

ДЖОН ФОН НЕЙМАН

*ВЫЧИСЛИТЕЛЬНАЯ МАШИНА
И МОЗГ*



*ИЗДАТЕЛЬСТВО АСТ
МОСКВА*

УДК 004.8
ББК 32.81
Н38

Серия «Эксклюзивная классика»

John von Neumann,
Ray Kurzweil

THE COMPUTER AND THE BRAIN, 3rd EDITION

Originally published by Yale University Press.
First Edition 1958. Second Edition published
as a Yale Nota Bene book in 2000. Third Edition 2012.

Перевод с английского *А. Чечиной*

Серийное оформление *А. Фереца, Е. Ферец*

Компьютерный дизайн *В. Воронина*

Печатается с разрешения Yale University Press при содействии
литературного агентства The Van Lear Agency LLC.

Нейман, Джон фон.

Н38

Вычислительная машина и мозг / Джон фон Нейман ;
[перевод с английского *А. Чечиной*]. — Москва : Изда-
тельство АСТ, 2023. — 192 с. — (Эксклюзивная классика).

ISBN 978-5-17-148015-8

Один из величайших ученых XX века и один из «отцов-основателей» сов-
ременных информационных технологий, американец венгерского происхож-
дения Джон фон Нейман прожил недолго (1903–1957), но за этот короткий
период успел отметить значительными достижениями в области информати-
ки, разработать теорию игр, внести существенный вклад в квантовую механику,
логику и концепцию клеточных автоматов и даже принять участие в Манхэт-
тенском проекте по разработке ядерного оружия.

Информационные технологии кардинально изменили жизнь человека,
затронув каждую ее сторону – от бизнеса и науки до политики и искусства.
Искусственный интеллект перестал быть фантастикой и сделался реальностью.
И именно эта всесторонняя компьютеризация общества поставила перед нами
новый фундаментальный вопрос: какова разница между интеллектом искусст-
венным и интеллектом человеческим и есть ли она вообще?

Джон фон Нейман в своей поистине пророческой, выдержавшей испытание
временем и до сих пор регулярно переиздающейся книге утверждает: такая
разница минимальна. Несмотря на все различия в архитектуре и строительных
блоках мозга и вычислительной машины, искусственный интеллект тем не ме-
нее способен имитировать работу мозга.

УДК 004.8
ББК 32.81

© Yale University Press, 1958
© Renewed by Marina V. N. Whitman, 1986
© Foreword to the Second Edition, Yale University, 2000
© Foreword to the Third Edition,
Ray Kurzweil, 2012



Школа перевода В. Баканова, 2018
© Издание на русском языке AST Publishers, 2022

ISBN 978-5-17-148015-8

Силлимановские лекции

На средства, пожертвованные в память о Хепсе Эли Силлимане, президент и сотрудники Йельского университета ежегодно представляют цикл лекций, призванных обозначить присутствие Божие и Промысел Божий в том виде, в коем они проявляются в мире природы и нравственности. По искреннему убеждению завещателя, всякое упорядоченное изложение явлений природы или фактов истории способствует достижению этой цели более успешно, нежели догматическое или полемическое богословие, которое посему должно быть исключено из материалов лекций. Темы преимущественно выбираются из областей естественных наук и истории; особое внимание уделено астрономии, химии, геологии и анатомии.

Рэй Курцвейл

Предисловие к третьему изданию

Информационные технологии кардинально изменили жизнь человека, затронув каждую ее сторону — от бизнеса до политики и искусства. Сегодня, в силу стремительного роста мощности компьютеров, с одной стороны, и одновременного улучшения соотношения «производительность — цена» — с другой, информационная эра неумолимо продолжает расширять сферу своего влияния. Учитывая данное обстоятельство, можно утверждать, что важнейшим информационным процессом, требующим всестороннего изучения, является сам человеческий интеллект. Вероятно, эта книга представляет собой первое серьезное исследование взаимосвязи между мышлением человека и машиной — исследование с точки зрения математика, сформулировавшего фундаментальную архитектуру компьютерной эры.

