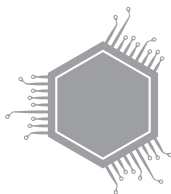


**МАРВИН  
МИНСКИЙ**



# МАРВИН МИНСКИЙ

## СООБЩЕСТВО РАЗУМА



Издательство АСТ  
Москва

УДК 004.8  
ББК 32.813  
М62

Серия «Наука: открытия и первооткрыватели»

Marvin Minsky  
THE SOCIETY OF MIND

Перевод с английского *В. Желнинова*

Серийное оформление *Г. Смирновой*

Компьютерный дизайн *В. Воронина*

Печатается с разрешения издательства Simon & Schuster, Inc.  
и литературного агентства Andrew Nurnberg.

**Минский, Марвин.**

М62 Сообщество разума / Марвин Минский ; [пер. с англ. В. Желнинова]. — Москва : Издательство АСТ, 2018. — 592 с. — (Наука: открытия и первооткрыватели).

ISBN 978-5-17-106639-0

В конце 70-х годов XX века Марвин Минский выдвинул неожиданную идею, согласно которой человеческий интеллект не так уж сильно отличается от искусственного, как это было принято считать. Со временем эта революционная и во многом обидная для человечества идея получила безупречное логическое обоснование в его программной книге «Сообщество разума».

Книга была опубликована в 1986 году и с тех пор многократно переиздавалась на разных языках. За прошедшие годы в области искусственного интеллекта была сделана масса открытий; количество умов, занятых в этом сегменте, с каждым годом растет, но странная вещь — «Сообщество разума» Марвина Минского по-прежнему представляет собой кладезь идей, не утративших новизны и оригинальности.

УДК 004.8  
ББК 32.813

© Marvin Minsky, 1985, 1986  
© Издание на русском языке AST Publishers, 2018

# Глава 1

## ПРОЛОГ

*Все следует упрощать до тех пор, пока это возможно, но не более того.*

Альберт Эйнштейн

Эта книга пытается объяснить, как работает разум. Способен ли интеллект возникнуть из состояния неразумности? Чтобы ответить на этот вопрос, мы покажем, что разум возможно создать, сконструировать из множества малых составных частей, каждая из которых сама по себе бессмысленна и неразумна.

«Обществом разума» я буду впредь именовать такую схему, в которой каждое сознание представляется состоящим из множества мелких процессов. Указанные процессы мы будем называть *агентами*. Каждый ментальный агент по отдельности выполняет некое простое действие, для чего не требуется ни разум, ни мышление вообще. Тем не менее, когда мы объединяем указанных агентов в сообщества — посредством весьма своеобразных способов, — это ведет к возникновению подлинного интеллекта.

В моей книге нет специфических технических подробностей. Она также представляет собой сообщество — общество многих малых идей. Каждая из этих идей сама по себе олицетворяет лишь то или иное проявление здравого смысла, но при объединении достаточного их количества мы обретаем возможность объяснить загадки разума.

Проблема в том, что эти идеи тесно взаимосвязаны друг с другом. Мои объяснения поэтому крайне редко выстраиваются в аккуратной прямой логической после-

довательности, так сказать, от начала до конца. Мне хотелось бы, конечно, соблюсти такую последовательность и проложить читателю путь к вершине, вырубить ментальные ступени в склоне горы познания. Однако вместо того я вынужден вести читателя по прихотливой паутине взаимодействий и связей.

Возможно, вина за это целиком лежит на мне, поскольку я не сумел выявить и описать некую базу неоспоримых и упорядоченных принципов. Впрочем, лично я склонен возлагать упомянутую вину на природу разума: в значительной степени его могущество проистекает, как кажется, из той самой взаимозависимости, опираясь на которую, вынуждены взаимодействовать ментальные агенты. Если это в самом деле так, с данным осложнением придется смириться; перед нами, увы, очередной из бесчисленных трюков эволюции.

Что можно сделать, когда понимаешь, что некое явление с трудом поддается описанию? Начинаешь набрасывать грубые рисунки, контуры, подобия очертаний, которые послужат строительными лесами для всего остального; к слову, не имеет значения, как много этих рисунков в итоге окажутся не соответствующими действительности. Затем мы принимаемся прорисовывать детали, дабы придать этим наброскам реалистичность и плоть. Наконец, на последнем этапе, нужно будет отвергнуть любые первоначальные догадки и выводы, которые перестанут укладываться в общую картину.

