

Одушевленные машины

От научной фантастики к научной проблеме



Андрей КОЛЕСНИКОВ,
кандидат философских
наук, доцент

Каждый из нас – личность – несет в себе свое неповторимое «я», то, что принято называть душой. Мы пока не понимаем до конца природы наших субъективных ощущений, что такое наше «я», где находится, на чем основано, но совершенно четко знаем, что оно есть. Весь мир мы делим на неодушевленные предметы и одушевленные сущности. Человек и даже комар – это «кто», а машина, даже самый мощный и совершенный современный суперкомпьютер – это «что». Можно ли изменить такой порядок вещей и превратить машину в одушевленный предмет?

Организм и механизм

Мы привыкли относить все, что связано с душой, к сфере действия высших непознаваемых сил. Но так ли это? Коль скоро существует природное явление (доподлинно известно, что каждый живой субъект ощущает себя «я»), то, безусловно, есть и его научное объяснение. А если существует причинный механизм, природная детерминация «я», то и нет каких-либо принципиальных преград для овладения и использования этого природного явления при создании машин.

До настоящего момента проблема создания одушевленных машин (обозначим их термином «пси-машины») в науке обозначена недостаточно определенно. Следует подчеркнуть, что она представляет собой нечто иное и существенным образом отличное от

традиционной области работ по искусственному интеллекту. В данном случае в центре внимания находится не столько решение плохо формализуемых, человекомерных задач или имитация мыслительных процессов, сколько моделирование субъекта – «я» [1; 2]. Новизна постановки проблемы создания одушевленных пси-машин состоит в ее тесной увязке с необходимостью раскрытия материальной природы феномена естественного «я» и его рассмотрения как фундаментального физического свойства, изначально присущего материи.

С началом промышленной эры дух и машина, организм и механизм в общественном сознании резко противопоставляются друг другу. Это противопоставление находит отражение в философской литературе [3] и в культуре в целом. Схожие негативные тенденции в восприятии мира техники современным гуманитарным сознанием проявляются и в отношении компьютеров и информационной революции [4]. Компьютер представляется полной противоположностью, антиподом живого человеческого сознания. Высказываются мысли о том, что он разрушает гармонию духовного бытия человека, уничтожает его внутренний мир, лишает радости истинного человеческого общения, примитивизирует мышление.

При всем очевидном различии человек и машина, организм и механизм имеют и поразительное сущностное сходство. Части организма, как и части машины, функционируют по одинаковым принципам механики и законам физики. Просто столь заметные различия между машинами и организмами обусловлены весьма разной

ОБ АВТОРЕ

КОЛЕСНИКОВ Андрей Витальевич.

Родился в 1961 году в г. Рига. В 1984 году окончил Белорусский технологический институт имени С.М. Кирова (ныне – Белорусский государственный технологический университет), в 1991 году – аспирантуру при Институте философии НАН Беларуси.

С 1984 года работал в Институте генетики и цитологии АН БССР. В 1995–2000 годах – учитель информатики гимназии № 5 г. Минска.

С 2000 года – преподаватель информатики в Белорусском институте правоведения. С 2010 года – доцент кафедры управления информационными ресурсами Академии управления при Президенте Республики Беларусь.

С 2012 года – докторант Института философии НАН Беларуси.

Кандидат философских наук (1998), доцент (2005).

Автор свыше 130 научных работ, в том числе одной монографии.

Сфера научных интересов: философские вопросы естествознания и техники; синергетика; междисциплинарные исследования; проблемы образования; информатика и компьютерное моделирование; проблемы биологической эволюции; искусственный интеллект.

степенью совершенства. Организмы, прошедшие несоизмеримо более долгий путь эволюции по сравнению с машинами, пока намного совершеннее последних. Однако организмами движет и управляет также не божественное провидение, но вполне определенные материальные силы.

Существенным шагом на пути познания этих сил явилось открытие «животного электричества». Следовательно, сила, которая теперь движет и управляет нашими машинами, была некогда заимствована у природы, а именно – у животных организмов. Они буквально пропитаны электричеством, которое циркулирует по нервам и заставляет сокращаться мышцы, приводя в движение тело.