В рамках грандиозного проекта по изучению механизмов работы мозга достигнуты значительные успехи: мы достаточно далеко продвинулись в обратном проектировании парадигм человеческого мышления и сегодня применяем их для создания более умных машин. Разумно ожидать, что искусственный интеллект (ИИ), построенный на базе данных принципов, в итоге воспарит над человеческим разумом. Однако, я полагаю, цель этого дерзкого предприятия состоит не в том, чтобы потеснить нас, а в том, чтобы расширить границы того, что уже сейчас можно смело назвать человеко-машинной цивилизацией. Именно это делает человеческий род поистине уникальным.

Итак, каковы же ключевые идеи, лежащие в основе информационной эры? По моему мнению, их пять. Джон фон Нейман выдвинул три из них, а также внес весомый вклад в четвертую. Клод Шеннон решил фундаментальную проблему надежности информации. Алан Тьюринг сформулировал идею универсальности машинных вычислений. Опираясь на Тьюринга и Шеннона, фон Нейман создал машину фон Неймана, которая была и остается фундаментальной архитектурой всех вы-

числительных машин, существующих на сегодняшний день.

В обманчиво скромном томике, который вы держите в руках, фон Нейман описывает предложенную им модель машинных вычислений и постулирует принципиальную эквивалентность живого мозга и искусственной машины. Он признает глубокие структурные различия, однако затем обращается к принципу эквивалентности всех вычислений, предложенному Тьюрингом, и предлагает рассматривать методы работы мозга как совокупность вычислений. Книга в высшей степени провидческая, особенно если учесть, что она была написана более полувека назад, когда нейронаука располагала лишь самыми примитивными инструментами. Фон Нейман предчувствовал быстрое развитие технологий и видел его неизбежные последствия в грядущей сингулярной трансформации человеческого существования.

Рассмотрим эти пять основных идей более подробно.

В 1940 году, услышав слово «computer», люди понимали, что речь идет об аналоговой вычислительной машине. В аналоговых машинах числа выражались разными уровнями

напряжения, а специализированные компоненты могли выполнять такие арифметические функции, как сложение и умножение. Главный недостаток подобных машин заключался в весьма низкой степени точности. Числа могли быть представлены с точностью до одной сотой, а так как уровни напряжения, представляющие эти числа, обрабатывались большим количеством арифметических операторов, ошибки имели тенденцию накапливаться. При выполнении достаточно длинных вычислений результаты оказывались в высшей степени неточными, а то и вообще бессмысленными.

Это явление знакомо всякому, кто хоть раз записывал музыку на магнитную ленту. Первая копия была немного шумнее оригинала (шум представляет собой случайные неточности). На второй копии (т. е. копии копии) шумов было еще больше, а десятая содержала один только шум.

Это наводит на мысль, что с той же проблемой столкнется и мир цифровых вычислительных машин. Рассмотрим передачу цифровой информации по некоему каналу. Поскольку идеальных каналов не существует, любому из них изначально будет присущ определенный

уровень ошибок. Допустим, мы выбрали канал с вероятностью правильной передачи каждого бита 0,9. Если я отправлю сообщение длиной в один бит, вероятность точной передачи его через этот канал будет равна 0,9. Допустим, я отправил 2 бита. Теперь точность составит $0,9^2 = 0,81$. А если я отправлю 1 байт (8 битов)? В таком случае вероятность, что я получу его правильно, окажется менее 0,5 (0,43, если быть точным). Вероятность точной отправки пяти байтов составит около 1 процента.

