



ИПМ им.М.В.Келдыша РАН

К 95-летию С.П.Курдюмова

К 95-летию
со дня рождения
С.П. Курдюмова



Горизонты
математического моделирования
и теория самоорганизации

В.С. Смолин

**Нейросетевые алгоритмы –
передовой край процессов
самоорганизации**

Рекомендуемая форма библиографической ссылки

Смолин В.С. Нейросетевые алгоритмы – передовой край процессов самоорганизации // Горизонты математического моделирования и теория самоорганизации. К 95-летию со дня рождения С.П. Курдюмова. — М.: ИПМ им. М.В.Келдыша, 2024. — С. 176-206.

<https://doi.org/10.20948/k95-11>

<https://keldysh.ru/e-biblio/k95/11.pdf>

Нейросетевые алгоритмы – передовой край процессов самоорганизации

В.С. Смолин

Институт прикладной математики им. М.В. Келдыша РАН

Аннотация. С созданием сильного ИИ не только человечество, но вся эволюция жизни вступает в новую эпоху. Жизнь как самый сложный и самый удивительный процесс самоорганизации плохо поддается точному описанию, но попытки выявить основные условия ее развития делаются давно. Как стало понятно примерно 70 лет назад, главным условием развития жизни является накопление знаний в ДНК, позволяющих воспроизводить прошедшие эволюционный отбор формы жизни. Выдвигалось много других важных условий развития «живой» самоорганизации, которые лучше, чем знания, подходили в качестве объектов исследования. Хотя с началом нейросетевой революции в машинном обучении появились возможности исследовать на моделях процессы получения знаний из окружающего мира без участия человека, инерция исследователей не затрагивать процессы познания при изучении самоорганизации остается. Проводится сравнение традиционных и нейросетевых подходов к описанию процессов самоорганизации. Т.к. только нейросети явно описывают процессы получения знаний, делается вывод, что нейросетевые алгоритмы составляют передний край развития процессов самоорганизации.

Ключевые слова: формальные нейроны, сильный искусственный интеллект (СИИ), цивилизация, конкуренция, декомпозиция, знания

Neural network algorithms as the cutting edge of self-organization processes

V.S. Smolin

RAS Keldysh Institute of Applied Mathematics

Abstract. With the creation of strong AI, not only humanity, but the entire life evolution enters a new era. Life, as the most complex and most amazing self-organization process, is difficult to accurately describe, but attempts to identify the basic conditions for its development have been made for a long time ago. As it became clear about 70 years ago, the main condition for the life development is the knowledge accumulation in DNA, making it possible to reproduce life forms that have undergone evolutionary selection. Many other im-

portant conditions for the development of “living” self-organization were put forward, which were better suited as research objects than knowledge. Although with the beginning of the neural network revolution in machine learning, it became possible to study on models the processes of acquiring knowledge about the surrounding world without human participation, the researchers inertia not to include the cognition processes in self-organization studies remains. A comparison of traditional and neural network approaches to describe self-organization processes was made. Since only the neural nets describe the acquiring knowledge processes, it is concluded that neural network algorithms constitute the cutting edge of the self-organization processes development.

Keywords: formal neurons, artificial general intelligence (AGI), civilization, competition, decomposition, knowledge

Введение

Чудо самоорганизации на физическом, химическом и особенно биологическом уровнях всегда было предметом удивления и восхищения. Выдвигались различные теории объяснения чуда: сакральные, атомистические, эволюционные, синергетические и другие, но все они описывали только отдельные аспекты процессов самоорганизации. Относительно простыми способами удалось объяснить только самоорганизацию на физическом и химическом уровнях – там, где сложные процессы идут исключительно под управлением физических законов. Для самоорганизации биологического уровня сложности необходимо привлекать значительные объемы знаний, которые, до открытия ДНК, было невозможно представить как сохранять и передавать в материальном мире. Даже расшифровка кода генома не дала ответа на вопрос: откуда берутся знания? Естественный (и направленный) отбор в эволюции позволяет оптимизировать передаваемые по наследству знания, но откуда взялись начальные знания?

Наконец, самое удивительное в процессах самоорганизации – способность человека к познанию окружающего мира. Человек, как и всё живое, является продуктом самоорганизации, и его способность к познанию – тоже. Откуда появилась эта способность и как она реализуется? За время развития цивилизации было разработано много методик, позволяющих ускорить процесс познания, сделать знания более достоверными. Но самый первый этап познания – выделение отдельных объектов и их свойств в окружающем мире, которые в дальнейшем будут исследоваться, – остается загадочным. Произносятся слова вроде «Выделение объектов и их свойств – это способность человеческого сознания», которые, вне зависимости от отнесения сознания к свойствам мозга или души, ничего не объясняют.

Успехи развития технологий в древности и в Средние века привели к повышению уровня производительности труда, к возможности всё большему числу людей заниматься наукой и получать новые знания. На основе которых быстрее создавались новые технологии, которые дали импульс

началу промышленной, а затем и информационной революциям в развитии цивилизации. Но практически все научные достижения, полученные за период развития цивилизации, несут знания про свойства окружающего мира, а не о процессе получения этих знаний. Конечно, процесс получения новых знаний имеет свою мифологию и историографию. Например, миф, что Д.И. Менделееву его периодическая таблица приснилась, или рассказ самого Менделеева, что таблица – результат 25-летнего труда, представляют интерес, но не проясняют механизмы получения новых знаний.

Для получения новых знаний (теоретических или практических) необходимо сформулировать (относительно) простую задачу, тогда ее можно решить (и получить новые знания). В то же время как выделение простой задачи из свойств окружающего мира – сложная проблема, и, как многие продолжают считать, решению не поддающаяся. Научное сообщество следует конструктивному (и, в целом, правильному) девизу: «Давайте решать простые задачи, сложные мы всё равно решать не умеем!».

И сам человек, и живая природа, и цивилизованное сообщество, и материальная база цивилизации, да и весь физический мир в целом имеют сложное строение. Образование, несущее нам результаты науки, дает представления о свойствах отдельных частей сложных систем, с которыми мы сталкиваемся в своей жизни. Получая разнородные знания, мы формируем свою субъективную картину мира в целом. Но мы не способны освоить весь багаж знаний, накопленных человечеством, и можем использовать только доставшуюся нам часть знаний и личный опыт (которые у каждого свои). И несмотря на усилия политиков, философов, ученых и педагогов, стремящихся привить нам единые взгляды на природу мироздания, субъективные картины мира каждого человека имеют свои, индивидуальные особенности.

Тем не менее, физический мир допускает построение своего объективного описания, и научный подход направлен как раз на создание объективных описаний окружающего нас мира. И наша неспособность решать сложные задачи неоднократно приводила к идее самоорганизации: мы можем познать простые принципы, на основе которых происходит формирование сложных объектов и явлений. Это, конечно, не объясняло процесс познания (его просто использовали), но позволяло двигаться в основном русле научной парадигмы: «Решаем простые задачи!».

Идея самоорганизации

Самоорганизации – идея, которая волнует человечество с древних времен. Воспроизводимость природных свойств, явлений и форм живых организмов всегда вызывала удивление человека и, при отсутствии других объяснений, приводила к выводу о «воле Божьей».

Одними из первых, кто стал развивать рациональные объяснения воспроизводимости природных форм, были античные философы Левкипп

и Демокрит (V в. до нашей эры). Согласно их теории число форм атомов ограничено, а число атомов – безгранично, что, аналогично текстам, составляемых из небольшого набора букв, объясняет не только разнообразие наблюдаемых объектов и явлений, но и повторяемость ряда их форм и свойств.

Такое объяснение свойств мира было материалистическим, не требовало привлечения Бога, что вызывало гнев ряда современников Демокрита (про которого, по ссылкам в трудах других философов, известно, что он написал порядка 70 работ). Сочинения Демокрита не дошли до нас, поскольку «существует гипотеза, которая утверждает, что труды Демокрита уничтожены сторонниками его философского противника Платона, который утверждал, что наука Демокрита вредна для общества и государства, потому что она атеистична, и призывал карать сторонников Демокрита "хлыстом и огнем"» [2].

В Античности и в Средние века не было приборов, позволяющих увидеть атомы, представления Демокрита оставались «игрой ума», недоказанной гипотезой. Современные успехи физики и химии подтвердили справедливость положений атомистической теории Демокрита, сделали ее одной из основных идей, на которых развивается современная наука. Ограниченность числа видов атомов позволяет объяснить повторяемость свойств веществ и ряд явлений неживой самоорганизации. Таких, например, как рост кристаллов, воспроизводимое (в огромных количествах!) образование дождевых капель и снежинок и многих других. При наличии множества явлений самоорганизации в неживой природе своим разнообразием они значительно уступают свойствам живых организмов и созданным в результате развития цивилизации техническим устройствам.

Идеи самоорганизации прошли сквозь Средние века, но их расцвет начался в XIX-XX вв. Наибольшее внимание исследователей самоорганизации привлекает теория биологической эволюции, основные положения которой были сформулированы еще в XIX в. Ж.Б. Ламарком и Ч. Дарвином. Есть и много других, более современных, получивших известность и в разной степени признанных теорий самоорганизации. Это и обобщенно-сознательный подход к тектологии (организации систем) А.А. Богданова [3], системные представления У.Р. Эшби [4] и П.К. Анохина [5], диссипативные структуры И. Пригожина [6], синергетика Г. Хакена [7], теория универсального эволюционизма Н.Н. Моисеева, гиперциклы М. Эйгена [8], автопоэзис У. Матураны и Ф. Варелы [9], эволюционная концепция развития вселенной Э. Янча [10], единая трансдисциплинарная теория Э. Ласло [11], теория самоорганизованной критичности П. Бака [24], теория самоорганизации С.П. Курдюмова [12], основанная на базе режимов с обострениями решений дифференциальных уравнений, сетевая, бутстрапная теория Ф. Капры [13]. Можно перечислять еще десятки имен, внесший свой вклад в теории развития и существования

сложных систем, так или иначе связанными с идеями самоорганизации, и обсуждать степень важности их вклада.