Вскоре после открытия животного электричества публично проводились бесчеловечные, но весьма показательные опыты, заключавшиеся в подсоединении электрического тока к трупам только что казненных преступников. При этом части мертвого обезглавленного организма вновь приходили в движение и создавали иллюзию оживления. Наблюдение за одним из подобных опытов и вдохновило писательницу Мэри Шелли на создание образа ужасного и несчастного чудовища, собранного из фрагментов различных мертвых тел и оживленного при помощи электричества.

Растительные организмы в отличие от животных статичны. Электричество в их жизнедеятельности играет не столь заметную роль. Растительный организм в основном структура, в то время как животный – процесс, причем процесс электрический (выражаясь точнее – электрохимический). Животный организм активно перемещается в пространстве и должен быстро реагировать на внешние события в реальном времени. Поэтому в основу его функционирования заложен некоторый быстрый процесс, а именно – спонтанная циркуляция электричества по телу. В организме первых самых примитивных животных уже циркулировали спонтанные электрические токи, характер течения которых видоизменялся, в том числе в зависимости от внешней среды, что, в свою очередь, приводило к конвульсивным спонтанным движениям частей примитивного животного организма. Движение тела под действием спонтанных токов позволяло переместиться, например, в направлении градиента концентрации

питательных веществ или в зону температурного комфорта.

Спонтанные электрические процессы в животных организмах постепенно эволюционировали, все тоньше взаимодействуя с событиями внешней среды, обеспечивая все более адекватные реакции на них и обработку информации. Процесс эволюции, сопровождающийся самонастройкой спонтанной электрической активности, в конечном счете перешел в мыслящий разум, осознающий себя уникальным, единичным «я». Конкретное течение электрических токов абсолютно уникально для каждого отдельного живого существа, хотя и имеет определенные параллелизмы на различных уровнях рассмотрения. История этого уникального электрического процесса протоколируется памятью и в совокупности отражает неповторимую траекторию судьбы каждого единичного живого существа.

Электрические процессы в структурах, составляющих тела живых существ, имеют кооперативную, самоорганизационную природу. Они протекают в системах, содержащих большое количество взаимодействующих элементов. Процессы, происходящие в больших коллективах взаимосвязанных элементов, способны порождать удивительно разнообразные макроскопические динамические сценарии. Они могут протекать плавно, порождать волны и динамический хаос, придавая поведению животного характерную спонтанность. Характеристики этого электрического процесса зависят от параметров и характеристик внешней среды, в которой обитает организм.

Переоткрытие времени

В традиционной технике делается ставка на общие, воспроизводимые свойства материи, на закономерно повторяющиеся явления. Например, при проектировании электронных ламповых приборов основное внимание уделялось именно воспроизводимым свойствам электронных ламп, а возникающие в них спонтанные помехи рассматривались как исключительно вредные явления, влияние которых необходимо было свести к минимуму. При разработке одушевленных пси-машин будущего, напротив, индивидуальные спонтанные «помехи» должны играть важнейшую роль в процессе их функционирования.

«Я» – это фиксация сугубой уникальности, единичности некоей вещественной структуры. Явление «я» основывается на фундаментальных физических (квантовых) свойствах материи, о которых мы не всегда вспоминаем, изучая ее воспроизводимые повторяющиеся свойства. С точки зрения квантовых представлений всякий атом во вселенной сугубо индивидуален. Во Вселенной нет двух одинаковых атомов водорода, и каждый из них имеет свою собственную историю. С позиции современного синергетического подхода уникальны не только все атомы во Вселенной, но и каждый миг ее истории также абсолютно неповторим. История Вселенной разворачивается во всем ее объеме, и каждое событие в ней сугубо однократно и необратимо.

