 (499) 1355211, e-mail: vie01@rambler.ru

Дмитрий Витальевич Чернов – младший научный сотрудник Института машиноведения РАН, Россия, Москва, 101990, Малый Харитоньевский пер., д. 4. Тел. – 8 (499) 1355211, e-mail: chernovdv@inbox.ru

УДК:004.896: 51-74

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ: ПРОБЛЕМЫ, ПЕРСПЕКТИВЫ, РИСКИ

Доктор техн. наук *Р.А. Дурнев*, кандидат психол. наук *К.Ю. Крюков*,
кандидат техн. наук *А.Е. Титов*

ФГБУ РАРАН

И.В. Жданенко

ФГБУ ВНИИ ГОЧС

Рассмотрено состояние развития слабого и сильного искусственного интеллекта. Определено, что методы слабого искусственного интеллекта являются частью автоматизации и применяются только в алгоритмизированном пространстве. Сильный искусственный интеллект должен решать творческие задачи в неалгоритмизированном пространстве. Установлены основные проблемы и ограничения, перспективы реализации и риски направлений создания сильного искусственного интеллекта.

Ключевые слова: искусственный интеллект, алгоритм, автоматизация, мышление, неокортекс, нейроны, биокибернетические системы.

ARTIFICIAL INTELLIGENCE: PROBLEMS, PROSPECTS, RISKS

Dr. (Tech.) *R. Durnev*, Ph.D. (Psychol.) *K. Kryukov*, Ph.D. (Tech) *A.E. Titov*
Federal State Budgetary Institution "Russian Academy of Rocket and Artillery Sciences"

I. Zhdanenko

All-Russian Research Institute for Civil Defense and Emergencies

The state of development of weak and strong artificial intelligence is considered. It is determined that the methods of weak artificial intelligence are part of automation and are used only in algorithmic space. Strong artificial intelligence should solve creative problems in non-algorithmic space. The main problems and limitations, the prospects for implementation and the risks of creating strong artificial intelligence are established.

Keywords: artificial intelligence, algorithm, automation, thinking, neocortex, neurons, bio-cybernetic systems.

Модная, в очередной раз, идея искусственного интеллекта (ИИ) находит всё больше сторонников в среде журналистов и ученых, обывателей и управленцев, военных и людей мирных профессий. Часто приводятся многочисленные рассуждения о проблемах и перспективах его применения в различных областях, выгодах и рисках, с ним связанных. При этом, как и с понятиями "жизнь", "знание", "управление" и т.п., "искусственному интеллекту" трудно дать конструктивное определение. Это связано, в том числе и с тем, что составляющие его части "искусственный", "интеллект", а также близкие "разум", "мыслительные процессы" и т.п. отданы на откуп философии. А философы в затруднении в связи с дефицитом знаний о мозге и когнитивном (познавательном) аппарате биологических систем. Поэтому и существует огромное количество определений ИИ, ни одно из которых не является общепринятым. Один из апологетов ИИ Марвин Мински называл такие понятия "словами-чемоданами" со многими смыслами. Нам больше нравится аналогия с облаком, которое на расстоянии имеет четкие очертания, а вблизи - туман.

Но несмотря на это, многочисленные разработчики заявляют об интеллектуальных гаджетах, бытовой технике, транспортных системах, комплексах вооружения. Насколько это правомерно и как к этому относиться - мы попробуем разобраться в этой статье.

Первое, о чем нужно сказать - в основном всё, что сейчас разрабатывается в рассматриваемой области - это только "слабый ИИ". Слабым его назвали потому, что, с одной стороны, он далеко "не дотягивает" до человеческого, но с другой - это не вполне обычное программное обеспечение на компьютере. Хотя большинство задач, решаемых им - это оптимизация решений (нахождение наилучших в каком-то смысле), классификация объектов (разбиение множества на классы в соответствии с критериями), экстраполяция данных (нахождение промежуточных значений) и для этого существуют традиционные методы. Но сами алгоритмы "слабого ИИ" приобрели антропоморфные (очеловеченные) черты. Они либо копировали биологические процессы (например, сеть нейронов в мозгу человека, как у Мак Каллока-Питса, или эволюционные алгоритмы, как у Д. Холланда), либо позволяли оперировать с рассуждениями на естественном языке (например, теория нечётких множеств Л.Заде). И хотя процесс получения результатов стал более эффективным, но их основу составляли те же алгоритмы, которые возможно запрограммировать на компьютере в виде последовательности нулей и единиц. Недаром многие специалисты говорят о том, что методы "слабого ИИ" - это не "подглядывание" за природой, а формализмы машинного обучения [1].

Поэтому методы "слабого ИИ" работают только в ситуациях, которые входят в алгоритм. Даже глубокое обучение нейросетей - это просто отладка и коррекция алгоритма. И в целом все методы "слабого ИИ", наряду с исследованием операций, системным анализом, теорией принятия решений, кибернетикой, являются составляющими автоматизации, т.е. процессов, выполняемых без участия человека.

Нет смысла говорить о тех успехах, которые достигла автоматизация - почти всё вокруг нас либо работает в автоматизированном режиме либо создано автоматизированным способом. Однако у автоматизации есть принципиальная особенность (судя по её успехам - не хочется говорить недостаток) - она не работает в неалгоритмизированном пространстве, в таком, где четко не определены все шаги и их порядок. Поэтому число ситуаций, доступных ей, ограничено, и для того, чтобы осваивать всё их бесконечное разнообразие - автоматизации нужно беспредельно долго развиваться.

