



ранении ускорит переход к персонализированной медицине, введет интеллектуальные системы в практику постановки диагнозов, прогноза распространения эпидемий и т. п.

В то же время все чаще раздаются голоса ученых, которые считают ИИ угрозой людям. Так, незадолго до своей смерти знаменитый Стивен Хокинг сказал, что развитие ИИ может стать как наиболее позитивным, так и самым страшным фактором для человечества, а Илон Маск

предлагает следующую формулировку: искусственный интеллект - это свойство искусственных систем решать интеллектуальные задачи, для которых отсутствует алгоритм решения. Отсюда следует важный вывод: как только задача решена на компьютере, она перестает быть интеллектуальной, поскольку это означает, что для нее уже создан алгоритм решения (как известно, компьютер может работать только по алгоритму). Следовательно,

## Развитие искусственного интеллекта может стать как наиболее позитивным, так и самым страшным фактором для человечества.

считает искусственный интеллект наибольшим риском, с которым мы столкнемся как цивилизация.

Чтобы оценить обоснованность различных мнений, необходимо первым делом разобраться, что вообще следует понимать под искусственным интеллектом. Сейчас существует тьма различных определений ИИ, большинство из которых имеет расплывчатый и неконкретный характер. И.Каляев привел ряд из них: «ИИ - научное направление, в котором решается задача аппаратного и программного моделирования тех видов человеческой деятельности, которые традиционно считают интеллектуальными»; «Свойство систем выполнять функции, которые традиционно считаются прерогативой человека». И с юмором прокомментировал, что если исходить из этих определений, можно сделать вывод, что стиральная машинка обладает искусственным интеллектом, ведь она выполняет функции, которые традиционно считаются прерогативой человека, поскольку до сих пор не менее двух третей населения Земли продолжают стирать вручную. А может, опереться на мнение английского математика Алана Тьюринга? В середине XX века он заявил, что вычислительная машина сравняется по интеллекту с человеком тогда, когда... сможет ввести его в заблуждение. В 2015 году программа «Соня Гусева», созданная Иваном Голубевым из Санкт-Петербурга, смогла обмануть судей в 47% случаев. Выходит, согласно Тьюрингу, что ИИ уже на 50% создан? Отнюдь. Программа «Соня Гусева» - просто хороший чат-бот, способный именно «вести людей в заблуждение» относительно того, с кем они общаются. И все.

По мнению докладчика, при определении искусственного интеллекта надо отталкиваться от понятия «естественный интеллект», то есть способности мозга человека решать интеллектуальные задачи путем приобретения, запоминания и целенаправленного преобразования знаний, применения этих знаний для управления средой. При этом под интеллектуальной задачей понимается задача, решение которой не было до сих пор известно, то есть отсутствовал алгоритм ее решения. Исходя из этого, И.Каляев

большинство из того, что сегодня мы называем ИИ, никакого отношения к интеллекту не имеет, поскольку это просто различные компьютерные программы, работающие в соответствии с уже созданным человеком алгоритмом. Это орудия труда, такие же, как молоток. Только он усиливает физические возможности человека, а компьютерные программы - интеллектуальные, умственные. Калькулятор тоже усиливает умственные способности человека, но при этом же никто не называет его искусственным интеллектом. «Поэтому все, что сегодня называется искусственным интеллектом, - говорит И.Каляев, - было бы корректно обозначить как интеллектуальные компьютерные технологии или псевдоинтеллектуальные компьютерные технологии. Тем не менее термин «искусственный интеллект» по отношению к компьютерным программам устоялся, принят во всем мире, и мы тоже вынуждены его использовать».

Чуть позже вокруг этой части доклада возник спор, стоит ли уделять терминам особое внимание? «Нужно, - ответил докладчик. - Если с чиновников будут спрашивать отчет за реализацию национальной программы в области ИИ, важно, чтобы не было разночтений в понятиях, иначе дойдет до абсурда. Условно говоря, накопят «стиралок» и их числом отпортуют о повсеместном внедрении ИИ».