На первый взгляд наилучший способ обойти эту проблему — сделать канал более точным. Предположим, канал совершает всего одну ошибку на миллион битов. Если я отправлю файл размером в полмиллиона байт (примерно столько «весит» скромная программа или база данных), вероятность правильной его передачи составит менее 2 процентов, хотя изначально точность канала была весьма высока. Учитывая, что ошибка в один бит может испортить всю компьютерную программу, это не лучший выход. Вне зависимости от исходной точности канала вероятность ошибки стремительно возрастает с увеличением размера сообщения. Выходит, эта проблема неразрешима?

Аналоговые вычислительные машины тоже аккумулируют неточности, но если ограничиться небольшим набором вычислений, они оказываются весьма полезными. Цифровые машины, с другой стороны, требуют наличия постоянной связи не только между двумя разными машинами, но и между собственными составными частями. Так, память связана с центральным процессором, внутри которого происходит постоянный обмен данными между регистрами, а также между регистрами и арифметическим устройством. Внутри арифметического устройства информация передается от одного битового регистра к другому. Если считать, что частота ошибок стремительно возрастает с увеличением количества таких связей и что ошибка в один бит может нарушить весь процесс, то необходимо признать: цифровые машины обречены.

Таково было общее мнение, пока Шеннон не провозгласил первую ключевую идею информационной эры. Он продемонстрировал, что мы можем создавать произвольно точные сообщения, используя самые ненадежные каналы передачи информации. В своей знаковой статье «Математическая теория связи», опубликованной в журнале *Bell System Technical*

Journal в июле и октябре 1948 года, Шеннон предложил теорему кодирования для каналов с шумами, которая гласила: если у вас есть доступный канал с любым коэффициентом ошибок (за исключением 50 процентов на бит, поскольку это означает, что канал передает чистый шум), вы можете передать сообщение с любой желаемой степенью точности. Другими словами, частота появления ошибок может составлять один бит из n бит, где n может быть сколь угодно большим. Например, если у вас есть канал, который правильно передает биты информации только в 51 проценте случаев (т. е. передает правильный бит чуточку чаще, чем неправильный), вы тем не менее можете передавать сообщения таким образом, что неправильным окажется только один бит из миллиона, один бит из триллиона или один бит из триллиона триллионов.

Как такое возможно? Ответ: через избыточность.

Сейчас подобное решение может показаться элементарным, но в то время оно было далеко не очевидным. Рассмотрим простой пример. Если я передам каждый бит три раза и применю «принцип большинства», то я значительно повышу надежность результата. Если

этого будет недостаточно, я могу увеличить избыточность, пока не получу необходимую надежность. Повторение информации — самый простой способ достичь высокой точности в каналах низкой точности. Однако данный подход не самый эффективный. В своей статье Шеннон не только заложил основы теории информации, но и предложил оптимальные методы обнаружения и исправления ошибок. Эти методы позволяли добиться любой желаемой точности через любой неслучайный канал.

Читатели более старшего возраста наверняка помнят телефонные модемы, в которых постоянно что-то шипело и шелкало, так как они передавали информацию по аналоговым линиям с высоким уровнем помех. Однако те же модемы могли передавать цифровые данные с очень высокой точностью — благодаря теореме Шеннона для канала с шумами.

Аналогичная проблема и аналогичное решение существуют и для цифровой памяти. Вы когда-нибудь задумывались, почему CD, DVD и программные диски продолжают работать даже после того, как упали на пол или были поцарапаны? Этим мы тоже обязаны Шеннону. Процесс вычислений состоит из трех элементов: связи (которая, как я уже го-

ворил, имеет место как внутри, так и между машинами), памяти и логических вентилях (которые выполняют арифметические и логические функции). Точность логических вентилях можно сделать произвольно высокой с помощью особых кодов для обнаружения и исправления ошибок. Именно благодаря теореме Шеннона мы можем обрабатывать большие и сложные цифровые данные, используя для этого достаточно длинные алгоритмы.