Автоматизации не доступны не только многие ситуации, но и различные действия, связанными с ними. Например, основной проблемой автоматизации военного управления, наряду с разнородностью, слабой согласованностью и взаимосвязанностью различных её звеньев, является то, что современными средствами практически невозможно исключить человека в сугубо творческих процессах планирования действий и тем более формирования целей действий, целеполагании.

Методы "слабого ИИ" (принадлежащие, как мы определили, к автоматизации) существуют без серьезных изменений уже много лет. Даже популярные сегодня в техническом зрении сверточные нейросети были предложены тридцать лет назад. И весь современный бум их использования связан не столько с научными достижениями, сколько с инновационными технологиями, т.е. коммерциализацией ранее полученного знания. А сама эта коммерциализация является следствием закона Мура, в соответствии, с распространенной трактовкой которого каждые два года (это верно для сегодняшнего времени) в два раза увеличивается производительность компьютеров.

Если говорить более корректно, то речь идет не столько о росте производительности, сколько о миниатюризации микроэлектроники, о росте количества транзисторов на кремниевой микросхеме заданных размеров. Но так как из года в год мы являемся свидетелями того, что по размерам наши персональные ЭВМ практически не меняются (или меняются не так значительно - не на порядки, что связано, в том числе и с эргономикой), а их характеристики постоянно улучшаются в разы, то и возникает иллюзия роста только производительности.

Аналогично и для методов "слабого ИИ" весь прогресс связан, в основном, с ростом возможностей микроэлектроники. Так, появление на рынке компьютерных игр мощных видеокарт подвигло специалистов по нейросетям к их использованию в своих нуждах. Это определялось тем, что в основе нейросетевых технологий лежат матричные вычисления - необходимо постоянно перемножать и складывать огромные таблицы сигналов, приходящих к нейронам, их весов, ошибок и других элементов. И если для обычных центральных процессоров это достаточно трудоемкие операции, то для графических процессоров, изначально работающих в матричном пространстве пикселей, это "родная" задача.

Еще одна иллюзия прогресса "слабого ИИ" связана с увеличением компактных устройств, наделенных элементами «разумного поведения» (гаджеты, «умная» бытовая техника, системы навигации и управления современных автомобилей). Причиной этого является увеличение их мощности до таких величин, которая раньше была доступна только большим ЭВМ, на которых отрабатывались ранее разработанные алгоритмы.

Но даже эмпирический закон Мура, продолжающий "выкачивать" из кремния все возможности, уже не так очевиден. К примеру, возникновение "многоядерности" процессоров говорит о том, что дальнейшая миниатюризация микроэлектроники близка к пределу. Аналогично, недалеко от естественной границы и длина волны ультрафиолетового излучения, с помощью которого "наносят трафарет" на кремниевой пластине. Чем меньше эта длина, тем меньше получаются транзисторы. Так вот, используется практически

минимально-возможная длина волны, дальше свет перейдет в рентгеновский или гамма-диапазон. А в этом диапазоне уже меняются свойства кремния, делающие невозможным его использование в микроэлектронике.

Еще более категоричные доводы сводятся к тому, что в ближайшее время один из размеров транзистора на кремнии будет составлять всего 10-20 атомов. А в этих условиях наибольшую роль будут играть уже квантово-механические эффекты, связанные с различными неопределенностями (в траектории электронов, их импульсе, вероятности нахождения в определенной области и др.), которые невозможно будет разрешить на современном уровне развития науки и техники. Поэтому очередное крутое развитие "слабого ИИ" (а значит и автоматизации) в недалёкой перспективе сменится более пологими изменениями.

Мы достаточно подробно поговорили об очередной грани автоматизации - "слабом ИИ", а как же быть с "сильным ИИ"? Ведь именно он и должен решать творческие интеллектуальные задачи, являющиеся прерогативой человека, работать в неалгоритмизированном пространстве, с различными неформализованными ситуациями.

Для решения таких задач разработка "сильного ИИ" может вестись по следующим направлениям:

разработка ИИ в виде программ на ЭВМ;

усовершенствование человека с целью повышения его когнитивных способностей;

создание биокибернетических систем, служащих подспорьем человеку в его мыслительных процессах.

Для реализации первого направления активно разрабатываются структурные модели мозга, в основном коры его больших полушарий, так называемого "неокортекса". В этих моделях, часто именуемых "коннектомами", делаются попытки определить место всех нейронов (а их у человека около 100 млрд.) со всеми их связями (порядка 10-15 тыс. на один нейрон). Для этого сейчас применяются методы магнито-резонансной и позитронно-эмиссионной томографии, электроэнцефалографии, делаются тончайшие срезы мозга, изучаемые под микроскопами. В будущем планируется цифровизация нейронов и их связей с помощью нанороботов и даже нейрохирургическая замена нейронов в мозгу транзисторами в компьютере. Несмотря на масштабность таких проектов, громкие заявления исследователей серьезных успехов пока не наблюдается. "Коннектомы" некоторых простейших существ известны уже много лет. Так, структурная модель мозга нематоды - маленького полупрозрачного червячка - существует уже около 30 лет. В ней учтены все 302 нейрона этого существа со всеми их связями [2]. Однако за такой период ученые так и не научились "оживлять" эту модель, которая больше напоминает словарь без значений слов.

