

Содержание

Предисловие	7
Введение	23
Часть I. Понимать, но не предвидеть	35
Понимать, не зная законов?	36
Малоизученные причины с непредсказуемыми следствиями	41
Причины известны, но следствия непредсказуемы	42
Детерминированный хаос: четыре примера ...	45
Бильярд Синая	46
Логистическое уравнение	48
Метеорология	56
Движение планет	58
В заключение немного философии	62
Часть II. Предвидеть, но не понимать?	65
Достаточно ли наблюдать, чтобы знать?	67
Падение тел	69
Движение Земли (<i>E pur si muove</i>)	71
О математике	75
Корреляция — это не причинность	79
Действовать, не понимая?	89

Зачем нужны большие данные.	91
Искусственный интеллект.	94
Социализация данных и расчетов?	108
Разум машины и человека.	114
О соперничестве человека и машины	116
Нужны ли нам доказательства?	121
Наука не развивается, если просто накапливать данные	126
Ограничения больших данных	128
Будущее, которое предстоит открыть	135
Список использованной литературы	153
Алфавитный указатель	157

Предисловие

Некоторые книги словно интересные попутчики. Они дают надежду отыскать новые пути и вершины. Они заставляют нас идти все дальше и самостоятельно продолжать начатое вместе путешествие, чтобы выйти за пределы достигнутых с ними границ.

Одна из них — книга Юбера Кривина «Земля: от мифов к знанию». Эта книга, опираясь на историю идей, объединяет исследования разных эпох и подробные описания разнообразных экспериментов и открытий, подтверждающих то, о чем ученые лишь догадывались, то, что они когда-то предсказывали, а потом изучали; она дает нам возможность оживить прошлые успехи и неудачи ученых, противоречия, с которыми они сталкивались, примкнуть к их блужданиям в мире предположений, но при этом не потеряться там самим. Она предлагает полную и подробную картину того, как зарождались и развивались современные физика и география. Она позволяет увидеть как красоту и эффективность, так и тщетность научных достижений. Она показывает, насколько часто ответы, которые открывают нам путь к пониманию мироздания, оказываются далеки от обычного здравого смысла. Эта книга со страстью передает нам не только знания, но и невероятный опыт человечества на бесконечном извилистом пути, который к ним приводит, — пути научных исследований.

Книга «Понимать, но не предвидеть. Предвидеть, но не понимать», по страницам которой Юбер Кривин ведет нас шаг за шагом, оставляет впечатление легкости и ясности мысли. Но вопросы, которые в ней поднимаются, очень непростые. Что такое «понимать»? И что такое «предвидеть»?

Итак, что же означает «понимать»?

Разработка новой теории подразумевает выход за пределы той области, которой принадлежит изучаемый вопрос, расширение границ размышлений — выход в космос неизвестности. Из этого путешествия мы возвращаемся не только с решением конкретной задачи — мы получаем в свое распоряжение общий принцип, который коренным образом изменяет наше видение того, что мы называем реальностью, и который позволяет ее прогнозировать, выявляя неизвестные до того компоненты. Кривин показывает нам это вместе с Коперником, Галилеем, Ньютоном, Больцманом, Эйнштейном...

Творчество в науке часто связано не только с интеллектом, но и с эмоциями, фантазиями, созерцанием и размышлением, интуицией, случайными встречами и неожиданными выводами. С воображением. И со многими другими, еще неясными процессами, ускользающими от нашего осознания. В книге «Глаз разума» невролог и писатель Оливер Сакс цитирует Эйнштейна: «Физические сущности, служащие элементами мышления, являются определенными символами и более или менее ясными образами, которые можно произвольно вызывать в воображении и комбинировать. Некоторые из этих образов (в моем случае) являются зрительными, некоторые — мышечными. Обычные слова или прочие символы с трудом подбираются уже на второй стадии»¹. Слова, символы и образы — это

¹ Цит. по: Сакс О. Глаз разума. — М.: АСТ, 2018.

