

УДК 004.021
ББК 32.973.3
В97

Вэй Ци Ма, Кёрдинг К., Голдрайх Д.

В97 Байесовские модели восприятия и действия / пер. с англ. В. С. Яценкова. – М.: ДМК Пресс, 2023. – 458 с.: ил.

ISBN 978-5-93700-229-7

Многие формы восприятия и действий можно математически смоделировать с помощью вероятностного (байесовского) вывода – метода, используемого для получения выводов на основе неопределенных данных. Согласно этим моделям, сталкиваясь с зашумленными и неоднозначными данными, человеческий мозг ведет себя как талантливый специалист по обработке данных или следователь на месте преступления. Данная богатая примерами и иллюстрациями книга представляет собой введение в методологию построения и использования вероятностных моделей перцептивного принятия решений и действий.

Издание адресовано широкому кругу читателей, интересующихся нейробиологией, когнитивными науками, машинным обучением, психологией, лингвистикой и математикой.

УДК 004.021
ББК 32.973.3

Copyright © MIT Press 2023. Права на русское издание получены через агентство Александра Корженевского

Все права защищены. Любая часть этой книги не может быть воспроизведена в какой бы то ни было форме и какими бы то ни было средствами без письменного разрешения владельцев авторских прав.

ISBN 978-0-262-04759-3 (англ.)

ISBN 978-5-93700-229-7 (рус.)

© 2023 Massachusetts Institute
of Technology

© Перевод, оформление, издание,
ДМК Пресс, 2023

Мы посвящаем эту книгу памяти Дэвида Книла (1961–2014). Большая часть того, что все мы трое знаем о байесовском моделировании восприятия и действия, – его заслуга. Он был заботливым и терпеливым наставником и отличным учителем и сделал изучение этой темы комфортным и приятным. Область байесовского моделирования восприятия и действия не была бы такой, какой она есть, без него, и эта книга, вероятно, никогда бы не была написана

Содержание

От издательства	14
Благодарности	15
Сокращения	18
Введение	19
Глава 1. Неопределенность и вывод	24
1.1. Цель восприятия	25
1.2. Гипотезы и их вероятности	26
1.3. Сенсорный шум и неоднозначность восприятия.....	31
1.4. Байесовский вывод в зрительном восприятии	33
1.5. Байесовский вывод в слуховом восприятии.....	38
1.6. Исторический обзор: восприятие как бессознательное умозаключение.....	45
1.7. Заключение.....	46
1.8. Рекомендуемая литература	47
1.9. Задачи	48
Глава 2. Применение правила Байеса	52
2.1. Этапы байесовского моделирования	52
2.2. Альтернативная форма правила Байеса	57
2.3. Площадное представление	57
2.4. Ошибка прокурора.....	59
2.5. Смена априорного убеждения: пример багажной карусели	61
2.6. Плоское априорное распределение: пример гештальт-восприятия	65
2.7. Оптимальность, эволюция и мотивы байесовского моделирования.....	68
2.8. Заключение	69
2.9. Рекомендуемая литература	70
2.10. Задачи	70
Глава 3. Байесовский вывод в условиях зашумленных измерений	78
3.1. Этапы байесовского моделирования	79
3.2. Этап 1: порождающая модель	80

3.2.1. Измерение: абстрактное сенсорное представление	81
3.2.2. Графовая модель	82
3.2.3. Распределение стимулов	83
3.2.4. Распределение измерений	84
3.2.5. Совместное распределение	86
3.3. Этап 2: вывод	86
3.3.1. Априорное распределение	87
3.3.2. Функция правдоподобия	88
3.3.3. Апостериорное распределение	90
3.3.4. Апостериорное среднее	94
3.3.5. Ширина апостериорного распределения	96
3.3.6. Оценка апостериорного среднего	96
3.3.7. Оценка MAP	97
3.4. Неопределенность и уверенность	97
3.4.1. Неопределенность	97
3.4.2. Байесовская уверенность	99
3.5. Несоответствие модели в выводе	101
3.5.1. Априорное несоответствие	101
3.5.2. Неудовлетворительные априорные распределения	102
3.6. Гетероскедастичность	103
3.7. Амплитудные переменные	103
3.8. Применение байесовских моделей	105
3.9. Восприятие	106
3.10. Заключение	107
3.11. Рекомендуемая литература	108
3.12. Задачи	108
Глава 4. Распределение отклика	114
4.1. Унаследованная изменчивость	114
4.2. Распределение отклика	116
4.3. Распределение убеждений или распределение отклика?	118
4.4. Оценка максимального правдоподобия	120
4.5. Смещение и среднеквадратичная ошибка	121
4.5.1. Перспектива «обратного смещения»	123
4.5.2. Все расходы оплачены	125
4.6. Другие оценки	127
4.7. Шум принятия решения и шум отклика	128
4.8. Заблуждения	128
4.9. Размышления о байесовских моделях	131
4.10. Заключение	132
4.11. Рекомендуемая литература	132
4.12. Задачи	132
Глава 5. Комбинация признаков и накопление свидетельств	140
5.1. Что такое объединение сигналов?	141
5.2. Формулировка байесовской модели	144

