



Фото Натальи Домриной.

БУДЕТ ЛИ ТЕРЯТЬ СОЗНАНИЕ ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ?

Доктор технических наук Дмитрий РОГАТКИН.

Наука начинается с измерений, утверждал Дмитрий Иванович Менделеев. А наука об измерениях, как известно, начинается со строгой и однозначной терминологии. И когда мы говорим о разработке искусственного интеллекта, ключевой вопрос заключается в чётком и однозначном определении этого термина. Без терминологической ясности создать «то — не знаю что», очевидно, проблематично. Однако единого ответа на вопрос, что такое искусственный интеллект (далее — ИИ), не существует.

Почти каждый автор, пишущий об ИИ, отталкивается от какого-либо частного, иногда своего собственного определения. Доктор технических наук Дмитрий Александрович Поспелов в книге «Фантазия или наука: На пути к искусственному интеллекту», вышедшей в 1982 году, дал такое определение: ИИ — «наука по имитации психики человека в технических системах». Сейчас распространено мнение, что это раздел научного знания, компьютерных наук и информационных технологий (см. статью А. Лагутенкова в

№ 3 «Науки и жизни», 2018 г.). Но являются ли нейросетевые программы распознавания, классификации, настроенные человеком на решение каких-то конкретных задач, которые сегодня часто именуют ИИ, реальным искусственным разумом, интеллектом? Нет, конечно.

Интуитивно под ИИ мы понимаем нечто большее — некий «сильный искусственный интеллект» (см. статью Я. Верова в № 6 «Науки и жизни», 2018 г.), то есть высокоуровневую техническую систему, наделённую, помимо сенсоров для восприятия информации (как минимум речевой), такими «умственными» способностями, таким опытом, знаниями, эрудицией, которые позволяют ей аналитически оценивать любую ситуацию, любую входную информацию, находить решения любых проблем, предсказывать ход событий в самых разных условиях и для различных бытовых сценариев не хуже, а то и лучше человека. Такие возможности позволили бы ИИ на равных общаться с человеком, понимая его и помогая ему. Это не строгое определение, но вполне приемлемое для нашей дискуссии. А не строгое оно потому, что для строгости необходимо уточнять, как, в каких ситуациях и с каким человеком проводится сравнение.

Продолжение обсуждения. Начало см. «Наука и жизнь» №№ 2, 3, 5—7, 9, 2018 г.

В 1950 году Алан Тьюринг — один из пионеров ИИ — в статье под названием «Вычислительные машины и интеллект» описывает процедуру, с помощью которой можно определить, сравнима ли машина в плане разума с человеком. Процедура, получившая впоследствии название «тест Тьюринга», это игра в имитацию человеческого разума. Идею подхватили, считая, что разумная машина обязательно будет тянуться к уровню цивилизованного человека и подражать ему (интересно, а разумной машине это действительно надо?). Но здесь, как и с определением Д. А. Поспелова, напрашивается один маленький вопрос: имитация — это всё, что мы хотим получить? Не запутываем ли мы сами себя?

Сегодня все специалисты в области ИИ делятся на два лагеря: функционалисты и коннективисты. Первые считают, что для решения проблемы сильного ИИ не надо изучать живые системы, не обязательно знать устройство мозга. Просто следует создавать функционально адекватно работающие машины, решающие задачи на уровне, сравнимом с уровнем человека, чего можно достичь, например, с помощью логических знаковых систем. Не обязательно копировать человека, следовать за природой, достаточно разрабатывать «интеллектуальные» компьютерные программы.

Представители второго направления считают, что необходимо воссоздать техническими средствами структуры мозга, повторить в точности все процессы, в них протекающие. Для этого они создают нейропроцессоры и нейрокомпьютеры.

И те и другие так или иначе опираются на аналогию с живой природой, с человеком. Даже если вы функционалист, примеры функциональностей всё равно приходится черпать из живой природы. Больше неоткуда. Следовательно, без изучения примеров живой природы, без изучения строения нервной системы, головного мозга, без знания их эволюции проблему сильного ИИ никак не решить. Но при попытке сравнения ИИ по функциональности с человеком необходимо задать всё тот же простой вопрос: а о каком человеке идёт речь? И сразу всплывает проблема терминов и определений...

Человек живёт на Земле более 100 тысяч лет. Корректно ли утверждать, что первобыт-

ный человек обладал интеллектом? А обладает ли интеллектом современный годовалый ребёнок? Нет? А трёхлетний? А пятилетний? С какого конкретно возраста человек им обладает? Давно известен «эффект Маугли»: ребёнок, не получивший обучения от сородичей, проживший ранние годы вне сообщества, не становится интеллектуально развитым. Дело, очевидно, не в строении мозга (у всех позвоночных мозг, кстати, имеет одинаковое строение), а в обучении, в уровне развития.

