

1.1. Искусственный интеллект от Адама до Франкенштейна



Первый профессор... сказал, что меня, быть может, удивит его работа над проектом усовершенствования умозрительного знания при помощи технических и механических операций... Затем он подвёл меня к раме, по бокам которой рядами стояли все его ученики... Поверхность её состояла из множества деревянных дощечек... Все они были сцеплены между собой тонкими проволоками. Со всех сторон каждой дощечки приклеено было по кусочку бумаги, и на этих бумажках были написаны все слова их языка в различных наклонениях, временах и падежах, но без всякого порядка... По его команде каждый ученик взялся за железную рукоятку, которые в числе сорока были вставлены по краям рамы, и быстро повернул её, после чего расположение слов совершенно изменилось. Тогда профессор приказал тридцати шести ученикам медленно читать образовавшиеся строки в том порядке, в каком они разместились в раме; если случалось, что три или четыре слова составляли часть фразы, её диктовали... Ученики занимались этими упражнениями по шести часов в день.

Джонатан Свифт. Путешествия в некоторые отдалённые страны света Лемюэля Гулливера, сначала хирурга, а потом капитана нескольких кораблей (пер. А. Франковского)

1.1.1. Ранняя история

Увидеть, как идея искусственного интеллекта развивалась в истории человеческой мысли, не так-то просто. Дело в том, что собственно искусственный интеллект как наука в её современном понимании сложился совсем недавно — обычно этим моментом считается создание теста Тьюринга, о котором мы поговорим ниже. До этого, конечно, были и другие рассуждения о мыслящих машинах, и попытки решать задачи, которые мы могли бы сейчас назвать задачами искусственного интеллекта, но грань между AI и другими науками провести сложно. Тем не менее давайте попробуем кратко пройти по истории того, как люди мыслили себе искусственный разум (рекомендую также уже переведённую на русский язык книгу Клиффорда Пикова [469, 688]).

Начнём с самого начала, прямо с книги Бытия. Привычные нам авраамические религии учат, что Бог создал человека по своему образу и подобию из глины, смешанной с разнообразными жидкостями; по всей видимости, эта идея восходит к вавилоно-аккадскому эпосу, в котором бог Мардук сотворил человека из глины, смешанной со слюной и кровью. Получается, все мы — пример искусственного интеллекта. И не так уж важно, создал ли ветхозаветный Бог людей сразу готовыми из глины, лично и самостоятельно разработав архитектуру нейронных сетей, или вольтеровский часовщик¹ создал и «завёл» механизм, а люди получились из него сами собой, — мы ведь сегодня не называем результат генетических алгоритмов или автоматического поиска нейросетевых архитектур² «естественным интеллектом».

В древнейшей авраамической традиции, иудейской, есть и другой яркий пример искусственного интеллекта. Как известно хотя бы из романа Густава Майринка, особо мудрые раввины могли создавать и оживлять *големов*, искусственных людей; в самой известной такой легенде творцом искусственного интеллекта выступает главный раввин Праги Йехуда бен Бецалель. Големы делались, что неудивительно, тоже из глины, и даже само слово «голем» означает «глина» или любой «необработанный материал».

Любопытно, что иудейская традиция различает задачи *распознавания речи* и её *синтеза*: големы не слишком умны, но могут слушать и понимать команды, хотя понимают их порой слишком буквально (примерно как в некоторых более современных историях о роботах). А вот отвечать и вообще говорить на человеческом языке — не могут, для этого, видимо, нужна душа, которую даже самый мудрый раввин вложить не может.

Греческие боги были не слишком склонны к научным занятиям, и разработку методов искусственного интеллекта в Древней Греции можно было ожидать

¹ *Франсуа-Мари Аруэ*, или *Вольтер* (François Marie Arouet, Voltaire, 1694–1778) — французский писатель и философ, один из главных деятелей Просвещения. В чисто художественной прозе Вольтер двигался от классицизма — он начинал как автор классической трагедии («Эдип», «Брут») и классической эпопеи («Генриада», «Поэма о битве при Фонтенуа») — к романтизму и сентиментализму («Танкред»), а порой и просто к сатире и фарсу («Орлеанская девственница»). Ещё более известна его философская проза: романы «Простодушный», «Кандид», «Задиг» и другие внесли важный вклад в Просвещение, популяризируя философские рассуждения и развивая новые интересные аргументы. Вольтера всегда хорошо знали в России: Екатерина II сначала вела с ним переписку, а после смерти философа выкупила его огромную библиотеку: 6814 книг и 37 томов с рукописями привезли в Петербург на специальном корабле; а просветительское движение, направленное против традиционной религии и феодализма, на защиту прав человека и равенства перед законом, в России во второй половине XVIII века стали буквально называть «вольтерьянством».