Еще в большей степени это относится к мозгу более сложных животных, а тем более к мозгу человека. Даже простая его структуризация потребует десятков лет кропотливого труда, не говоря уже об использовании такой гигантской модели для имитации сложных мыслительных процессов.

Другим путем создания "сильного ИИ" в виде программы на ЭВМ является разработка функциональных моделей мозга. Человечество уже давно проявляет огромный интерес к своей мыслительной деятельности. И за неимением технологий (диагностических, сканирующих, компьютерных и других) исследователи создавали то, что им было доступно - мысленные модели мыслительных процессов, адекватность которых в более позднее время стали проверять различными психо-физиологическими экспериментами, тестами, опросами и т.п. И таких моделей было разработано огромное количество практически во всех сферах человеческой деятельности. Нет смысла приводить их даже самую грубую классификацию, это будет многотомное издание. Из последних таких моделей наиболее системной, на наш взгляд, является запоминающе-прогностическая модель Дж. Хоккинса [3], в которой рассматриваются алгоритмы восприятия сенсорных сигналов, способы запоминания окружающего мира, принципы формирования прогнозов мышления и действий и даже то, что отличает человека от животного - интроспекция или самоанализ, т.е.

осознание самого себя, своих собственных мыслей. Об этой модели упоминает такой известный футуролог, специалист по ИИ, как Рэй Курцвейл [4]. На рубеже 2010 года он говорил о том, что данная модель переводится в машинный код и её планируется запустить на компьютере. Но после этого никакой информации, ни об успехах этого процесса, ни о его неудачах, к сожалению, не было.

В целом, говоря о создании "сильного ИИ" в виде программы на ЭВМ следует отметить, одной из проблем являются вычислительные ограничения при решении даже слабо интеллектуальных задач, таких, какие легко решают маленькие дети. Так, для распознавания кошачьих мордочек Cat Detector от Google использовал 1000 серверов с 16 000 ядер [5]. "Умный" вертолет, который умел бы выполнять эту крайне узкую функцию, с таким набором аппаратуры не смог бы подняться в воздух. Аналогично трудно себе представить компактное устройство (например, андроидный робот), который может играть с человеком в шахматы или игру Го. В первом случае в этот робот пришлось бы запихнуть несколько шкафов суперкомпьютера Deep Blue компании IBM, выигравшего у Г. Каспарова в 1997 г., во втором - 1920 CPU (центральных процессоров) и 280 GPU (графических процессоров), победивших Ли Седоля в 2016 году, и это практически на предельных возможностях кремния! При этом и компьютеры и соответствующее программное обеспечение узкоспециализированы - умеют только играть в свои дискретные игры с полной информацией, игры, для которых важна только текущая позиция и не значима предыстория. Для них крайне эффективны различные переборные методы с дополнительными правилами, как эти переборы сокращать. В более сложных играх, например, в покер, уже требуется память ходов, ставок, возможностей игроков, да ещё и умение блефовать, поэтому здесь ИИ не так силен.

Для логического вывода и доказательства математических утверждений также нет алгоритмов существенно лучше полного перебора возможных логических цепочек. Но если для доказательства теоремы, состоящего из 5 строк, требуется перебрать не так уж и много вариантов, то для 50 строк этих переборов нужно сделать столько, что может не хватить и всего времени существования Вселенной [6].

Ведущие игроки микроэлектроники (Intel, AMD и др.) давно уже озабочены "пределом кремния", поэтому, активно занимаются поиском путей решения этой проблемы. Это создание трехмерных транзисторов, использование не-кремниевых материалов (например, графена), развитие обратимых и квантовых вычислений и т.п. Но все эти пути, как это часто бывает при эволюционном развитии техники, помимо очевидных преимуществ, например в быстродействии, обладают и серьезными недостатками. Так, трехмерные транзисторы, работают как обогревательные приборы, не-кремниевые материалы чувствительны к параметрам тока, квантовые вычисления критичны к нарушению когерентности состояний элементарных частиц.

Но даже если указанные недостатки будут преодолены, то все равно через некоторое время мы подойдем к пределу трехмерных транзисторов, графена или нанотрубок и опять начнется поиск очередных путей повышения производительности или более точно - соблюдение массо-габаритных параметров компьютеров, соответствующих нуждам человека. Ну а если при наступлении пределов каких-то материалов не будут найдены приемлемые решения - компьютеры будут увеличивать свою массу, расти вширь, все больше потреблять энергии. А так как человек решает все более сложные задачи и в своем любопытстве никогда себя не ограничит, вычислительные машины, для вычисления невообразимого числа знаков после запятой в числе пи, могут, по шутливо-серьезному сценарию того же Р. Курцвейла, поглотить энергию всей Вселенной...

Но, пожалуй даже это не самое главное - все указанные разработки будущего (трехмерные транзисторы, графен и др.) работают только с алгоритмами, состоящими из нулей и единиц или большего числа таких состояний в квантовом компьютере (за счет различных сочетаний нулей и единиц, их суперпозиций). И именно алгоритмы (вернее - их предельность) и являются камнем преткновения на пути создания "сильного ИИ".