5.2.1. Этап 1: порождающая модель.....	144
5.2.2. Этап 2: вывод.....	146
5.2.3. Этап 3: оценка распределения.....	149
5.3. Искусственный конфликт сигналов	149
5.3.1. Различение распределений.....	151
5.4. Обобщения: априорные распределения, множественные сигналы	151
5.5. Накопление свидетельств	153
5.6. Объединение сигналов при неоднозначности	154
5.7. Применение объединения сигналов	155
5.8. Заключение	157
5.9. Рекомендуемая литература	157
5.10. Задачи	158
Глава 6. Обучение как вывод	163
6.1. Множество форм обучения.....	163
6.2. Изучение вероятности бинарного события	165
6.2.1. Прогноз.....	168
6.2.2. Уравнения обновления	170
6.2.3. Неопределенность	170
6.2.4. Биномиальное распределение.....	170
6.2.5. Неравномерное априорное распределение.....	171
6.3. Связь между байесовским обучением и обучением с подкреплением	172
6.4. Изучение точности нормального распределения.....	174
6.4.1. Почему бы не вывести дисперсию?	177
6.4.2. Прогноз.....	177
6.5. Изучение наклона линейной зависимости	178
6.6. Изучение структуры причинно-следственной модели.....	180
6.7. Другие формы обучения.....	182
6.8. Заключение	183
6.9. Рекомендуемая литература	183
6.10. Задача.....	184
Глава 7. Различение и обнаружение	188
7.1. Примеры задач	188
7.2. Различение	189
7.2.1. Этап 1: порождающая модель	189
7.2.2. Этап 2: вывод	190
7.2.3. Гауссова модель	194
7.2.4. Решающее правило с точки зрения измерения	195
7.2.5. Несколько задач могут иметь одно и то же байесовское решающее правило.....	196
7.2.6. Этап 3: распределение отклика	196
7.3. Обнаружение	199
7.4. Уверенность в решении.....	200
7.5. Дополнительные характеристики распределения откликов.....	203
7.5.1. Рабочая характеристика приемника.....	203

7.5.2. Различимость.....	205
7.6. Связь между байесовским выводом и теорией обнаружения сигналов ...	207
7.7. Промежуточные варианты	208
7.8. Заключение.....	208
7.9. Рекомендуемая литература.....	209
7.10. Задачи.....	209
Глава 8. Бинарная классификация	214
8.1. Примеры задач бинарной классификации	214
8.2. Порождающая модель	217
8.2.1. Зеркально симметричные распределения стимулов, обусловленные классами.....	217
8.3. Маргинализация.....	218
8.3.1. Сумма двух бросков кубика.....	219
8.3.2. Непрерывные переменные	221
8.3.3. Условная маргинализация.....	221
8.3.4. Использование порождающей модели.....	222
8.4. Вывод	223
8.5. Распределение отклика.....	226
8.6. Другие распределения стимулов, обусловленных классами.....	228
8.7. «Следуйте по стрелкам».....	231
8.8. Заключение	232
8.9. Рекомендуемая литература	232
8.10. Задачи	233
Глава 9. Мешающие переменные верхнего уровня и неоднозначность.....	239
9.1. Примеры задач	239
9.2. Размер как мешающая переменная верхнего уровня в восприятии глубины сцены.....	241
9.3. Маргинализация.....	246
9.4. Цветовое зрение.....	247
9.5. Распознавание объектов	250
9.6. Заключение	252
9.7. Рекомендуемая литература.....	252
9.8. Задачи	253
Глава 10. Одинаковый или разный?	256
10.1. Примеры задач	256
10.2. Бинарные стимулы	257
10.2.1. Этап 1: порождающая модель.....	257
10.2.2. Этап 2: вывод.....	258
10.2.3. Этап 3: оценка распределения.....	260
10.3. Непрерывные стимулы	262
10.3.1. Этап 1: порождающая модель.....	262