² Neural architecture search (NAS) — один из важных современных методов разработки архитектур глубоких нейронных сетей, в котором архитектуры получают «сами собой» посредством направленного поиска либо генетическим алгоритмом, либо обучением с подкреплением. При помощи NAS уже было разработано множество архитектур, которые применяются сплошь и рядом, — функции активации Swish, семейства свёрточных сетей EfficientNet и EfficientDet и так далее — и не так уж ясно, чем наша роль в их создании принципиально отличается от роли того самого вольтеровского часовщика...

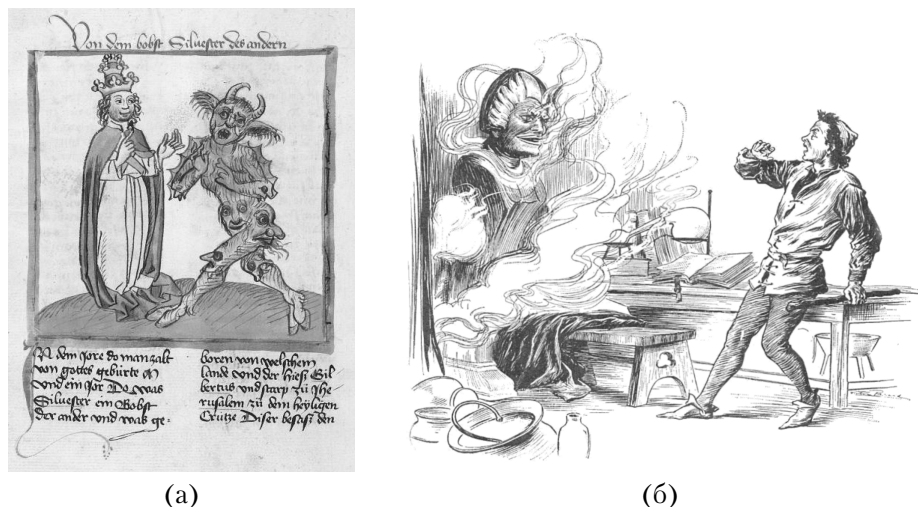


Рис. 1.1. Средневековая магия: (а) папа Сильвестр II договаривается с дьяволом; (б) говорящая голова Роджера Бэкона пугает его слугу

разве что от великого кузнеца и изобретателя Гефеста. Гефест не подвёл — он создал гигантского человекоподобного робота по имени Талос ($\Gamma\alpha\lambda\omega\varsigma$)¹. У Талоса была одна вена, по которой тек божественный ихор; вена была заткнута гвоздём, который находился в пятке, по греческой традиции единственным уязвимом месте Талоса. Зевс подарил Талоса Европе для охраны острова Крит; андроид обегал остров по периметру трижды в день и бросал гигантские камни во вражеские корабли.

Здесь для нашего современного взгляда любопытно, что метко бросать гигантские камни человечество научилось довольно скоро: собственно, те же древние греки ещё и до Архимеда умели создавать вполне рабочие катапульты; а вот отличать вражеские корабли от своих, да и вообще различать корабли на фоне воды и понимать, где они находятся, — это непростая задача компьютерного зрения, решение которой до сих пор, наверное, можно улучшать.

Средневековые алхимики часто заключали договоры с дьяволом (рис. 1.1), по которым получали свои сверхъестественные способности; иногда эти способности были весьма похожи на искусственный интеллект.

¹ В наше время в популярной культуре Талос появляется реже, чем другие персонажи греческой мифологии, но забавно, что это обычно связано с компьютерными играми. Игра *Talos Principle* — интересная головоломка с философским сюжетом, в которой вы играете за искусственный интеллект... Пожалуй, не буду спойлерить дальше. А армия США с 2013 до 2019 года безуспешно разрабатывала *Tactical Assault Light Operation Suit* (TALOS), роботизированный экзоскелет, который должен был защищать носителя при помощи «жидкой керамической брони» из наночастиц. Как тут не вспомнить бессмертное «Nanomachines, son!» от Хидео Кодзимы (кстати, *Metal Gear Rising: Revengeance* вышла в том же 2013 году).