Основной посыл этих теорий: понять все тонкости работы сложных систем нет возможности, но они некоторым образом существуют и без нашего понимания. Значит – есть простые принципы, которые позволяют сложным системам развиваться. Выделить эти принципы на основе наблюдения за сложными системами нельзя, но можно их постулировать и соотносить постулаты с наблюдаемыми свойствами сложных систем.

Почти никто из теоретиков (кроме, например, У. Матураны и Ф. Варелы [9]) не отрицает, что функционирование сложных систем основано на знаниях. Но процесс получения и обмена знаниями сложный, и попытка его исследования (была) малоперспективна. Под лозунгом стремления к конструктивной простоте анализировались свойства физических систем, которые можно соотнести с процессами самоорганизации, и находились аналогии таких процессов в сложных биологических и социальных системах. Наиболее радикальные исследователи систем (например, Ф. Капра [13]) не видят особой разницы, с чем соотносить явления в сложных системах – с физическими процессами или восточными духовными практиками, поскольку на точные количественные прогнозы развития процессов в сложных системах всё равно никто из «системщиков» не претендует.

При этом в физических системах было найдено и изучено много интересных свойств, таких как бифуркации, странные аттракторы, автоколебательные и неравновесные процессы и многое другое, которые можно рассматривать как соответствующие идеям самоорганизации. Не вызывает сомнения, что полученные результаты полезны уже сейчас и могут помочь в понимании сложных процессов познания, но явно на объяснение процессов получения и использования знаний эти результаты не направлены.

Чем знания отличаются от данных и информации?

Поскольку термин «знания» уже использовался в тексте и будет часто встречаться далее, необходимо описать, что под ним понимается.

В бытовом понимании данные, знания и информация – это синонимы, их все можно представить в бинарной форме (последовательности нулей и единиц). Дело не столько в форме, сколько в готовности (пригодности) описаний для использования с различными целями. Будем различать:

- *данные* – описания объектов, явлений и действий в любой форме;
- *знания* – описания объектов, явлений и действий в форме, пригодной для формирования действий сложных систем;
- *информация* – описания, позволяющие выбрать знания для формирования действий.

Данные могут иметь различную степень упорядоченности и обобщения, но если люди (или сложные системы) не могут их использовать для

своих действий, это не знания. Знания могут потребоваться в разных ситуациях, в будущем, это продукт длительного хранения. А информация позволяет выбрать знания для формирования действий в конкретный момент, а затем может быть полезна только как данные для получения знаний.

Одно и то же описание может быть использовано и как данные, и как знания, и как информация, в зависимости от процесса, в котором это описание используется. Так, например, коды ДНК, содержащиеся в геноме, при их секвенировании (расшифровке) являются данными. Когда те же самые коды ДНК находятся в ядре живой клетки – это знания, которые определяют действия, выполняемые клеткой в процессе ее развития и воспроизводства. Наконец, та часть кодов, которые используются для реализации непосредственно выполняемых действий – это информация о том, каким образом выполнять действия в данной ситуации.

Данные могут храниться на любом носителе [15], в том числе и изолированно, тогда как знания и информация – обязательно на носителе, связанном с устройствами и/или системами, способными их использовать.

Чтобы данные стали знаниями, нужно чтобы их форма была согласована с системой, которая их использует, и носитель данных был связан с устройствами, формирующими свои действия на основе этих знаний. Будем использовать следующее определение знаний: ***«Знания – описания объектов, явлений и действий в виде, согласованном со свойствами системы и пригодной для формирования действий сложных систем».***

Если знания, используемые системой точны, достоверны и соответствуют сложившейся ситуации, то это повышает вероятность, что действия, сформированные на таких знаниях, позволят достичь поставленной цели. И что поставленная цель адекватна требованиям и возможностям системы. А если это не так, то такие знания имеют низкую вероятность сохранения по ряду причин: от неиспользования недостоверных знаний для распространения, через простое «забывание» таких знаний, и до уничтожения вместе с системой, которая их использовала.

Достаточно широко распространено отношение к самоорганизации как к чуду, которое должно происходить само, безо всяких усилий (и без использования знаний). Иначе это не фокус, не чудо, а технологический результат производственного процесса или деления клетки. Но давайте не будем верить в сказки, а просто определим, будет ли воспроизводиться чудо жизни без кода ДНК, содержащегося в геноме? Или технические чудеса удастся реализовать, если строго не выполнять требования технологий?

«Самоорганизация» – это не волшебная палочка, которая выполняет любые прихоти ее владельца после произнесения правильных заклинаний (хотя и для этого, согласно сказкам, тоже нужны знания). Самоорганизация – это создание условий (возможно, с приложением некоторых усилий) для выполнения рациональных в некотором смысле действий. Смыслы и цели назовем ниже, а здесь обратим внимание, что человек – тоже продукт

самоорганизации. И всё, что он делает, включая не всегда разумные действия, – это самоорганизация как на уровне одного организма, так и локального сообщества и цивилизации в целом. И для успешной реализации самоорганизации действия следует формировать на основе знаний.

Чудеса живой природы и замечательные технические достижения человечества все основаны на использовании знаний. Знания – это не абстрактная нематериальная субстанция. Как справедливо отмечается, например в [15], знания всегда имеют материальный носитель. Это относится и к ДНК, и к структурам мозга, и к книгам, и к машинной памяти. Причем просто физического носителя знаний (которые на носителе ничем не отличаются от данных) недостаточно. Нужны еще системы, которые позволяют преобразовывать сохраненные знания в действия. В живой природе это клетки, которые на основе ДНК и РНК осуществляют процессы деления, роста и дифференциации. В цивилизации это человек, который преобразует книжные знания и практические действия. В ЭВМ это компьютерное железо, которое преобразует программы и данные, хранящиеся в памяти, в исполняемые алгоритмы. Без преобразователей в действия знания останутся данными, которые могут стать полезными только в том случае, если для них найдется соответствующий преобразователь, который имеет источник энергии или способен использовать знания для получения энергии из окружающей среды.

Почему именно знания, а не другие (из множества предложенных другими авторами, например, [3-14]) причины являются основой самоорганизации? Знания – главная причина успешности самоорганизации только в живой природе, физические и химические процессы самоорганизации в неживой природе не требуют накопления дополнительных знаний сверх тех физико-химических свойств компонентов процесса, которые им и так присущи. Работы, в которых знания не рассматриваются как причина самоорганизации, посвящены исследованию более общих свойств самоорганизации, которые можно относить как живой, так и неживой природе. В живой природе тоже выполняются физико-химические законы, в том числе и относящиеся к процессам самоорганизации. Но без использования накопленных знаний вероятность успешной реализации «живых» и «цивилизационных» процессов самоорганизации практически равна нулю.

Самоорганизация в живой природе и цивилизации характеризуется (по сравнению с неживой природой) большим разнообразием форм и сложностью проявляемых свойств. Реализация «живых» процессов самоорганизации построена не только на законах природы, но и на знаниях, накопленных жизнью и человеческой цивилизацией (рис. 1).



Рис. 1. Отличие процессов самоорганизации, основанных на знаниях

Знания, как основа самоорганизации жизни и цивилизации

Знания – основа цивилизации, на них построены большинство результатов человеческой деятельности. Но описание способности человека к получению и использованию знаний просто как результата Творения не выглядит научным. Выделение человека из животного мира принято объяснять эволюционными теориями, а креационизм обычно рассматривают как реакционное течение, препятствующее развитию науки и идей СИИ.

Исходя из эволюционных взглядов, Homo Sapience возник как продолжение развития возможностей живых организмов приспосабливаться к изменяющимся условиям, которые реализуются с использованием накопленных в процессе эволюции знаний. Человек научился сохранять и передавать свои знания в речи, текстах и теперь посредством информационных технологий, а у животных знания сохраняются и передаются в формах: РНК и ДНК, строении тела и примерах поведения. Человек как биологический вид тоже передает знания в таких формах, как и все животные. Наличие дополнительных, «небиологических» форм передачи и хранения знаний позволили ему перейти к новому этапу самоорганизации – построению цивилизации на основе методов получения знаний, не только превосходящих способы, используемыми животными, но и позволяющих постоянно увеличивать разрыв по скорости и качеству получения знаний.

Нейросетевая революция в машинном обучении сделала неубедительным тезис о невозможности понять процесс познания. Современные системы ИИ на нейросетевых алгоритмах позволяют реализовать всё больше «интеллектуальных» функций, которые еще несколько лет назад были недостижимыми для машин. Настройка огромного числа параметров нейросетей – это, по сути, и есть процесс получения знаний об окружающем мире. Мы можем развивать и совершенствовать технические способы получения знаний о мире, но есть ли граница между техническими и человеческими знаниями, которую невозможно преодолеть?