В общем смысле под алгоритмом понимается конечный набор упорядоченных шагов. Следует обратить внимание на то, что хотя набор шагов и конечен, но процесс их выполнения может не иметь предела (так называемая проблема "останова") или просто непонятно - возможно ли получить данным алгоритмом результат за конечное время? В качестве примера можно привести алгоритм, пытающийся доказать, что все чётные числа являются суммой двух простых (которые делятся без остатка только на себя или единицу). Не понятно, завершит ли компьютер когда-нибудь такое доказательство?

Если решение задачи может быть выражено в виде алгоритма, она является разрешимой (вычислимой). Именно для таких задач и была разработана концепция универсальной машины Тьюринга, которая наряду с принципами архитектуры Дж. Фон Неймана, привела к бурному развитию электронно-вычислительных машин. Вообще считается, что компьютеры – это физическое воплощение концепции алгоритма, которое было введено для формализации мышления в процессе решения математических и иных задач [6].

Но является ли мыслительный процесс вычислимым, можно ли его представить в виде алгоритма? Знаменитые теоремы Гёделя о неполноте свидетельствуют, что нет. Даже в арифметике, оперирующей только натуральными числами и являющейся, наверное, самой строгой частью математики, которая в свою очередь является самой строгой областью человеческих знаний, не всегда можно доказать теоремы исходя из системы аксиом. Сами по себе аксиомы - исходные положения, от которых "пляшет" математик. Они элементарны в том смысле, что не доказываются, принимаются на веру, интуитивно. Например, одной из аксиом евклидовой геометрии является следующая: через две точки можно провести единственную прямую. Трудно придумать какие-то более элементарные положения, которые позволят доказать эту аксиому. Можно, правда, углубиться в строгие определения "точки" и "прямой", но таким образом мы, пожалуй, в принципе ничего не сможем познать в окружающем бесконечно разнообразном мире нашим ограниченным разумом.

Так вот, в упрощенной формулировке теоремы Гёделя утверждают, что в рамках заданной системы аксиом найдутся положения, которые невозможно ни доказать, ни опровергнуть. А следствием этому является вывод о том, что не существует алгоритма, который мог бы во всех случаях проверить истинность или ложность арифметического высказывания. Ну а если продолжить эти рассуждения применительно к нашей тематике можно утверждать, что нельзя запрограммировать компьютер так, чтобы можно было доказать произвольную теорему. А произвольной является великая теорема Ферма или гипотеза Пуанкаре, про которые компьютер, в отличие от людей (Э. Уайлса, Г. Перельмана), ничего "сказать" не может в рамках алгоритмического пространства, т. е. заданной ему программистом или найденной им самостоятельно в Интернете системы аксиом. Не способствует и самостоятельное формирование им аксиом, например с помощью генетических алгоритмов. Ведь аксиом можно придумать огромное (бесчисленное) множество, но чтобы доказать теорему - его нужно сузить до полной и непротиворечивой системы. Это сделать алгоритмически нельзя и прежде всего потому, что простым перебором за конечное время эту задачу не решить. Например, для автоматического доказательства теорем в пять строк (примерно одно правило и пять аксиом) требуется перебрать 3125 вариантов, а в пятьдесят строк – $8,9 \times 10^{34}$ вариантов [6]! Что же говорить о доказательстве Великой теоремы Ферма, занявшей несколько сот страниц?

Человек же творчески, интуитивно (непознаваемо) создает эвристики, правила, упрощающие перебор, снижающие размерность задачи. Но сам процесс нахождения этих эвристик - тоже безразмерный в связи с их бесчисленным числом. В этом случае компьютеру необходимо научиться находить сверх-эвристику, число которых тоже бесконечно. Для упрощения нахождения нужных сверх-эвристик необходимо найти сверх-сверх-эвристику и так до бесконечности...

Повторимся, всё это в полной мере относится даже к "строгой" математике, где, казалось бы, всё и вся можно доказать. Ну а тем более это применимо к более размытой об-

ласти естественных наук, зыбкой области гуманитарных наук и совсем уж неопределенной житейской области (которая включает в себя и математику и химию и социологию отношений и традиции и ненормативную лексику и т.п.). Пример из естественнонаучной дисциплины - для физики притяжение между двумя телами уменьшается согласно квадрату расстояния. Несмотря на признание этого положения законом и облечение в аналитическую зависимость, по сути - это всего лишь предварительное утверждение, поскольку невозможно проверить силу гравитационного притяжения для всех пар тел, существующих во Вселенной, на всех возможных расстояниях. Т.е. это утверждение истинно, пока не найдена будет ситуация, где оно не работает.

Что уж говорить о том, что компьютеры совсем не приспособлены для придумывания идеальных миров, не имеющих аналогов в прошлом и возможностей в будущем, например движение в мире без силы трения (Ньютон) или погоня за лучом света (Эйнштейн)?

Обо всем этом много говорил Р. Пенроуз [8], приводя примеры хорошо структурированных математических или физических задач, не имеющих общего аналитического решения, но о которых человек, в отличие от компьютера, может сделать вполне определенные выводы. Это, к примеру, решение диофантовых уравнений, замощение плоскости правильными плитками-многоугольниками, гравитационное взаимодействие трех тел и др. Он связывает это с невычислимостью сознания, которая, в свою очередь, зависит у него от неопределенности квантово-механических эффектов, происходящих в микротрубочках цитоскелета нервных клеток головного мозга...