Традиционно в научном познании основное внимание уделяется воспроизводимым повторяющимся при определенных условиях событийным последовательностям. На них основано функционирование техники сегодняшнего дня, машин, создаваемых человечеством до настоящего времени. Существование организмов, в отличие от функционирования механизмов и машин, включает в себя и невоспроизводимые, единичные, уникальные, спонтанные физические явления и процессы. С возникновением и развитием синергетики в современной науке появился математический аппарат для исследования и описания этих явлений. Лауреат Нобелевский премии по

химии И. Пригожин символически назвал новый этап в развитии науки переоткрытием времени [5].

Рассмотрим следующий компьютерный эксперимент с использованием языка Лого и традиционного для данного языка виртуального объекта – робота-черепашки. В модели предполагается, что черепашка случайным образом блуждает по прямоугольной плоской (логически замкнутой в тор) поверхности, на которой разбросана еда (например, кочаны капусты, обозначенные белыми кружками). Для того чтобы продолжать свою виртуальную жизнь, черепашке необходимо питаться. Чувство голода в модели определяется числом в диапазоне от 1 до 4. Если голод = 1, это означает, что черепашка сыта. Однако с каждым моментом времени чувство голода получает небольшое приращение. В свою очередь, голод влияет на уровень возбуждения, который изменяется в пределах от 0 до 1 и характеризует долю возбужденных нейронов в мозгу черепашки. В модели предполагается, что голод и возбуждение связаны между собой простой рекуррентной формулой:

$$\text{возбуждение}' = \text{голод} * \text{возбуждение} * (1 - \text{возбуждение}) \quad (1),$$

где

возбуждение' – доля возбужденных нейронов в следующий момент времени;

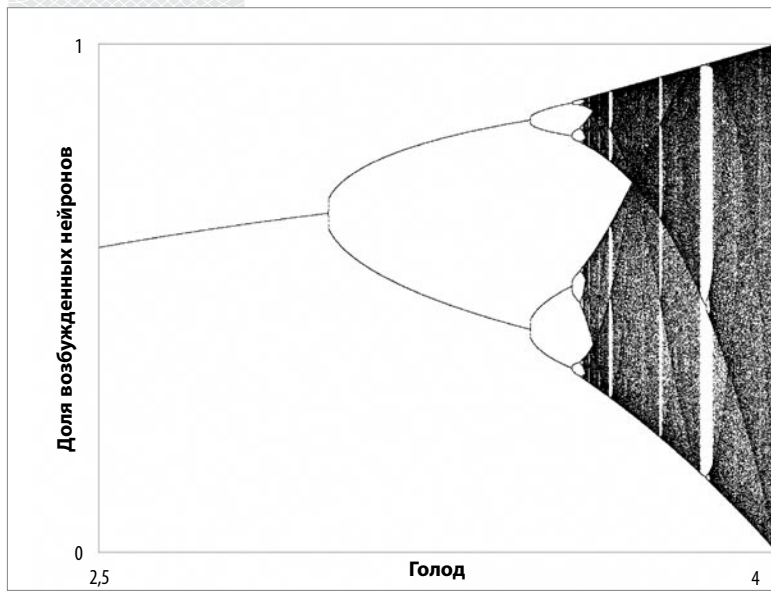
голод – уровень голода (может принимать значения между 1 и 4);

возбуждение – доля возбужденных нейронов в текущий момент времени.

В результате кооперативного спонтанного процесса возбужденные нейроны могут передавать импульсы не возбужденным клеткам и таким образом их соотношение постоянно динамически изменяется. В обсуждаемой модели эти изменения можно охарактеризовать, используя дерево Фейгенбаума (рис. 1). Алгоритм его построения сводится к следующему. На график наносятся значения, вычисленные путем последовательного итерирования уравнения (1) при разных значениях параметра (в данном случае это уровень голода, заданный в диапазоне от 2,5 до 4). При этом ряд первых итераций пропускается для того, чтобы исключить переходной процесс и отобразить на графике лишь точки, соответствующие установившемуся режиму.

