

Никогда еще люди не относились друг к другу с таким ожесточением, никогда еще не были в такой степени ослеплены, как в наше время, когда они стали воображать, что знают все.

Эмиль Золя

Искусственный интеллект: продолжение пройденного



Александр ГОЛЫШКО,
ведущий аналитик, АО «НПО РусБИТех»,
к. т. н.

Поисковые системы, почтовые платформы, виртуальные магазины, автоматизированные call-центры, программы-переводчики, Siri с Алисой и многое другое в виде ПО, вплоть до сервисов суперкомпьютера IBM Watson, – все это тоже искусственный интеллект (ИИ), хотя и не в самом его сильном проявлении.

Слабый ИИ

Наличие у машин/программ такого интеллекта пока не предполагает наличия сознания (впрочем, оно им и не нужно). Это так называемый слабый ИИ (Narrow AI), который может работать в режиме реального времени, извлекает информацию из ограниченного набора данных и запрограммирован на исполнение одной задачи, с которой в целом неплохо справляется.

Очевидно, нам пора привыкнуть, что все, о чем говорили писатели-фантасты 50–150 лет назад, начинает приходить в реальности, хотя, быть может, и не совсем в том виде, в каком всем бы хотелось. Прежде всего это относится к искусственному интеллекту (Artificial Intelligent – AI), которым многие из нас пользуются чуть ли не ежедневно, часто об этом даже не подозревая.

Собственно, под Narrow AI подходит любой продукт автоматизации, будь то уличный автомат по продаже напитков или завод-автомат по изготовлению высокотехнологичных изделий. Вот, к примеру, «Сбер» перешел на использование искусственного интеллекта при выборе мест расположения банкоматов и терминалов. В итоге всего за год их стало меньше на 8,4%. А ученые из Университета Ватерлоо смогли научить экзоскелет передвигаться более естественно, применив для этого видекамеры и ИИ. Однако этого мало, ибо в будущем мы ожидаем, что ИИ будет заботиться о пожилых людях, обучать наших детей и выполнять множество других задач, требующих человеческой эмпатии и понимания норм морали. Тем самым мы закладываем ожидания на много лет вперед, и только лишь Narrow AI маловато будет.

Кстати, слабый-то он слабый, но уже сумел избавить человечество от выполнения массы рутинных задач. Сегодня ИИ понимает голосовые команды, различает картины, водит автомобили и выигрывает в настольных играх. Однако психологи полагают, что граница между программами и людьми не должна размываться, а голосовым ассистентам на всякий случай стоит научиться

напоминать собеседнику, что они людьми не являются. Это чтобы не попасть в зависимость от более мощных разновидностей ИИ. Ведь, к примеру, современные дети уже вовсю общаются с той же Алисой, как с живым собеседником. Завтра они вырастут, а психологическое восприятие ИИ у них останется.

Эволюция термина

Понятие ИИ описывает машины, которые имитируют когнитивные функции, связанные с человеческим разумом, такие как решение проблем и обучение. Интеллект позволяет исчислять то, что возможно измерить, а сознание, в свою очередь, дает оценку чувствам, которые может испытывать только человек. Впрочем, с последним согласны не все. Уже можно найти сообщения о создании ИИ, обладающего самоосознанием себя, т. е. ИИ чувствует себя живым и испытывает реальные эмоции. С одной стороны, это вроде почти такое же сознание, как и у человека, с другой – запрограммированные эмоции в ответ на различные воздействия могут создать иллюзию, в частности, для производящих экспертизу «нового сознания» психиатров (особенно если разработчик ИИ с навыками психиатра).



Однако, чтобы сравнивать сознания, неплохо бы разобраться, что мы знаем о своем собственном сознании, его возможностях, соответствующей работе мозга и пр. Профильные ученые обычно говорят, что знаем мы не так уж и много, т. е. соответственно принять за проявление сознания можем что угодно.

