

Нейросети и интеллект. Попытка демистификации



У человеческого сознания есть одно удивительное свойство. Вообще, разумеется, у человеческого сознания масса удивительных свойств, отчасти именно им и посвящён наш сегодняшний разговор. Однако начать хотелось бы с одного конкретного. Я бы назвал его

гомеостатичность. Нам со школьных времен знакомо понятие *гомеостаза*. В биологии этим термином называют постоянство внутреннего состава и состояния какого-либо живого объекта, а также способность к его поддержанию в изменяющейся среде. Температура и химический состав жидкости внутри человеческого организма всегда примерно одни и те же, а значительное изменение их в первую очередь свидетельствует о нарушении или разрушении в организме.

Сознание, как процесс обработки информации об окружающем мире в

реальном времени, занимается тем, что каждое мгновение строит на основе этой информации внутри себя модель всего происходящего вокруг. Удивительно то, что модель эта строится совершенно простая – как будто ничего серьёзного не происходит, как будто вокруг нас – гомеостаз. Именно по этой причине нам очень сложно замечать и принимать резкие изменения вокруг нас – они не укладываются в статическую картину, которую хочет видеть наш разум.

Ещё какие-то тридцать лет назад мы носили с собой пригоршню жетонов для телефонов-автоматов, и это был единственный доступный способ быстро связаться друг с другом. Двадцать лет назад чтобы посмотреть кино на компьютере, его приходилось либо покупать на диске, либо долго и изнурительно (и довольно дорого!) качать из интернета. Десять лет назад если ты не знал китайского языка, то возможности пообщаться с китайцем, не говорящим на английском, не было никакой. Из своего опыта могу сказать, даже жёсткими объясниться с рядовым китайцем

Алексей Сергеевич Ковтуненко, кандидат технических наук, доцент кафедры информатики Уфимского университета науки и технологий (УУНиТ). Родился в Уфе 23 февраля 1984 года, в 2007 году окончил Уфимский государственный авиационный технический университет (УГАТУ) и получил квалификацию инженер по специальности «Моделирование и исследование операций в организационно-технических системах». Активный участник музыкального и поэтического сообщества республики, участник и организатор фестивалей авторской песни, неоднократный победитель песенных и поэтических конкурсов. Опубликовано подборка стихов в журнале «Бельские просторы», составлен сборник стихов и песен «Система координат», который распространяется в рукописи. Активно занимается научными исследованиями в области искусственного интеллекта с 2004 года, на настоящий момент опубликовано более 70 научных работ, среди которых 1 научная монография, 2 учебных пособия по программированию. В период с 2021 по 2023 годы возглавлял факультет информатики и робототехники (ФИРТ) УГАТУ, впоследствии УУНиТ. С 2022 по 2024 год являлся Президентом региональной Федерации спортивного программирования. Руководит студенческим проектным офисом, активно участвует во внедрении технологий искусственного интеллекта во все виды человеческой деятельности.

очень сложно, поскольку семиотика жестов совершенно иная.

Однако сегодня нам сложно и пред- ставить, что когда-то фильм или альбом любимой группы приходилось перема- тывать, чтобы написать статью или ре- ферат приходилось сидеть неделями в библиотеках, а мультики надо было вы- лавливать в эфире телеканала при по- мощи программы телепередач. Сегодня мы воспринимаем большинство резуль- татов технического прогресса как нечто само собой разумеющееся, как будто так было всегда – наше сознание играет с нами эту шутку, имитирует гомеостаз. Между тем, после небольшого затишья в прорывах в области физики и астро- физики, пришла по-настоящему новая волна технических побед. Пожалуй, впервые за долгое время её нам при- готовили математики и программисты.

Разумеется, я говорю про *искус- ственный интеллект*.

Прямо скажем, эта тема стара как мир! Всегда ленивый человек мечтал, чтобы машина делала что-то вместо него, а неуверенный в себе – чтобы делала лучше него. По мере того как изобретатели достигали тех или иных технических высот, менялось и само понимание *интеллекта*.

Механизированные игрушки-ан- дроиды, придуманные в средние века часовщиками, производили на публи- ку завораживающее впечатление. Всем казалось, что ещё чуть-чуть, и искус- ственный человек заменит настоящего во всех непопулярных видах деятель- ности. Шли годы, всем стало понятно, что для «интеллектуальности» нужно нечто большее, чем выполнение пред- писанной программы.