Так мы поступаем в повседневной жизни, так решаем головоломки, вызывающие у нас изрядные затруднения. Не важно, имеем ли мы дело с черепками разбитых горшков или с винтиками громадных машин. Пока не разглядишь конструкцию целиком, пускай бегло, понять назначение какой-либо ее части не удастся.

## 1.1. Агенты разума

Внятные и здравые теории разумности должны учитывать по крайней мере три различных шкалы времени, а именно: медленное течение времени, ибо человеческий

мозг эволюционировал на протяжении миллиардов лет; быстрое течение времени, ибо мозг стремительно развивается в первые недели и месяцы жизни, от младенчества до детства; а еще, так сказать, промежуточное состояние, то есть те столетия, которые потребовались на формулирование и развитие идей.

Чтобы объяснить, что такое разум, мы должны показать, как он возникает из бессмысленного, неразумного «материала», из своих составных частей, которые намного меньше и проще, чем любое проявление разумности. Если у нас не получится объяснить разум посредством его сведения к тем элементам, которые не имеют собственных мыслей и чувств, это будет означать, что мы ходим по кругу. Но каковы же эти простейшие элементы, эти «агенты», обеспечивающие разумность человека? Ответу на данный вопрос посвящена настоящая книга; памятуя об сказанном, давайте оценим нашу задачу. Итак, перед нами множество частных вопросов.

**Функция:** *Как действуют агенты?*

**Воплощение:** *Из чего они состоят?*

**Взаимодействие:** *Как они коммуницируют?*

**Происхождение:** *Откуда взялись первые агенты?*

**Наследственность:** *Все ли люди от рождения обла-  
дают одинаковыми агентами?*

**Обучение:** *Как мы создаем новых агентов и изменя-  
ем прежних?*

**Характер:** *Каковы наиболее важные разновидности  
агентов?*

**Конфликты:** *Что происходит, когда агенты проти-  
воречат друг другу?*

**Намерение:** *Откуда берутся желания и побуждения?*

**Компетенции:** *Почему группы агентов способны на  
то, чего не могут делать агенты по отдельности?*

**Личность:** *Что наделяет агентов единообразием или  
индивидуальностью?*

**Смысл:** *Откуда возникает способность что-либо  
понимать?*

**Восприимчивость:** *Откуда у агентов чувства и эмоции?*

**Осознание:** *Могут ли агенты обладать сознанием и осознавать самих себя?*

Насколько теория разумности способна объяснить столь многое, если учесть, что каждый отдельный вопрос представляется слишком трудным для ответа? Да, эти вопросы и вправду выглядят заковырыстыми, особенно когда мы пытаемся вычленить тот или иной из них из общей картины и разорвать его связи с другими. Но стоит нам вообразить разум как сообщество агентов, каждый ответ мгновенно начинает проливать свет на все прочие вопросы.

## 1.2. Разум и мозг

*Никем и никогда не допускалось [изрек поэт Им-лак], что мышление исконно присуще материи, что всякая частица является мыслящим существом. Но ежели какую часть материи лишить мысли, кто или что, по нашему мнению, будет за нее думать? Один материальный предмет отличается от другого только формой, размерами, плотностью, своим движением и направлением движения; чему из перечисленного, сколь угодно видоизмененному или дополненному чем-либо иным, можно приписать наличие сознания? Быть круглым или квадратным, твердым или текучим, большим или малым, двигаться медленно или быстро — все это так или иначе свойства материальных предметов, и они все одинаково чужды природе мышления. Если материи суждено лишиться мысли, это означает, что она всего-навсего научится думать как-то иначе, под воздействием иных сил, однако все эти новшества, какими она сможет обладать, все равно не будут иметь ни малейшего отношения к мышлению.*

Сэмюел Джонсон

Каким образом столь цельный орган, как мозг, может производить нечто столь призрачное и эфемерное, как мысли? Этим вопросом задавались многие мыслители



прошлого. Мир мыслей и мир вещей казались слишком далекими друг от друга, чтобы взаимодействовать хоть в какой-то степени. До тех пор пока мысли воспринимаются как бесконечно отличные от всего остального, мы не располагаем точкой отсчета для наших рассуждений.