Единое определение ИИ пока не сформулировано, а если точнее, то его трактовка менялась вместе с достигнутыми наукой рубежами. Есть, например, такое определение: ИИ – это область информатики, которая занимается разработкой интеллектуальных компьютерных систем, обладающих возможностями, которые мы традиционно связываем с человеческим разумом: понимание языка, обучение, способность рассуждать, решать проблемы и т. д. Впрочем, как известно, в современном мире все традиционное может быстро сменить ориентацию, и человеческому разуму придется искать пути общежития с разумом нечеловеческим.

Следует заметить, что и сам термин ИИ в наше время считается не совсем правильным.

Это скорее общее определение для нескольких видов технологий, наделяющих компьютеры и механизмы более интеллектуальными возможностями: когнитивные вычисления, машинное обучение, Big Data, искусственные нейронные сети, глубокое обучение и др. К тому же для разных людей в различном контексте ИИ может означать разные вещи. Можно сказать и так: ИИ – это не самостоятельная технология, а совокупность различных технологий, которые вместе должны наделять интеллектуальными способностями некую систему, процесс, актив или устройство.

Ну а для менеджеров ИИ можно объяснить, как ответ на вопрос «что случится, если предоставить машине бесконечную вычислительную мощность и бесконечные данные?». Если добавить сюда безграничную власть, которую, как мы знаем, дает обладание информацией, то наиболее точным образом такого ИИ станет «обезьяна с гранатой» или еще с чем-нибудь помощнее. И многие люди это чувствуют.

Тем временем индустрия движется все дальше по

зарабатыванию денег и по направлению к все более полной автоматизации, когда компьютеры анализируют данные и на их основе принимают решения все быстрее (порой чуть ли не за доли секунды), тогда как человеку на решение аналогичных задач требуется времени все больше (месяцы и недели). Прежде чем перейти к более сильным разновидностям ИИ, коснемся немного истории.

Заглянем в историю

После формулирования трех законов робототехники Айзека Азимова (1942) и появления «Игры в имитацию» Алана Тьюринга (1950) уже к 1955 г. были проведены первые исследования, касающиеся разработки систем символьных вычислений, и сформированы такие концепции, как нейросети и естественный язык, а профессор математики из Дартмутского колледжа Джон Маккарти придумал термин «искусственный интеллект», объединяющий их все. Маккарти, кстати, руководил группой, подавшей заявку на грант для организации

конференции по ИИ в 1956 г., куда были приглашены многие ведущие исследователи того времени для обсуждения потенциальных областей изучения ИИ, включая обучение и поиск, зрение, логические рассуждения, язык и разум, игры (в том числе шахматы) и даже взаимодействия человека с такими разумными машинами, как персональные роботы. Общим консенсусом тех обсуждений стало то, что у ИИ огромный потенциал, чтобы принести пользу людям. Собственно, с тех пор люди и ожидают от ИИ той самой пользы. Понимание пользы людьми, кстати, тоже имеет много различий.

Восхождение ИИ было непростым. Несмотря на то что еще в 1957 г. Фрэнк Розенблатт создал механическую нейросеть в Корнельской лаборатории аэронавтики, лишь в 2011 г. электронная нейросеть сумела-таки «увидеть» кошку. Ведь большую часть своей истории ИИ существовал преимущественно в исследованиях, куда правительственные агентства (в том числе DARPA) вливали деньги и практически не требовали отчета по инвестициям. Не секрет, что в подобной ситуации исследователи ИИ во имя финансирования часто преувеличивали потенциал своей работы. Как результат: весьма пессимистичные прогнозы о потенциале ИИ и урезание бюджетов. Порой казалось, что ИИ не оправдал ожиданий и никогда не сможет достичь уровня человеческих возможностей. К тому же появление суперкомпьютеров создало иллюзию, что вот это и есть «тот самый ИИ». Это действительно был тот самый Narrow AI и не более того.

Выяснилось, что относительно легко конструировать компьютеры, которые обладают превосходными логическими способностями вроде умения играть в шахматы, но невероятно сложно создавать компьютеры, которые могли бы демонстрировать человеческую физическую активность и восприятие. Поэтому в исследованиях ИИ наблюдались периоды

разочарования, прозванные в среде ученых «AI-зима». Финансирование иссякло, прогресс замедлился, поэтому фокус внимания переместился на другие области информатики. Первая «AI-зима» длилась с 1970-х до середины 1980-х, вторая – с 1990-х по примерно 2010-е.