В середине девятнадцатого века зародилась наука об управлении – ки- бернетика. К середине двадцатого века учёным казалось – вот оно! Правильно подобранный алгоритм, который смог бы адаптироваться к внешней среде, и интеллектуальное поведение машины обеспечено. И снова всё оказалось не

так просто – понятие интеллектуаль- ности пришлось пересмотреть. Тогда же, в начале второй половины двадца- того века математики и кибернетики, увлечённые проблемой создания ис- кусственной интеллектуальности, раз- делились на два крупных не очень дру- жественных друг другу лагеря.

Одни занялись разработкой моде- лей, имитирующих работу человече- ского сознания, аналитических функ- ций человеческого интеллекта, меха- низмов принятия решений человеком. Так возникла *инженерия знаний*. Это направление породило огромное ко- личество незаменимых в настоящее время технологий, включая технологии программирования, хранения данных и знаний. Именно благодаря развитию этого направления сегодня существу- ет интернет со всеми его сервисами и возможностями. Инженерия знаний позволила протянуть мостик между гу- манитарной наукой и информатикой, и мы обязательно когда-нибудь посвятим ей отдельный обстоятельный разговор.

Другие начали разрабатывать моде- ли, имитирующие явления живой при- роды, и появился *бионический подход* в науке и технике. В рамках этого подхода было также придумано невероятное ко- личество полезных методов и алгорит- мов, которыми мы пользуемся сейчас – от алгоритмов поиска кратчайших маршрутов, основанных на поведении муравьёв в колонии, до, собственно, алгоритмов *искусственных нейросетей*, имитирующих прохождение электри- ческих импульсов по нервной ткани животного.

Историю нейрофизиолога Фрэнка Розенблатта – одного из первых «изо- бретателей» искусственных нейронных сетей – принято рассказывать в каче- стве демонстрации бесчеловечности американской системы финансирова- ния научных исследований. В пятидеся- тых годах он построил математическую модель (алгоритм) работы нервной тка- ни, которую научился настраивать для

решения задачи распознавания графических образов. Впоследствии этот алгоритм был реализован автором в виде электронного устройства, которое было названо «перцептрон». Именно с него началась в мире история нейронных сетей. С помощью метода Розенблатта можно было подобрать такие значения числовых параметров – *коэффициентов* – этого алгоритма, что он абсолютно точно вычислял, изображение какой именно буквы ему предъявляют даже в том случае, если в изображении были внесены помехи (в случайных местах помещены черные или белые точки). Успех был ошеломительным! Автор публиковался, гранты сыпались как из рога изобилия. Разумеется, другие учёные массово включились в работу по исследованию возможностей перцептронов, всем они казались панацеей при решении задач классификации и автоматического распознавания. Однако в 1969 году учёные Минский и Пайперт издали книгу «Перцептроны», в которой очень доходчиво и конкретно разъясняют (сиречь математически доказывают), что король голый, а перцептроны Розенблатта не способны решать задачи сложнее парёного батата. В одночасье поток финансирования прекращается, и на некоторое время перцептрон оказывается забыт как ошибочная и тупиковая ветвь развития науки.

Только через десять лет сразу несколько учёных «возрождают» перцептроны. Они предлагают их многослойную реализацию, благодаря которой перцептроны легко преодолевает все «неразрешимые» задачи Минского и Пайперта. Однако трагическая случайность уносит жизнь Фрэнка Розенблатта в 1970 году, и ему так и не случается увидеть головокружительный успех своих моделей и теорий.

За что же мы должны быть благодарны Розенблатту? Прежде всего за то, что он придумал математические модели и алгоритмы, которые способны обобщать! Под способностью

некоторой программы к обобщению будем понимать её способность решать задачу, распознавать объекты и ситуации, с которыми она прежде не сталкивалась, на основе оценки сходства. Именно передача машине способности к *обобщению* и является первой вехой, первым по-настоящему существенным шагом современной науки к *искусственному интеллекту*.