Далее академик напомнил коллегам, что исследования в области ИИ развиваются уже более 60 лет по двум основным направлениям: логическому и нейрокибернетическому. Логический подход направлен на создание прикладного (слабого) искусственного интеллекта, то есть компьютерных программ, предназначенных для решения какой-либо одной задачи или их небольшого множества. Нейрокибернетический подход направлен на создание универсального (сильного) искусственного интеллекта - аналога человеческого мозга, способного решать любые интеллектуальные задачи.

И.Каляев привел примеры основных достижений, полученных в рамках обоих подходов. Запомнилось, что в 1957 году была создана первая логическая программа для игры в шахматы, а уже в 1997-м про-

Из первых уст

Елизавета ПОНАРИНА

# Камо грядеши?

Пути развития искусственного интеллекта обретают огромную значимость



Игорь КАЛЯЕВ, глава Совета по приоритету научно-технологического развития РФ, академик РАН

Под занавес 2020 года состоялось заседание рабочей группы 01 Технического комитета РФ «Искусственный интеллект», одной из главных задач которого является стандартизация терминов и определений в области искусственного интеллекта.

Основным докладчиком на этой онлайн-встрече выступил академик Игорь Каляев, глава Совета по приоритету научно-технологического развития РФ, в сферу деятельности которого попадает и ИИ. Предваряя его выступление, руководитель группы, генеральный директор Международного НИИ проблем управления Александр Агеев отметил, что Игорь Анатольевич много десятилетий занимается данной тематикой в приложении к кон-

кретным техническим системам, причем «с очень убедительной эффективностью». Свою презентацию академик И.Каляев озаглавил «Искусственный интеллект: камо грядеши?», как бы подчеркивая фразой на старославянском великую, почти библейскую значимость ответа. По его мнению, мир сегодня разделен надвое в своем отношении к ИИ: одни видят в нем чуть ли не мессию, который спасет наш мир от всех бед и напастей, а другие, наоборот, воспринимают сродни всадникам Апокалипсиса, которые грозят уничтожить человеческую цивилизацию и превратить нас в рабов машин.

Безусловно, в настоящее время ИИ является одним из основных трендов мирового научно-технического развития, и поэтому каждая

уважающая себя страна старается сформировать национальную программу в области искусственного интеллекта. Лидер в этом деле - Китай, который вкладывает около 20 миллиардов долларов в год, а в смежные отрасли - более 150 миллиардов. Значительные суммы выделяются на эту сферу и в США: федеральное ассигнование - около 1 миллиарда долларов в год, суммарные расходы корпораций и венчурных фондов - почти 25 миллиардов. В Европе цифры скромнее, последние три-четыре года финансирование НИОКР в области искусственного интеллекта - около 2 миллиардов долларов, но планируемые целевые инвестиции до 2022 года - около 30 миллиардов. Россия тоже пытается не отстать в этой гонке, и в 2019 году указом Президента РФ была утверждена Национальная стратегия развития ИИ.

Считается, что к 2025 году ИИ обеспечит удвоение темпа роста ВВП ведущих государств и даст прирост мирового ВВП почти на 15 триллионов долларов. Промышленности он поможет в деле оптимизации логистики поставок комплектующих, при планировании промышленных процессов, создании безлюдных производств, беспилотного транспорта; в энергетике позволит точнее прогнозировать сетевые перегрузки, повысить отказоустойчивость энергосистем; в сельском хозяйстве даст шанс роботизировать сбор, переработку, транспортировку и хранение продукции; в здравооо-



грамма Deep Blue победила чемпиона мира по шахматам. В 2016 году программа Alpha Go обыграла чемпиона мира по игре Го (число позиций в игре в  $10^{125}$  раз больше, чем в шахматах), а в 2017-м программа Libratus победила четверых профессиональных игроков в покер, взяв банк в 1,7 миллиона долларов. Турнир длился 20 дней, сыграли более 120 тысяч раздач. Это особенно впечатляет знающих, что покер в отличие от шахмат, шашек и Го является игрой с неполной информацией. Машина не в курсе, какие карты на руках игроков. Более того, игроки имеют право блефовать.

Примерно в то же время, в 1956 году, в Массачусетском технологическом институте создали компьютерную модель нейронной сети. Спустя сорок лет, в 1995-м, японская компания Fujitsu представила нейрокомпьютер Neuro-Turbo, который управлял взаимодействием 1000 нейронов. В 2009 году ученые смоделировали работу 1 миллиона нейронов и 10 триллионов синапсов, что примерно равно мозгу кошки или 4% человеческого мозга (проект SynAPSE). В 2014 году был создан процессор с 1 миллионом программируемых нейронов и 256 миллионами синапсов.