Успехи развития нейросетевых алгоритмов позволяют дать положительный ответ на данный вопрос. Ясного понимания процессов обработки данных в глубоких нейросетях пока достичь не удалось, но знания, кото-

рые нейросети накапливают в процессе своего обучения, уже позволяют решать широкий список впечатляющих своей сложностью задач.

Представления, что успеха в исследованиях можно достичь только в случае, если проблему познания не затрагивать, очень живучи. Многие исследователи предпочитают изучать другие, не связанные с получением новых знаний, причины самоорганизации потому, что только в этом случае их теории найдут понимание у широкой аудитории. Рассмотрим некоторые из традиционных и относительно новых взглядов, отмечая, что все эти теории вполне научны и полезны. Но игнорирование проблемы получения именно новых знаний, а использование полученных человеком знаний резко снижают перспективную полезность таких теорий при создании СИИ.

Познание и энтропия

Условия для реализации маловероятных условий для протекания процессов самоорганизации в живой природе и цивилизации требует не только наличия знаний и возможности их использовать при формировании действий, но и некоторой энергии как для выполнения действий, так и для преобразования знаний в действия и данных в знания.

Вопрос о необходимости энергии для осуществления процессов самоорганизации, пусть в несколько вычурной форме, но сыграл значительную роль в развитии теории самоорганизации. Приблизительно через 100 лет после исследований Карно о свойствах тепловых машин, опубликованных в 1824 г., биологи и философы обратили внимание, что получение энергии из окружающей среды живыми организмами противоречит второму закону термодинамики. Этот факт рассматривался чуть ли не как основное свойство жизни, причем биологов [16] поддерживали в данном вопросе физики [17], и весьма известные. Тогда, 80-100 лет тому назад, рассмотрение вопросов противодействия увеличению энтропии было вполне прогрессивным, оно позволяло лучше понять свойства самоорганизации.

Но сейчас, когда вопросы энергообеспечения жизнедеятельности клеток (в том числе и для использования знаний) хорошо изучены, попытки вернуться к представлениям, что целью развития жизни является именно противостояние увеличению энтропии и необходимость для этого «свободной энергии» (о чем писал еще Бауер в 1935 г. [16]), выглядят ретроградными, обращенными в прошлое. Следует отметить, что на основе накопленных знаний все живые системы и цивилизация действительно умеют противодействовать вредоносным воздействиям и хаосу. Измерение энтропии может служить некоторой мерой эффективности этой борьбы. Но если этими измерениями заменять исследование путей получения и использования знаний, то на современном этапе такие взгляды (как, например, в [18,19]) только тормозят развитие теории самоорганизации.

Приведем простой пример. Холодильник тоже борется с энтропией, которая стремится выровнять температуру внутри и снаружи холодильника. И использует для этого «свободную энергию» (из розетки). А что нам важнее: измерения изменения энтропии или знания о технологии производства холодильников? С созданием СИИ ситуация аналогичная, только задача построения СИИ несколько сложнее производства холодильников. Тем не менее, «энтропийные» представления о свойствах СИИ пользуются популярностью у руководителей научных подразделений, которые сформировали свои взгляды на возможности объяснения процесса познания задолго до начала нейросетевой революции в машинном обучении.

Байесовские подходы

Некоторым прогрессом по отношению к «энтропийным» взглядам является переход к идее получения «новых знаний» на основе теоремы Байеса [20]. Хотя доказана эта теорема была почти на сто лет раньше развития представлений об энтропии (и на двести – информационной интерпретации энтропии), прогрессивной в байесовском подходе является идея получения новых знаний, а не просто констатация необходимости противостоять хаосу. За почти 250 лет развития теории применения байесовских подходов в этой области достигнуты значительные успехи. Анализируются не только вероятности событий, но и распределения вероятностей (как правило, с использованием дивергенции Кульбака–Лейблера [21]), при этом успешно прогнозируются результаты уникальных событий (таких, как исход спортивных поединков, что необходимо для успешной работы тотализаторов). Более того, байесовский подход применяется и для обучения нейросетевых алгоритмов [22], и даже сторонники «энтропийного» подхода к самоорганизации используют его в своих теориях [19].

Как энтропийный, так и байесовский подходы имеют вполне научную основу и развивают правильные взгляды на некоторые аспекты процессов самоорганизации. Но, что свойственно и другим авторам теорий самоорганизации [3-14], и не только им, чтобы сделать популярной теорию, необходимо заявить, что именно на ее основе можно решить все проблемы. Если переводить “evidence” не как «доказательство», а как «очевидность», а “likelihood” не как «вероятность», а как «правдоподобие», то сразу становится «очевидным», что получать «правдоподобные» знания можно только на основе теоремы Байеса (см. рис. 2). Хотя практически все учебники (по разным предметам) содержат задачи, по условиям которых, поняв материал учебника, можно получить новые знания путем решения задачи (без теоремы Байеса). Да, обычно в задачах находят детерминированные ответы, а теорема Байеса – для вероятностных, но и теория вероятности – это не только Байес. И «парадокс Монти Холла», который принято объяснять байесовскими методами [23], вполне разрешим и без теоремы Байеса. А «чудесное» свойство, что формула Байеса позволяет «переста-

вить причину и следствие» основано только на философских представлениях о наличии причины у каждого события. А математика позволяет только посчитать корреляцию между событиями, что подчеркивается полной симметричностью «причины» и «следствия» в формуле Байеса.

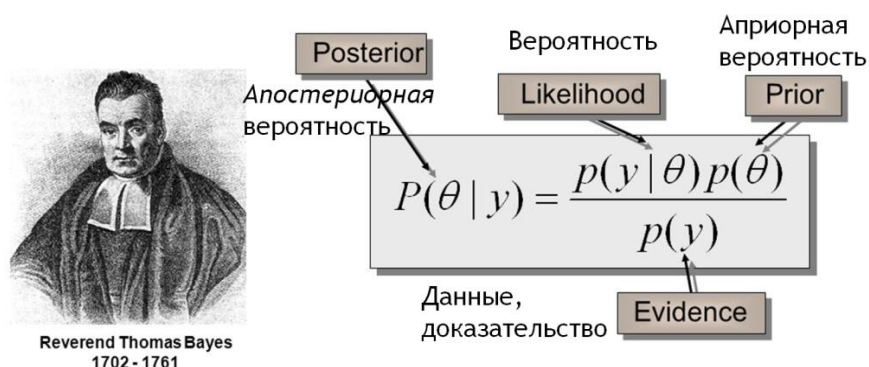


Рис. 2. «Очевидность» и «правдоподобие» в теореме Т.Байеса

Ни в коей мере не отрицая конструктивность и полезность теорий, разработанных на основе байесовских подходов, следует отметить, что они (как и все другие теории самоорганизации) используют данные наблюдения параметров, выбранных человеком для описания свойств объектов, которые тоже выделяются из сложного мира не теоремой Байеса. Для создания «узкого» ИИ, использующего знания человека про объекты и их параметры, теорема Байеса вполне годится. Но СИИ должен уметь самостоятельно, не полагаясь на знания человека и человечества в целом, выделять объекты из сложного мира и описывать их существенными параметрами. А эту проблему теорема Байеса не решает.

Жизнь 3.0




		Способна ли приспособливаться и размножаться?	Способна ли проектировать свои действия?	Способна ли проектировать своё строение?
	Жизнь 1.0 (просто биологическая)	✓	✗	✗
	Жизнь 2.0 (культурная)	✓	✓	✗
	Жизнь 3.0 (технологическая)	✓	✓	✓

Рис. 3. Три стадии эволюции жизни по М. Тегмарку [37]

Но можно ли, пользуясь словами, соответствующим понятиям, созданным человеком, описать свойства СИИ без таких понятий? Наверное, нельзя, но среди созданных человеком понятий можно выделить те, которые относятся к человеческой деятельности напрямую, и те, роль человека в которых ограничивается формулировкой этих понятий. Можно определять СИИ как «сильный искусственный интеллект (СИИ, AGI) – это вид

искусственного интеллекта (ИИ), который может выполнять широкий спектр когнитивных задач *так же хорошо или лучше, чем люди*» [24,25], а можно определить СИИ без отношения к человеческой деятельности, как новый этап самоорганизации, как предложил М. Тегмарк [37] (см. рис. 3).

В своей книге Тегмарк предполагает разделить всех живые организмы на 2 уровня, а создание третьего уровня – СИИ – пока только ожидается. Жизнь 3.0 по Тегмарку сможет не только выживать и воспроизводить себе подобных (что может любой живой организм на основе знаний, передаваемых в ДНК). И не только проектировать свои действия (что в разной мере могут все животные, обладающие нервной системой и внешними сенсорами). А целенаправленно улучшать свое строение, используя постоянно совершенствуемые методы проектирования таких улучшений.

Хотя сам М. Тегмарк не выделяет роль накопления знаний в процессе эволюции жизни, легко сопоставить предложенные им этапы развития жизни с тем, что следующий этап требует больших (по объему и качеству) знаний. Каждый этап характеризуется появлением новых методов получения, сохранения и передачи знаний. Новый этап эволюции позволяет получать большие объемы более актуальных знаний. Главное, постоянно возрастают темпы получения новых знаний, позволяющие выживать во все более быстро изменяющейся среде. При этом все этапы жизни используют знания, накопленные на предыдущих этапах, никогда (кроме момента начала) процесс познания не начинается с нуля.