Но, пожалуй, это слишком далеко уведет от нашей темы, поэтому можно просто сделать вывод о том, создание сильного ИИ в виде программы на ЭВМ в принципе не возможно. Хотя когда-нибудь человечество на каких-нибудь супермощных вычислительных устройствах запустит программу, которая будет объединять практически всё, что возможно алгоритмизировать, например, в военной сфере - и расчет траектории полета снаряда, и распознавание комплексов вооружения противника и принятие решения командиром и допрос пленного и многое другое. И судя по тому, что на многие вычислительные вопросы у такой программы будут ответы, то возникнет иллюзия искусственного интеллекта. Но в действительности это будет автоматизированная система, которая попадая в новую, ранее не встречавшуюся и, поэтому не алгоритмизированную ситуацию, не сможет с ней справиться. И здесь человеку опять придется потрудиться и над самой ситуацией и над включением её в алгоритм.

Конечно, можно дальше продолжить совершенствование такой глобальной автоматизированной системы и научить её самостоятельно создавать новые алгоритмы под новые ситуации. Но каждый раз целевая установка алгоритмизации будет исходить от человека, он будет являться работодателем у послушного, очень производительного, многофункционального, но не способного к творчеству работника.

Вторым направлением создания "сильного ИИ", т.е. интеллекта, достигаемого не естественными, природными причинами, и который будет сильнее обычного человеческого - это повышение когнитивных способностей Homo sapiens. И здесь нужно сказать о параметрах мозга человека, которые изменялись в период эволюционного развития, и сейчас достигли физического предела. Во-первых, это скорость вычислительных элементов. Для биологических нейронов пик составляет около 200 Гц, что на семь порядков медленнее современных микропроцессоров. Во-вторых, скорость внутренних коммуникаций. Аксоны передают сигнал со скоростью менее 120 м/с, в то время как оптические линии ограничены скоростью света. В третьих, количество вычислительных элементов и связанный с ним объем памяти. Для человека это, как мы говорили ранее, около 100 млрд. нейронов, для компьютеров существуют только ресурсные ограничения (по площади или пространству, доступной энергии и т.п.). При этом в кратковременной памяти человек может удержать одновременно 5-9 блоков информации, а объем его долгосрочной памяти ограничен примерно миллиардом битов. Очевидно, что у компьютеров аналогичные показа-

тели намного превышают людские. Можно привести ещё и параметры надёжности, точности, работоспособности, восстанавливаемости и т.п.

По всем указанным параметрам "углерод" будет значительно уступать "кремнию". И улучшение этих параметров естественным путем не предвидится. Для условий окружающего мира, которые сопровождали человека до последнего времени, это был своеобразный "оптимум финишного изделия", причем оптимум с технико-экономической точки зрения. Природа всегда рациональна и действует по принципу бритвы Оккамы - "не плодит сущностей без надобности". Поэтому скорость передачи электрохимических сигналов нервными клетками в высокой степени соответствует скорости изменения самого внешнего мира (быстроты изменения погоды, реакции опасных для человека животных и т.д.). Для эволюционных механизмов не было смысла ускорять мышление во много раз при изначальной инертности, статистической устойчивости окружающей среды.

Также емкость памяти и размеры мозга определяются изначальным размером черепной коробки младенцев и связанных с ней размером родовых путей, а также большими метаболическими затратами на поддержание мозга. Вес последнего составляет всего около 2% веса тела при потреблении 20% вдыхаемого кислорода. Кроме того, большие размеры мозга привели бы задержке обмена сигналами между различными его зонами, потребовали изменения системы охлаждения и питания (кровеносной системы) и т.п.

Поэтому человек, достигнув преимущества перед другими видами живых существ, остановился в эволюционном развитии. Но так было до недавнего времени, когда его окружала относительно статичная, прогнозируемая природа. И даже антропогенное воздействие на неё до последнего времени практически не изменяло скорости происходящих в ней процессов.

Мы уже несколько раз упомянули "до последнего времени" и это, безусловно, преднамеренно. Ведь в последнее время всё меняется настолько быстро, что мы сами стали способны эмпирически, опытно подмечать эти изменения, не сравнивая различные абзацы всевозможных исторических хроник. Мало того, что нас стал окружать чужеродный техногенный мир, напичканный шагающими экскаваторами и атомными электростанциями, автомобилями и самолетами, гаджетами и миксерами, синтезированными веществами и электронными микросхемами. Даже родной, устойчивый природный мир стал флуктуировать под натиском новой геологической силы современности - человека. Мы, к сожалению, являемся свидетелями глобального потепления, увеличения частоты и масштабов стихийных бедствий, скорости изменения погоды. В этой связи человеку для выживания и устойчивого развития также необходимо меняться, чтобы соответствовать скорости изменения окружающего мира. Но естественный процесс эволюционных изменений длителен, и сопряжен с "выбраковкой не эффективных особей". Поэтому и возникла идея искусственного усовершенствования человека, который будет адекватен современным реалиям. Мы не будем говорить обо всем спектре таких возможных изменений, а поговорим только об улучшении когнитивных, познавательных, мыслительных, а значит и интеллектуальных наших способностей. Путь этот долгий, трудоемкий и не очень эффективный - серьезного повышения уровня интеллекта достичь, по видимости, не удастся. Но, в то же время, более интеллектуальные ученые могут быстрее создать "сильный ИИ".