и есть первоначальные инструменты мысли? Или же существуют формы мысли, предшествующие им, такие, которым сложно дать конкретную характеристику и точно их описать? О таких мыслях, принимающих форму чистого значения, говорил русский психолог Лев Выготский. Мыслеформы, предшествующие образам и словам и не поддающиеся описанию...

Иногда некоторые интуитивные мысли посещают нас в странных сновидениях. Химик Август Кекуле говорил, что как-то заснул во время поездки в омнибусе в Лондоне и именно во сне увидел химические связи между атомами, что позволило ему выдвинуть свою теорию химической структуры. «Хотя химические связи в природе обычно не видны, — пишет Оливер Сакс, — они так же реальны и видимы, так же визуально вообразимы для Кекуле, как для Фарадея линии напряженности вокруг магнита, электромагнитные поля». Но самым знаменитым сном Кекуле стал тот, благодаря которому он открыл формулу бензола. Это случилось в Генте, в Бельгии, во время работы над учебником по химии. Ему никак не удавалось сосредоточиться. «Работа не двигалась, мои мысли витали где-то далеко. Я повернул стул к камину и задремал. Атомы снова запрыгали перед моими глазами. На этот раз небольшие группы скромно держались на заднем плане. Мой умственный взгляд мог теперь различить длинные ряды, извивающиеся подобно змеям. Но смотрите: одна из змей схватила свой собственный хвост и в таком виде, как бы дразня, завертелась перед моими глазами. Как будто вспышка молнии разбудила меня. И я провел всю ночь, разрабатывая следствия из гипотезы». Змея с хвостом в пасти — Уроборос, что в переводе с греческого значит буквально «кусающий хвост». Именно в таком виде во сне, по словам Кекуле, ему предстала неизвестная прежде структура бензола.

Шестиугольник, сформированный шестью атомами углерода, каждый из которых связан с атомом водорода. «Будем учиться видеть сны, господа, и, возможно, тогда нам откроется истина», — сказал он, выступая перед Немецким химическим обществом в 1890 году по случаю празднования двадцатипятилетия своего открытия. «Будем учиться видеть сны! Но лучше воздерживаться от обнародования своих снов до тех пор, пока их содержание не будет проверено в состоянии бодрствования».

А что же такое «предвидеть»?

«Наше воображение, — говорит Франсуа Жакоб, — разворачивает перед нами постоянно обновляющуюся картину того, что может произойти, того, что возможно. Мы не можем думать о себе вне текущего момента, но нам не дано знать, чем станет этот момент. Увы, мы не можем узнать того, что интересует нас больше всего на свете: что произойдет завтра».

Хрустального шара, показывающего будущее, не существует. Любое предвидение — это экстраполяция на основе прошлого. Понимание причинно-следственных связей и ограничений, известных по прошлым ситуациям, позволяет нам не «узнать, что будет завтра», а прогнозировать развитие конкретных событий — чаще всего очень приблизительно. А умение предвидеть события, которые могут произойти, часто означает также и умение их предотвращать, ограничивая поле определенных возможностей. Нельзя предотвратить заболевание краснухой. Но можно предположить, что прививка значительно снизит риск заболеть.

При этом понимание не обязательно ведет к предвидению, говорит нам Кривин. И знакомит нас с теориями, позволяющими понимать и формулировать математический процесс детерминированного хаоса. Детерминированный хаос — очевидный пара-