Таким образом, напрашивается первый вывод: интеллект — это уровень развития системы, а не сама система. В каком-то смысле правы те, кто утверждает, что для появления интеллекта искусственной системе необходимо наличие «плоти» (см. статью А. Мелихова в № 5 «Науки и жизни», 2018 г.). Системе необходимо участвовать в событиях, жить, учиться у других и на своих собственных ошибках, набивая «шишки», — тогда у неё и возникнет, возможно, интеллект. Система — одно, её интеллект — другое. Так ли это происходит в живой природе?

Современные учёные не сомневаются в наличии у большинства животных начиная с определённой ступени эволюции элементов сознания, психики и рассудочной деятельности.

Систематически изучать психику животных начали в XIX веке биологи Фредерик Кювье, Чарльз Дарвин, Джордж-Джон Роменс. Наблюдения Дарвина за животными в естественных условиях и в неволе позволили ему уже в те годы чётко выделить три основные категории поведения — инстинкт, способность к обучению и элементарную способность к мышлению (рассуждению). В 1888 году его соратник Дж. Роменс опубликовал книгу «Ум животных», где показал непрерывность развития психики у животных на всех уровнях эволюционного процесса. В 1940-х годах зоопсихолог Конрад Лоренц обнаружил, что поведение животных, помимо инстинктов, во многом определяется их внутренними мотивами, характером и настроением. Он пришёл к выводу, что позвоночные животные имеют элементы психоэмоционального восприятия и поведения, хотя и более элементарные, чем у людей. А в середине 1950-х отечественный зоолог Леонид Викторович Крушинский из МГУ им. Ломоносова чётко доказал серией уникальных опытов наличие у животных



Леонид Викторович Крушинский (1911—1984), лауреат Ленинской премии, профессор МГУ им. М.В. Ломоносова. Оставил богатое научное наследие в различных областях биологии: феногенетике, физиологии высшей нервной деятельности, патофизиологии, генетике поведения, этологии. Создал учение об элементарной рассудочной деятельности животных как предыстории интеллекта человека.

элементарной рассудочной деятельности. Животные умеют рассуждать и осознают себя как отдельный выделенный субъект во внешнем мире, то есть обладают самосознанием. Без этого невозможно избегать опасности, строить жилище, охотиться, спариваться. Во всех процессах такого рода есть особь и есть окружающий её мир, который она обязана осознавать как нечто внешнее.

Бурные дискуссии возникают именно по поводу наличия сознания у животных. Говорят, что так и до интеллекта у животных дойти недолго... Доля правды в этом есть. Ветеринарам хорошо известно, что животные могут терять сознание. Можно ли потерять то, чего нет? Здесь снова наблюдается прямая связь с проблемой терминов и определений. Что понимать под сознанием?

Одна и та же живая интеллектуальная система может находиться как в состоянии сознания, так и без него, в состоянии потери сознания. Может она находиться и в состоянии сна (состояние изменённого сознания). Следовательно, сознание — режим работы системы, её текущее состояние, а не нечто извне вложенное, как программа в систему. Равно как и мышление, рассудочная деятельность. Это процессы, происходящие внутри системы, а не отдельные сущности. Другое

дело, что это за процессы и состояния? Как они образуются и организуются внутри живой системы?

Оказывается, в общих чертах это давно известно. Управление в живых системах, все процессы такого управления основаны на явлении раздражимости, на рефлексах. Одноклеточные реагируют на внешние раздражения всей клеткой в целом. У многоклеточных появляются две специальные системы — нервная и эндокринная. Последнюю разработчики ИИ часто упускают из вида, но она не менее важна для развития интеллекта, чем первая. Примеров раздельного функционирования нервной и эндокринной систем в живой природе нет. Чем более развиты эти системы, тем разумнее поведение вида и отдельных его особей.

Реакции обычной речной пиявки с простой узловой нервной системой на раздражения достаточно примитивны, унифицированы и легко могут быть смоделированы несложным техническим устройством. Пиявка способна лишь уплощаться, вытягиваться, сокращаться и изгибаться. Когда кто-то задевает кожу или надавливает на тело пиявки, она в этом месте выгибается в сторону от воздействия, разворачивается и уплывает прочь. Для охоты пиявка использует тепловой датчик, нацеливаясь на движущееся нагретое тело. Базовый ритм движений пиявки задаётся рядом возбуждающих и тормозящих взаимодействий в её нервной системе. Периферические рецепторы служат для регуляции, усиления или полной остановки движений. Особую роль в управлении пиявкой играет гормон эндокринной системы — серотонин. В крови неподвижных, вялых пиявок уровень серотонина ниже, чем у активных особей. Стимуляция же секреторных серотонин клеток способствует повышению активности животного. Можно полностью удалить серотонин из крови эмбриона пиявки при помощи специального реактива, который избирательно разрушает серотонинергические нейроны в развивающихся ганглиях. Такое животное само по себе не способно плавать, однако плавательные движения сразу появляются при добавлении в среду серотонина. Всё просто.