10.3.2. Этап 2: вывод.....	264
10.3.3. Этап 3: вероятность отклика.....	265
10.3.4. Пересмотр этапа 2: вывод стимулов.....	266
10.4. Оценка сходства нескольких элементов.....	268
10.5. Организация восприятия.....	270
10.5.1. Формирование контура.....	272
10.5.2. Пересекающиеся линии.....	274
10.6. Заключение.....	277
10.7. Рекомендуемая литература.....	277
10.8. Задачи.....	278

Глава 11. Поиск..... 281

11.1. Разнообразие форм зрительного поиска.....	282
11.2. Локализация цели при маскировке.....	283
11.3. Локализация цели с помощью зашумленных измерений.....	287
11.4. Обнаружение цели при маскировке.....	290
11.5. Обнаружение цели с помощью зашумленных измерений.....	293
11.6. Применение.....	295
11.7. Заключение.....	296
11.8. Рекомендуемая литература.....	296
11.9. Задачи.....	297

Глава 12. Вывод в меняющемся мире..... 302

12.1. Отслеживание непрерывно меняющегося состояния мира.....	302
12.1.1. Этап 1: порождающая модель.....	303
12.1.2. Этап 2: вывод.....	305
12.2. Обнаружение точки изменения.....	308
12.2.1. Единственная точка изменения.....	308
12.2.2. Случайные точки изменения.....	310
12.2.3. Более реалистичное обнаружение точки изменения.....	312
12.3. Заключение.....	312
12.4. Рекомендуемая литература.....	313
12.5. Задачи.....	313

Глава 13. Сочетание вывода с полезностью..... 317

13.1. Примеры задач.....	317
13.2. Выбор между двумя действиями.....	319
13.3. Выбор между несколькими действиями.....	320
13.4. Математическое определение ожидаемой полезности.....	323
13.4.1. Бинарная классификация.....	324
13.4.2. Непрерывная оценка.....	325
13.5. Функции затрат «чистого восприятия».....	327
13.5.1. Дискретные задачи.....	327
13.5.2. Непрерывная оценка.....	328
13.5.3. Восприятие и действие.....	329

13.6. Что значит принятие оптимального решения	330
13.7. Усложненные ситуации	331
13.7.1. Функции стоимости для неопределенных результатов	331
13.7.2. Нелинейная зависимость между вознаграждением и полезностью	332
13.7.3. Искажения вероятности	334
13.7.4. Шум принятия решения	334
13.8. Применение	334
13.8.1. Зрительное различие	335
13.8.2. Целенаправленные движения	336
13.8.3. Поощряемые доверительные интервалы	336
13.8.4. Вывод функций полезности	338
13.9. Заключение	339
13.10. Рекомендуемая литература	340
13.11. Задачи	340
Глава 14. Нейронная функция правдоподобия	346
14.1. Порождающая модель активности одиночного нейрона	347
14.2. Нейронная функция правдоподобия для одного нейрона	349
14.2.1. Случай 1: колоколообразная настроечная кривая	351
14.2.2. Случай 2: монотонная настроечная кривая	352
14.3. Нейронная функция правдоподобия, основанная на популяции нейронов	354
14.4. Упрощенная модель	357
14.5. Связь между концепциями абстрактной и нейронной моделей	360
14.6. Использование нейронной функции правдоподобия для вычислений	362
14.7. Нейронная реализация байесовских вычислений	364
14.8. Применение байесовских вычислений	364
14.9. Заключение	365
14.10. Рекомендуемая литература	366
14.11. Задачи	367
Глава 15. Байесовские модели в контексте	370
15.1. Байесовское и оптимальное поведение	370
15.2. Чрезмерно сильные заявления об оптимальности	371
15.3. Почему одни модели поведения оптимальны, а другие нет	372
15.4. Байесовские модели не являются механистическими	373
15.5. Байесовский перенос	374
15.6. Вероятностные вычисления и гибридные модели	375
15.7. Сложность реального мира	376
15.8. Как стать настоящим байесовцем	378
15.9. Рекомендуемая литература	379
Приложение А. Обозначения	381