Об улучшении природы человека говорится уже давно. И также давно предлагается широкий диапазон способов такого улучшения - от духовно-нравственного воспитания практически во всех основных верованиях человечества до селекции *Homo sapiens* в рамках различных теорий неodarвинизма. Так, в мировых религиях в том или ином виде говорится о том, что человек создан по образу и подобию Творца. Но если образ дан ему изначальное, то подобие достигается в течении всей его жизни. Подобие во всех свойствах и качествах, в том числе в разуме и мудрости. И чтобы быть разумным человек должен познавать окружающий мир во всем своем многообразии и красоте. Для этого и дано ему любопытство, настолько сильное, что это привело к запретному познанию добра и зла и изгнанию из рая. Может быть с библейским событием и связан (не исторически, не

в плане хронологии, а скорее метафорически, аллегорично) всплеск интеллектуальных возможностей далеких прародителей людей. В более близкий к нам период уже апостол Павел в своих посланиях говорит о том, что надлежит быть разномыслиям дабы выявились искуснейшие. Искуснейшие - в способности познания, анализе фактов, в споре за истину, т.е. в значительной степени в интеллектуальной деятельности.

Поэтому уже очень давно известно, что познание, т.е. получение нового знания, обучение является основным способом улучшения интеллекта. В ходе познавательных процессов мозг меняется. Хотя новых нейронов в нём не образуется, но изменяются связи между ними. В отличие от компьютера мозг способен заново прокладывать нервные пути всякий раз, когда усваивается новая информация. Поэтому развить, или более осторожно - скорректировать, когнитивные возможности можно.

Другим направлением повышения интеллектуальных способностей, характерным для некоторых стран, является улучшение питания, особенно детей, устранение различных нейротоксичных соединений. Это будет влиять на здоровье, в том числе умственное, людей, на улучшение генофонда.

Более технологичные способы связаны с деактивированием различных частей мозга, что позволяет улучшить некоторые интеллектуальные способности. Примером для реализации такого способа явились "саванты" - люди, получившие, к сожалению, какие-то мозговые травмы, приобретя при этом сверхгениальные способности в очень узких интеллектуальных областях. Метод транскраниальной магнитной стимуляции позволяет заглушить левую височную долю и орбитофронтальную кору. На сегодняшний день это позволяет привести к незначительным улучшениям отдельных функций интеллекта на очень непродолжительное время, как утверждают исследователи - из-за грубого метода, не развитых технологий и других причин.

Следующим направлением является усиление памяти лекарственными препаратами. Считалось, что забывание - это пассивная, самопроизвольная деградация воспоминаний. Но сейчас установлено, что забывание - тоже активный процесс, требующий, в том числе, участия нейромедиатора дофамина.

Следует отметить, что у компьютера этот процесс тоже активен - в ячейках памяти старая информация стирается и заменяется новой. Это ведет к упорядочиванию системы "компьютер - источник питания", т.к. мир из-за наших расчетов становится более предсказуемым, менее хаотичным, что ведет к понижению энтропии информации. Но в замкнутой системе энтропия не может снижаться, поэтому понижение информационной энтропии ведет к повышению энтропии термодинамической, т.е. к увеличению тепла. Такой процесс приводит к серьезному нагреву микросхем и поэтому современный компьютер в большей степени работает, как обогревательное устройство, рассеивающее тепло. Желая с большей пользой использовать его вычислительные мощности, даже разработана теория обратимых вычислений, реализация которой позволит превратить хаотичное тепло в упорядоченные расчеты, но это, пожалуй, отдельная тема.

Так вот, есть лекарственные препараты, подавляющие процесс забывания. Это улучшает когнитивные способности человека за счет возможности оперирования большими объемами информации.

Лекарственное улучшение когнитивной сферы касается не только памяти, но и других свойств интеллекта, стимулировать который можно ноотропными средствами, нейрометаболическими стимуляторами. Но действие их кратковременно, не стабильно, обладает множеством побочных эффектов, и не позволит обеспечить устойчивый рост мыслительных способностей. Всё-таки для улучшения функциональности мозга требуется "бережный баланс, тонкая настройка и тщательная культивация", а не подкормка чудодейственным зельем.

Известно так же то, что введение стволовых клеток в гиппокамп способствует улучшению памяти, мыслительных процессов. Но эти волшебные клетки способны, к сожалению, неожиданно превращаться в злокачественные, приводящие к онкологическим заболеваниям.

Кроме того, в результате многолетних генетических исследований, ученые обнаружили существование участков генов, отвечающих за гигантское ускорение эволюции человека за последние 6 млн. лет, т.е. период, когда выявилось коренное отличие людей от человекообразных приматов. Эти исследования породили различные идеи о том, что вызвав мутации генов, отвечающих за интеллект, можно увеличить количество серого вещества, увеличить площадь неокортекса и повысить интеллект человека. Но всё это не так очевидно, т.к. площадь поверхности и объем новой коры полушарий головного мозга у дельфинов, слонов, кашалотов и некоторых других млекопитающих больше, чем у человека, чего не скажешь об уровне мыслительных способностей (или может мы просто не можем по-настоящему измерить этот уровень?).