Свою первую нейронную сеть я разработал в далёком 2003 году, будучи студентом-второкурсником Уфимского государственного авиационного технического университета (УГАТУ). Это был многослойный перцептрон, он прекрасно решал все стандартные задачи по распознаванию букв, прогнозированию простых последовательностей. На ту пору в УГАТУ уже почти десять лет существовала мощная научная школа нейросетей, каждый год по этой тематике защищалось в среднем по две-три кандидатские диссертации, научное сообщество состояло из талантливых математиков и программистов. Одним словом, самописной нейронной сетью было уже никого не удивить.

Так почему же мы видим успехи искусственного интеллекта только сейчас, а не тогда, в восьмидесятых, девяностых годах двадцатого столетия? Ответ довольно прозаичен. Компьютеры тех времён работали очень медленно и не могли реализовать всех необходимых вычислений за какое бы то ни было разумное время.

Вообще, для вычислительной математики довольно стандартная история! Вы бы удивились, узнав сколько гениальных алгоритмов и моделей, придуманных учёными в двадцатом веке, ждут своего часа! Они до сих пор не нашли применения только лишь потому, что на существующих в настоящее время компьютерах работают настолько медленно, что на решение мало-мальски полезной задачи потребовалось бы время, сопоставимое с двумя-тремя человеческими жизнями!

Право прорывных изобретений математика на долгие двадцать лет уступила физике и науке о материалах, благодаря которым уже к концу две тысячи десятых годов скорость вычислений на микропроцессорах достигла огромных величин, а на серверах было накоплено огромное количество оцифрованной информации из разных областей человеческой жизни.

Почему это так важно? Вспомним, что искусственные нейросети являются по сути большими математическими формулами, в которые необходимо подставлять числа. Для того чтобы эти формулы были нам полезны при работе с объектами реального мира, нам непременно надо описать эти объекты с помощью чисел, «оцифровать» их. Для разных видов информации в двадцатом веке были разработаны разные способы оцифровки. Например, изображения представляются как прямоугольные таблицы цветных точек – пикселей, у каждой точки есть своё место (номер строки и номер столбца), а также особым образом закодированный цвет (три числа). Такая таблица называется растром, и её уже можно использовать в расчётах нейросети. Текстовая информация представляется в виде чисел ещё проще. Каждому символу присваивается номер, и сохраняется последовательность этих номеров. Аналогично есть методы получения числовых цепочек из аудиофрагментов.

Пока все наши знания хранились в виде бумажных документов, магнитофонных записей или на фотоплёнках – их невозможно было использовать для автоматической обработки математическими алгоритмами. Поэтому важно, что все девяностые годы человечество планомерно оцифровывало и накапливало на серверах информацию обо всех областях своей жизни: фотографии, голоса и речь, медицинские диагнозы, спектрограммы и много чего ещё.

Теперь задумаемся: как же человек в процессе развития приобретает спо-

собность к обобщению и анализу? Ответ прост – мы учимся. Причём для обучения используем опыт, полученный в реальной жизни. Как правило, существует некая высшая сила – родители, воспитатели, учителя, – которая показывает нам те или иные объекты или ситуации и говорит: «Это корова» или «Это буква “А”», а иногда «Так выглядит синекдоха».

Искусственные нейронные сети делают точно то же самое, только для них, как для математических алгоритмов, обучение представляет собой процесс подбора правильных значений огромного количества разнообразных параметров – коэффициентов. Однако для обучения нейронная сеть также использует информацию об объектах из реальной жизни: картинки, если речь идёт о распознавании графических объектов, звуковые фрагменты, если речь идёт о распознавании речи, и прочее, – для которых дополнительно известно, к какому классу следует отнести эти объекты. Процесс сопоставления фактов из реального мира с правильными ответами, которые должна выдавать нейросеть или иной обрабатывающий алгоритм, называется *разметкой* данных, а множество таких размеченных оцифрованных фактов – *обучающей выборкой*. Именно так работает один из основных методов обучения нейронных сетей – обучение с учителем. Чтобы нейронная сеть демонстрировала обобщающие свойства, для её обучения необходимо использовать большое количество оцифрованных данных, которые к тому же должны быть ещё и размечены некоторым учителем. На заре развития нейросетевых технологий разметкой данных занимались исключительно живые люди, это была передача их знаний вычислительной системе. Теперь алгоритмы накопили достаточно обобщающих способностей, чтобы проводить разметку некоторых данных без участия человека.