Тегмарк классифицирует этапы жизни по возможностям, которыми они обладают, не рассматривая вопрос, чем обеспечивается наличие возможностей. Но на такой вопрос легко найти ответ: всякие возможности реализуются на основе знаний и материальных структур, позволяющих преобразовать знания в действия. Как и Шредингер [17], астрофизик Тегмарк придерживается взглядов, что проблемы познания находятся за пределами человеческого понимания и предпочитает обсуждать свойства поведения, а не недоступные для наблюдения субъективные знания. Вне зависимости от того, что является носителем знаний, ДНК, структура весов связей в нервной системе или большие данные в компьютерах. Принято, и не без основания, считать, что разобраться в знаниях в любой форме очень трудно, к успеху ведут не бесплодные попытки разобраться в сложных зависимостях, а выделение и изучение простых свойств окружающего мира.

Но времена меняются, и развитие компьютерной техники, вычислительной математики и нейросетевых алгоритмов (как ее части) позволяют решать всё более сложные задачи обработки больших данных. А химики и биофизики разрабатывают всё более тонкие методы получения таких данных. Которые, в числе прочего, начинают выводить процесс познания из числа безнадежных для исследования.

Генетический отбор

Физика и химия позволяют из свойств атомов выводить свойства чистых веществ и различных соединений, а для объяснения повторяемости форм жизни знания законов физики и химии недостаточно. Наследование у живых организмов, похожесть детей на своих родителей никогда не отрицались, но научно описать законы наследования признаков длительное время считалось невозможным. Законы наследования Менделя, доложенные им 8 февраля 1865 г. брюннскому Обществу естествоиспытателей, были опубликованы и разосланы в 120 библиотек европейских университетов. Мендель дополнительно разослал свои результаты 40 крупным исследователям-ботаникам. Но работы Менделя не вызвали интереса.

Только через 35 лет, с развитием представлений о генах, в 1900 г. К. Корренс, Х. де Фриз и Э. Чермак, независимо друг от друга, заново открыли выведенные Менделем законы наследования [27].

Генетика, как наука, основанная на экспериментальных наблюдениях, бурно развивалась в первой половине XX в., но прорыв случился в 1953 г., когда Джеймс Уотсон и Фрэнсис Крик выполнили работы по расшифровке рентгенограмм ДНК. Датой открытия двойной спирали ДНК считается 28 февраля 1953 г. Статья Уотсона и Крика (ставших впоследствии нобелевскими лауреатами) была опубликована в журнале *Nature* 25 апреля 1953 г. [28]. Понимание строения ДНК открыло широкие перспективы для изучения и направленного вмешательства в структуру ДНК.

К середине 1970-х гг. появились первые методы секвенирования ДНК – экспериментального определения последовательного расположения оснований нуклеиновых кислот (А, Т, G и С), кодирующих различные белки в живой клетке. За почти 50 прошедших с тех пор лет успело смениться 3 поколения методов секвенирования ДНК [29], а стоимость одной процедуры снизилась более, чем в 10 тыс. раз. Полный набор кодирующих и некодирующих последовательностей в ДНК человека называется геномом.

В 1990 г. стартовал проект «Геном человека» (The Human Genome Project, HGP) – международный проект, главной целью которого было определение последовательностей пар оснований ДНК генома человека. В обосновании проекта говорилось: «знание человеческого генома так же необходимо для прогресса медицины и других наук о здоровье, как знание анатомии необходимо для ее нынешнего состояния». В 2003 г. было объявлено о полной расшифровке генома человека, хотя и с добавкой «в основном» (85%). Вопросы о структуре и функциях ряда компонент геномов человека и животных остаются полем для исследований [30], но при этом успехи в данной области с каждым годом всё более впечатляющие.

Возможность контролировать содержание участков ДНК позволила перейти к изменению структуры и управлению функциями генома. В 2012-13 гг. появились принципиально новые методы манипулирования генетическим материалом, основанные на применении технологии CRISPR-

Cas [31], в которых в дальнейшем стали применять белок Cas9. Параллельно с этим развиваются методы управления экспрессией генов, позволяющие изменять синтез белков, менять дифференциацию клеток и получать стволовые клетки из соматических или иных клеток путем эпигенетического перепрограммирования.

Вопросы управления экспрессией генов (геном у организма один, но разные гены проявляют свою активность – экспрессию – в разное время и в соответствии с расположением в структуре организма) рассматривались со второй половины XX в., но за последнее десятилетие продвижение вперед особенно впечатляет [32]. Практически любые клетки организма (содержащие полный геном) могут за счет регуляции экспрессии генов быть преобразованы в любую из широкого (и продолжающий расширяться) списка узкоспециализированных, дифференцированных клеток. Это открыло путь к восстановлению поврежденных тканей отдельных органов животных и человека и позволяет сильно ускорить процесс селекции. Больше не требуется ждать достижения организмом зрелости, любая клетка может быть регуляцией экспрессии генов превращена в яйцеклетку или сперматозоид.

Современные успехи секвенирования, анализа, редактирования и регуляции экспрессии генов дают «белковое» направление создания «Жизни 3.0». Генетические эксперименты на человеке запрещены законодательством ряда стран, но возможности – уже есть. И эти запреты уже не нарушались [33] и будут нарушаться в дальнейшем.

Вне зависимости от запретов технологии генной инженерии продолжают развиваться, и человек, как биологический вид их освоивший, вроде бы начинает соответствовать определению Жизни 3.0 Тегмарка – вот же они, возможности проектировать свое строение! Причем даже не обязательно экспериментировать на человеке, можно развить шимпанзе, дельфина или осьминога до уровня «разумности», превышающий человеческий, и передать «брату по разуму» имеющиеся технологии. Но проще и эффективнее всё-таки развивать человека – его структура тела уже сейчас хорошо согласуется с материальной базой построенной нами цивилизации.

Учение о селекции применительно к человеку – это евгеника, которая в современном его понимании зародилось в Англии, его лидером был Френсис Гальтон – двоюродный брат Чарльза Дарвина. Евгеника призвана бороться с явлениями вырождения в человеческом генофонде, которые всё сильнее проявляются по мере развития цивилизации. Одной из наиболее важных причин вырождения является развитие медицины и снижение на ее основе смертности, в первую очередь – детской. Если еще полтора-два столетия назад в младенчестве и детстве погибало от болезней и травм более 50% родившихся детей, чем осуществлялся естественный отбор, выживали наиболее приспособленные. Сейчас детская смертность в развитых странах составляет менее 1%, выживают и дают потомство практически все, в том числе те, кто без серьезного медицинского вмешательства не

произвел бы потомства, что создает дополнительную нагрузку на здравоохранение. И без борьбы с явлениями вырождения (которая уже сейчас активно ведется) ситуация будет только ухудшаться.

Далеко не все методы осуществления селекции людей можно признать цивилизованными. Так, «отбраковка» слабых детей, осуществлявшаяся, согласно преданиям, в древней Спарте и преступные нацистские практики расовой гигиены, эксперименты над людьми и уничтожение «нежелательных» социальных групп как «методы селекции» в современном обществе неприемлемы. Тем не менее, существуют и цивилизованные подходы борьбы с вырождением, и они не только применяются сегодня, но и имеют тенденцию к расширению круга решаемых ими задач.

В Европе действуют «Запрет евгенической практики в Европейском союзе» (2000) и «Конвенция о биомедицине и правах человека» (2005), запрещающие, например, «создание эмбрионов человека в исследовательских целях». Тем не менее, в большинстве цивилизованных стран, в том числе и в РФ, практикуются и приветствуются пренатальная диагностика – дородовая диагностика плода для обнаружения патологии на стадии внутриутробного развития. И преимплантационная генетическая диагностика (до имплантации в матку) эмбриона при экстракорпоральном оплодотворении (рис. 4 а). Законодательство большинства стран допускает прерывание беременности (любых сроков) по медицинским показаниям при наличии информированного согласия родителей.

Экстракорпоральное оплодотворение (ЭКО) получило распространение как используемая в случае бесплодия репродуктивная технология. С развитием методов генной инженерии область ее применения может значительно расшириться. При ЭКО яйцеклетку извлекают из организма, оплодотворяют искусственно «в пробирке» и полученный эмбрион в течение 2-5 дней содержат в инкубаторе, где он развивается, затем эмбрион помещают в полость матки [34]. Поскольку не все эмбрионы выдерживают процедуру переноса, обычно оплодотворяют и пересаживают в матку несколько зародышей, что иногда приводит к рождению 2, 3 и более детей.

На 3-4 день после оплодотворения эмбрионы содержат достаточное число клеток, чтобы было доступно взятие генетической пробы без вреда для эмбриона. Удешевление процедуры секвенирования генома позволяет провести сравнительный анализ геномов нескольких зародышей, а диагностика по коду ДНК дать пусть вероятностные, основанные на накопленных корреляциях, но прогнозы характеристик зародышей после их созревания. Эти данные позволят родителям осуществлять «информированный выбор» не только по медицинским показаниям, но и на основе эстетических и социальных представлений о желательных свойствах своих детей.