Много сказано и об итеративной селекции эмбрионов, позволяющей их "конструирование" с идеально точным сочетанием генетических данных родителей, и геномной модификации соматических клеток взрослого человека и многим другим направлениям создания генетически модифицированного *Homo sapiens*. Сейчас современное общество негативно относится даже к генетической инженерии сельскохозяйственных продуктов, не говоря уже о животном царстве. Но такое отношение может перемениться. Ведь эволюция человека, закончившись на биологическом уровне, продолжается на социальном, изменяя мораль общества, нормы поведения, устои и мировоззрение. Вспомним про изменение отношения к пыткам и смертной казни или к сексуальной ориентации и процедуре экстракорпорального оплодотворения.

Также может измениться и отношение к генетическим изменениям эмбрионов. Так, к примеру, демографическая политика ряда стран, в которых наблюдается высокие значения показателя воспроизводства населения и не очень высокие уровни доходов, может привести к тому, что их граждане начнут стремиться к рождению меньшего количества детей. В связи удешевлением генетических технологий многие могут пожелать, чтобы это меньшее количество детей было бы более интеллектуальным. При этом если на начальном этапе и возникнет расслоение населения на тех, кто пойдет таким путем и кто отвергнет его, то в будущем оно довольно быстро может исчезнуть в связи размыванием строгих этических норм. Ведь последствием этого расслоения будут успеваемость в школе и университете, профессиональная успешность, размеры доходов и многие другие составляющие базовых потребностей людей. А повышение среднего интеллектуального уровня граждан какой-либо страны отрицательно скажется на балансе различных геополитических игроков (более умная нация будет обладать более совершенным вооружением). А это, в свою очередь, вместе с желанием создать, в том числе, совершенных солдат будущего, вызовет гонку вооружений в когнитивно-физической сфере человека, и искусственно переведет его из класса млекопитающих (*Mammillae*) уже в совершенно новый класс "технородящих" (*Artificially borning*) [9].

Последствиями такого негативного сценария может быть радикальное изменение нашего мира. Даже с помощью традиционных технологий он стал неузнаваем, превратился во "вторую природу". А ведь они, в основном, касались только внешней среды. Изменение же внутренней среды самого человека наверно даже нельзя будет выразить никакой метафорой типа "ящик Пандоры" или "новый Армагеддон". Изменится сама суть человека, поэтому и говорить в рамках рассматриваемого пути о создании искусственного интеллекта, подобного человеческому, будет бессмысленно.

И так, у нас остается третий путь - создание биокибернетических систем, служащих подспорьем человеку в его мыслительных процессах и совокупность которых можно будет назвать его экзокортексом.

Уже сегодня существуют различные устройства, позволяющие облегчить страдания людей с различными невротическими, психическими и другими расстройствами. Они позволяют, например, мысленно набирать буквы на экране, синтезировать голос парализованного человека, снизить интенсивность и частоту изнурительных припадков эпилепсии, шевелить пальцами обездвиженной руки и многое другое. Для анализа

мыслительной деятельности людей применяются различные сканирующие устройства, томографы, вживляемые в мозг электроды. Безусловно, достижения нанотехнологий и микроэлектроники приведут в будущем к массовому применению на коммерческой основе подобных устройств. Причем не только для больных, но и для здоровых людей, например, солдат будущего.

Фантасты, футурологи и ученые, пытающиеся заглянуть за ближайшие горизонты развития технологий, предлагают большое количество таких устройств, предназначенных, в том числе и для развития когнитивных функций человека, например специальные шлемы с миниатюрным магнито-резонансным томографом и процессором для распознавания мыслей и мгновенной передачи в машинных кодах на любые расстояния.

Конечно, сегодня мы не можем распознать сами мысли - только их отражения в мозге: движение электрических зарядов, изменение химического состава крови или клеток и т.п. Сравнивая такие отражения (паттерны) с мыслями людей, можно устанавливать определенные взаимосвязи и применять их для "расшифровки" мыслей человека, а затем передавать, и принимать эту информацию без помощи звуков, жестов или письма.

При этом возникает естественный соблазн охватить подобными технологиями все когнитивные и эмоциональные способности человека. Прямой нейрокомпьютерный интерфейс позволит человеку использовать всю мощь машинных вычислений - идеально хранить, и получать информацию практически неограниченных объемов, с высокой скоростью проводить точные расчеты, практически мгновенно передавать данные на значительные расстояния. Такая биокибернетическая система должно превзойти мозг современного человека по всем характеристикам, стать своего рода сильным искусственным (или полуискусственным) интеллектом. Люди смогут общаться напрямую, передавая партнеру свои неискаженные мысли, образы, знания и опыт. Ну а следующий шаг - планетарная мозговая сеть Brain-net, в которой взаимодействие между людьми на этих принципах осуществляется уже на глобальном уровне...

Но, как всегда, у всех смелых мечтаний человека есть какие-то ограничения. Они могут нарушать законы природы (во всяком случае, открытую их часть, от слова "открытие") или быть тяжело воплощаемы с различных точек зрения. Например, техническая реализация прорывных устройств может натолкнуться на отсутствие материалов, сложности в переходе с лабораторного на производственный уровень (как программируемое вещество или лазерный термоядерный синтез), и, наконец, не экономическую целесообразность.