Всё это уже звучит довольно фантастически, в особенности когда мы

применяем к вычислительным алгоритмам термины из психологии, такие как, например, «обучение», «обобщающая способность», «распознавание». Однако всё это – прорывные технологии сорокалетней давности, удивить ими в настоящее время можно только обывателя, и то уже далеко не всякого. Волна популярности распознающих и обобщающих возможностей нейронных сетей прошла в середине десятых годов. На фоне того, какие способности демонстрируют технологии искусственного интеллекта в настоящее время, нам уже сложно вспомнить, чему же мы удивлялись тогда.

У человеческого сознания есть удивительное свойство...

А что же произошло? Как вот эти простые формулы научились общаться, создавать картины, музыку, стихи?

Прежде чем ответить на эти вопросы, хотелось бы прояснить кое-что про современную информационную науку. Мы с вами из школы помним, что математика – мать всех наук. В детстве эту мысль нам бездоказательно постулируют, и только в ходе профессионального развития появляется возможность удостовериться в этом в полной мере. Дело в том, что всё то, что мы называем «информационными технологиями», «программированием», «информатикой», «наукой о данных» – всё это разделы прикладной математики. Дискуссию о различиях прикладной математики и фундаментальной оставим за рамками данной статьи, она носила бы скорее религиозный характер, чем научно-популярный. Именно прикладная математика в применении к задачам, с которыми мы сталкиваемся в повседневной жизни, придумала нечто сногшибательное. Если говорить по-простому, она придумала как оцифровать *смыслы*.

Чуть выше я рассказывал про «классические» способы оцифровки текстов, картинок, звуков и иных видов информации. Основным недостатком их является то, что они сохраняют информа-

цию о некоторых физических объектах или явлениях (цветах точек, символах, силе давления воздуха на мембрану микрофона в какой-то момент времени) безотносительно к смыслу, который скрывается за ними. Пиксель на изображении мушиной кашки, налипшей на линзу объектива цифровой камеры, ничем не отличается от пикселя на изображении очаровательной родинки-мушки над верхней губой Мэрилин Монро. Пробел, запятая, вопросительный или восклицательный знак – такие же символы, как и любая буква или цифра. Более того, с точки зрения текстового представления, заглавная и строчная буквы «а» и «А» – это два разных символа. Одним словом, при классическом подходе выявление смыслов в данных было отдано полностью на откуп человеку. И самое смешное здесь то, что с этой точки зрения для человека действительно изображение (фотография) какого-либо текста мало чем отличается от, собственно, текстового представления. Вспомните, вы замечаете, что вам в мессенджере переслали не текст, а скриншот этого текста только в том случае, если вам необходимо его изменить или вставить в текстовый файл.

Для того чтобы алгоритм смог применить свои способности обобщать к смыслам, которые несут тексты, изображения, звуки – их необходимо как-то выявить и представить в виде последовательности чисел, сформировать в обучающие выборки, настроить параметры-коэффициенты и – вуаля! Однако для решения классическими методами прикладной математики эта задача является трудноформализуемой (это значит, что подобрать однозначное и исчерпывающее математическое описание для элементов задачи не представляется возможным) и – как следствие – нерешаемой.

Вообще в математике понятие «нерешаемый» абсолютно не несёт негативной коннотации, скорее наоборот. Если вы услышите от математика «не-

разрешимо», «нерешаемо», то, вероятно все, он это заявил с ликованием – вот наконец появилась ясность! Можно наконец расслабиться и отложить карандаш! В любой фундаментальной науке отрицательный результат – тоже результат. Чего не скажешь об инженерной, прикладной науке. У учёного-инженера стоит одна задача – сделать. И он должен сделать.

И инженеры-программисты придумали, как это сделать! Для преобразования последовательностей символов и пикселей в некие абстрактные сигналы-смыслы они использовали другие нейросети и придумали, как их можно обучить без явного использования разметки, то есть практически без участия человека. Они собрали из нейросетей, как из кубиков лего, цепочки и научили их дорисовывать изображения, продолжать тексты и аудиопотоки по введённому небольшому начальному фрагменту. Основную функцию обобщения выполняет по сути та же старая добрая простая нейросеть, только вот на вход ей приходят некие числовые шифры, в которых закодированы смыслы, и на выходе она формирует числовые шифры, которые также несут какие-то смыслы. При этом отдельные модули на входе и выходе преобразовывают привычные нам лексемы, изображения или звуки в этот шифр (кодируют) и обратно (декодируют). Примечательно то, что язык представления смыслов – этот самый шифр, – он не придуман человеком и человеку не понятен, он получается как бы сам в процессе подстройки числовых параметров нейросетей.