докс, напоминающий об идее слияния случайного и необходимого, которую приписывают Демокриту. Хаос — иначе говоря, нарушение порядка, откровенно беспорядочные проявления, невозможность ничего предугадать, — и одновременно детерминированный, то есть подчиняющийся законам природы и причинно-следственным связям. При этом теория хаоса гласит, что сложные системы чрезвычайно зависимы от первоначальных условий и что небольшие изменения в окружающей среде могут быть чреваты непредсказуемыми последствиями. «То, что бабочка вдруг взмахнет крылышками в Бразилии, — метафорически объяснял Эдвард Лоренц, — может привести к возникновению торнадо в Техасе». Это называют чувствительностью к начальным условиям. Научные модели, ориентирующиеся на теорию хаоса, не позволяют предсказывать то, что случится вне заданного временного отрезка. И неважно, идет ли речь о будущем или о прошлом. В 2011 году астроном Жак Ласкар с коллегами показал, что хаотичное движение двух гигантских астероидов, Цереры и Весты, расположенных в пространстве между Марсом и Юпитером, мешает прогнозировать траектории этих планет далее чем на 40 000 лет. И что эта нестабильность влияет на траекторию Земли: орбиту, по которой Земля движется вокруг Солнца, можно просчитать только на временном отрезке в 60 миллионов лет. И в прошлое, и в будущее. «Церера и Веста для Солнечной системы и есть тот самый эффект бабочки, — говорил Седрик Виллани. — Палеонтологи, возможно, никогда не узнают, какой была орбита Земли, когда по ее поверхности ходили динозавры».

Вместе с тем «знание об ограничениях не является ограничением знания», говорит Кривин. Определить, на каком этапе, вне какого временного отрезка прогнозирование невозможно, — уже само по себе

прогноз. Оценить, подсчитать степень неясности — значит установить пределы нашего неведения относительно будущих событий. Это — воспользуюсь термином из трактата Николая Кузанского, гуманиста эпохи Возрождения, — есть *docta ignorantia*¹. Мы всегда лишь предполагаем, поскольку наука беспрестанно открывает не только то, о незнании чего нам было известно, но и то, о незнании чего мы даже не подозревали.

Понимание не обязательно позволяет прогнозировать. И наоборот. Как утверждает Кривин, для того чтобы предвидеть, просчитывать, угадывать и делать выводы, вовсе не обязательно обладать пониманием происходящего. Быть способным, опираясь на громадную массу разнородных данных (большие данные²), устанавливая надежные, не вызывающие сомнений корреляции между заведомо разрозненными элементами, событиями или процессами и, отталкиваясь от этих корреляций, приходиться к определенным выводам, открытиям или прогнозам не всегда означает полное понимание происходящего. Достаточно открыть все причинно-следственные связи, стоящие за этими корреляциями. И выработать новые теории.

Начав с исследования принципов научного подхода, Кривин приходит к сложному вопросу о различии между человеческим интеллектом и тем, что мы сегодня обозначаем двусмысленным термином «искусственный интеллект».

¹ Ученое незнание (лат.).

² Большие данные (англ. *big data*) — общее название для структурированных и неструктурированных данных огромных объемов, которые эффективно обрабатываются с помощью масштабируемых программных инструментов. — *Прим. пер.*

Искусственный интеллект открывает перед нами новый мир, очертания которого еще четко не определены. Может быть, мы являемся свидетелями рождения новых методов научного исследования? Заменителей человеческого подхода к науке? Или же просто нового, очень мощного инструмента?

Одно из величайших различий между нами и умными машинами, по мнению Кривина, состоит в том, что любая из них пока может фокусироваться лишь на узком круге задач. Одни лучше людей играют в шахматы, или в го, или в покер, или в «Рискуй!»¹. Другие диагностируют злокачественные поражения кожи столь же эффективно, как целая команда дерматологов.

Некоторые программы, как сообщается в статье, опубликованной в журнале *Nature* в конце марта 2018 года, не хуже профессиональных химиков выстраивают последовательности реакций, позволяющие синтезировать органические молекулы из составляющих. А есть и такие машины, которые, взаимодействуя с нами, начинают все лучше и лучше различать человеческие эмоции. И чем больше данных получает искусственный интеллект, способный к самостоятельному обучению, тем правильнее он обрабатывает информацию и тем эффективнее становится. Он работает нечеловечески методично и сфокусированно, но переключаться на другие задачи не умеет. И в погоне за столь невероятной эффективностью мы рискуем лишиться того, что могли бы получить с помощью исследований, воображения, интуитивных изысканий и попыток понять, если бы не были так сфокусированы

¹ «Рискуй!» (англ. *Jeopardy* — «рисковать») — популярная американская телевикторина, в которой участники должны подобрать к имеющимся утверждениям правильный вопрос. — *Прим. пер.*

на узконаправленных задачах, — в этом убежден специалист в области информационных технологий и вычислительной техники Эдвард Теннер. «Воображение важнее знания, — говорил Эйнштейн. — Знание ограничено, а воображение охватывает весь мир». Но что такое воображение?