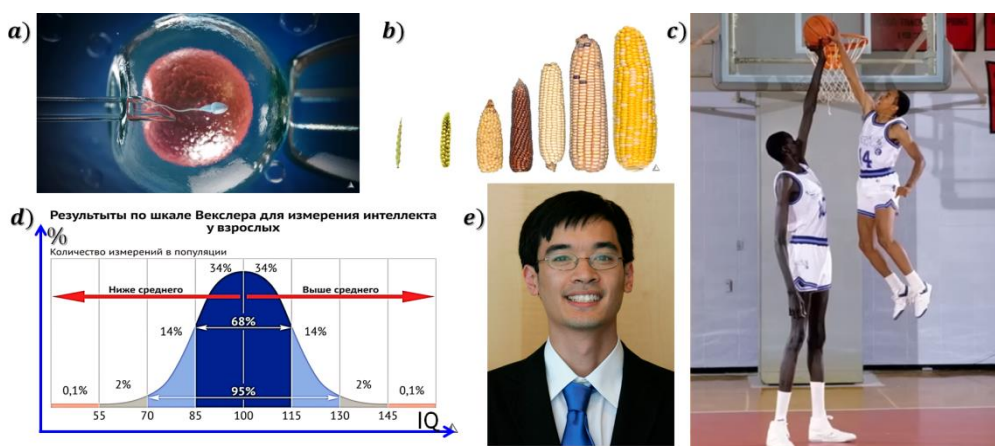


Рис. 4. Возможности создания «углеродной» жизни 3.0:
 а) ИКСИ (интрацитоплазматическая инъекция сперматозоида); б) селекция кукурузы; в) рост и прыгучесть, как наследственные свойства; д) распределение коэффициента IQ; е); Теренс Чи-Шен Тао, американский математик австралийского происхождения с китайскими корнями, IQ=220-230

Регуляция экспрессии генов позволяет пойти дальше: клетки эмбрионов, давшие лучшие (по оценке родителей и врачей) прогнозные характеристики плода можно опять преобразовать в яйцеклетки и еще раз оплодотворить. Процедуру можно повторить несколько раз, выбирая эмбрионы с лучшими (с точки зрения анализа их кодов ДНК) характеристиками. Практика показывает, что многоступенчатая селекция позволяет достичь более высоких показателей, чем однократный отбор (который, впрочем, тоже дает некоторый эффект). Использование регуляции экспрессии генов не требует десятилетий для нескольких ступеней отбора – даже при 5-7 ступенях отбора речь может идти о неделях. Это даже быстрее, чем генетический отбор мух-дрозофил, цикл репродукции которых (с учетом метаморфоз, как у всех насекомых) составляет порядка 30 дней.

Пока что такой многоступенчатый отбор на практике невозможен (по ряду причин, приводимых ниже), но ни одна из них не является принципиально неустранимой и уже в ближайшие годы можно ожидать коммерческих предложений по направленному формированию свойств планируемых для зачатия детей. Сначала методом одноступенчатого отбора эмбрионов по анализу кода ДНК, а затем и целенаправленной многоступенчатой селекции с использованием донорского материала. Причины, почему это (открыто) не делается сейчас, следующие:

1. Недостаточно изучен геном человека, что не позволяет делать надежные оценки жизнеспособности плода по его ДНК-кодам. И, если вероятность появления здорового ребенка на основе геномов родителей, доказавших свою жизнеспособность фактом достижения зрелости, стремится к 100%, то с каждой следующей ступенью, пусть получившие хорошие

оценки в тестах, но не проверенные на жизнеспособность экспериментально (жизнью хотя бы до 16-25 лет) геномы будут иметь всё более низкую вероятность выживания. По мере накопления знаний про геном человека этот риск можно будет всё лучше контролировать, и появятся желающие рискнуть.

2. Одно из направлений борьбы с генетическими дефектами состоит в исключении близкородственного скрещивания. Это означает, что родители при многоступенчатом отборе смогут поучаствовать своим генетическим материалом не более одного раза для каждого из формируемых («в пробирке», *in vitro*) эмбрионов. Неважно при создании эмбрионов для начальной, промежуточной или конечной ступени отбора. Остальной генетический материал – от доноров, не являющихся родственниками ни отцу, ни матери. Родители могут использовать свой генетический материал на первой ступени, но тогда доля «донорского» генофонда с каждой ступенью отбора будет увеличиваться, а родительского – уменьшаться. Если на первой ступени родительские гены составят 100%, то на второй – 50%, затем 25% и далее уменьшение в два раза на каждой ступени. Если же использовать родительский генофонд в конце, то, как минимум, один из родителей должен смириться с ролью генетического «дедушки» (или «бабушки») с вкладом в 25% в геном ребенка. С одной стороны, чем меньше вклад родителей в генофонд, тем больше влияние «рекомендованных» доноров с желательными свойствами генофонда. А готовность родителей быть «далекими» (в генетическом смысле) родственниками своих детей пока что скорее исключение, чем правило. Но на начальном этапе много желающих не потребуется, а успешность экстракорпорального «отбора» может изменить взгляды на важность «генетического» родства.

3. При том, что стоимость полного секвенирования генома упала в десятки тысяч раз, цена прочтения ДНК-кода одного генома продолжает измеряться в сотнях долларов. Если оплатить анализы 2-3 геномов уже сейчас доступно многим, то многоступенчатый отбор потребует несколько тысяч, возможно – десятков тысяч долларов, что сильно ограничивает круг родителей, для которых данная операция по выбору свойств детей будет финансово посильна. Но если процесс снижения цены секвенирования сохранится, то это ограничение тоже отпадет.

4. Возможности выбора свойств детей родителями нуждаются в регулировании, иначе возможен дисбаланс состава населения, как полового, так и различающегося по социальной активности, интеллектуальным и физическим способностям. Если все матери родят голубоглазых и кудрявых блондинов, физически крепких и с неудержимым стремлением к лидерству, то это приведет к социальному взрыву. Ограничения не обязательно должны носить запретительный характер, возможно стимулирование родительского выбора для формирования требуемого государству распределения людей с разнообразными способностями. Такого, чтобы не возника-

ло диспропорции числа мужчин и женщин, руководителей и подчиненных, «ученых» и «спортсменов», хотя последний выбор не антагонистический: хорошие умственные способности не мешают физическому развитию. Пока общество не готово к такому регулированию, но управляемое прогнозирование может помочь сглаживать социальные диспропорции. Страны, внедрившие такое регулирование получают конкурентные преимущества. Следовательно высоки шансы, что родительский выбор свойств детей под государственным контролем в перспективе имеет хорошие шансы распространения.

Направленный генетический отбор (при условии, если удастся свести к минимуму его отрицательные последствия) может помочь в ускорении накопления необходимых для развития самоорганизации знаний. Отбирать можно не только физические свойства организмов (см. рис. 4 b и c), но умственные качества, которые тоже возможно (вероятностно) определять на основе корреляций кода ДНК со способностями людей. Умственные, как и физические качества, можно развивать, и тренируя память, получая образование, и просто живя в цивилизованном обществе, можно научиться решать значительно более сложные задачи, чем те, которые были доступны без развития интеллекта. Но играют роль и врожденные способности, разный начальный уровень которых позволяет развить интеллект до разных высот в различных областях.

Одной из наиболее популярных характеристик умственных способностей человека является уровень IQ (количественная оценка уровня интеллекта человека) [35]. По своему построению он определяется как отклонения от среднего уровня ($IQ = 100$) по качеству и количеству выполненных заданий за ограниченное время. Естественно, что индивидуальные результаты могут быть как лучше ($IQ > 100$), так и хуже ($IQ < 100$) средних значений. Как показано на рис. 4 d, в идеале распределение значений IQ должно быть нормальным и по шкале Векслера 68% результатов должно попадать в диапазон $IQ = 100 \pm 15$. Но при построении шкалы разными авторами используются различные задания, и статистика набирается по разным популяциям. Это приводит к тому, что оценки IQ не носят абсолютно характера и для одного человека, при оценке по разным шкалам IQ, могут отличаться на $5 \div 15$ пунктов и более. Тем не менее, для умственной деятельности IQ является достаточно хорошо формализованным количественным параметром, позволяющим оценивать (и сравнивать) способности к решению не очень сложных «интеллектуальных» задач.

Как и физические способности, которые характеризуются далеко не только физической силой, но и гибкостью, выносливостью, прыгучестью и рядом других параметров, умственные способности тоже не сводятся только к IQ. Не менее важны память, упорство в достижении целей, способность договориться и многое другое. Причем если физические параметры, как правило, удастся измерить количественно, то с умственной деятельно-

стью, за исключением коэффициента IQ (который тоже измеряется приблизительно, но всё-таки количественно), большинство оценок носят качественный характер. Например, память может быть хорошей или плохой, а у некоторых личностей встречается выдающаяся память. Так наш соотечественник Николай Пржевальский и создатель одной из первых вычислительных машин Джон фон Нейман не просто помнили содержание прочитанных ими книг, но могли дословно цитировать их с произвольного места. Это, естественно, не единственные примеры людей с выдающейся памятью в истории, но в целом их заметно меньше, чем обладателей $IQ > 115$.

Есть и много различных характеристик умственной деятельности, таких как целеустремленность, способность находить компромиссы, эмоциональный интеллект, лидерские качества и ряд других. Выраженное наличие данных свойств не менее полезно для умственной деятельности, чем способность быстро решать задачи (что оценивается IQ) и хорошая память. Сочетание нескольких качеств усиливает эффект от каждого по отдельности, но их проявление напрямую не взаимосвязано, они могут проявляться у одного человека в произвольном наборе. Так, Ян Лекун в своем докладе о перспективах СИИ [36] отмечал, что «Интеллект не коррелирует со стремлением доминировать, даже у людей».