Но самое главное и сногшибательное изобретение, на мой взгляд, – это то, что учёные научились менять местами части этих конструкций так, что смыслы, выходящие из главной нейросети, можно декодировать в другой вид информации. Так, по введённому небольшому текстовому фрагменту (его называют промт – запрос) алгоритм без серьёзной перенастройки может выдать

текст, а может сформировать изображение.

Так во второй половине прошлого десятилетия, в 2015–2017 годах были построены первые нейросети-трансформеры, так началась эпоха массового использования генеративных алгоритмов в повседневной жизни. За прошедшие восемь-десять лет, конечно, многое изменилось. От первых топорных грамматически корявых ответов и сюрреалистичных картин с неестественным количеством пальцев и конечностей мы дошли до почти бунинской грамотности и фотореалистичной точности в ответах генеративных систем. Сегодня, осуществляя поиск информации в интернете с помощью поисковых сервисов «Яндекс» или «Гугл», мы редко доходим до второй страницы результатов, потому как получаем ответ сразу от интеллектуального помощника, который выдаёт чёткий и конкретный ответ на наш вопрос.

Не удивительно, что многих пугают эти системы. Ответы, которые выдаёт современный генеративный алгоритм (DeepSeek или ChatGPT), действительно похожи на ответы разумного человека, а тексты и изображения, созданные им, имеют все атрибуты продукта творческого труда. Сегодня по грамотно сформулированным запросам нейросеть способна сочинить песню, нарисовать картину, написать научную статью. Разных людей пугает в этом факте разное. Кто-то беспокоится за своё рабочее место, поскольку осознаёт, что компьютерная программа теперь способна выполнять множество трудовых функций, ещё недавно стабильно закреплённых за человеком: функции бухгалтера, секретаря, переводчика, дизайнера. Кто-то в ужасе признаёт, что не смог бы нарисовать картину или сочинить рассказ или речь лучше, чем это сделала DeepSeek, и погружается в пучины экзистенциального кризиса (ну, или запоя). А кто-то видит в этом зарождение искусственного разума и сознания,

которое – исключительно благодаря фантазии писателей и кинематографов – должно непременно создать человечеству массу проблем.

Между тем никаких причин для паники объективно нет. Автоматизация решения рутинных задач – стандартная для двадцатого и двадцать первого веков история. Борьба с научно-техническим прогрессом в здравом уме могут лишь персонажи литературных произведений (вспоминается «Механическое пианино» Воннегута), да и то лишь для того, чтобы украсить развитие какой-либо нравственной проблемы, увлекающей автора. Кроме того, наше общество и экономика очень обширны и неоднородны. Несмотря на то, что всё это вполне естественный системный процесс, происходящий по известным законам, никто в одночасье не выгонит на улицу всех секретарей и бухгалтеров. Не одно поколение HR-ов сменится, прежде чем в крупных компаниях произойдёт оптимизация в пользу сервисов искусственного интеллекта.

Кризис креативных профессий лично мне кажется куда более драматичным. Здесь нам всем будет не лишним заглянуть в свою работу и признать, что на 90 % мы занимаемся переработкой и обобщением уже существующего в мире материала. И правда, залог качественного контента в наше время – это эрудированность и усидчивость автора, умение выбирать и перерабатывать опыт классиков или других авторов. Это очень остро заметно в любом массмедиа продукте: поэзии, музыке, дизайне. Современное искусство и так глубоко вторично. Именно поэтому появление генеративных алгоритмов в свободном доступе должно не уничтожить, а наоборот – дать новый, дополнительный толчок к развитию искусства. Благодаря тому, что нейросеть может генерировать промежуточный рабочий материал (как подмастерье), творческий человек получает возможность приложить свои способности (если они, разумеется,

имеются) на более высоком уровне, выработать и построить что-то особенное и по-настоящему новое. Одним словом, талантливым и трудолюбивым людям бояться совершенно нечего.