Пока чувство голода не слишком сильно, доля возбужденных нейронов остается ста-

▼ Рисунок 1. Дерево Фейгенбаума





◀ Рисунок 2. Рабочее окно программы

бильной и относительно невысокой (на изображении дерева Фейгенбаума этот режим соответствует начальному участку «ствола» дерева, до его первого разветвления, или, как принято называть это в синергетике, бифуркации). Однако по мере усиления чувства голода возбуждение нарастает. Затем численность возбужденных нейронов начинает испытывать колебания. Режим с простым одночленным циклом, когда возбуждение попеременно принимает два фиксированных значения, соответствует участку между первым разветвлением и последующим удвоением дочерних ветвей графика. При дальнейшем усилении голода колебания становятся все более сложными и, в конечном счете, через серию бифуркаций переходят в динамический хаос, который можно определить как состояние паники.

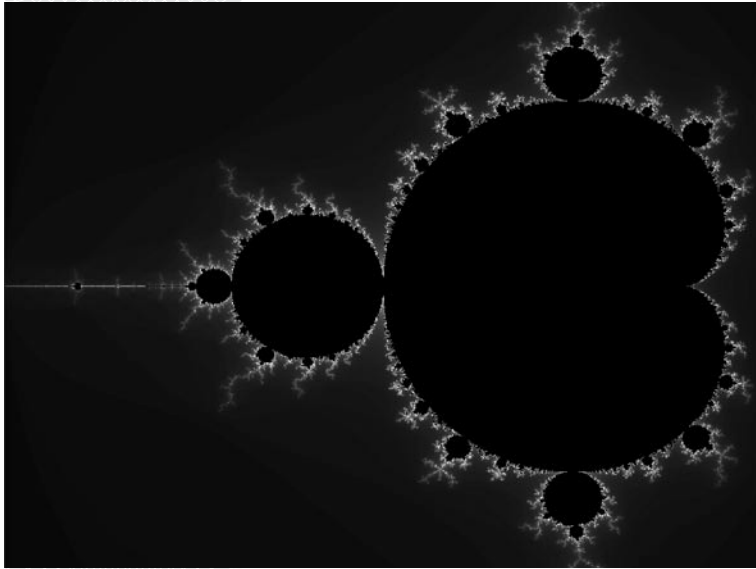
Уровень возбуждения отражается на двигательной активности черепашки. Чем выше доля возбужденных нейронов в ее мозгу, тем быстрее черепашка движется и тем чаще случайным образом меняет направления своего движения в поисках пищи (предполагается, что черепашка не видит

капусту и съедает кочан в том случае, если непосредственно на него наткнется). То есть поведение голодной черепашки разительным образом отличается от поведения ее сытой, когда она движется неторопливо и достаточно прямолинейно. Голодная же быстро и беспорядочно рыщет по игровому полю в поисках еды. После того как очередной кочан найден и съеден, черепашка на некоторое время успокаивается. Если еда долго не находится, то ее поведение становится вовсе лихорадочным и паническим. Когда же в результате капусту найти так и не удастся либо вся пища уже съедена, то голод нарастает до своего критического значения, равного четырем, и существование виртуальной черепашки прекращается.

Внешне поведение виртуальной черепашки поразительно напоминает поведение живого существа (рис. 2).

За работой программы можно наблюдать часами, несмотря на то, что в основе лежит предельно простой алгоритм и всего лишь элементарное разностное уравнение. Порождаемое данным уравнением дерево Фейгенбаума имеет определенную взаимосвязь с еще одним нелинейным отображе-