У каждой эпохи есть свои «технические» метафоры для жизни и мозга, которые неизменно сравнивают с самыми сложными механизмами из тех, что сумел создать человек, — сегодня это цифровые технологии и искусственный интеллект. Достижения нейрофизиологии — науки, изучающей человеческий мозг, — помогают совершенствовать искусственный интеллект. А тот, в свою очередь, помогает нам лучше понимать, как функционирует наш мозг. То, что мы называем интеллектом, возникает в результате взаимодействия множества процессов, из которых ни один по отдельности создать интеллект не может. Аналогичным образом за психические и мыслительные процессы отвечает множество взаимодействующих друг с другом нейронов, объединенных в нейронную сеть, причем ни один нейрон в отдельности управлять психической активностью не может. И аналогичным образом живая клетка возникает из взаимодействия множества молекул, из которых ни одна в отдельности не обладает свойствами живого организма. Этот феномен известен как эмерджентность: на основе определенных комбинаций и отношений между элементами формируются системы, обладающие новыми свойствами, которых образующие их элементы сами по себе лишены. Так излагает это Филип Андерсон в своей статье *More is different* («Количество изменяет свойства»), о которой рассуждает Кривин. Это одна из причин, по которым сложно предсказать, сможет ли нынешний искусственный интеллект, узконаправленный и сфокусированный

на одной цели, за счет постоянного совершенствования и взаимодействия с внешним миром однажды превратиться в другой — общий, универсальный искусственный интеллект, сверхинтеллект, благодаря чему появятся машины, одаренные качественно новыми умениями, сравнимые с нами по широте сознания, субъективности, эмоциям, способности мечтать... «Что такое сознание и может ли машина его иметь?» — так поставили вопрос Станислас Деан и его коллеги в статье, опубликованной в журнале *Science* в октябре 2017 года. Идея заключалась в том, что нейронные сети, отвечающие за глубокое обучение систем искусственного интеллекта, действуют схожим образом с нашими бессознательными ментальными процессами; ученые размышляли, при каких условиях эти нейронные сети смогли бы достичь уровня, сопоставимого по сложности с человеческим сознанием, и обрести субъективное восприятие.

«Мы состоим из странной смеси нуклеиновых кислот и воспоминаний, снов и протеинов, клеток и слов», — писал Франсуа Жакоб. Одно из фундаментальных различий между человеческой психической деятельностью и искусственным интеллектом машин, говорит нам Кривин, заключается в том, что у нас есть тело. И в том, что тело и разум постоянно взаимодействуют. Более того, как утверждал Спиноза, «тело и разум суть одно и то же в двух различных аспектах»¹.

У машин тела нет — в том смысле, в каком оно есть у нас. Но разве обязательно машинам во всем походить на нас, чтобы нас «превзойти»? Обязательно ли здесь ориентироваться на идею антропоморфности?

¹ Цит. по: Панасюк В. Ю. История зарубежной философии: М.: 2001 (http://www.gumfak.ru/filos_html/ponas/pon13.shtml).