Но уже поведение речного рака, поведение других членистоногих, не говоря уже о позвоночных и млекопитающих, более разнообразно. Это следствие развития их нервной и эндокринной систем, хотя такой базовый принцип функционирования систем, как ответ

на раздражение, направленный на исключение раздражения (отдёргивание руки) или на его минимизацию (сужение зрачка при ярком свете), остаётся неизменным. Это и есть ключ к проблеме.

Работа головного мозга, мышление, расщудочная деятельность, равно как и выброс гормонов при эмоциях, стрессе, суть процессы рефлекторного ответа на возникающие раздражения. Просто исходным раздражителем, стимулом для мозга может быть не только внешнее физическое воздействие, но и внутренние проблемы, переживания, эмоции (действие гормонов). Великий физиолог Иван Михайлович Сеченов (всем «создателям» ИИ рекомендуется внимательно изучить его книгу «Рефлексы головного мозга», вышедшую в 1866 году, спустя всего четыре года после выхода в свет книги «Происхождение видов» Ч. Дарвина) отметил, что все мыслительные процессы инициируются внутренними чувственными раздражениями (возбуждениями). Они порождаются как разница между желаемым и действительным. Мозг в ответ на чувственное раздражение в результате внутренней нервно-психической деятельности начинает вырабатывать решение для возникшей проблемы, а для часто встречающихся проблем — выдавать за счёт выработки условных рефлексов уже готовые, опробованные решения. Условный рефлекс — тоже рефлекс, но более высокого уровня.

Для чего нужен условный рефлекс? Идеологически — это механизм оптимизации и минимизации затрат (энергетических, вычислительных) на предстоящую деятельность. Выработка слюны у собак готовит их к максимально быстрому и эффективному усвоению пищи. Для роботов на простых примерах также можно показать пользу и даже естественную необходимость выработки условных рефлексов.

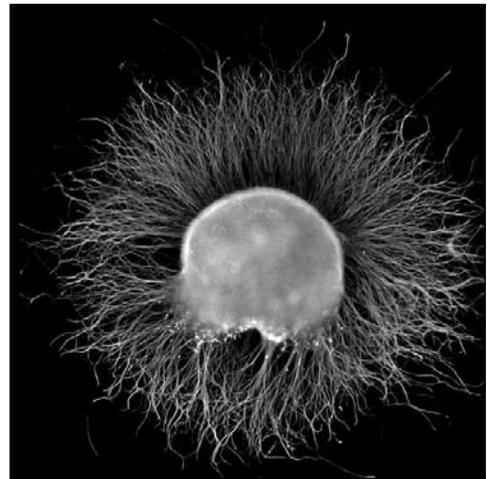
Например, представим ситуацию, что медицинский сервисный робот входит в комнату через дверь, расположенную по центру стены. У противоположной стены, также по центру, установлена кровать, на которой лежит больной человек. Робот его обслуживает и должен к нему периодически подходить и измерять температуру. Он мог бы прямо двигаться по центру комнаты к кровати, однако в центре комнаты стоит стол, слева от стола — шкаф, причём так, что робот не может по своим габаритам пройти между ними, а справа от

стола проход свободен. Действия робота представим по шагам:

1. Проход через дверь.
2. Движение прямо до стола.
3. Оценка ситуации (как обойти?).
4. Поворот направо.
5. Проход к кровати.
6. Измерение температуры.

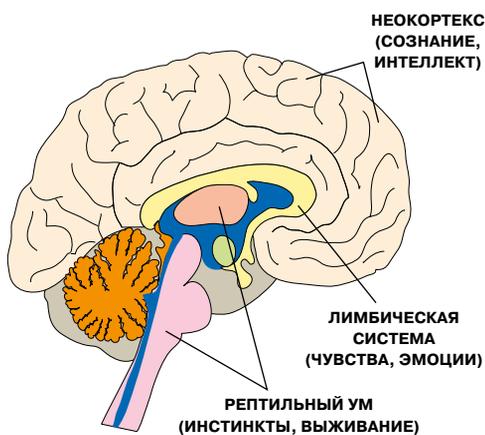
Если эта процедура единичная, то ничего не изменить. Да и нет смысла менять. Но если робот выполняет процедуру ежедневно, то с какого-то этапа, запомнив решение на основе предыдущего опыта, можно исключить п. 3 и объединить шаги 2 и 4 в один новый шаг: движение по диагонали. Вот и выработался условный рефлекс — то, что иногда мы называем «двигаться на автопилоте». Если путь многократно испытан, мы часто не задумываемся, как действовать. Причём если существует робот меньших габаритов, который способен пройти слева от стола, то он может использовать разные стороны обхода стола. Вот мы и пришли к интересному новому выводу о неизбежной индивидуальности выработки рефлексов поведения даже у роботов и об обретении ими персонального жизненного опыта.