Так, через всё Средневековье проходит бродячий сюжет о говорящей голове, обычно бронзовой, которая могла самостоятельно отвечать на вопросы; видимо, магическая часть этого сюжета опиралась как раз на иудейскую традицию. В «Истории английских королей» Уильяма Мальмсберийского, опубликованной в 1125 году, говорящая голова упоминается в описании чудес папы Сильвестра II¹, который привёз из аль-Андалуса книгу с секретными знаниями арабов и смог избежать гнева владельца книги лишь с демонической помощью (см. рис. 1.1, а). Говорят, что известный алхимик Альберт Великий изготовил говорящую голову, чем немало расстроил своего учителя, Фому Аквинского, который в приступе гнева сломал дьявольскую машину; по одной из версий, эта говорящая голова скрывала в себе длинную трубку, посредством которой сам Альберт мог вещать из соседней комнаты. Подобные истории ходили о Боэции, Арнольде из Виллановы, Роберте Гроссетесте и многих других, но самой известной стала голова, созданная Роджером Бэконом² (см. рис. 1.1, б).

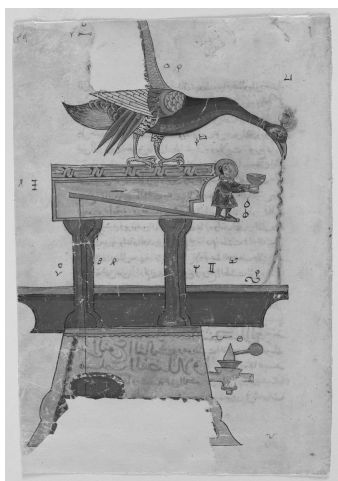
1.1.2. Механические автоматы Средних веков и Нового времени

В Средние века мы уже можем и перейти от мифов и легенд к настоящим созданиям человеческого гения. Конечно, об искусственном интеллекте в его современном понимании говорить было рано, но разные механические и гидравлические приспособления становились всё более изощрёнными и изобретатели часто придавали своим механизмам человекоподобный вид. Уже в XII веке арабский изобретатель аль-Джазари³, которого иногда называют «отцом роботики»,

¹ *Сильвестр II* (Silvester PP. II, ок. 946–1003) — средневековый учёный и церковный деятель, папа римский с 999 по 1003 год, известный до избрания папой как Герберт Орильякский (Gerbertus Aureliacus, Gerbert d'Aurillac). Герберт долго жил и работал в Барселоне, где благодаря связям с Кордовой получил доступ к знаниям арабского мира; так, Герберт одним из первых европейцев познакомился с арабскими цифрами, осознал их удобство и стал пропагандировать их использование в Европе. Он популяризировал и другие достижения арабской науки, вернув в Европу абак, армиллярную сферу (модель небесной сферы, использовавшуюся для определения координат светил) и астролябию, которые не просто скопировал, а усовершенствовал. Разумеется, Герберту приписывали изучение магии и астрологии в Кордовском халифате, общение с дьяволом и связь с сукубом по имени Меридиана, а бронзовая голова, умевшая отвечать на вопросы «да» или «нет», якобы помогла ему подняться к папскому трону.

² *Роджер Бэкон* (Roger Bacon, ок. 1214–1292) — английский философ, теолог и естествоиспытатель, монах-францисканец и профессор богословия в Оксфорде. Его главная работа, *Opus majus* (1268), содержит новые результаты в математике, оптике, алхимии и астрономии, предлагает реформу календаря, которую через 300 лет проведёт папа Григорий XIII, а также первой в Европе описывает порох. История с говорящей головой — это взгляд на Бэкона из раннего Нового времени, когда его считали фаустианской фигурой, магом, обладающим тайными знаниями. Однако на самом деле Роджер Бэкон был провозвестником современной науки: за 350 лет до своего однофамильца Фрэнсиса он уже критиковал схоластику и закладывал основы эмпиризма, не боясь противоречить даже Аристотелю, а за 250 лет до Мартина Лютера критиковал распущенность духовенства, из-за чего 12 лет просидел в тюрьме.

