

Искусственный интеллект - это машина для свободы человека



О творчестве, свободе, сложностях и перспективах разработки искусственного интеллекта рассказывает профессор, директор Института компьютерных наук и технологий, заведующий научно-исследовательской лабораторией нейросетевых технологий и искусственного интеллекта СПбПУ, доктор технических наук Лев Владимирович Уткин.

Лев Владимирович, когда появилась идея создания искусственного интеллекта?

Можно сказать, что идея возникла примерно в 50-е годы прошлого века, когда был изобретен перцептрон (от слова perception – англ., восприятие), математическая модель мозга Френка Розенблатта. Это была первая попытка смоделировать нейронную сеть головного мозга. Розенблатт тогда создал электронное устройство, которое должно было имитировать процессы мышления человека. Это не было прорывом, но идея уже возникла и над созданием искусственного интеллекта стали работать ученые разных стран. Эти работы знали несколько «зим» и «вёсен», периодов, когда исследования то прекращались, то вновь возобновлялись и вдохновляли новых ученых.

Вообще, можно выделить два больших периода и две противоположных парадигмы исследования искусственного интеллекта. Первый период, он и исторически первый, когда центральной идеей было создание так называемого искусственного интеллекта, реализованного в экспертных системах. Машину обучали правилам по принципу «если – то», и эти правила были основанием определенной экспертной системы, позволяли получать знания. Это был дедуктивный подход, принцип «от общего к частному». Эта волна исследований «замерзла», главным образом потому, что трудно найти экспертов в разных областях знания, формализовать все имеющиеся знания, и к тому же сами знания все время меняются, появляются все новые и новые открытия, системы, и такой экспертный искусственный интеллект попросту не успевал за реальным развитием науки.

Второй период, вторая волна началась в 70-80-е годы, когда стали разрабатывать противоположный, статистический подход. Это были разработки в сфере машинного обучения, индуктивный подход от частного к общему, от примеров к зависимостям. Как ребенок изучает мир, например, ему показывают листья клена и дуба и объясняют их отличия и сходство: и то и другое – лист, но вот это – клен, а это дуб, и вот такие различия. И ребенок учится узнавать и распознавать листочки. Так и машину стали обучать распознавать частности и выводить из них закономерности. Это была идея нейронных сетей. Ведь что такое нейрон? Это очень узкоспециализированная клетка, аддитивный классификатор, который проводит электрическое возбуждение и химически взаимодействует с другими нейронами, принимая извне, обрабатывая, проводя, храня и выводя наружу информацию. Разрабатывая искусственный интеллект, хотели сделать аналог нейронов головного мозга. Аналога не получилось, но то, что сделали, назвали нейронными сетями, потому что они тоже способны проводить и классифицировать информацию, принимать ее и выдавать вовне.

В чем, на ваш взгляд, привлекательность идеи искусственного интеллекта для человека? Почему ученые к ней возвращаются снова и снова, существуют то «зимы», то «вёсны»?

Трудно сказать, в чем именно привлекательность идеи... Человеку свойственно искать новое, двигаться вперед. Была идея переложить некоторые функции человека на компьютер. Так же, как до этого машины заменяли человека во многих других областях тяжелого или рутинного труда. Например, врачебная диагностика. Есть отличные врачи, но далеко не во всякой клинике они есть. И тогда машина при отсутствии врача может осуществлять диагностику, принимать решение направить на операцию или назначить лечение. Ту же цель преследует разработка беспилотного автомобиля: отсутствие водителя или его присутствие только как оператора. То есть идея – облегчить труд. Как когда-то калькулятор заменил человека в сложных вычислениях, так теперь компьютер заменил, например, работу по обработке медицинских анализов. Когда-то врачи-лаборанты сидели часами, сгорбившись над микроскопом и вручную считали, изучая капельку крови, а теперь все это почти мгновенно делает машина. Сегодня искусственный интеллект может заменить человека во многих областях, Давид Ян как-то сказал, что «искусственный интеллект – это новое электричество». Представьте себе мир без электричества сейчас!

Есть ли профессии, которые искусственный интеллект заменить не сможет?