Для того чтобы машины сравнялись с нами, скорее необходимы навыки и умения, определенная степень автономности и умение ставить себе цели, способные привести машины к равным с нами возможностям и затем к бегству от нашего контроля. И совершенно не обязательно, чтобы «существа», способные ставить такие цели, были во всем похожи на нас. Мне кажется, в основе опасений, связанных с возникновением «суперразумных» машин, лежат два страха. Первый — страх того, что машины слишком «очеловечатся» и мы утратим нашу эксклюзивную «гуманистичность», превратившись лишь в одну из возможных форм разумной жизни (что уже происходило в прошлом, когда существовало несколько человекообразных видов). А второй страх, напротив, состоит в том, что машины, не обладающие ничем из того, что мы считаем гуманистическим, выйдут из-под нашего контроля и нам придется противостоять чему-то нечеловеческому или даже античеловеческому.

Но «есть в этих страхах слепая зона, — написала Кейт Кроуфорд, научная сотрудница компании «Майкрософт», в комментарии, опубликованном в журнале *Nature* осенью 2016 года. — Этот страх будущего заставляет забыть о реальном риске при работе ныне существующих систем». По мнению исследователя в области компьютерных технологий Педро Домингоса, «люди беспокоятся о том, что компьютеры станут слишком умными и захватят мир, тогда как настоящая проблема в том, что сами они абсолютно глупы и уже захватили все, что можно».

Машинные алгоритмы, большие данные, искусственный интеллект каждый день предоставляют нам и будут предоставлять завтра огромное количество услуг. Но человеческие предвзятость и предрассудки могут повлиять на алгоритмы работы и программы

обучения, которыми мы снабжаем искусственный интеллект. И есть риск, что эти предвзятость и предрассудки только углубят неравенство за счет отдельных лиц и социальных групп, которые и так уже маргинальны, дискриминированы и лишены права голоса.

Одна из проблем, связанных с использованием алгоритмов и искусственного интеллекта, — чрезмерное доверие к полученным ими результатам. Исследование, опубликованное в журнале *Science Advances* в январе 2018 года, показало, что использование алгоритмов судьями в Соединенных Штатах для прогнозирования рецидивов преступного поведения у подсудимых давало не лучший результат, чем умозаключение группы людей, полностью лишенных подготовки в данном вопросе. При этом уровень слепого доверия к результатам работы алгоритмов оказывается тем выше, чем недоступнее исходные коды программ, чем недоступнее данные для обучения искусственного интеллекта и чем сложнее отследить этапы самообучения машин и последовательность шагов, приводящих к ответам.

Может быть, в нашем сегодняшнем отношении к возможностям прогнозирования, которыми обладает искусственный интеллект, есть что-то наивное. Следовательно, одним из прогнозируемых последствий его развития станет основательное потрясение сферы трудоустройства в большинстве (если не во всех) профессиональных областей. Тем не менее, как ни странно, использование искусственного интеллекта для изучения этого вопроса в настоящее время, насколько я знаю, не является основной задачей. Инструменты машинного прогнозирования до сих пор почти совсем не использовались, для того чтобы определить последствия их массового применения...

Пока речь идет о процессах, правила и цели которых нам заведомо известны, потому что мы сами

их установили — как в случае с шахматами, го и так далее, — адекватность результатов, полученных искусственным интеллектом, не вызывает вопросов. То же самое касается и ситуаций, когда у нас есть возможность быстрой проверки результата, как в случае с этапами химического синтеза. Но если цель в том, чтобы добиться от искусственного интеллекта решения задач, справиться с которыми мы сами никак не можем, и/или чтобы получить от него инструкции по их решению, проблема становится глубже. Нам бы хотелось, чтобы искусственный интеллект неожиданно и непредсказуемо предложил такое решение, о котором мы и не думали. Но как проверить достоверность этого непредвиденного решения? Чем неожиданнее и, соответственно, интереснее результаты, тем больше потребуется времени и исследований для их проверки, определения их последствий и даже просто для их понимания. Таким образом, чтобы развивать искусственный интеллект, нужно тратить время и усилия не только на предварительные исследования, но и на последующие.