Приложение В. Основы теории вероятностей	384
V.1. Объективная и субъективная вероятности	384
V.2. Смысловое толкование вероятности	385
V.3. Дополняющее событие	386
V.4. Представление в виде диаграммы Венна	386
V.5. Случайные величины и их распределения	387
V.5.1. Сравнение дискретных и непрерывных случайных величин	388
V.5.2. Суммарная вероятность всегда равна 1	389
V.5.3. Дискретные распределения вероятности	389
V.5.4. Непрерывные распределения вероятности	390
V.5.5. Формальное определение функции плотности вероятности	392
V.5.6. Нормирование	393
V.6. Среднее значение, дисперсия и математическое ожидание	394
V.7. Нормальное распределение	395
V.7.1. Определение	395
V.7.2. Центральная предельная теорема	395
V.7.3. Произведение двух нормальных распределений	396
V.7.4. Произведение нескольких нормальных распределений	397
V.7.5. Интегральное нормальное распределение	398
V.7.6. Распределение фон Мизеса	398
V.8. Дельта-функция	399
V.9. Распределение Пуассона	400
V.10. Выборка из распределения вероятностей	400
V.11. Распределения, включающие несколько переменных	401
V.11.1. Совместная вероятность	401
V.11.2. Маргинализация	402
V.11.3. Условная вероятность	403
V.11.4. Независимость	405
V.11.5. Правило Байеса	406
V.12. Функции случайных величин	408
V.12.1. Функции одной переменной: изменение переменных	408
V.12.2. Пример с яблоками	410
V.12.3. Обязательство неведения	412
V.12.4. Математическое представление маргинализации	412
V.12.5. Функции нескольких переменных	414
V.13. Задачи	416
Приложение С. Подбор параметров и сравнение моделей	420
C.1. Что такое модель?	420
C.2. Свободные параметры	421
C.3. Вероятность параметра	421
C.4. Оценка максимального правдоподобия	422
C.5. Подбор параметров по данным из задачи оценки	423
C.5.1. Простая модель	424
C.5.2. Более удачная модель	429
C.5.3. Сравнение моделей	430

С.5.4. Тест отношения правдоподобия	431
С.5.5. Перекрестная проверка	431
С.5.6. Сопоставление методов сравнения моделей	432
С.5.7. Восстановление параметров и восстановление модели.....	433
С.5.8. Ограничения сравнения моделей.....	433
С.6. Абсолютное качество подгонки.....	434
С.7. Подбор параметров модели по данным в задаче различения.....	435
С.7.1. Простая модель	435
С.7.2. Более удачная модель?.....	437
С.8. Подбор параметров модели по данным в задаче классификации.....	437
С.9. Правильный план эксперимента для байесовского моделирования	438
С.10. Рекомендуемая литература.....	439
С.11. Задачи.....	440
Библиография.....	441
Предметный указатель.....	454

Благодарности

Мы долго шли к написанию этой книги и в долгу перед многими людьми. Впервые эта идея пришла нам в голову в июне 2009 г., когда мы вместе с Аланом Стокером и Джонатаном Пиллоу читали курс вычислительной нейробиологии в Instituto Gulbenkian de Ciência в Оэйраше, Португалия. В то время неудержимый оптимист Конрад предсказывал, что мы закончим книгу к декабрю 2009 г. И вот, спустя всего четырнадцать лет, книга у нас в руках. Однако отсрочка имела свои преимущества: на протяжении многих лет мы использовали черновики глав и идеи книги для обучения байесовскому моделированию сотен студентов бакалавриата, аспирантов и докторантов на наших курсах в Университете Макмастера, Медицинском колледже Бейлора, Северо-Западном университете, Нью-Йоркском университете, Университете Пенсильвании, а также в учебных пособиях на конференциях и летних школах. Многие из этих студентов, а также наши ассистенты, в частности Рональд ван ден Берг, Анна Кучирайтер, Люси Лай, Дженнифер Лаура Ли, Джули Ли, Хорхе Менендес, Сашанк Писупати, Анне-Лене Сакс, Шан Шен, Бэй Сяо и Хёрмет Йилтиз и члены лаборатории (невозможно перечислить их всех), внесли многочисленные исправления и предложили различные задачи и их решения. Мы благодарим Нувана де Сильва за проверку всех задач в предыдущей версии книги. Мы благодарим читателей черновиков рукописи, в частности Луиджи Ачерби, Роберта Джейкобса, Майкла Лэнди, Зили Лю и Хавьера Трэвера, за предоставление глубоких и полезных отзывов о содержании и изложении; мы также благодарим Роберта и Зили за то, что они были двумя нашими самыми стойкими сторонниками на протяжении многих лет. Мы глубоко благодарны Бреннану Кляйну, постдоку из Северо-восточного университета, за профессиональные иллюстрации, сопровождающие книгу, а также за множество забавных рисунков и за то, что мы стали более ответственными авторами. Эта книга никогда не была бы закончена без помощи 1,3,7-триметилксантина¹. Мы благодарны Роберту Прайору из MIT Press, которому, судя по фамилии², самой судьбой было велено стать нашим редактором. Он неоднократно устанавливал для нас жесткие сроки и каждый раз терпеливо соглашался их продлевать, он сделал возможной бесплатную онлайн-версию. Наконец, мы хотели бы поблагодарить наши семьи, которые оказывали незаслуженно безграничную поддержку все эти годы.