Мы, конечно, не знаем, как будет развиваться искусственный интеллект. Точно не заменит физический труд. И, скорее всего, он не заменит творческие процессы. Сегодня искусственный интеллект – это статистика. И у машины точно нет того, что есть у человека, – интуиции. Мы пока совсем не знаем, как работает человеческая интуиция. А пока не знаем, не сможем и смоделировать. Там, где нужна интуиция – в исследованиях, науке, искусстве, в человеческой коммуникации, в любом творчестве, – нет возможностей и для искусственного интеллекта.

То есть сейчас искусственный интеллект моделирует лишь когнитивные процессы?

Должен признать, что тут есть некоторая путаница понятий. Когда мы говорим «интеллект», мы, действительно, думаем о когнитивной сфере. Но «искусственный интеллект» – это в некотором смысле маркетинговый ход. Искусственный интеллект на самом деле сегодня – процессы машинного обучения на основе статистики. Они быстро развиваются, радуют своими достижениями, но они не моделируют когнитивную функцию.

Но, как минимум, уже пытаются моделировать эмпатию, эмоциональный интеллект? Сегодня искусственный интеллект умеет распознавать жесты и выражение лица, интонацию, переводить язык китов и дельфинов, даже дописывать симфонии Баха на основе анализа его творчества... Казалось бы, это движение в сторону творческой составляющей человеческой деятельности. Каковы перспективы в области эмоционального искусственного интеллекта?

У машины нет и не может быть эмоций. Модель можно научить распознавать эмоции человека, изменения высоты тона голоса, мимику, жесты, а также частоту пульса, температуру, расширение зрачков и так далее. То есть мы просто обучаем программу на большой выборке, какие параметры будут соответствовать какой человеческой эмоции. Также можно научить машину генерировать определенные звуки и выражения, которые будут восприниматься человеком как проявление тех или иных эмоций. Но не более.

И фильмы, которые так романтично показывают нам эмоциональных роботов – плачущего Электроника, страдающего от любви робота Вертера, даже Терминатора, принимающего благородные решения на основании эмпатии – это всего лишь человеческие фантазии?

Может быть, когда-то и будет иначе, но сегодня и в обозримом будущем точно нет. Творчество, свобода, эмоции – сфера специфически человеческого.

Мы знаем, что научные разработки стимулируются творческим воображением и интуицией. Многие научно-технические достижения вначале появились в фильмах и книгах, а затем, действительно, воплотились в реальность. Так же человеческое воображение рисует и страхи, связанные с развитием искусственного интеллекта. Есть ли риски, связанные с самообучением машин? Может ли искусственный интеллект выйти из-под контроля?

Самообучение машины сильно отличается от самообучения человека. Если ребенку достаточно нескольких примеров и он научится самостоятельно – благодаря воображению – достраивать свой опыт, то машине нескольких примеров недостаточно, машине нужны тысячи и тысячи примеров, огромные массивы размеченных данных. Если машина сталкивается с неразмеченными данными, она не знает, что это, и не сможет домыслить или вообразить. Научная фантастика переносит на искусственный интеллект черты интеллекта человеческого, и страхи связаны с непониманием сути процессов машинного обучения. Так что самообучение – один из мифов, связанных с искусственным интеллектом. Вот, например, когда мы говорим о беспилотном автомобиле. Что делает искусственный интеллект? Он непрерывно, каждую долю секунды анализирует и классифицирует состояние окружающей дорожной среды и учитывает текущие изменения. Но если на дороге встретится новый знак, не учтенный в исходной классификации, машина его просто не распознает. Она выдаст ошибку и примет решение в рамках той ситуации, в которой находится: либо передаст информацию человеку, либо просто заблокирует все текущие процессы, остановится.

Сейчас во многих странах, в том числе и в России, разрабатываются этические кодексы по применению искусственного интеллекта. Расскажите, какие этические проблемы возникают в связи с новейшими разработками?

Во-первых, это проблемы открытости данных. Машине для обучения, повторяюсь, нужны огромные выборки данных. А где их брать? Например, для разработок в области медицины. Насколько этично раскрывать данные о ДНК для использования их для обучения машины? Законодательства многих стран не позволяют использовать эти данные. Во-вторых, это проблемы, возникающие, когда машина выдает решения, связанные с расовыми или гендерными особенностями.