³ *Абу аль-Из ибн Исмаил ибн аль-Раззас аль-Джазари* (1136–1206) — механик, математик, изобретатель и астроном из государства династии Артукидов. Он изобрёл коленчатый вал, конструировал клапанные насосы, машины для подъёма воды, фонтаны и многое другое. В 1206 году аль-Джазари составил трактат под названием «Книга знаний об остроумных механических устройствах»,



(а)



(б)

Рис. 1.2. Автоматы аль-Джазари: (а) фонтан в виде павлина;
(б) слуга, подающий воду для мытья рук

создал фонтан в виде павлина с автоматизированными человекоподобными слугами и робота-слугу, который подавал воду для мытья рук механизмом, подобным устройству современных унитазов (рис. 1.2). А самым известным его изобретением был музыкальный автомат в виде четырёх роботов-музыкантов, который считается одним из первых программируемых роботов: вероятно, музыканты могли играть в разных ритмах в зависимости от положения рычагов автомата.

Наука в Европе была тогда не так развита, но люди не оставляли мысли об искусственном интеллекте: так, в «Книге о Ланселоте Озёрном», созданной около 1230 года, Ланселот сражается с медными рыцарями-роботами. И вскоре европейцы тоже стали создавать механические автоматы. Например, в 1292 году Роберт II д'Артуа привёз в свой замок Эден (Hesdin) итальянского рыцаря Рено Куанье, которому поручил создать сад с механическими диговинками.

Куанье справился; местные мастера под его руководством создали механических птиц, обезьян, говорящую механическую сову, автоматический орган и массу увеселительных автоматов: статуи, поливающие гостей водой, мост, который разрушался, когда по нему шли люди, и тому подобное [603]. Были в эденском саду и просто шутки, отвечавшие, видимо, чувству юмора Роберта II; например, там было зеркало с надписью, приглашающей посмотреть, как вы будете выглядеть, если обсыпать вас мукой... и оно не обманывало!

где описал около 50 новых механизмов, в том числе конструкцию своих роботов. Исследователи считают, что музыкальный автомат аль-Джазари, представлявший собой фигуры четырёх музыкантов, мог содержать программируемую драм-машину.

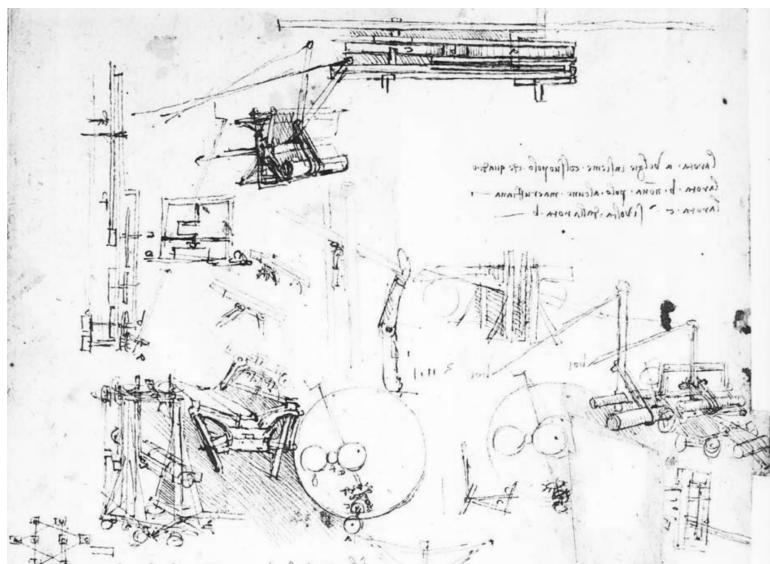


Рис. 1.3. Наброски Леонардо да Винчи из *Codex Atlanticus*

Спустя двести лет Леонардо да Винчи тоже придумал механического рыцаря, который работал на системе тросов и умел садиться, двигать руками и головой и открывать забрало [502]. На рис. 1.3 изображены, по всей видимости, наброски к устройству руки этого робота, которая должна была уметь поднимать тяжести, но разобраться в происходящем не так просто — известно, что Леонардо трепетно относился к своим изобретениям и часто намеренно запутывал записи, добавлял лишние, не относящиеся к делу рисунки и так далее.

Наверное, нет смысла подробно описывать историю чисто технического развития различных человеко- и звереподобных автоматов, которые были весьма популярны в Новое время, но, по сути, представляли собой всего лишь механические игрушки. Упомяну лишь несколько примеров. Механическая утка, которую в 1738 году выставил в саду Тюильри французский часовщик Жак де Вокансон, стала одним из самых знаменитых «роботов» своего времени. Она могла крутить головой, хлопать крыльями, заглатывать еду из рук зрителя и даже через несколько минут выводить её остатки в «переваренном» виде (на самом деле, конечно, это был всего лишь трюк, но давайте не будем останавливаться на его устройстве подробнее). Де Вокансон также создал автоматического флейтиста, который был уже полноценным «андроидом», человекоподобным роботом.