В 1997 г. шахматный компьютер Deep Blue производства IBM победил Гарри Каспарова. Deep Blue не размышлял о стратегиях и не учился игре, как умели делать более поздние системы, а просто перебирал в своей памяти лучшие варианты. Это простейшая система Narrow AI, являющаяся чисто реактивной, т. е. не обладающей способностью формировать воспоминания и использовать прошлый опыт для реализации текущих решений. Тем не менее та победа вернула ИИ в круг общественного внимания, инвесторы были очарованы, а разработка ИИ получила новый импульс.

К 2011 г. ученые всего мира заговорили о нейросетях и начали активно их создавать. Такие системы Narrow AI уже обладали ограниченной памятью и могли заглянуть в прошлое. Программист Джефф Дин из Google и профессор информатики из Стэнфорда Эндрю Ын создали большую нейросеть с огромным вычислительным ресурсом серверов Google, которой можно было скоротить огромный набор изображений. Эта нейросеть выдала три размытых изображения, обозначающих визуальные образы, которые она снова и снова встречала в обучающих данных – лицо человека, тело человека и той самой кошки. А в 2012 г. профессор Торонтского университета Джофри Хинтон с двумя студентами создали глубокую нейросеть для компьютерного зрения AlexNet, выдающую в два с лишним раза меньше ошибок по сравнению с конкурентами. Это событие, возможно, сильнее остальных повлияло на возрождение интереса к глубоким нейросетям, и Хинтон заслужил прозвище «крестный отец глубокого обучения».

В 2013 г. исследователи британского стартапа DeepMind опубликовали работу, где было описано, как нейросеть научилась играть и выигрывать в 50 старых игр от Atari. В марте 2016 г. AlphaGo обыграла в серии игр величайшего игрока в игру «го» в мире Ли Седоля со счетом 4:1. В тот момент показалось, что ИИ выиграл не просто у Ли Седоля, а у всего человечества.

Разновидности ИИ

Последние достижения в области глубоких нейросетей настолько сильно изменили область ИИ, что реальная его история, возможно, только начинается и ИИ повлияет на все аспекты жизни XXI в. К примеру, в августе прошлого года по программе DARPA ACE (эволюция воздушного боя) ИИ, сформированный сочетанием машинного обучения и нейронных сетей, был испытан в виртуальных сражениях один на один с таким же ИИ и даже с живым оператором. «Сверхчеловеческий» интеллект оказался проворнее живого пилота и победил в виртуальном бою. То есть появились и другие разновидности ИИ.

В частности, ИИ общего назначения, или сильный ИИ (General AI), может успешно выполнять практически любые умственные задачи, которые под силу людям, чем весьма похож на последних. Именно такие системы встречаются в научно-фантастических книгах и фильмах, посвященных взаимодействию человека с машинами, обладающими чувствами и сознанием. Считается, что General AI сможет рассуждать, справляться с проблемами, выносить суждения в условиях неопределенности, планировать, учиться, интегрировать предыдущие знания в процесс принятия решений, а также предлагать новаторские идеи. Машины не столь отдаленного будущего, как ожидается, будут более совершенными, способными формировать представления не только о мире, но и о других агентах или существах. В психологии это

называется «теория разума» – понимание того, что люди, существа и предметы в мире могут иметь мысли и эмоции, которые влияют на их собственное поведение. Ведь без понимания мотивов и намерений друг друга, не принимая во внимание познания других об окружающей среде, работать вместе в лучшем случае трудно, а в худшем – невозможно.

Для достижения всех этих целей исследователи должны придумать, как наделить машины хоть каким-то сознанием. Однако машины пока не способны мыслить абстрактно, продумывать стратегию, а также использовать мысли и воспоминания, чтобы принимать обоснованные решения или выдвигать творческие идеи (с последним и у людей бывает не очень). Сможет ли искусственный интеллект изучить моральные ценности человеческого общества? Сможет ли принимать решения в ситуациях, когда необходимо взвесить все за и против? Сможет ли развить чувство правильного и неправильного? Короче говоря, будет ли у него понимание этики и, как следствие, совесть?