Тем не менее, способность выбирать и принимать наиболее эффективные решения (даже сравнивая предложения консультантов) требует наличия развитых умственных способностей. Одной из основных причин разрушения империй называют вырождение правящей касты, когда на руководящие посты начинают попадать люди не по своим способностям руководить, а по принадлежности к руководящей касте. Такое общество начинает принимать менее эффективные решения, чем соседние страны, и проигрывает в конкуренции, что ведет его к краху. Институт дуэлей в древности и в Средние века поддерживал наличие хоть какого-то отбора в правящих кастах, с его утратой процесс деградации поколений правителей стал более заметен. Но даже в Средние века идею «просвещенного монарха» (то есть обладающего знаниями и развитым интеллектом) реализовать в полном объеме не удалось.

Экстракорпоральный многоступенчатый отбор эмбрионов мог бы позволить правящей касте создавать своих потомков, обладающих выдающимися способностями, как физическими, так и умственными. Это позволило бы им выбирать из своих рядов успешных руководителей и реализовать пресловутую идею «просвещенных монархов голубой крови». С одной стороны, это было бы именно то, за что критикуют евгенику – разделение человечества на расу господ и расу рабов «обоснованием добровольного классового компромисса капиталистов и наемных рабочих». Такое разделение только закрепляло бы уже существующую поляризацию населения Земли на бедных и богатых. С другой стороны, имущественное

разделение на планете и без евгеники постоянно увеличивается, и, если нельзя этому противостоять, то хотя бы уменьшить вред, наносимый руководящей кастой планете, можно путем повышения качества человеческого материала этой касты и увеличения способности принимать взвешенные, в большей степени направленные на прогресс цивилизации (что выгодно всем) решения.

Надо при этом понимать, что идея «просвещенного монарха» является утопической, поскольку правящие касты и так прекрасно живут и заинтересованы не в экспериментах на себе, а в сохранении своего привилегированного положения. Более того, способность проводить в жизнь решения, не отвечающие интересам широких масс, рассматриваются ими как особая доблесть. Именно поэтому правящие касты возглавляют движение за совместимость СИИ с человеком (читай – с правящими кастами) [37]. Их пугает, что развитие вышедшего из-под контроля СИИ может выявить неэффективность проводимой правящими кастами политики для развития цивилизации.

Более вероятно развитие и применение технологии экстракорпорального многоступенчатого отбора эмбрионов в странах, недавно вступивших в борьбу за лидирующее положение в мире, таких как Китай, Индия, Бразилия и др. Их не устраивает сложившаяся в мире ситуация. Для победы в конкурентной борьбе с ведущими технологическими державами им необходимо принимать лучшие, чем у соперников, решения. Отчасти это основным странам «неприсоединения» удается уже сегодня, но повышение качества «человеческого материала» руководителей разного уровня в этих странах будет важным аргументом в борьбе за мировое лидерство.

На рис. 4 е приведен портрет американского математика австралийского происхождения с китайскими корнями Теренса Чи-Шен Тао (1975 г.р.), оценка значения IQ которого составляет 220-230. «Характерной особенностью работы Теренса Тао является очень интенсивная совместная работа со многими математиками, одновременная погруженность в самые разные разделы современной математики» [38]. Его способности к решению математических проблем ценятся коллегами настолько высоко, что если где-то возникают неразрешимые проблемы, то считается, что надо просто пригласить Теренса Тао и проблема будет снята. Наличие таких «универсальных решателей проблем» в разных областях деятельности крайне важно для успеха развития страны. В советское время роль такого «решателя» научных проблем в СССР выполнял М.В. Келдыш, который «мог разобраться в любой научной проблеме». Руководство М.В. Келдышем АН СССР в качестве ее президента в течение ряда лет было заметным вкладом в становление СССР в качестве одной из ведущих мировых держав. Конечно, успехи СССР определялись многими факторами, в том числе сталинской кадровой политикой, основанной на отборе руководителей по проявленным ими способностям. И Келдыш был в числе

большой плеяды достойных руководителей, обеспечившей подъем науки и промышленности в СССР. Наметившееся сегодня отставание РФ в значительной степени связано с «утечкой мозгов» и недостатком в стране способных руководителей науки и экономики различных уровней.

Возможности улучшить «человеческий материал» на основе технологий экстракорпорального многоступенчатого отбора эмбрионов есть. Повышать надо не только IQ, но и другие свойства умственной деятельности. Но остановимся на оцениваемом на основе анализа кода ДНК генома эмбрионов IQ. Уже сейчас на основе корреляции возможно различать геномы с прогнозом IQ, отличающимся на 10÷20 единиц. Если начинать с эмбрионов от родителей с IQ = 130 (разброс IQ потомства которых в норме лежит в пределах 110÷150), то за 5÷6 ступеней экстракорпорального отбора можно выявить эмбрионы с прогнозным IQ в районе 160÷190 (что уже очень высокие показатели), а 10÷12 ступеней отбора – порядка 200÷250. Это соответствует уровню Теренса Тао (и, вероятно, М.В. Келдыша, хотя данных по его уровню IQ нет). Наличие сотен, а тем более тысяч ученых и руководителей с такими способностями может резко изменить положение страны в мировом сообществе, способствуя более быстрому и эффективному созданию и внедрению перспективных технологий и развитию страны в целом.

При всей важности влияния (как положительного, так и отрицательного), которое смогут оказать технологии экстракорпорального многоступенчатого отбора эмбрионов, по ряду причин они не составляют передний край развития процессов самоорганизации:

1. Основная проблема состоит в том, что, хотя сам процесс многоступенчатого отбора эмбрионов может быть произведен за недели и месяцы, созревание плода и затем взросление ребенка занимает в сумме многие годы. Даже если в процессе развития не будет выявлено никаких катастрофических побочных эффектов, и прогнозные свойства будут достигнуты, то формирование зрелого специалиста, способного эффективно развивать экономику страны, произойдет не раньше, чем через 20-25 лет. И вопрос совместимости такого «искусственного» интеллекта (с правящей кастой) остается открытым.

2. Вторым недостатком «углеродного» ИИ на основе развития человеческого генома является отсутствие возможности «вставить свои мозги» в чужую голову. В условиях научно-технического прогресса постоянно появляются новые отрасли науки и технологий, в которых не хватает профильных специалистов. Знания же специалистам передаются только путем переподготовки, которая занимает месяцы и даже годы. Это большая задержка – в современном мире изменения происходят быстрее.

3. Наконец, такой «углеродный» ИИ не будет соответствовать Жизни 3.0 по Тегмарку в полной мере, по крайней мере, на начальном этапе его развития. Поскольку мы пока не до конца понимаем зависимости, связы-

вающие коды ДНК со свойствами взрослого организма, возможности «проектирования строения своего организма» есть, но они ограничены имеющимся биологическим материалом.

По сути, пока возможно только совершенствование генетического кода уже существующих организмов, о создании принципиально новых решений речи не идет. Хотя нельзя не отметить, что и это – выдающееся достижение, важность которого трудно переоценить. С развитием генной инженерии всё большее число ограничений на редактирование геномов будет сниматься, но пока они есть и их много.

Кремниевый СИИ – не отбор, а проектирование!

Большинство из перечисленных выше недостатков «углеродного» интеллекта отсутствуют у «кремниевого» ИИ:

1. После разработки изготовление микросхем и печатных плат занимает дни и даже часы, после сборки системы из комплектующих (что тоже измеряется днями и часами, а на конвейере – минутами) и загрузки исполняемых программ (минуты и секунды) система готова к обучению. Процесс обучения более длительный, может составлять недели и месяцы, но не годы и, тем более – не десятилетия.

2. Возможности создания цифровых копий «кремниевого» интеллекта ограничено только числом носителей передаваемых знаний. Время «вставки уже обученных мозгов» в чужую голову при перезаписи на заводе может составлять минуты и секунды, при скачивании «мозгов» по сети – часы, но это не идет ни в какое сравнение с месяцами и годами переподготовки «углеродного интеллекта».

3. При проектировании «кремниевого» СИИ нет генетических ограничений, разработчикам не нужно обеспечивать выживание и воспроизводимость эмбрионов, устойчивость их развития в системе и выполнение ряда других требований, важных для «углеродного» интеллекта.

Наличие некоторых потенциальных преимуществ у «кремниевого» СИИ не означает, что они уже реализованы. На сегодняшний день нет общепризнанной теории мышления, функция сознания вызывает споры. Вне зависимости, описываем ли мы высшую нервную деятельность человека или пытаемся смоделировать мышление и сознание нейросетевыми алгоритмами. Но человек способен мыслить и обладает сознанием (и многими другими, важными психическими функциями), а когда нейросетевые алгоритмы достигнут хотя бы человеческого уровня – вопрос открытый.

Но если еще 10-15 лет назад СИИ был просто темой в научной фантастике, а прогноз Курцвейла [39] о создании СИИ к 2045 г. казался слишком оптимистическим, то теперь созданием СИИ, как основной целью исследований в области ИИ, занимаются такие компании, как OpenAI, DeepMind и Anthropic PBC. В Китае в 2020 г. создан Пекинский институт СИИ (Beijing Institute for General Artificial Intelligence), в РФ вопросы со-

здания СИИ курирует Сбер [40]. И это – только крупные проекты с финансированием в сотни миллионов долларов в год. Более мелких проектов по созданию СИИ насчитывается уже не десятки, а сотни.