Столь же индивидуальны память о прошедших событиях и картина внешнего мира. Их нельзя скопировать и передать от одного робота другому. Их нельзя запрограммировать и сформировать заранее. Для большого



Спинной ганглий семидневного зародыша цыплёнка, выращенный в искусственной среде. Видны расходящиеся от ганглия аксоны. Ганглий — нервный узел — скопление нервных клеток, состоящее из тел, дендритов и аксонов нервных клеток, и глиальных клеток.

Фото: commons.wikimedia.org /CC BY-SA 2.5.



Неокортекс — новая кора мозга человека. Она покрывает однородным серым плащом всё белое вещество больших полушарий головного мозга, которое представляет собой проводящие пути, «провода», лишённые самостоятельной возможности генерировать импульсы. Когда употребляют термин «brain» — мозг, то имеют в виду именно эту часть центральной нервной системы.

работа сформированная интеллектуальная стратегия поведения маленького робота неприемлема. Более того, если оба робота обслуживают больного, один измеряет температуру, а другой, скажем, кормит, то в модели мира каждого робота будет присутствовать другой робот, других габаритов и с другими функциональными обязанностями. В памяти каждого робота будут храниться его собственные действия, а также — отдельно — действия другого робота. В памяти каждого из них неминуемо формируется индивидуальная история и индивидуальное восприятие мира. Если же в «мозг» одного робота бездумно взять и переписать историю и модель мира другого робота, то коллизии при движении и переобучение неизбежны.

Мировосприятие оказывается индивидуальным и уникальным даже у простейших роботов. Выходит, разум, интеллект — вещь штучная, персональная, формируемая в процессе индивидуальной деятельности и обучения. Это система сложных и взаимосвязанных условных рефлексов, вырабатывающихся в ответ на повторяющиеся раздражения, на внешние и внутренние вызовы.

Эволюционно головной мозг развивался у животных путём объединения нервных узлов (ганглиев) в единый функционирую-

щий орган, а также путём надстройки новых отделов над старыми. Причём, что важно, новые отделы головного мозга всегда берут на себя наиболее сложные функции, интегрально объединяющие и анализирующие сигналы от старых структур. В частности, с развитием у млекопитающих неокортекса (новой коры) как наивысшего интегративного центра многие сложные ассоциативные функции перешли к нему, оставив древней и старой коре участие в формировании эмоций и памяти, а среднему мозгу — формирование быстрых, экстренных реакций на внезапные внешние раздражения. Здесь прослеживается, скажем так, не «игра в имитацию», а «игра в ассоциации». Последующее ассоциативно объединяет предыдущее. Ассоциативный условно-рефлекторный ответ на раздражения — основной принцип работы мозга. Отсюда и стереотипы, и косность мышления, аналогии при решении разных задач.

Не исключено, что для построения «сильного ИИ» необходимо в первую очередь, моделируя и компоная разные нейронные сети, создать систему, которая будет автономно работать на принципах иерархической ассоциации условных и безусловных рефлексов. Надо попытаться не настроить какую-то отвлечённую нейронную сеть вручную на решение конкретной задачи, а создать специализированную сеть с автономной выработкой условных рефлексов на внешние и внутренние раздражения. Представляется правдоподобным, что, начиная с определённого уровня организации такой сложной ассоциативно-рефлекторной структурированной системы, в ней естественным образом возникнет сознание как единственно возможный и эффективный режим её работы в условиях необходимости интегрировать и «обрабатывать» мириады раздражений одновременно. Интеллект же, вероятно, появится у неё тоже эволюционно закономерно, как наивысший уровень развития, но только при условии её обучения, активного существования и общения в социуме аналогично развитых систем (человек здесь не исключение). Причём это будет уникальный и персональный, как у каждого человека, уровень интеллекта. Не исключено, что в бессознательном режиме работы уровень интеллекта недостижим в принципе. Тогда ответ на вопрос, вынесенный в название статьи, становится очевидным и утвердительным. Имеющий сознание в любой момент может его потерять.