Несколько столетий назад, между прочим, казалось совершенно невозможным объяснить происхождение жизни, поскольку живые существа воспринимались как принципиально отличные от всего прочего. Казалось, что растения появляются буквально из ничего. Мнилось невероятным, что животные могут двигаться и учиться. Вдобавок те и другие обладали способностью к воспроизводству, тогда как мир вокруг был лишен этой способности. Однако постепенно эта удивительная пропасть между живым и неживым начала смыкаться. Было обнаружено, что каждое живое существо состоит из мельчайших клеток, а сами клетки, как выяснилось, состоят из сложных, но вполне постижимых химикалий. Вскоре было установлено, что растения не порождают каких-либо веществ и просто извлекают большую часть своего состава из газов в воздухе. Загадочно пульсирующие сердца превратились благодаря исследованиям в механические насосы, представляющие собой совокупности мышечных клеток. Но лишь в текущем столетии\* Джон фон Нейман теоретически показал, что клеточные автоматы способны самовоспроизводиться, а Джеймс Уотсон и Фрэнсис Крик, почти независимо друг от друга, обнаружили, что клетки на самом деле копируют собственный наследственный код. Образованному человеку больше нет необходимости искать некую особую жизненную силу, которая якобы питает всякое живое существо.

Аналогичным образом еще сто лет назад мы не располагали, по существу, никакими возможностями для объяснения принципов работы мышления. Затем психотерапевты — выделю среди них Зигмунда Фрейда и Жана Пиаже — разработали теории умственного развития ребенка. Несколько позже, если обратиться к «механической» стороне процесса, математики наподобие Курта

---

\* Имеется в виду XX век. — *Здесь и далее примеч. ред.*

Геделя и Алана Тьюринга начали создавать модели, показывающие доселе неизвестные способности программируемых машин. Эти два направления исследований стали смыкаться лишь в 1940-х годах, когда Уоррен Маккалох и Уолтер Питтс\* показали, что машины можно побудить видеть, рассуждать и запоминать. Современная научная дисциплина исследований в области искусственного интеллекта возникла только в 1950-х годах, и ее появлению способствовало изобретение современных компьютеров. Вследствие этого стали формироваться новые идеи по поводу способности машин делать то, что прежде признавали исключительной особенностью человеческого разума.

Большинство людей до сих пор убеждены, что машина никогда не обретет полноценного сознания, не сможет испытывать честолюбие или зависть, не оценит юмор и не будет в состоянии приобщиться к иному ментальному жизненному опыту. Конечно, мы чрезвычайно далеки от того, чтобы создавать машины, умеющие делать все, что делают люди. Но отсюда следует, что нам нужны новые, полнее продуманные теории относительно принципов работы мышления. В настоящей книге будет показано, что крошечные машины, которые мы называем «агентами разума», вполне могут оказаться теми самыми «крупницами истины», которых не хватало авторам упомянутых теорий.

### 1.3. Общество разума

Нам известно, что все, о чем мы думаем и что делаем, мыслится и делается именно нами. Но кто эти «мы»? Какие малые сущности сотрудничают в нашем сознании ради того, чтобы обеспечить мысли и дела? Чтобы увидеть

---

\* Маккалох (также Мак-Каллок) У. — американский нейропсихолог, один из основоположников кибернетики, предложил модель сети электронных «нейронов». Питтс У. — американский лингвист, логик и математик, разработчик теории искусственных нейронов, которая легла в основу теории искусственного интеллекта.

схожесть разума с сообществом, попробуйте следующее: возьмите чашку с чаем.