А проблемы, связанные с правами человека? Например, система «Окулус», которая обучена делать выводы о чертах характера

человека на основании анализа его поисковых запросов в Интернете, – насколько этична подобная тотальная слежка, которую делает возможной именно искусственный интеллект?

Согласен с вами. Здесь, действительно, есть проблемы и необходимы законодательные действия. Сейчас проводятся конференции по этике искусственного интеллекта, где эти проблемы выносятся на обсуждение. Надеюсь, будут и приемлемые решения.

Ведутся ли разработки в области соединения этического подхода и, условно говоря, эмпатии, вернее, возможностей учета машиной так называемого «человеческого фактора»? Ну, например, когда машина диагностирует онкологическое заболевание и рекомендует пациенту сложную операцию, не учитывая какие-то специфические обстоятельства, связанные с возможным стрессом пациента от самой информации...

С этим как раз связано наиболее актуальное направление в разработке машинного обучения – объяснительный искусственный интеллект. Это модели, которые позволяют объяснять решения. Не просто выдать диагноз, но объяснить, почему и по каким признакам он поставлен. Если говорить о медицине, для врача важно получить именно объяснение диагноза, машина оказывает врачу помощь в исследовании, например в обработке рентгеновских снимков, мониторинге множества показателей жизнедеятельности и так далее. А решение и диагноз – сфера ответственности человека, врача.

В вашей лаборатории одно из направлений называется «аномальное поведение машинных процессов». Насколько эти машинные процессы соответствуют процессам человеческого аномального поведения? В фильмах нам показывают различных сошедших с ума роботов, которые ведут себя вполне как люди-психопаты. Но это человеческие фантазии, а как на самом деле?

На самом деле речь идет о процессах внутри машины, это просто технический термин и любые аналогии тут неуместны. Например, лежит пациент в больнице под кучей мониторов, которые отслеживают его пульс, давление, температуру и другие показатели. И вот температура пациента повышается, а пульс, наоборот, замедляется. Это процесс нетипичный, аномальный с точки зрения человеческого тела. Машина обнаруживает нечто новое, неизвестное ей – и это аномальный машинный процесс. Мы можем обучить машину нормативным процессам. А то, что выходит за пределы нормы, она распознает как нечто неизвестное. В этом случае машина должна выдать сигнал аномального процесса и сообщить о нем человеку. Объяснительный искусственный интеллект может даже найти причину аномального состояния – это одна из задач, над которыми ведется работа.

Вы возглавляете магистерскую программу по искусственному интеллекту в Политехе. Чему учите студентов?

Мы открыли программу для магистров два года назад. Могу сказать, что я доволен тем, как работает эта программа, хотя и трудности есть. Конкурс к нам большой, и это свидетельствует об интересе к проблеме искусственного интеллекта у молодежи. Многие студенты участвуют в реальных проектах, один из плюсов программы – практическое обучение. Есть и дисциплины, например, «Философия искусственного интеллекта», которые изучают общественно значимые проблемы. Мы вообще стараемся дать разносторонние знания. Я включил в программу изучение направления, связанного с экспертными системами в разработке искусственного интеллекта, этого предмета нет почти нигде больше. Но я считаю, что это необходимо, потому что, с моей точки зрения, будущее за комбинированным дедуктивно-индуктивным подходом. У человека есть априорные знания, и машина должна иметь некие априорные установки, с которыми она будет соотносить свои статистические данные.

Какие вы видите междисциплинарные перспективы развития исследований искусственного интеллекта?

Прежде всего, это биология, медицина, техника. Но здесь есть проблема, связанная с качеством математического образования в смежных областях. Так что лучше, чтобы механик оставался механиком, биолог биологом, а data scientist (специалист, который с помощью математических алгоритмов и программных инструментов анализирует данные. – Примеч. Ред.) умел находить с ними общий язык. Например, когда мы ведем разработку искусственного интеллекта для помощи в лечении и диагностике рака, мне нужно объяснить врачу на простом языке, какие именно данные я хочу от него получить. А затем, получив размеченный data set, я буду с ним работать и давать задачу системе. Поэтому междисциплинарность – это не основная задача, прежде всего, нужны хорошие специалисты в области машинного обучения. Пока их мало.

С моей точки зрения, искусственный интеллект – это новая технологическая революция. И необходимо говорить об этом, популяризировать это направление, уделять больше внимания в университетском обучении.