Гораздо более совершенных андроидов спустя 50 лет создал другой француз, тоже часовщик, Пьер Жаке-Дро (Jaquet Droz — такая марка часов существует и по сей день) и его сыновья Анри-Луи и Жан-Фредерик. Три их анроида —



Рис. 1.4. Автоматоны Жаке-Дро: (а) рисовальщик; (б) каллиграф

«Музыкант Марианна» в виде девушки, играющей на органе посредством механизма, состоящего из 2500 элементов, «Рисовальщик Анри» из 2000 элементов (см. рис. 1.4, а) и «Каллиграф Шарль» из 6000 элементов (см. рис. 1.4, б) — до сих пор хранятся в Музее искусства и истории в Невшателе. Марианна умела играть несколько мелодий, нажимая на клавиши миниатюрного органа, и музыка не была записью, а действительно извлекалась из инструмента. Анри мог создать четыре разных рисунка: портрет Людовика XV, портрет королевской четы, собаку с подписью «мой пёсик» и едущего на колеснице Купидона. Но самым сложным и интересным андроидом был Шарль. Он мог писать тексты длиной до 40 символов, распределённые в четыре строки. Это был фактически первый в истории компьютер: текст сначала нужно было запрограммировать на специальном зубчатом диске, и программа транслировалась в движения мальчика при помощи специального заводного механизма, то есть у Шарля был самый настоящий механический программируемый контроллер. Существует даже версия, что слово «андроид», которое появилось в XVIII веке, происходит от имени Анри-Луи Жаке-Дро. Впрочем, это вряд ли: греческая этимология слишком уж очевидна, да и первые употребления слова «андроид» встречаются раньше, чем семья Жаке-Дро принялась за дело.

1.1.3. Ранняя математическая логика

Но это всё-таки чистая механика, и хотя на неподготовленного зрителя такие механизмы могли произвести сильное впечатление, ни о каком разуме для этих автоматов речи не шло¹. Встречалось среди таких автоматов и откровенное мошенничество. Так, шахматный автомат «Турок» Вольфганга фон Кемпелена гремел на всю Европу и обыграл даже самого Наполеона I, но в конце концов

¹ Хотя и сейчас ситуация со многими достижениями искусственного интеллекта примерно та же, не исключено, что положение дел вскоре может измениться...

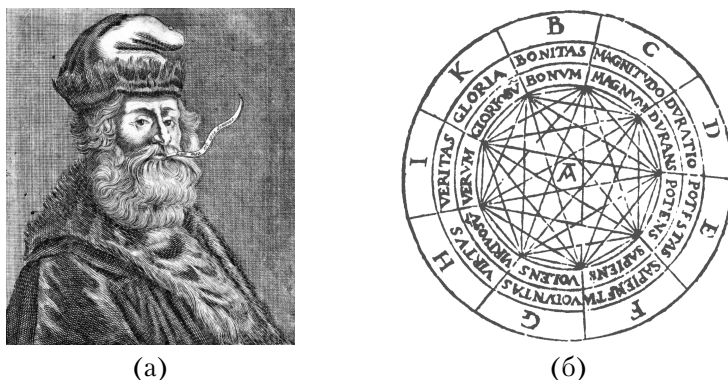


Рис. 1.5. Раймунд Луллий: (а) портрет самого философа; на ленточке во рту написано *Lux mea est ipse dominus* («Мой свет — сам Господь»); (б) простейшая логическая машина Луллия: атрибуты Бога и всевозможные связи между ними

выяснилось, что искусственный интеллект не такой уж искусственный: внутри автомата сидел сильный шахматист, который и вёл партию посредством действительно довольно сложного механизма управления механическим турком¹.