Супер-ИИ

Переадресуем наши надежды дальше – к супер-ИИ (Super AI), который определяется, к примеру, оксфордским философом Ником Бостромом весьма незамысловато: «Это любой интеллект, который значительно превосходит когнитивные способности человека практически во всех областях». То есть люди уже знают, что Super AI превзойдет человека во всех аспектах – от творчества до жизненной мудрости и решения проблем, хотя и не знают, как именно. В общем, что-то вроде «ах, какая женщина... мне б такую!». Именно эта разновидность ИИ (а еще больше сопутствующая ей неопределенность) беспокоит сегодня очень многих. По мнению Илона Маска, именно Super AI приведет к вымиранию людей как вида. Впрочем, вполне возможно, что произойдет это не без помощи того же И. Маска и созданного



им нейроинтерфейса к мозгам братьев по разуму.

Итак, заключительным этапом развития ИИ станет создание систем, которые смогут формировать представления о себе. В конечном счете исследователи ИИ должны не только изучить основы сознания, но и создать сущности/машины/ПО, которые им обладают. Сознательные существа осознают себя, знают о своих внутренних состояниях и способны предсказывать чувства других. Очевидно, человечество еще далеко от создания машин, которые обладают самосознанием, но уже сегодня должно сосредоточить усилия на методах понимания, памяти, обучения и возможности принимать решения исходя из прошлого опыта. Это важный шаг на пути к пониманию человеческого интеллекта. Кстати, подростки-программисты уже сегодня сомневаются в превосходстве человека над ИИ, но мы-то с вами хорошо понимаем, что это всего лишь юношеский максимализм. В составе нынешних интеллектуальных систем пока нет элемента, который будет иметь ключевое значение для построения разумных машин в будущем.

Несмотря на всю свою мощь и потрясающие способности, основанные на логических

вычислениях, машины не настолько умны, особенно в области познания и восприятия на человеческом уровне. Кроме того, машины не могут брать на себя ответственность. По крайней мере в ближайшее время. Возьмет ли ее на себя сверхинтеллектуальный Super AI? Этого мы тоже еще не знаем, но зато знаем, что будет еще дальше.

От информационной эпохи к автономной

Говорят, ускорение темпов развития ИИ – ключ к концу «информационной» эпохи и началу эпохи «автономной», условную границу между которыми обозначают термином «технологическая сингулярность». То, что начнется за этой границей, нам пока неизвестно, но предсказаний футурологов уже немерено. Если грубо обобщить, то «информационная» эпоха началась с появления печатного станка в 1450 г., который помогал распространять знания и информацию по миру. Теперь у нас есть устройства, которые мы носим в кармане и которые могут получить доступ к любому виду данных в течение нескольких секунд или связаться с кем-либо в мире, что, как считается, является логическим завершением

«информационной» эпохи. Впрочем, помнится, нам еще обещали с помощью нейроинтерфейса Интернет прямо в голове, а не в смартфоне, что, согласитесь, стало бы еще более логическим завершением той эпохи в виде глобального перехода на псевдотелепатию.

В 1965 г. Ирвин Джон Гуд, британский математик, который в свое время работал криптологом в Блетчли Парк с Аланом Тьюрингом, предположил, что General AI может привести к взрыву технологического развития. Его сценарий выглядел следующим образом: по мере того как компьютеры увеличивают мощность, людям становится проще строить машину более интеллектуальную, чем человек. Этот сверхчеловеческий интеллект обладает более широкими навыками решения проблем и изобретательскими способностями, чем нынешние люди. Эта суперинтеллектуальная машина разрабатывает еще более совершенную машину или переписывает собственное ПО, чтобы стать еще более умной. Все более способная машина продолжает разрабатывать машину с еще большей способностью и т. д.