Некоторые исследователи продолжают не верить в способности человека осмыслить процесс познания и рассчитывают повторить путь эволюции, в которой процесс определялся естественным отбором, без понимания смысла возникновения таких психических явлений, как мышление и сознание. Эволюционные алгоритмы [41] позволяют, как и в генной инженерии, производить оптимизацию свойств системы, не понимая глубоко, каким образом эти свойства реализуются. Но таким путем оптимизация происходит не быстро (время занимает в основном не создание вариантов реализации системы, а проверка эффективности ее работы), и выиграть у генной инженерии «углеродного» интеллекта будет сложно, поскольку там используются хорошо проверенные эволюцией «детали».

Не просто лучшим, а единственно правильным путем представляется создание СИИ методами проектирования. Целенаправленное создание СИИ из компонент с полностью понятными функциями – вот прямой путь, который дал человечеству преимущества над животными при решении почти всех остальных проблем, на порядки превосходящий по скорости все остальные способы достижения прогресса.

На что надеются разработчики СИИ? Да, сегодня нейросетевые алгоритмы позволяют решать удивительно сложные задачи, но ведь это всё равно бездушная машина, которая просто выполняет алгоритмы и не понимает, что она делает. Как преодолеть грань, отделяющую неживое от живого, научить машину мыслить, дать ей сознание? Ответы на это дает анализ эволюции сложных систем и наука их изучающая – синергетика.

Синергия и эмерджентность – новые свойства, проявляемые при совместном использовании нескольких компонент

Потрясающие успехи техники, которые в значительной мере основаны на научных открытиях, сформировали общественное мнение, что именно открытие базовых законов природы ведет к пониманию происходящих процессов и позволяет осуществлять прогресс технологий и развитие цивилизации. Это действительно так, но отражает только одну сторону прогресса. Получение знаний о законах природы (например, выявление электромагнитных явлений) только открывает новое поле деятельности по развитию технологий. Создание и совершенствование новых технологий требует значительно больших усилий и времени. Проектирование самых простых устройств (к примеру, мышеловки или доводчика двери) требует применения междисциплинарных знаний, соединения деталей с различными свойствами для получения требуемого эффекта. При этом фактические знания (о наличии, параметрах, возможностях обработки и пр.) комплектующих и материалов изделия не менее важны, чем теоретические.

Открытие новых законов расширяет возможности инженеров, позволяет решать задачи, которые до этого были трудны или совсем недоступны. Это, естественно, ускоряет прогресс, но он, пусть медленнее, но шел и до эпохи научных открытий. Объединяя различные детали и явления, человеку удавалось получить новые свойства предметов и последовательностей действий. Так, привязав к палке острый камень, древний человек получил каменный топор, который обладал свойствами, заметно превосходящими те, которые были у палки и камня по отдельности. Подавляющее число наименований производимой продукции путем соединения частей и/или придания им специальной формы позволяет достичь усиления имеющихся или получения новых свойств по отношению к необработанному материалу.



Рис. 5. Эволюция природы, жизни и информационной техники: в процессе усложнения систем появляются новые свойства, жизнь и техника эволюционирует на основе накопления знаний

В теории самоорганизации (синергетике) явления усиления имеющихся или получения новых свойств системы по отношению к имеющимся у ее компонент свойствам носят названия *синергии* и *эмерджентности*. Часто (ошибочно) трактуют данные термины, как проявляющиеся исключительно случайно (неожиданно), в результате развития процессов самоорганизации. Случайность синергии и эмерджентности обычно относится только к возникновению ситуаций ненаправленного, естественного создания условий протекания процессов самоорганизации, а сами процессы протекают вполне закономерно. Знания как раз для того и нужны, чтобы целенаправленно, детерминировано создавать условия для закономерного выполнения процессов самоорганизации.

Если известны условия проявления синергии и эмерджентности, то их можно и нужно целенаправленно создавать для достижения желаемого результата. Дело за малым – надо только понять, как наличие мышления и сознания у животных и человека создают эмерджентно новые возможности по формированию более рационального поведения, в сравнении с использованием механизмов условно-рефлекторной деятельности.

Суть парадокса

Наука стремится познать законы, описывающие свойства реального мира. В своей книге [26] М. Тегмарк предсказывает, что после создания СИИ прогресс науки ускорится и возможно достижение исчерпывающего знания всех основных законов природы. Будет ли это означать завершение процесса познания?! Астрофизик Тегмарк полагает что да, даже описывают возможность встречи двух цивилизаций, достигших предела познания и делает вывод, что они будут равны по уровню развития.

Но это не так... Как люди, так и страны с одинаковым уровнем знаний обладают очень сильно различающимися возможностями (которые и определяют уровень развития). Понимание основных законов природы без наличия знаний и средств их использования для достижения своих интересов является просто игрой ума и влияет на уровень развития общества только потенциально, если такие знания и средства удастся создать.

Появившиеся в сети виртуальные миры (и даже настольные логические игры) позволяют смоделировать ситуацию «исчерпывающего» знания о мире. Объяснение правил игры в шахматы ребенку, ранее их не знавшему, займет несколько минут. Но полное знание правил шахмат совсем не делает из него хорошего игрока, а тем более гроссмейстера.

В чем состоит проблема? Все фигуры ходят просто (и известно как), но число вариантов расположения фигур на доске превосходит пределы, позволяющие описать все варианты и выбрать для каждого наилучший ход. Зато есть возможность для стоящей на доске позиции проанализировать, какие ходы приведут к ее улучшению. Если перебирать все возможные ходы, то глубоко проанализировать варианты не удастся даже суперкомпьютеру. Следует анализировать только лучшие продолжения [42].

Но, если нам известны лучшие ходы, почему сразу не сделать ход без дополнительного анализа? Дело в том, что даже в шахматах (в миттельшпиле) позиции не повторяются, прогноз, какие ходы будут лучшими, позволяет только уменьшить список анализируемых продолжений. Выяснить, какой из сокращенного списка ход действительно является лучшим, можно только путем анализа продолжений.



Рис. 6. Виртуальный мир Майнкрафт (Minecraft) и настольная игра

В реальном мире ситуация аналогичная. Дело не в том, что мы не знаем свойства отдельных объектов и явлений и способов их изменять («двигать фигуры»). А в том, что ситуации в сложном мире не повторяются и для выбора лучшего «хода» приходится анализировать его близкие и дальние последствия (с целью улучшения выбора последовательности действий).

Не вызывает сомнения, что реальный мир сложнее шахмат и статистически достоверными методами мы не можем для каждой ситуации запомнить наилучшую последовательность действий. Но еще одна причина более высокой сложности мира состоит в том, что если в шахматах мы можем знать свойства всех фигур (и другие не появятся), то в реальном мире новые «фигуры» (объекты и явления) встречаются довольно часто.

И дополнительной проблемой преодоления сложности реального (или виртуального) мира является необходимость выделения из мира отдельных «фигур», свойства которых могут быть статистически достоверно изучены за ограниченное время, для включения их в процесс анализа позиции.

Таким образом, суть парадокса состоит в невозможности статистически достоверного изучения свойств мира «как целого», что приводит к необходимости выделять в нем отдельные «фигуры» (объекты и явления), свойства которых возможно познать и использовать при уточнении планирования последовательности действий.

Возможность декомпозиции

Успехи современных нейросетевых алгоритмов удивляют именно потому, что даже огромных массивов данных, используемых для обучения нейросетей не должно хватать (согласно расчетам) для накопления достоверной статистики про преобразования, которые ими затем выполняются. Как же нейросетям удастся обучаться и успешно решать задачи?!

Рациональным объяснением удивительных свойств нейросетей является предположение, что пусть значительно продвинутый по сравнению с

[43], но всё равно универсальный алгоритм оптимизации параметров нейросетей поддерживает любые способы приближения к оптимальным значениям. И декомпозиция сложных сигналов – один из таких способов. А «глубина» (большое число слоев, в современных нейросетях – сотни) нейросетям нужна именно для того, чтобы повысить вероятность возникновения декомпозиции сложных сигналов.

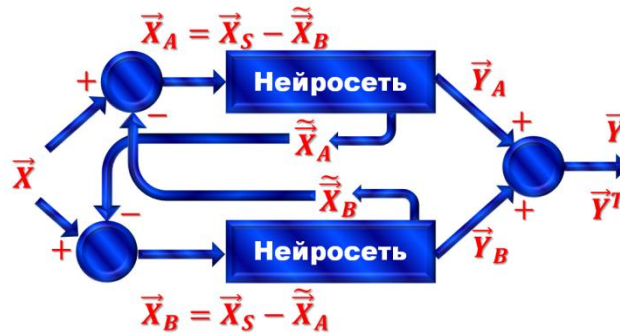


Рис. 7. Декомпозиция сложного сигнала на основе разностной схемы

Важным шагом в развитии нейросетевых алгоритмов станет развитие структур, в которых декомпозиция формируется не вероятностным образом, а направленно, как это показано на рис. 7. Если нейросеть, обученная компоненте A , воспроизводит ее по сложному сигналу X (состоящему из компонент A и B) и затем воспроизведенная A вычитается из X , то вторая нейросеть сможет значительно быстрее выявить свойства компоненты B . Обучение можно будет считать успешным, когда $\tilde{X}_B = \tilde{X}_S - \tilde{X}_A$ позволит воспроизводить \tilde{X}_B такое, что $\tilde{X}_A = \tilde{X}_S - \tilde{X}_B$ станет равным \tilde{X}_A .

Декомпозицию можно применять и к более сложным ситуациям, содержащим намного больше компонент, тогда

$$\tilde{X}_i(t) = \tilde{X}(t) - \sum_{j \neq i}^N \tilde{X}_j(t).$$

Важным свойством приводимой на рис. 7 разностной схемы декомпозиции является нелинейность суммирования и вычитания, в связи с чем эти операции тоже необходимо осуществлять нейросетевыми средствами.