При этом академик утверждает, что большинство достижений ИИ связано не с каким-либо повышением «интеллекта» компьютеров, а просто с ростом их быстродействия, за счет чего компьютер успевает за ограниченный промежуток времени перебрать все большее число вариантов развития ситуации на большее число шагов вперед и, как следствие, выбирать наиболее оптимальный вариант действий в текущей ситуации. В качестве доказательства Игорь Анатольевич привел график роста производительности наиболее мощных суперкомпьютеров мира по годам и наложил на него достижения в области ИИ. Наглядно видно, что все значимые достижения в области ИИ напрямую коррелируют с ростом производительности используемых компьютеров. При этом возникает вопрос, когда компьютеры смогут достичь производительности, достаточной для создания «сильного» искусственного интеллекта, то есть «железного» аналога человеческого мозга? Кажется, еще чуть-чуть подождем - и достигнем успеха.

Однако ситуация не столь радужная, наоборот, не столь критичная, как кажется на первый взгляд. По оценкам академика, для моделирования работы человеческого мозга в реальном времени требуется суперкомпьютер с производительностью  $10^{20}$ - $10^{21}$  флопс (операций с плавающей запятой в секунду). И теоретически такой суперкомпьютер может появиться после 2025 года. Если же опираться на существующие ныне технологии, то он будет занимать объем, эквивалентный зданию с основанием 300 на 300 метров и высотой 50 метров, а для обеспечения его работы потребуются 15 ГВт электроэнергии, что сопоставимо с мощностью трех Саяно-Шушенских ГЭС. Сравните с мозгом обычного человека, который занимает объем порядка 1,5 дм<sup>3</sup> и потребляет всего лишь около 20 ватт электроэнергии. Очевидно, что мозг устроен и функционирует по совершенно другим и до сих пор непонятым нам принципам,



пам, и поэтому попытки создания его аналога на базе традиционных компьютерных технологий мало перспективны. В качестве примера академик привел попытки создания «железного» аналога мозга простейших живых организмов, предпринятые еще в 80-е годы прошлого века, которые закончились ничем, - «железный» мозг даже близко не работал как естественный.

Далее Игорь Анатольевич обратил внимание слушателей, что в настоящее время наблюдается переход от ноосферы по Вернадскому, то есть биосферы, управляемой человеческим разумом (интеллектом), к киберсфере - биосфере, управляемой компьютерным «ис-

кусственным» интеллектом. С одной стороны, конечно, переход к киберсфере несет целый ряд неоспоримых благ для человечества, освобождая людей от монотонного и тяжелого труда, повышая эффективность производства и сельского хозяйства, уровень медицинского обслуживания и т. п. С другой стороны, этот переход порождает целый ряд глобальных проблем, которые могут привести к опасным непрогнозируемым последствиям для всего человеческого сообщества, таких как деградация естественного интеллекта, отсутствие доверия к решениям, формируемым ИИ, возможность искусственного управления человеческой эволюцией.

## “ Подмена человеческих и культурных ценностей ведет к возникновению нового типа войн - эволюционных - целью которых являются уничтожение самосознания и вмешательство в естественную эволюцию противника.

Докладчик вывел на экран картину конца XIX века художника Богданова-Бельского, изображающую обучение устному счету в сельской школе. Приглядевшись к школьной доске, изображенной на картине, можно попробовать испытать себя, вычислив значение написанной там дроби, в числителе которой сумма квадратов чисел - 10, 11, 12,

13 и 14, а в знаменателе - 365. Сельские дети, обутые в лапти, решают этот пример в уме, а все ли участники обсуждения могут справиться с этим без помощи гаджета или хотя бы калькулятора? Повсеместное использование технологий ИИ и рост зависимости людей от компьютерных устройств сопряжены с рисками «умственной гиподинамии» человеческого общества.

