

МАРВИН

МИНСКИЙ

МАРВИН
МИНСКИЙ

МАШИНА ЭМОЦИЙ



Издательство АСТ
Москва

УДК 159.95
ББК 88.3
М62

Серия «Шляпа Оливера Сакса»

Marvin Minsky
THE EMOTION MACHINE

Перевод с английского
Виктории Дегтяревой и Алины Курышевой

Компьютерный дизайн *Галины Смирновой*

Печатается с разрешения автора и литературного агентства
Brockman, Inc.

Минский, Марвин.

М62 Машина эмоций / Марвин Минский ; [перевод с английского В. Дегтяревой, А. Курышевой]. — Москва : Издательство АСТ, 2020. — 512 с. — (Шляпа Оливера Сакса).

ISBN 978-5-17-114660-3

Марвин Минский — американский ученый, один из основоположников в области теории искусственного интеллекта, сооснователь лаборатории информатики и искусственного интеллекта в Массачусетском технологическом институте, лауреат премии Тьюринга за 1969 год, медали «Пионер компьютерной техники» (1995 год) и еще целого списка престижных международных и национальных наград.

Что такое человеческий мозг? Машина, — утверждает Марвин Минский, — сложный механизм, который, так же, как и любой другой механизм, состоит из набора деталей и работает в заданном алгоритме. Но если человеческий мозг — механизм, то что представляют собой человеческие эмоции? Какие процессы отвечают за растерянность или уверенность в себе, за сомнения или прозрения? За ревность и любовь, наконец? Минский полагает, что эмоции — это всего лишь еще один способ мышления, дополняющий основной мыслительный аппарат новыми возможностями.

УДК 159.95
ББК 88.3

© Marvin Minsky, 2006
© Перевод. В. Дегтярева, 2018
© Перевод. А. Курышева, 2018
© Издание на русском языке AST Publishers, 2020

*Посвящается Глории, Маргарет,
Генри и Джулиане*

Прошлые соавторы

Пуш Сингх
Джон Маккарти
Сеймур Пейперт
Оливер Селфридж
Р. Дж. Соломонофф

Импраймеры

Джордж А. Миллер
Дж. К. Р. Ликлайдер
Эндрю М. Глисон
Соломон Лефшец
Уоррен С. Мак-Каллох
Клод Э. Шеннон

Поддержка

Джеффри Эпштейн
Кацухико Ниси
Николас Негропonte
Гарвардское общество научных сотрудников
Управление военно-морских исследований
Корпорация *Toshiba*

Оглавление

Введение	11
Глава 1. ВЛЮБЛЕННОСТЬ	21
1.1. Увлечение	21
1.2. Море психических загадок	25
1.3. Настроения и эмоции	31
1.4. Эмоции младенцев	33
1.5. Разум как облако ресурсов	37
1.6. Эмоции взрослых людей	45
1.7. Каскады эмоций	49
1.8. Теории чувств, значений и машин	52
Глава 2. ПРИВЯЗАННОСТЬ И ЦЕЛИ	57
2.1. Игры в луже	57
2.2. Привязанность и цели	59
2.3. Импраймеры	65
2.4. Обучение, основанное на привязанности, повышает статус цели	68
2.5. Научение, удовольствие и коэффициент релевантности	72
2.6. Совесть, ценности и самоидеалы	74
2.7. Привязанности у младенцев и животных	79
2.8. Кто они, наши импраймеры?	87
2.9. Модели себя и самодисциплина	89
2.10. Публичные импраймеры	93

Глава 3. ОТ БОЛИ К СТРАДАНИЮ	98
3.1. Ощущение боли	98
3.2. Как боль приводит к страданию?	101
3.3. Механизмы страдания.	104
3.4. Преодоление боли	111
3.5. Психические корректоры, подавители и цензоры	118
3.6. Фрейдистский бутерброд	125
3.7. Контроль над нашими настроениями и состояниями	127
3.8. Эмоциональная эксплуатация	131
Глава 4. СОЗНАНИЕ	135
4.1. Что за штука такая — Сознание?	135
4.2. Распаковываем чемодан сознания	139
4.3. Мозг А и мозг Б	142
4.4. Почему важность сознания так переоценивают .	153
4.5. Как порождается сознание?	159
4.6. Загадочный феномен «переживания».	168
4.7. Модели себя и самоосознание.	170
4.8. Картезианский театр	174
4.9. Последовательный поток сознания	178
Глава 5. УРОВНИ ПСИХИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ	182
5.1. Инстинктивные реакции	185
5.2. Заученные реакции	187
5.3. Сознательность	189
5.4. Рефлексивное мышление	198
5.5. Саморефлексия	201
5.6. Самоосознанная рефлексия	204
5.7. Воображение	208
5.8. Представление воображаемых сцен	215
5.9. Машины предсказаний	221
Глава 6. ЗДРАВЫЙ СМЫСЛ	225
6.1. Что мы понимаем под «здоровым смыслом»? . .	228
6.2. Житейские знания и логика	240