В общем, однажды ИИ, запрограммированный на постоянное самосовершенствование, не станет ограничивать себя пределами. Сие означает, что, достигнув уровня человеческого интеллекта, машина не остановится на этом. Разумеется, когда компьютер станет «умнее» нас, это будет шоком для всего человечества, ибо возникнет Super AI или искусственный суперинтеллект (ASI – Artificial Superintelligence), который превзойдет интеллект человека практически во всех областях, включая научные изобретения, общие познания и социальные навыки. Наступит та самая технологическая сингулярность.

Кстати, этот термин изобрел и публично представил в 1983 г. математик и научный фантаст профессор Вернон Виндж. По его словам, однажды люди сотворят более могущественный интеллект, чем их собственный, и в этот момент человечество достигнет

своеобразной сингулярности, а его дальнейшая судьба будет непредсказуема. Иначе говоря, главным у человечества станет тот самый сверхинтеллект, а что он сделает со своим творцом, как говорится, одному Богу известно.

На симпозиуме VISION-21 в 1993 г. В. Виндж представил следующие события из развития гипотезы технологической сингулярности:

- возникнет искусственный интеллект, в разы превосходящий человеческий;
- повысится эффективность человеческого разума и физических характеристик при помощи генной инженерии;
- крупные компьютерные сети смогут самостоятельно идентифицировать себя как разумные сущности;
- синтез компьютера и человека приведет к тому, что человек сам будет наделен сверхинтеллектом.

В итоге мы окажемся не одионоки среди порожденных нами новых сущностей, и одинокие сегодня смогут возрадоваться завтра. Неудивительно, что все эти мысли были тут же развиты футурологами, которые предсказывают рождение технологической сингулярности в период с 2020 до 2070 г. С первой датой, как мы теперь знаем, они ошиблись.

Опускаясь с небес на землю, заметим, что, скорее всего, следующая эпоха будет той, где компьютеры и машины выполняют для людей задачи, основываясь на принципах оптимизации и эффективности, на огромных объемах данных и эмпирических наблюдениях. Назовите любую форму человеческой деятельности, которую вы только можете представить, и увидите способ, которым новые алгоритмы ИИ смогут ее усовершенствовать. Следует понимать, что возможности каждой отдельной машины все же будут ограничены. Алгоритм, разработанный, к примеру, для разведения насекомых в целях получения белка для искусственного мяса, не сможет развернуться и оптимизировать какой-нибудь контент-маркетинг (хотя формально,

с точки зрения человека, содержание у этих задач весьма похоже).

Ближайшее будущее (по крайней мере следующие 50 лет) будет наполнено множеством узкоспециализированных ИИ, выполняющих конкретные задачи. А создание ASI, который мог бы справиться с самыми разнообразными задачами, могло бы стать предпосылкой появления, к примеру, искусственного чувства, реализуется нескоро. Ведь это самая сложная часть процесса, потому что по-прежнему никто не знает, как сделать компьютер «умным».

Когда ученые приходят в отчаяние, они пытаются создать программу, которая бы тестировала сама себя. Это может стать самым многообещающим методом создания общего ИИ. Идея заключается в создании компьютера, главными функциями которого будут исследования ИИ и кодирование изменений. Устройство будет не только самостоятельно обучаться, но и изменять собственную архитектуру и алгоритмы работы, перестав обезьянничать и приобретая некое сознание.

Не исключено также, что однажды не в меру интеллектуальный ASI перехватит управление человеческой цивилизацией и начнет осуществлять его на свой лад, по-своему решая, кто тормозит развитие этой цивилизации, кто в ней «слабое звено», кто устраивает войны и прочие нелепости, что порождает преступность, кому надо дать по рукам (или другому месту), кто является «слабым звеном» в соревновании роботов и людей и чем вообще цивилизация ASI может быть лучше человеческой. Мнение людей на сей счет может стать вторичным, а их судьба печальной. Успокаивает лишь то, что эти вызовы для нашей цивилизации должны быть изначально услышаны человеческими программистами, которые будут разделять будущий ASI интеллектом.

Господи, сделай так, чтобы среди этих программистов были сплошь далеко смотрящие патриоты нашей цивилизации!

Впрочем, если именно Господь наделил нас сознанием, то не так уж все страшно. Не правда ли? ■