Вторая важная идея, на которую опирается наш информационный век, — универсальность машинных вычислений. В 1936 году Алан Тьюринг описал «машину Тьюринга» — абстрактную вычислительную машину, которая состоит из бесконечно длинной ленты, разделенной на клетки с цифрами 1 или 0. Машина считывает одну клетку за другой и содержит набор правил в виде пронумерованных состояний, фактически представляющих собой хранимую в памяти программу. Каждое правило предписывает машине совершить одно действие, если в считываемой клетке стоит 0, и другое действие, если в считываемой клетке стоит 1. Возможные действия включают запись 0 или 1, перемещение ленты на одну клетку вправо или влево, оста-

новку ленты. Каждое состояние содержит номер следующего состояния, в которое должна перейти машина. Завершив алгоритм, машина останавливается; выход процесса остается на ленте. Хотя лента теоретически бесконечна, любая программа (которая не подразумевает бесконечный цикл) использует конечную часть ленты; следовательно, если мы ограничимся конечной памятью, машина по-прежнему сможет решать широкий круг задач.

В машине Тьюринга нет ничего сложного, верно? На самом деле именно этого и добивался ученый. Он хотел, чтобы его машина была максимально простой (но не проще, перефразируя Эйнштейна). Позже Тьюринг и его бывший учитель, Алонзо Черч сформулировали тезис Черча — Тьюринга, согласно которому задача, которая не может быть решена машиной Тьюринга, не может быть решена никакой *другой* машиной. Хотя собственно машина Тьюринга способна выполнять крайне ограниченное количество команд и одновременно обрабатывает всего один бит, она может вычислить все, что может вычислить любая вычислительная машина.

Строгие интерпретации тезиса Черча — Тьюринга предполагают принципиальную эк-

вивалентность того, что человек может думать или знать, и того, что может быть вычислено машиной. Основная идея заключается в том, что человеческий мозг подчиняется естественным законам; следовательно, его способность обрабатывать информацию не может превосходить таковую у машины (и соответственно у машины Тьюринга).

В своей статье Тьюринг заложил теоретические основы машинных вычислений. Хотя это целиком и полностью его заслуга, важно отметить, что большое влияние на него оказала лекция, прочитанная Джоном фон Нейманом в 1935 году в Кембридже (Англия). Лекция была посвящена идее программы, которую можно хранить в памяти — концепция, позднее воплощенная в машине Тьюринга. На фон Неймана в свою очередь произвела глубочайшее впечатление статья Тьюринга 1936 года, где были изложены принципы машинных вычислений и которую в конце 1930-х — начале 1940-х годов он включил в список обязательной литературы, составленный для своих коллег.

В той же работе Тьюринг сообщает о другом неожиданном открытии, а именно — о проблеме неразрешимых задач. Неразрешимые задачи — это хорошо описанные задачи с однознач-

ным ответом, который, однако, не может быть вычислен на машине Тьюринга (т. е. на *любой* машине). Это противоречит постулату XIX века, гласящему, что все задачи, которые могут быть описаны, в конечном счете будут решены. Тьюринг показал, что неразрешимых задач столько же, сколько и разрешимых. В своей «Теореме о неполноте» 1931 года Курт Гедель приходит к аналогичному выводу. Таким образом мы оказываемся в странной ситуации: с одной стороны, мы можем описать задачу и доказать, что однозначный ответ существует, а с другой — знаем, что ответ никогда не будет найден.

Гораздо больше можно сказать о философском значении работ Тьюринга, Черча и Геделя, однако в рамках данного предисловия достаточно отметить следующее. Тьюринг показал, что в основе всех машинных вычислений лежит очень простой механизм. Поскольку машина Тьюринга (и, следовательно, любая вычислительная машина) способна определять дальнейший образ действий, исходя из результатов предыдущих операций, она способна принимать решения и моделировать произвольно сложные иерархии данных.

К декабрю 1943 года Тьюринг спроектировал и построил «Колосс», который часто на-