*Вы БЕРЕТЕСЬ за чашку: агенты хотят взять чашку.*

*Вы ДЕРЖИТЕ РАВНОВЕСИЕ: агенты не хотят допустить, чтобы чай пролился.*

*Вы ИСПЫТЫВАЕТЕ ЖАЖДУ: агенты хотят, чтобы вы выпили чай.*

*Вы ДЕЛАЕТЕ ДВИЖЕНИЕ: агенты хотят поднести чашку к вашим губам.*

При этом ни одно из перечисленных действий не занимает ваш разум, пока вы перемещаетесь по комнате и разговариваете с друзьями. Вы вряд ли вообще задумываетесь о равновесии; оно никак не связано с первым движением; это первое движение не имеет отношения к жажде; а жажда никоим образом не подразумевает вовлеченности в общение. Почему так происходит? Потому что все вышеупомянутые агенты зависят друг от друга. Если каждый из них выполняет свою малую работу, то действительно большая работа делается ими всеми совместно, а в результате вы пьете чай.

Сколько процессов необходимо для того, чтобы указанное действие — питье чая — состоялось и чай не пролился из чашки? По меньшей мере добрая сотня, чтобы обеспечить нужное положение ладони, запястья и руки в целом. Вдобавок тысячи мышечных систем должны работать совместно, управляя расположением костей и суставов, благодаря чему тело может ходить. А чтобы удерживать равновесие, каждый из этих процессов вынужден взаимодействовать с некоторыми другими. Допустим, вы споткнулись и начали падать. В этом случае тело «запускает» множество других процессов, чтобы вы могли быстро выпрямиться. Некоторые из них отслеживают угол наклона и местоположение ваших ног. Прочие сосредотачиваются на том, как поступить с чаем, ведь вы не хотите обжечь собственную руку и не желаете облить горячим напитком кого-то из присутствующих в комнате. Вам необходимы быстрые решения.

Все это происходит, пока вы беседуете, и ни один из процессов, как представляется, не нуждается в серьезном осмыслении. Впрочем, если задуматься, то же самое можно сказать и о беседе. Какие агенты подбирают слова так, чтобы мы могли выразить именно то, что мы имели в виду? Как эти слова объединяются в фразы и предложения, каждое из которых связано с предыдущим и последующим? Какие элементы сознания следят за всем, нами сказанным, и контролируют, к кому обращены наши слова? Мы почувствуем себя глупо, если придется повторяться, но лишь тогда, когда у нас не прибавилось слушателей.

Мы всегда делаем несколько дел одновременно, например планируем, ходим и говорим; это кажется настолько естественным, что мы воспринимаем происходящее как должное. Однако данные процессы на самом деле опираются на такое число «машин», какое трудно себе вообразить. В нескольких следующих разделах настоящей книги мы сосредоточимся на единственном проявлении обычной деятельности — конструировании предметов из детских кубиков. Для начала мы разделим этот процесс на более мелкие этапы, а затем проследим, как каждый этап связан со всеми остальными.

Занимаясь этим, мы будем стараться подражать Галилею и Ньютону, узнавшим так много о природе посредством изучения простейших маятников и весов, зеркал и призм. Наше исследование того, как строить из кубиков, будет подобно фокусировке микроскопа на простейших объектах, которые возможно найти, чтобы «распахнуть дверь» в бескрайнюю и поразительную вселенную. По той же причине множество биологов в наши дни уделяют больше внимания крошечным микробам и вирусам, а не великолепным животным — львам или тиграм. Для меня и для целого поколения исследователей мир изучения детских кубиков стал призмой и маятником, опорой для изучения разума.

*В науке больше всего узнаешь, изучая то, что видится малым.*

## 1.4. Мир кубиков

Представим, что ребенок играет с кубиками и что разум этого ребенка содержит в себе множество малых разумов. Назовем их ментальными агентами. Прямо сейчас действиями ребенка руководит агент, именуемый Строителем. Его «специальность» — возведение башен из кубиков.

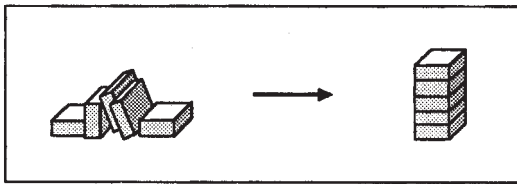


Рис. 1

Нашему ребенку нравится смотреть, как башня прирастает кубиками, которые накладываются один на другой сверху. Но строительство башни — слишком сложная задача для любого простого агента-одиночки, поэтому Строитель вынужден прибегать к помощи нескольких других агентов:

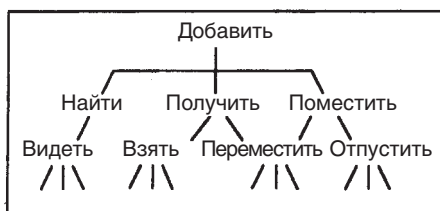


*Выбрать место для строительства башни.  
Добавить новый кубик к башне.  
Оценить высоту башни.*

Рис. 2

На самом деле даже взятие нового кубика и помещение его на башню сверху будет чрезмерно сложной задачей для любого отдельного агента. Поэтому операция добавления в свою очередь подразумевает помощь других агентов. Следовательно, нам нужно больше агентов, как следует из схемы (рис. 3).

Зачем разбивать операции на такие мелкие этапы? Потому что разум, как башня из кубиков, формируется подобным же образом, разве что состоит не из кубиков, а из процессов. Если же нагромождение кубиков кажется



*Для операции ДОБАВЛЕНИЯ сначала нужно НАЙТИ новый кубик. Затем рука должна ВЗЯТЬ этот кубик и ПОМЕСТИТЬ его на верх башни.*

Рис. 3

нам бессмысленным, стоит вспомнить, что так мы думали отнюдь не всегда. Когда в раннем детстве мы обнаруживали какие-либо «строительные» игрушки, то почти наверняка радостно осваивали их неделю за неделей, выясняя, для чего они годятся. Если теперь такие игрушки видятся нам скучными, спросите себя, в чем и как вы изменились. Прежде чем вы занялись более важными делами, вам ведь казалось диковинным и замечательным строить башни и дома из кубиков, верно? Пускай все взрослые знают, как это делается, никто не понимает, *как мы этому обучаемся!* Вот то, что нас интересует. Чтобы складывать кубики столбиками и рядами, нужны навыки, которые каждый из нас усвоил так давно, что мы напрочь забыли, как их изучали. Ныне они воспринимаются лишь как проявление здравого смысла; кстати, отсюда все трудности психологии. Эта забывчивость, эта детская амнезия заставляет нас предположить, что все наши чудесные способности присутствовали в нашем разуме изначально, и мы не перестаем спрашивать себя, откуда они взялись и как развивались.

### 1.5. Здравый смысл

*Нельзя мыслить о мышлении, не мысля о мыслях насчет чего-то.*

Сеймур Пейперт

Мы нашли способ разделить Строителя нашей башни на составные части. Но на этом со Строителем еще дале-

ко не все ясно. Чтобы возвести простой столбик из кубиков, агенты нашего ребенка должны выполнить следующие операции:

*Увидеть* — значит опознать кубики как таковые, независимо от их цвета, размера и местоположения, где бы они ни лежали, на свету или в тени, пускай даже частично скрытые за или под другими предметами.

Потом, когда это произошло, агент *движения* направляет руку и кисть по сложной траектории в пространстве, но так, чтобы не задеть макушку башни и не ударить ребенка по лицу.

*И насколько глупо будет выглядеть, если агенты поиска и хватания заставят выбрать кубик из основания башни!*

Если внимательно изучить эти требования, перед нами откроется изумительный мир сложных вопросов. Например, как агент поиска определяет, какие кубики по-прежнему доступны для использования? Он должен «понимать» сцену действия с точки зрения того, что мы пытаемся сделать. Это означает, что нам нужны теории относительно того, что такое понимание, и относительно того, способна ли машина сознавать цели. Рассмотрим все практические условия, к которым вынужден приноравливаться фактический Строитель. Он должен решить, достаточно ли в его распоряжении кубиков для достижения цели, насколько они крепки и широки, чтобы служить опорой для других кубиков, которые будут положены сверху.

Что если башня начнет раскачиваться? Настоящий строитель предугадывает такую возможность. Значит, какое-либо сочленение внутри постройки недостаточно ровное. Быть может, фундамент ненадежен или башня слишком высока для его ширины? А, быть может, причина в том, что самый последний кубик положили слишком небрежно.

Все дети учатся таким рассуждениям, но мы редко задумываемся обо всем этом впоследствии. Становясь взрослыми, мы привыкаем воспринимать подобные мысли как простые проявления «здравого смысла». Однако