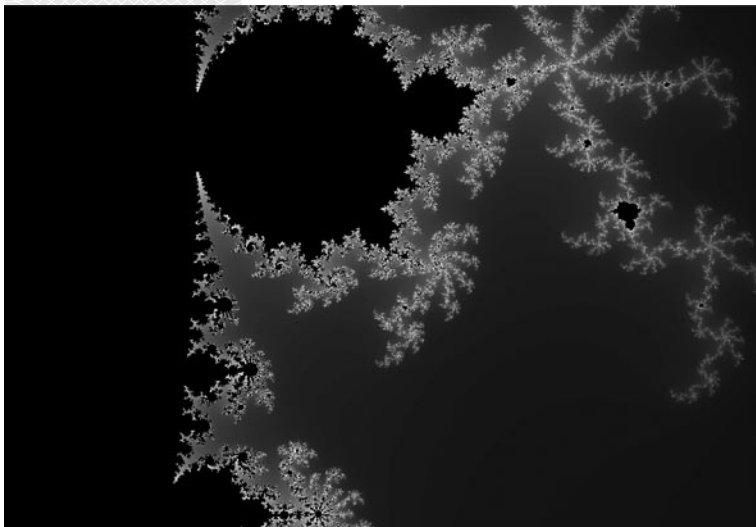
Мыслящие приборы



▲ Рисунок 3. Множество Мандельброта

нием – множеством Мандельброта (рис. 3). Структура множества Мандельброта имеет сложное строение, в основе которого лежит бесконечное рекурсивное повторение вариаций одной и той же формы (рис. 4). В этом угадывается сущностное родство фрактальных форм множества Мандельброта с рефлексивным принципом функционирования самосознания [6]. Приведенные примеры наглядно иллюстрируют то обстоятельство, что с развитием синергетики в современной науке появился инструментарий для рационального познания, моделирования и последующего технического воссоздания систем, функционирующих на принципах живых организмов [7].

▼ Рисунок 4. Увеличенный фрагмент границы множества Мандельброта



Чтобы машина могла претендовать на признание ее носителем истинного разума, помимо ощущения себя в качестве «я», она должна обладать личным специфическим мыслительным опытом (который должна уметь накапливать и обобщать), личной памятью и личной историей существования (неким аналогом судьбы). Ее мыслительная деятельность, увязанная с сугубо уникальными неповторимыми электрическими пульсациями, должна опираться на собственный опыт и формироваться на основе индивидуальных образов, аналогий, сопоставлений, иными словами, всего того, что можно назвать мыслительными протоконструктами [8]. База протоконструктов у каждого мыслящего субъекта индивидуальна. Из них он формирует свои субъективные внутренние мыслительные коды, которые впоследствии объективируются в форме текстов на естественном языке, математических формул, компьютерных программ и т.д.

Формирование базы собственных уникальных когнитивных протоконструктов должно быть неотъемлемой чертой функционирования одушевленных, мыслящих приборов. Их генерация, конструирование, апробация и отбор должны быть взаимосвязаны со спонтанным электрическим процессом, лежащим в основе самоидентификации «я» (для его краткого обозначения введем термин «я-процесс»). Таким образом, общая схема функционирования или жизнедеятельности пси-машины в общем виде включает в себя следующие элементы. Прежде всего, в основе функционирования такого гипотетического прибора должен лежать тот самый спонтанный самоподдерживающийся электрический я-процесс, обеспечивающий осознание прибором своей индивидуальности и своего существования. Его течение должно коррелировать с рядом внешних и внутренних событий, обеспечивая их осознание и реакцию на них. Мыслительная деятельность прибора может основываться на переборе и конструировании субъективных когнитивных протоконструктов, при помощи которых им будет осуществляться интерпретация усваиваемого и вновь генерируемого объективированного знания.

Именно способность машины переживать и ощущать свое «я», подобно настоящему живому организму, следует считать

главным критерием в процессе решения проблемы создания искусственной личности и искусственного интеллекта. Тест Тьюринга в этом смысле носит внешний характер и фиксирует лишь степень внешнего сходства поведения машины и живого существа. Даже абсолютное внешнее сходство не позволит сделать однозначный вывод о разумности такой машины, если не будет смоделировано главное – внутреннее ощущение «я». Сколь угодно похожая внешне на разумное существо машина, не имеющая настоящего собственного «я», останется всего лишь курьезным механизмом.

Таким образом, в качестве рабочей гипотезы можно высказать предположение, что создание одушевленных разумных пси-машин в принципе возможно. Данные перспективы уже начинают осознаваться научным сообществом, что выразилось в возникновении так называемой когнитивной информатики. В Канаде выходит специализированный журнал *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*. Ежегодно под эгидой *Institute of Electrical and Electronics Engineers (IEEE)* проводятся конференции соответствующей тематики.