6.3. Намерения и цели	254
6.4. Большие разницы	268
6.5. Принятие решений	277
6.6. Рассуждение с помощью аналогии	281
6.7. Положительная и отрицательная компетентность	288
Глава 7. МЫШЛЕНИЕ.	295
7.1. Как происходит выбор тем для размышления?	300
7.2. Модель разума, основанная на Критиках и Переключателях	302
7.3. Эмоциональное мышление.	307
7.4. Некоторые полезные способы думать	310
7.5. Некоторые полезные Критики	313
7.6. Воплощение эмоций	318
7.7. Бессознательные процессы Пуанкаре	322
7.8. Когнитивные контексты	331
7.9. Центральные проблемы человеческой психологии	343
Глава 8. НАХОДЧИВОСТЬ	348
8.1. Находчивость	348
8.2. Оценка расстояния	353
8.3. Паналогия.	356
8.4. Каким образом люди так быстро учатся?	365
8.5. Коэффициент релевантности	370
8.6. Творчество и гениальность.	376
8.7. Воспоминания и репрезентации	380
8.8. Иерархия репрезентаций	400
Глава 9. САМОСТЬ.	407
9.1. Как мы представляем себя?	411
9.2. Личные качества	422
9.3. Почему нам нравится идея Самости?	435
9.4. Что такое Удовольствие и почему нам оно нравится?	439
9.5. Почему чувства так сложно описать?	446
9.6. Ощущение переживания опыта	453

9.7. Как организован человеческий разум?	456
9.8. Величие в сложности	467
9.9. Некоторые из источников человеческой находчивости	469
БЛАГОДАРНОСТИ	474
БИБЛИОГРАФИЯ	477
ПРИМЕЧАНИЯ	495

Введение

Нора Джойс своему мужу Джеймсу:
Почему ты не пишешь книг, которые можно читать?

Я надеюсь, что эта книга будет полезна всем, кто хочет понять, как работает человеческий разум, ищет совета о том, как мыслить более продуктивно, или стремится создать более умные машины. Читатели, которые интересуются темой искусственного интеллекта, наверняка найдут в ней кое-что полезное. Также она, как мне кажется, представляет интерес для психологов, неврологов, философов и ученых в сфере теории вычислительных машин, поскольку в ней выдвигается много новых теорий по самым сложным вопросам, с которыми сталкиваются эти специалисты.

Мы все восхищаемся крупными достижениями в области искусства, а также естественных и гуманитарных наук, но редко отдаем должное тому, чего достигаем сами в ходе своей повседневной жизни. Мы узнаем вещи, которые видим, понимаем слова, которые слышим, и запоминаем то, что пережили, чтобы позже использовать эти знания в применении к новым проблемам и возможностям.

А еще мы делаем кое-что замечательное, на что не способны никакие другие существа: всякий раз, когда нам не удастся достигнуть цели, думая привычным способом,

мы начинаем думать о самих своих мыслях — и если такое «рефлексивное мышление» показывает, где мы поступили не так, это позволяет нам сформировать новый и более продуктивный способ думать. Однако мы пока очень мало знаем о том, как мозгу удастся все это делать. Как работает воображение? Чем порождается сознание? Что такое эмоции, чувства и мысли? Как мы вообще умудряемся думать?

Сравните это с прогрессом, который наблюдается в изучении материальных феноменов. Что такое твердые вещества, жидкости и газы? Что такое цвета, звуки и температуры? Что такое силы, напряжения и натяжения? Какова природа энергии? В наше время почти все подобные тайны уже объяснены с помощью очень небольшого числа простых законов, таких как уравнения, открытые физиками вроде Ньютона, Максвелла, Эйнштейна и Шредингера.

Естественно, психологи, стремясь подражать физикам, пустились на поиски некоего компактного свода законов, который мог бы объяснить, что происходит внутри нашего мозга. Однако в этой книге вы увидите аргументы в пользу того, что это начинание обречено на неудачу, потому что такого простого набора законов не существует, ведь мозг состоит из сотен участков, каждый из которых сформировался для выполнения определенных видов задач; одни распознают ситуации, другие приказывают мышцам выполнять действия, третьи формулируют цели и планы, а четвертые накапливают и используют огромное количество знаний. И хотя нам еще мало известно о том, как работают эти сотни мозговых центров, мы знаем, что их строение основано на информации, которая содержится в десятках тысяч унаследованных генов, — так что каждый участок мозга функционирует по-своему, опираясь на несколько иной набор законов.