Более подробно рассмотрение принципов работы «разностной» схемы (рис. 7) проводится, например, в [44].

Формирование моделей отдельных объектов и явлений на пространственно разнесенных нейросетях (которых для работы в сложном мире должно быть много) позволит использовать их не только для непосредственного формирования действий в ответ на входной сигнал, но и для моделирования вариантов развития ситуации при различных воздействиях на ее компоненты. Процесс моделирования, как отмечалось выше, необходим

для улучшения выбора последовательности действий и может быть сопоставлен с мышлением.

Наличие выбора: сразу действовать (на основе аппроксимации $\bar{X} \rightarrow \bar{Y}$) или предварительно подумать (провести варианты моделирования) приводит к необходимости этот выбор контролировать. Это алгоритмически вполне реализуемо и может служить аналогом сознания, которое у людей постоянно направлено на оценку сложившейся ситуации и решения проблемы: уже действовать или еще подумать.

Выводы

Человечество с доисторических времен начало создавать себе различные инструменты и средства для повышения возможностей выживания и воспроизводства потомства. Развитие этого процесса привело к созданию цивилизации, и на современном этапе инструменты создаются не только для выполнения физической, но и умственной работы. Это приближает нас к созданию сильного искусственного интеллекта. По оптимистичным оценкам ждать осталось недолго – 1-2 года.

Возможны различные пути реализации СИИ, включая «углеродный», на основе успехов генной инженерии. Но наиболее перспективным, доступным к реализации в ближайшие годы является путь развития нейросетевых алгоритмов. Среди причин, по которым нейросети сейчас являются передним краем процессов самоорганизации, можно выделить две: возможность их создания путем проектирования (а не отбора, как в генной инженерии) и соответствие теореме о необходимой сложности [45], поскольку количество параметров сетей формальных нейронов не ограничено генетикой.

Наиболее вероятной страной, которая создаст СИИ является Китай, поскольку там совпадает заинтересованность в его создании для укрепления своей позиции в борьбе за мировое лидерство с научными, технологическими и экономическими возможностями для разработки теории, конструирования и воплощения в жизнь систем СИИ.

Литература

1. Мифы о Менделееве. [https://ru.wikipedia.org/wiki/Менделеев, Дмитрий Иванович#Мифы о Менделееве](https://ru.wikipedia.org/wiki/Менделеев,_Дмитрий_Иванович#Мифы_о_Менделееве)
2. https://psyera.ru/filosofiya-atomistov-levkip-p-i-demokrit_15993.htm
3. Богданов А.А. Тектология – Всеобщая организационная наука. // Берлин – Санкт-Петербург, 1922. (Переиздание: – М.: Экономика, 1989)
4. Эшби У.Р. Принципы самоорганизации // Принципы самоорганизации / Под ред. д.т.н. А.Я. Лернера. – М.: Мир, 1966. С.314-343.

5. *Анохин П.К.* Принципиальные вопросы общей теории функциональных систем... // 1 Принципы системной организации функций. – М.: Наука, 1973. С.5-6.
6. *Пригожин И.* Введение в термодинамику необратимых процессов. – Спрингфилд, Иллинойс: издательство Чарльза К. Томаса, 1955.
7. *Haken H.* Synergetics: an introduction: Nonequilibrium phase transitions and self-organization in physics, chemistry, and biology. – Berlin New York: Springer-Verlag, 1978.
8. *Eigen M., Schuster P.* The hypercycle: A principle of natural self-organization. – New York: Springer-Verlag Berlin Heidelberg, 1979.
9. *Maturana H.R., Varela F.J.* Autopoiesis and cognition: The realization of the living. – Springer Science & Business Media, 1991.
10. *Jantsch E.* Technological planning and social futures. – London, SW: Associated Business Programmes Ltd., 1972
11. *Laszlo E.* Introduction to systems philosophy. Toward a new paradigm of contemporary thought. – New York London Gordon & Breach, 1972.
12. *Bak P, Tang C, Wiesenfeld K.* Self-organized criticality: An explanation of the 1/f noise // Phys. Rev. Lett. 1987, 59(4), 381-384.
13. *Князева Е.Н., Курдюмов С.П.* Основания синергетики. Режимы с обострением, самоорганизация, темпомиры. – СПб.: Алетейя. – 414 с.
14. *Капра Ф.* Дао физики. Исследование параллелей между современной физикой и восточной философией. – М.: Манн, Иванов и Фербер, 2017. – 419 с.
15. *Дубровский Д.И.* Проблема «Сознание и мозг» теоретическое решение. – М.: «Канон+» РООИ «Реабилитация», 2015. – 208 с.
16. *Бауэр Э.С.* Теоретическая биология. – М.-Л.: ВИЭМ, 1935. – 149 с.
17. *Шредингер Э.* Что такое жизнь? Физический аспект живой клетки / Издание третье, дополненное и исправленное. – Москва-Ижевск: НИЦ «Регулярная и хаотическая динамика», 2002 – 92 с.
18. *Галимов Э.М.* Феномен жизни: Между равновесием и нелинейностью. Происхождение и принципы эволюции. – М.: Едиториал УРСС, 2006. – 256 с.
19. *Friston K.* The free-energy principle: a unified brain theory? Nat Rev Neurosci 2010, 11, 127–138.
20. *Stone J.V.* Bayes' rule: A tutorial introduction. – University of Sheffield, England, 2013.
21. *Kullback S., Leibler R.A.* On information and sufficiency // Annals of Mathematical Statistics. 1951, 22(1), 79–86.
22. *Ветров Д.П., Кропотов Д.А.* Байесовские методы машинного обучения / Учебный курс. – МГУ имени М.В. Ломоносова, 2007. <https://bayesgroup.github.io/bmml/2016/BayesML-2007-textbook-1.pdf>
23. *Barbeau E.* Fallacies, flaws, and Flimflam: The problem of the car and goats // The College Mathematics Journal. 1993, 24(2), 149-154.

24. https://en.wikipedia.org/wiki/Artificial_general_intelligence
25. *Heaven W.D.* Google DeepMind wants to define what counts as artificial general intelligence // MIT Technology Review. Retrieved 1 March 2023.
26. *Tegmark M.* Life 3.0: Being human in the age of artificial intelligence / First ed. – New York: Knopf, 2017.
27. *Гайсинович А.Е.* Грегор Мендель (биографический очерк) // Грегор Мендель. Опыты над растительными гибридами / ред. Б.Л. Астауров. – М.: Наука, 1965. С.118-132.
28. *Crick F., Watson J.* Molecular structure of nucleic acids: A structure for deoxyribose nucleic acid // Nature. 1953, 171(4356), 737-738.
29. *Бородинов А.Г. и др.* Поколения методов секвенирования ДНК (обзор) // Научное приборостроение. 2020, 30(4), 3-20.
30. *DeLisi C.* Genomes: 15 years later // Human Genome News. 2001. 11, 5-6.
31. *Barrangou R.* The roles of CRISPR-Cas systems in adaptive immunity and beyond // Current Opinion in Immunology. 2015, 32, 36-41.
32. *Surat P.* A guide to understanding gene expression. <https://www.azolifesciences.com/article/A-Guide-to-Understanding-Gene-Expression.aspx>
33. *Greely H.* CRISPR'd babies: Human germline genome editing in the “He Jiankui affair” // Journal of Law and the Biosciences. 2019, 6(1), 111-183.
34. *Кузьмичёв Л.Н., Штыря Ю.А.* Экстракорпоральное оплодотворение. Только факты. Информация к размышлению. – М.: Специальное издательство медицинских книг, 2012. – 128 с.
35. *Kaufman A.S., Kaufman J.C., Liu X., Johnson C.K.* How do educational attainment and gender relate to fluid intelligence, crystallized intelligence, and academic skills at ages 22–90 years? // Archives of Clinical Neuropsychology. 2009, 24(2), 153-163.
36. *LeCun Y.* A path towards autonomous machine intelligence <https://yandex.ru/video/preview/398815152026420216>
37. *Russell S.* Human compatible: Artificial intelligence and the problem of control. United States: Viking, 2019.
38. *Теренс Т. (陶哲軒).* https://ru.wikipedia.org/wiki/Тao,_Теренс
39. *Kurzweil R.* The singularity is near. – New York: Viking Books, 2005.
40. *Ведяхин А. и др.* Сильный искусственный интеллект : На подступах к сверх-разуму. – М.: Интеллектуальная Литература, 2021. – 232 с.
41. *Саймон Д.* Алгоритмы эволюционной оптимизации. – М.: ДМК Пресс, 2020. – 940 с.
42. *Silver D., Hubert T., Hassabis D. et al.* Mastering chess and shogi by self-play with a general reinforcement learning algorithm. [arXiv: 1712.01815](https://arxiv.org/abs/1712.01815)
43. *Rumelhart D.E., Hinton G.E., Williams R.J.* Learning internal representations by error propagation // Parallel distributed processing. V.1, pp. 318-362. Cambridge, MA, MIT Press. 1986.

44. *Smolin V, Sokolov S.* AGI's hierarchical component approach to unsolvable by direct statistical methods complex problems // Engineering Proceedings. 2023, 33(1), 67.
45. *Ashby W.R.* An introduction to cybernetics. – Chapman & Hall, 1956.