Академик привел цитату из книги «Фактор понимания» известного русского философа А.А.Зиновьева: «Реальное будущее человечества представляется как господство высокотехнологичных, но примитивных существ, не имеющих ни малейшего понятия о том, как факти-

чески устроены и функционируют фундаментальные законы природы и человеческого сознания». Следующий вопрос, поставленный академиком, насколько мы можем и должны доверять решениям, предлагаемым ИИ? Он привел пример суперкомпьютера WATSON, который, по мнению разработчиков, должен был заменить врачей при диагностике и лечении онкологии. Для этого WATSON два года обучался на базе более чем 600 000 медицинских документов и 25 000 историй болезни, а затем его подключили к целому ряду медицинских центров США, ЕС и Азии. Но в 2018 году многие из них отказались от его услуг. Дело в том, что совпадение терапии, назначаемой WATSON и врачом-человеком одному и тому же пациенту, оказалось менее 50%. А ведь речь шла о жизнях людей. Медики не рискнули ставить на больных эксперимент подтверждения или опровержения правильности рекомендаций ИИ. Но есть и другой пример, вспомнил докладчик, когда программа AlphaGo, играя первую партию с

чемпионом мира по игре в Го, сделала ход, шедший полностью вразрез с канонами игры. Разработчики схватились за головы, сочтя, что создали ущербную программу, но через несколько ходов выяснилось, что именно этот странный ход привел программу к победе в данной партии. Возможно, что и WATSON предлагал правильную терапию онкобольных, но отсутствие доверия между ним и врачами привело к его отключению. Проблема доверия к решениям, формируемым ИИ, порождает целый ряд сопутствующих задач, в том числе разработки объяснительных механизмов функционирования ИИ, разработки принципов распре-

деления ответственности между ИИ и человеком за принимаемые решения, разработки мультимодальных интерфейсов общения ИИ и ИИ (в том числе не имеющих точных сенсорных интерпретаций, например, эмоциональных, интуитивных, нравственных, эстетических) и т. п. Но еще более опасным свойством киберсферы, по мнению академика, является появление возможностей целенаправленного (искусственного) управления человеческой эволюцией путем навязывания и распространения посредством медиа и интернет-пространства неестественных ценностей и потребностей; пропаганды и мотивирования низменных инстинктов и примитивизации; подмены человеческих и культурных ценностей; манипулирования сознанием, повсеместной лжи, дезинформацией, лицемерием и т. п. Все это ведет к возникновению нового типа войн - эволюционных - целью которых являются уничтожение самосознания и вмешательство в естественную эволюцию противника.

В заключение академик И.Каляев сослался на повесть братьев Стругацких «За миллиард лет до конца света», в которой, по его мнению, очень образно показаны перспективы создания «сильного» искусственного интеллекта. По сюжету некий ученый находится на пороге великого открытия, которое может «потрясти» основы мироздания. Но едва он к нему приближается, что-то происходит в его жизни, отвлекающее от работы, то сосед заходит с ящиком коньяка, то ошибается дверью красивая девушка, а то под его окном падает огромное дерево. В итоге ученый приходит к выводу, что все это не случайные события, а явления природы, - просто она не хочет, чтобы были нарушены основы ее мироздания. Это, конечно, фантастика, говорит И.Каляев, но здесь заложен очень глубокий смысл: именно физические законы природы вряд ли позволят людям создать «сильный» искусственный интеллект, сравнимый с естественным или тем более превосходящим его, поскольку целью такого «супермозга» будет именно изменение законов природы. Даже если гипотетически мы когда-нибудь создадим полный «железный» аналог человеческого мозга, он все равно останется грудой железа, не более того. Должна быть еще какая-то не понятная нам составляющая, назовите ее «душа» или как-то по-другому, с помощью которой эта железка сможет обрести человеческое сознание.

Доклад вызвал острую дискуссию, участники обсуждали едва ли не каждое его положение: ряд ученых высказал сомнение, что с нахождением алгоритма действия задача перестает быть интеллектуальной, переходя в разряд инструкций, последовательности действий. Но, безусловно, такого уровня дискуссии демонстрируют академическую глубину рассмотрения проблем и свидетельствуют, что настала пора создания тезауруса ИИ и ответственного осмысления путей его дальнейшего развития, направленного не на попытки противостояния природе и естественной эволюции человечества, а на усиление его умственных возможностей и развитие духовных качеств, сохранение природы и традиционных общечеловеческих ценностей. ■