Ближе к теме собственно науки искусственного интеллекта подбирались другие средневековые учёные, которые пытались создать машины для проведения логических рассуждений или для порождения новых идей. Здесь нельзя не упомянуть Раймунда Луллия², который ещё в XIII веке в своём *ars universalis* дошёл фактически до панлогизма, то есть отождествления реального порядка вещей с порядком логическим. А это вполне, простите за каламбур, логично привело его к созданию логической машины, состоявшей из концентрических кругов,

¹ В наши дни в честь знаменитого автомата называется сервис *Amazon Mechanical Turk*, на котором живые люди выполняют небольшие и дешёвые задания. Часто «туркеры» как раз и занимаются разметкой наборов данных для моделей, о которых мы будем говорить в этой книге; самый популярный русскоязычный аналог *Amazon Mechanical Turk* — это сервисы «Толока» и «Задания» от «Яндекса».

² *Раймунд Луллий* (Raymundus Lullius на латыни, Ramon Llull на каталонском, 1235–1315) — каталонский философ, поэт, теолог, логик и в целом один из важнейших мыслителей высокого Средневековья. Он родился в городе Пальма-де-Майорка, там же и учился и с юности, что логично для испанца, изучал арабский язык и культуру; говорят, Луллий купил для этого раба, который преподавал ему язык, логику и суфизм. Любопытно, что первые его работы — «Компендиум логики аль-Газали» и «Книга размышлений о Боге», которая считается шедевром христианской литературы, — были написаны на арабском. На 32-м году жизни Луллия во время сочинения эротической песни явился распятый Христос, после чего Луллий поселился на пустынной горе Мирамар и пришёл к мысли об особом методе или искусстве, которое бы могло служить основой поиска истины и обращать неверных в христианство. Свой логический метод Луллий называл *ars generalis*, *ars universalis* или *ars magna*, что позже привело к ошибке: алхимики тоже называли искусство превращения металлов и получения жизненного эликсира *ars universalis*, и многие источники приписывают Луллию массу алхимических трудов, что вовсе не соответствует действительности.

разными вращениями которых можно было получать разные комбинации понятий (Луллий был по совместительству и основателем комбинаторики). Простейшая такая машина изображена на рис. 1.5, б; там ничего не движется, схема просто иллюстрирует всевозможные связи между атрибутами Бога, что помогает получить множество ортодоксальных истин: слава Бога вечна, а вечность Его славна и так далее.

Конечно, это совсем не автоматический логический вывод в современном понимании, но Луллий сделал первую попытку автоматического порождения текста и даже смысла. Эту попытку позже будет высмеивать Джонатан Свифт в лапутянской машине для производства знаний, которая описана в эпиграфе к этому разделу, а до логического завершения эту идею доведёт Хорхе Луис Борхес в своей «Вавилонской библиотеке». Но смех смехом, а говорят, что, например, арифмометр Готфрида Лейбница¹ был вдохновлён не только огромными объёмами вычислений, о которых Лейбниц узнал после знакомства с Гюйгенсом², но и панлогическими идеями Луллия.

А сам Лейбниц в своей «Монадологии» описал так называемую «мельницу Лейбница», машину размером с мельницу, которая симулирует чувства и восприятие, но внутрь которой можно зайти и осмотреть её движущиеся части, не найдя ни в одной из них восприятия или чувств. Лейбниц этим, конечно, объяснял дуализм и необходимость существования души. Точнее говоря, не совсем

¹ *Готфрид Вильгельм Лейбниц* (Gottfried Wilhelm von Leibniz, 1646–1716) — немецкий философ, логик, математик и механик, физик, юрист и дипломат, историк и не только. Лейбниц считается «завершителем философии XVII века», идеи которой он развил в учении об анализе и синтезе и учении о монадах, неделимых простых сущностях. В математике Лейбниц заложил основы математической логики и комбинаторики, в частности одним из первых описал двоичную систему счисления, изобрёл конструкцию арифмометра (куда лучше, чем у Паскаля), в физике ввёл понятие «живой силы», подобно кинетической энергии (Лейбниц использовал именно произведение массы на квадрат скорости), а в истории первым обратился к вопросу о возникновении российской правящей династии и систематизировал источники о происхождении Рюрика. Кстати, Лейбниц был знаком с Петром I, который именно под влиянием Готфрида Вильгельма одобрил проект создания Академии наук в Петербурге. Но главной работой Лейбница нужно признать разработанное им дифференциальное исчисление. Вы наверняка слышали о долгом споре Ньютона и Лейбница о приоритете, который надолго омрачил отношения между английскими и европейскими математиками, но бесспорно то, что вариант Лейбница был гораздо доступнее (труды Ньютона знамениты своей сложностью для понимания), и именно он популяризировал новый подход к математике.