Важнейшим полем применения пси-машин может стать сфера управления самыми различными сторонами жизнедеятельности человечества будущего. Использование разумных машин – единственное эффективное средство для построения и обеспечения функционирования по-настоящему справедливого, устойчиво развивающегося общества [9, с. 12]. В управляющих разумных машинах должна быть воплощена добрая воля человеческого разума, его естественная тяга к законности и справедливости. Это вовсе не значит, что управление обществом будущего в целом следует передать машинам: стратегические решения, разумеется, должен принимать человек.

Одушевленные разумные пси-машины будут незаменимы и в процессе освоения Вселенной. Превращение человечества в великую космическую расу, расселение по Галактике будет сопряжено с преодолением огромных межзвездных расстояний. Реактивные приборы, по выражению К.Э. Циолковского, способны спасти семя человечества. Причем выражение это можно рассматривать не просто как метафору. Преодолевать немислимые расстояния, которые отделяют нас от

потенциально пригодных для освоения и заселения планет, люди действительно смогут только в виде семени, то есть в нерожденном состоянии. Транспортировать через космическую бездну можно будет, по всей видимости, лишь ДНК, но никак не живых людей. Родиться они должны будут непосредственно на пункте назначения. Обеспечить весь этот сложный процесс смогут лишь одушевленные разумные пси-машины.

В прекрасном фильме Станислава Ростоцкого «Дело было в Пенькове» есть кадры, когда главным героям грезятся фантастические, сверкающие серебристым, металлическим блеском трактора будущего, самостоятельно выходящие в поле и разумно выполняющие свои функции, а люди лишь наблюдают за происходящим, сидя за огромным экраном. В далеком детстве эта сцена фильма произвела на меня глубокое впечатление. В ней воплотилась мечта человека иметь в лице созданных им машин разумных и автономных помощников, способных трудиться самостоятельно, а не быть просто механическим продолжением его рук и ног. Сейчас, благодаря развитию таких отраслей науки, как синергетика, нанотехнологии, молекулярная биология, квантовая физика, создание подобных машин может рассматриваться уже не только как фантастическая идея, но как реальная научная проблема, решение которой в будущем позволит весьма существенно расширить горизонты познания и возможности человечества. ▀

ЛИТЕРАТУРА

1. Алексеев, А.Ю. Трудности и проблемы проекта искусственной личности / А.Ю. Алексеев // *Полигнозис*. – 2008. – № 1. – С. 74–87.
2. Колесников, А.В. Можно ли записать «Я» на DVD и одухотворить машину? / А.В. Колесников // *Полигнозис*. – 2008. – № 1. – С. 88–91.
3. Бердяев, Н.А. Дух и машина / Н.А. Бердяев // *Судьба России. Опыт по психологии войны и национальности* / Н.А. Бердяев. – М., 1990. – С. 200–207.
4. Колесников, А.В. Человек и машина в образовательном процессе XXI века / А.В. Колесников, С.Н. Сиренко // *Информационно-образовательные и воспитательные стратегии в современном обществе: национальный и глобальный контекст; материалы международной научной конференции, г. Минск, 12–13 ноября, 2009 г.* – Минск: Право и экономика, 2010. – С. 208–211.
5. Пригожин, И. Перетоткрытие времени / И. Пригожин // *Вопросы философии*. – 1989. – № 8. – С. 3–19.
6. Тарасенко, В.В. Фрактальная логика / В.В. Тарасенко. – М.: УРСС, 2008. – 116 с.
7. Хакен, Г. Тайны восприятия / Г. Хакен, М.М. Хакен-Крель. – М.: Институт компьютерных исследований, 2002. – 272 с.
8. Колесников, А.В. Когнитивные протококонструкты и проблема междисциплинарной интеграции научного знания / А.В. Колесников // *Проблемы управления*. – 2014. – № 1. – С. 105–109.
9. Колесников, А. Кризис и компьютерный мир будущего / А. Колесников // *Компьютерные вести*. – 2009. – 30 июля.