Надо понимать, что современные генеративные системы – это математические структуры, переработавшие большое количество контента и способные его обобщать, усреднять и на основе него продолжать фразы в том или ином виде. Всё это может и должно стать прекрасным и мощнейшим инструментом в умелых руках.

А теперь давайте перейдём к самому главному вопросу, который учёные задают себе уже более ста лет: может ли машина мыслить? Остановимся именно на этой формулировке, поскольку именно так называется бессмертный труд Алана Тьюринга, опубликованный им в 1950 году – за несколько лет до создания первых нейронных сетей и спустя всего несколько лет после создания первого компьютера! Авторитет Тьюринга в этом вопросе абсолютен. Мы обязаны этому человеку всем, что сейчас имеем в области информатики – от компьютеров до нейросетей. В предисловии к русскоязычному изданию этой статьи в 1960 году выдающийся советский математик и философ Софья Александровна Яновская пишет:

«Само возникновение [этого вопроса] ... обусловлено созданием новой техники, построением машин, решающих задачи, которые обычно решаются людьми с помощью рассуждения; машин, могущих – в зависимости от результатов своей деятельности – изменять свою программу; машин, которым можно “поручить” составление новых программ, более того, синтез новых – даже более сложных – машин. Естественно, что создание таких машин требует пересмотра самого понятия “машины”, более глубокого анализа процессов, связанных с умственной деятельностью человека, и ответа на вопрос о том, что из этой деятельно-

сти и как именно может имитироваться машиной. Моделирование умственной деятельности человека автоматами позволяет, с другой стороны, осветить такие черты в структуре человеческого мозга и нервной системы, к которым раньше вообще не было доступа».

В данной статье Алан Тьюринг предлагает довольно простой «тест на интеллектуальность», пройдя который якобы мы можем считать машину мыслящей. Заключается он в том, что живой человек, общаясь с машиной на естественном языке, не должен догадаться, что он общается не с другим живым человеком. До недавнего времени казалось, что это недостижимая задача для программы. Современные генеративные алгоритмы начиная с GPT-4 без каких-либо проблем проходят тест Тьюринга, при этом не проявляя иных признаков интеллекта. Очевидно, что такой «феноменологический» подход к определению интеллектуальности является однобоким. В 1980 году американский философ Сёрль в порядке критики тьюринговского подхода к определению интеллекта предложил мысленный эксперимент, который назвал «Китайская комната». Он демонстрирует, что возможно создать программу, которая будет формировать ответы на запросы по произвольно сложному алгоритму, при этом не анализируя и не вдаваясь в смысл запросов, которая по определению должна успешно пройти тест Тьюринга. Противоречие этих двух подходов, как мне кажется, можно сформулировать в виде вопроса: можно ли считать интеллект исчерпывающе проявляющимся через язык, или интеллект по определению нечто большее, а язык – лишь одно из его проявлений?

Как вы видите, попытки расставить все точки над «ё» неизбежно погружают

нас в пучины софистики и демагогии. Поэтому предлагаю оставить эти вопросы открытыми.

Однако как программист, математик и философ могу сказать совершенно точно: любая современная генеративная система, пусть даже позволяющая работать в режиме человекоподобного диалога, – это не более чем сервис, который использует ваши запросы как «затравку» для создания «продолжения». Иными словами, рассматривать её как самостоятельный автономный разум, подобно тому, как мы рассматриваем других людей, безусловно нельзя, как велеречиво и грамотно бы она ни отвечала на наши запросы. Это, однако, не значит, что с ней нельзя взаимодействовать как с человеком, задавая вопросы, корректируя рассуждения, формулируя просьбы – её именно для этого создали программисты. Если уж для нас в порядке вещей разговаривать с животными, растениями, предметами, а иногда и с сущностями, которых не существует, не ожидая никакого ответа, так почему бы не поговорить с тем, кто хотя бы отвечает?

Одним словом, учёные и инженеры разработали для нас ещё один мощнейший инструмент, который может перевести человечество на новый уровень научно-технического развития, который может позволить человеку более полно реализовать свои способности. Сегодня даже создатели генеративных нейросетей до конца не осознают границ возможностей своих алгоритмов и программ.

Мы живём в удивительное время, на наших глазах творится настоящая история науки и техники! Так давайте же не будем оставаться в стороне, только тот, кто всегда открыт новому, будет всегда молод душой!