Развитие биокибернетических систем тоже будет связано с большими трудностями. Например, на современном этапе существуют значительные медицинские риски, связанные с внедрением имплантов в мозг - инфекции, кровоизлияния, другие нарушения мозговой деятельности, связанные с вмешательством.

Также, сегодня нет понимания как добиться непосредственного взаимодействия между мозгом и компьютером. Ведь мозг имеет совсем иные, по сравнению с компьютером, методы хранения и представления данных. Кроме того, у каждого из нас существуют свои уникальные способы формирования и передачи смыслов, которые зависят от житейского опыта, социокультурных условий, генетических факторов, особенностей физиологии и т.д. Смысловое значение информации представлено не дискретными логическими массивами ячеек памяти компьютера, а ассоциативными образами, сочетающими в себе информацию различных, иногда перекрывающихся или взаимозаменяемых (синестезия) зон коры головного мозга. При этом, по-видимому, для понимания одного человека другим, информацию нужно подвергать декомпозиции и переводу в некий универсальный символичный язык. Как при этом установить соответствие между группами нейронов двух людей, чтобы такие образы могли переходить от одного человека к другому пока совершенно неясно.

Еще один вопрос - сможем ли мы на самом деле эффективно повысить уровень нашего интеллекта? Так, обеспечив доступ мозга к существенно большему объему информации, мы можем столкнуться с ограничением скорости ее обработки, которая, в свою очередь, зависит

от скорости прохождения нервного импульса. В итоге придется "апгрейдить" механизм такой обработки с учетом всех физиологических процессов организма.

И, наконец, если в будущем мы сможем создавать подобные нейроморфные интерфейсы, это очевидно изменит сложившуюся систему межличностных отношений, моральных устоев, переформирует социокультурную среду. Но когда это будет, как это будет, не изменимся ли мы сами? И готовы ли мы к такому преобразению? И судя по обилию подобных вопросов это ещё только начало пути...

Можно долго продолжать рассуждать об искусственном интеллекте, тема эта неисчерпаема как мироздание. Эйнштейн однажды остроумно заметил, что в мире есть две бесконечные вещи - Вселенная и человеческая глупость, хотя в первом он не вполне уверен. Ну а человеческая глупость - это какая-то лишь сторона интеллекта или, вернее, уровня его развития. И если уж сторона бесконечна, то что и говорить о всей многосторонней фигуре, которой является интеллект и его антропоморфизм - искусственный интеллект?

В заключении хотелось бы отметить, что авторы отдают себе отчет в том, что вышеприведенные рассуждения не претендуют на серьезное приближение к истине. Поэтому мы приветствуем любую критику и тем более дискуссию в этом вопросе. Только спор рождает истину, которая проверяется практикой.

Литература

1. Николенко С., Кадури А., Архангельская Е. Глубокое обучение. Погружение в мир нейронных сетей. СПб: Питер. - 2018.
2. Каку Митио. Будущее разума. Пер. с англ. М.: Альпина нон-фикшн. - 2018. - 502 с.
3. Дж. Хоккинс. Об интеллекте. Пер. с англ. М.: ООО «И.Д.Вильямс». - 2007. - 240 с.
4. Р. Курцвейл. Эволюция разума, Или бесконечные возможности человеческого мозга, основанные на распознавании образов. Пер. с англ. М.: Эксмо. - 2018. - 352 с.
5. Что мы думаем о машинах, которые думают: Ведущие мировые ученые об искусственном интеллекте / Джон Брокман; пер. с англ. М.: Альпина нон-фикшн. - 2017, - 324 с.
6. Н. Бостром. Искусственный интеллект. Этапы. Угрозы. Стратегии. Пер. с англ. М.: Манн, Иванов и Фербер. - 2016. - 496 с.
7. Потапов А.С. Искусственный интеллект и универсальное мышление. СПб.: Политехника. - 2012. - 711 с.
8. Пенроуз Р. Тени разума: в поисках науки о сознании. Москва-Ижевск: Институт компьютерных исследований. - 2005. - 688 с.
9. Новые технологии и продолжение эволюции человека? Трансгуманистический проект будущего. М.: Издательство ЛКИ. - 2008. - 320 с.

Сведения об авторах

Дурнев Роман Александрович, доцент, заместитель начальника научно-исследовательского управления ФГБУ РАРАН, тел.: 8 (903) 258-67-16, E-mail: rdurnev@rambler.ru, SPIN-код: 3267-1337

Крюков Кирилл Юрьевич, руководитель секретариата Совета главных конструкторов по системе вооружения сухопутной составляющей сил общего назначения ФГБУ РАРАН, тел.: 8 (499) 160-29-33, E-mail: raran@bk.ru

Титов Андрей Евгеньевич, директор научного центра авиационной техники и вооружения. Федеральное государственное бюджетное учреждение «Российская академия ракетных и артиллерийских наук» (ФГБУ РАРАН), г. Москва, 1-я Мясниковская ул., д.3, стр.3. e-mail: tigma_zver_tam@mail.ru, 8 (925) 193-78-56.

Жданенко Ирина Васильевна, старший научный сотрудник ФГБУ ВНИИ ГОЧС (ФЦ), тел.: +7(909)953-56-59, E-mail: izhdanenko@yandex.ru, SPIN-код: 7747-6337