² *Христиан Гюйгенс ван Зейлхем* (Christiaan Huygens, 1629–1695) — голландский физик, механик, астроном и изобретатель, первый президент Французской академии наук. Первая известность пришла к Гюйгенсу в астрономии: вместе с братом они усовершенствовали конструкцию телескопа, и Гюйгенс открыл и описал кольца Сатурна и его спутник Титан. Но главный вклад он оставил в математике и механике: Гюйгенс был тонким геометром и решил множество задач механики — описал циклоидальный маятник, разработал теорию эволюта и эвольвент, вывел законы равноускоренного движения — чисто геометрическими методами, ведь он работал ещё до изобретения математического анализа. И да, работы Гюйгенса сильно повлияли на Ньютона: «Маятниковые часы» были его настольной книгой. А для нас ещё интересно, что в 1657 году Гюйгенс написал приложение «О расчётах в азартной игре» к книге ван Схоотена «Математические этюды»; это было одним из первых изложений зарождавшейся теории вероятностей, и именно по этой книге с ней познакомился Якоб Бернулли.

так: Лейбниц и Декарт одними из первых вообще начали говорить о том, что в мире есть некая механистическая «материя», в которой всё должно объясняться на уровне чисто научном, в терминах, так сказать, соударения бильярдных шаров. Это потребовало от них поместить в какой-то другой тип сущего душу и мысли — сомневаться в существовании души для Лейбница и Декарта было бы немислимо. Но это всё равно был огромный шаг к материализму: до этих рассуждений, в аристотелевской парадигме, вполне нормально было бы объяснить, скажем, рост деревьев тем, что «деревья имеют естественную склонность расти».

Но заметьте, что по сути это в точности знаменитый аргумент Джона Сёрла¹ о «китайской комнате»! В мысленном эксперименте Сёрла человек или робот находится в комнате, на стенах которой описан очень сложный алгоритм перекладывания карточек с непонятными символами, и он перекладывает эти карточки согласно алгоритму. Герой эксперимента не понимает, что он делает, и не знает ни одного иероглифа, но в итоге оказывается, что вся комната в целом разговаривает на китайском языке: на поступающих на вход карточках написаны реплики на китайском, а после выполнения алгоритма на выходе получаются разумные ответы на них, тоже на китайском языке. Сёрл задаёт не имеющий пока однозначного ответа вопрос: возникает ли здесь понимание китайского языка и если да, то «кто» его понимает — человек, комната, алгоритм?.. Так что в истории искусственного интеллекта, как и любой другой науки, идеи переплетаются и мотивируют друг друга в весьма причудливых сочетаниях.

Не будем углубляться в историю искусственного интеллекта в литературе и других видах искусства: начавшись с романа Мэри Шелли «Франкенштейн, или Современный Прометей», тема искусственного интеллекта, его опасностей и возможностей, особенностей его взаимодействия с людьми уже никуда не исчезала. А настоящей наукой искусственный интеллект стал в середине XX века, когда в литературе уже давно начался золотой век научной фантастики, непредставимой без роботов и мыслящих машин.

Давайте обсудим, с чего эта наука начиналась.

¹ Джон Роджерс Сёрл (John Rogers Searle, р. 1932) — американский философ. Сёрл получил известность благодаря своим исследованиям языка: он развил теорию речевых актов, ввёл понятие иллокутивного акта как минимальной единицы языкового общения, разделил иллокутивное (то есть функциональное) и пропозициональное (фактическое) содержание высказываний и так далее [519, 524]. Но нам, конечно, интереснее его работы о философии сознания и искусственном интеллекте: сначала теория иллокутивных актов привела его к теории *интенциональности*, то есть, упрощённо говоря, того, как ментальные состояния могут быть *о чём-то*, направлены на что-то в окружающем мире [521], а затем Сёрл стал развивать подход биологического натурализма к проблеме сознания, утверждая, что дихотомия между объективной реальностью и субъективным опытом ложна: сознание — это субъективный опыт, вызванный физическими процессами в мозге [523]. Здесь уже видна и «китайская комната», которую Сёрл придумал ещё в 1980 году [520]. Развивая философский взгляд на искусственный интеллект, Сёрл ввёл понятие «сильного AI» как сознания, возникающего в процессе работы компьютерной программы [523], но сам считал, что этого не будет, а разум не компьютерная программа [522]; порекомендую также работы Штефана Харнада, который много спорил с Сёрлом и в процессе спора, как и подобает философу, прояснял и развивал его идеи [241–244].