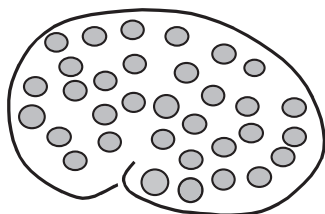
Вместе с осознанием того факта, что в нашем мозге действуют столь сложные механизмы, приходит понимание, что мы должны двигаться ровно в противополож-

ном направлении от физиков: вместо поиска простых объяснений нам нужно искать более сложные способы объяснять самые привычные психические феномены.

Например, значения таких слов, как «чувства», «эмоции» или «сознание», кажутся настолько естественными, ясными и прямыми, что мы даже не видим, с какой стороны начинать их обдумывать. Однако в данной книге утверждается, что каждое из этих слов пытается описать действие обширной сети процессов, происходящих внутри нашего мозга. Например, в главе четвертой будет показано, что термин «сознание» отсылает нас более чем к двадцати различным процессам такого типа!

Казалось бы, превращая то, что поначалу считалось простым и ясным, в проблемы, которые теперь представляются более сложными, мы только ухудшаем ситуацию. Однако если взглянуть на вещи масштабнее, это повышение сложности на самом деле упростит нам задачу. Ведь, разделив старую загадку на части, мы получим вместо большой задачи несколько новых и более мелких, каждая из которых по-прежнему может быть сложной, но уже не будет казаться неразрешимой. Кроме того, глава девятая разъяснит нам, что отношение к себе как к сложному механизму не должно умять в нас чувства собственного достоинства, а только усиливать чувство ответственности за свои действия.

Процесс деления старых больших вопросов на более мелкие мы в этой книге начнем с того, что изобразим типичный мозг как состоящий из огромного множества частей, которые будем называть «ресурсы».



Это изображение пригодится всякий раз, как нам захочется объяснить какую-либо психическую деятельность (например, гнев, любовь или смущение), для того чтобы продемонстрировать, как данное состояние ума становится результатом функционирования определенного набора умственных ресурсов. Например, состояние, называемое «гнев», судя по всему, пробуждает ресурсы, помогающие нам реагировать с необычайной скоростью и силой, одновременно подавляя ресурсы, которые, наоборот, заставляют планировать поведение и действовать более осмотрительно; иными словами, гнев заменяет осторожность агрессивностью, а сочувствие — враждебностью. Аналогичным образом чувство, называемое «страх», оперирует ресурсами так, чтобы заставить вас отступить.

Читатель: *Иногда у меня бывает такое состояние, когда все кажется радостным и ярким. А бывает иначе (хотя ничего не изменилось): все окружающее начинает казаться мрачным и темным, и друзья называют меня грустным или подавленным. Почему я становлюсь жертвой таких состояний — или настроений, чувств, предрасположенностей — и чем вызывается все их странное влияние?*

На это существует несколько распространенных ответов: «Такие изменения вызываются действием химических веществ в мозге», или «Это результат избытка стресса», или «Они появляются, если поддаваться депрессивным мыслям». Однако в подобных заявлениях почти ничего не говорится о том, как именно функционируют эти процессы, тогда как идея выбора определенных ресурсов способна более конкретно показать, каким образом меняется способ думать. Например, глава первая начнется с размышления об очень знакомом феномене:

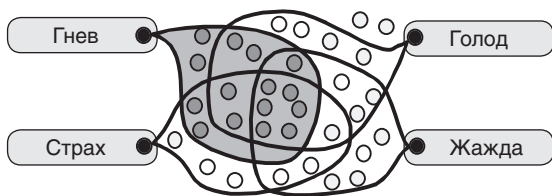
Когда кто-то, кого вы знаете, влюбляется, возникает ощущение, как будто перед вами почти что другой человек:

он думает иначе, у него появляются иные цели и приоритеты. Как будто где-то щелкнули выключателем и запустили новую программу.

Что такого может происходить внутри мозга, что объясняло бы подобные перемены в мыслительных процессах? Вот как подходит к вопросу эта книга:

Каждое из наших основных «эмоциональных состояний» вызывается включением определенных ресурсов и одновременным выключением других, в результате чего изменяется то, как ведет себя наш мозг.

Но что именно активирует такие наборы ресурсов? В последующих главах мы постараемся доказать, что нашему мозгу необходим и еще один тип ресурсов, которые мы будем называть «Критики». Каждый из них специализируется на распознавании определенного состояния, а затем активирует конкретный набор других ресурсов. Некоторые из наших Критиков заложены в нас с рождения и являются источником определенных «инстинктивных» реакций — таковы, например, гнев, голод, страх и жажда, которые развились, помогая выжить нашим предкам. Если точнее, гнев и страх развивались для защиты и обеспечения безопасности, в то время как голод и жажда — для добычи пропитания.



Однако по мере обучения и роста мы также развиваем умение активировать другие, новые наборы используемых ресурсов — они приводят к возникновению психических состояний того типа, который мы считаем