Еще одна проблема, как мне кажется, состоит в том, что при безнадзорном обучении искусственный интеллект — так же как и хаотические системы — очень зависим от первоначальных условий: даже минимальные изменения в обучающем наборе данных могут привести к непредсказуемым последствиям. И по этой причине возникает проблема воспроизводимости. Минимальная вариативность данных и разница в наборе итераций может привести к довольно-таки различным результатам. А во всех областях, связанных с общественным благом (здоровье, правосудие, защита частной жизни, общественная безопасность), воспроизводимость результатов совершенно необходима.

Искусственный интеллект позволяет нам сегодня и в еще большей степени позволит завтра выполнять задачи и добиваться целей, которые прежде казались недостижимыми, причем эта работа будет даваться быстрее и легче. Любые научные и технические революции сильно влияют на повседневную жизнь — несомненно, революция в развитии искусственного интеллекта, разворачивающаяся сейчас, тоже вызовет потрясения в культурной и общественной жизни. Поэтому всем нам необходимо как можно скорее задуматься о возможных преимуществах и недостатках этого развития с точки зрения не только эффективности, продуктивности и экономической выгоды, но и гуманистических ценностей.

Вдобавок к этому развитие искусственного интеллекта ставит перед нами важнейшие вопросы об истощении невозполнимых минеральных ресурсов и увеличении потребления энергии. К сожалению, мы принимаемся изучать и исправлять негативные последствия технологических прорывов лишь постфактум; попытки контролировать, корректировать и предотвращать нежелательные эффекты нашей деятельности начинаются только тогда, когда те достигают значительных масштабов. Загрязнение окружающей среды, утрата биологического разнообразия и изменение климата — лишь несколько примеров. Остается только надеяться, что в случае с развитием цифровых технологий и искусственного интеллекта дело будет обстоять по-другому.

Истинный прогресс человечества — не достижения в технической и научной областях сами по себе, а то, как они используются. При этом важно, соблюдаются ли права каждого человека, сохраняются ли ресурсы нашей планеты и доступны ли они всем в равной степени.

Наша задача — сделать так, чтобы достижения науки и техники, в том числе машинные алгоритмы, большие данные и искусственный интеллект, использовались всем во благо. Чтобы они не становились инструментами власти и обезличивания, помогая навязывать нормативы, насаждая все очевидно объективное и делая невозможным любое субъективное взаимодействие. Чтобы их использование не привело к «царству цифр», как говорил Ален Сьюпио¹, который предполагал, что целью в будущем может стать «самое эффективное достижение измеримых целей, а не следование справедливым законам». Царство это будет основано на точном прогнозировании, которое уничтожит будущее, борясь с ним в настоящем, стирая различия времен, отбирая те события, которые можно предсказать, и отбрасывая те, которые предсказать невозможно. Конформизм станет всеобщей философией, заставляя бояться всего необычного, редкого, уникального и нового, поскольку его нельзя ни с чем сравнить. Из соображений экономии времени и «объективной» пользы при принятии решений люди все меньше будут думать о гуманистических ценностях. Но ведь чем больше развиваются и используются эти технологии, тем более человечными мы должны становиться.

«Понимать, но не предвидеть. Предвидеть, но не понимать» — книга необходимая. Это источник интеллектуального обогащения и удовольствия, удивительная смесь глубины мысли, строгости рассуждений и юмора, путешествие к самому сердцу науки и размышления над неоднозначными вопросами, что и отличает человеческий разум от машины. Но поскольку она призвана помочь нам понять, чем является

¹ Директор Нантского института высших исследований. — *Прим. пер.*

сегодня — и чем не является — искусственный интеллект, то снабжает нас средствами самоанализа, заставляя спрашивать себя о том, кто такие мы сами и кем хотим быть. Она позволяет вообразить и исследовать поле своих возможностей, невзирая на неизвестность и сомнения, не отказываясь от ответственности. И в конце концов сделать демократичный выбор, осознанно желая сохранить все то редкое, драгоценное и исключительное — подлинно человеческое, — что есть в каждом из нас.

Жан-Клод Амейзен — ученый-медик, почетный президент Национального консультативного комитета по этике, ведущий радиопрограммы «На плечах Дарвина» (канал France Inter)