¹ Кофеин. – *Прим. перев.*

² *Prior* (англ.) в байесовской статистике означает априорное распределение, о котором говорится в этой книге. – *Прим. перев.*

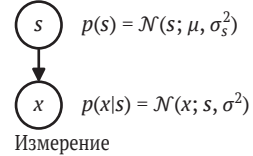
Четыре этапа байесовского моделирования

Примеры из глав 3–4

Этап 1: порождающая модель

- Нарисуйте диаграмму**, где каждый узел является переменной, а каждая стрелка – зависимостью. Наблюдения/измерения располагаются внизу.
- Для каждой переменной **напишите уравнение ее вероятностного распределения**. Для каждого наблюдения примите модель шума. Для других переменных получите распределение из экспериментального плана или из естественной статистики. Если есть входящие стрелки, то распределение является условным.

Воздействие



Этап 2: байесовский вывод (правило принятия решения)

- Вычислите апостериорное распределение состояния системы с учетом наблюдений**. Оптимальный наблюдатель делает это, используя распределения в генеративной модели. В качестве альтернативы наблюдатель может предположить другие распределения (например, ошибочных убеждений). Маргинализируйте (усредните) переменные, отличные от наблюдений и целевого состояния системы.
- Определите целевую переменную апостериорного распределения**. Задайте функцию полезности, затем максимизируйте ожидаемую полезность при апостериорной вероятности. (Альтернатива: выборка из апостериорного набора.) Результат: решающее правило (отображение наблюдений в решение). Когда полезность – это точность, целью является максимизация апостериорной вероятности (максимальное апостериорное решающее правило).

$$\mathcal{L}(s; x) = p(x|s)$$

$$p(s|x) \propto \mathcal{L}(s; x)p(s)$$

$$p(s|x) = \mathcal{N}\left(s; \frac{I_s \mu + Ix}{I_s + I}, \frac{1}{I_s + I}\right)$$

$$\hat{s} = \frac{I_s \mu + Ix}{I_s + I}$$

$$I_s \equiv \frac{1}{\sigma_s^2} \text{ и } I \equiv \frac{1}{\sigma^2}$$

Этап 3: вероятность отклика

Для каждого уникального испытания в эксперименте **вычислите вероятность того, что наблюдатель примет каждое возможное решение с учетом воздействий в этом испытании**. Для этого используйте распределение наблюдений с учетом воздействий (из этапа 1) и правила принятия решений (из этапа 2).

- Хороший метод: выборочные наблюдения согласно этапу 1; для каждого применить решающее правило; свести ответы в таблицу.
- Лучший метод: численное интегрирование по наблюдениям.
- Самый лучший (когда это возможно): аналитическое интегрирование по наблюдениям.

Необязательно: добавьте шум или провалы отклика.

$$p(\hat{s}|s) = \int p(\hat{s}|x)p(x|s)dx$$

$$= \mathcal{N}\left(\hat{s}; \frac{I_s \mu + Is}{I_s + I}, \frac{I}{(I_s + I)^2}\right)$$

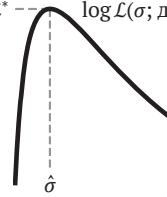
Этап 4: подгонка модели и сравнение моделей

- (a) Вычислите **логарифмическое правдоподобие параметра**, логарифмическую вероятность ответов субъекта во всех испытаниях для гипотетической комбинации параметров.
- (b) **Максимизируйте логарифмическое правдоподобие параметра**. Результат: оценки параметров и максимальное логарифмическое правдоподобие. Проверьте восстановление параметров и сводной статистики с использованием синтетических данных. Используйте более одного алгоритма.
- (c) **Добейтесь соответствия сводной статистике** путем повторного запуска обученной (подобранной) модели.
- (d) **Сформулируйте альтернативные модели** (например, измените этап 2). **Сравните максимальное логарифмическое правдоподобие по моделям**. Исправьте количество параметров (например, используя информационный критерий Акаике). Протестируйте восстановление модели, используя синтетические данные.
- (e) **Проверьте результаты сравнения моделей**, используя сводную статистику.
- (f) Необязательно: **оцените абсолютное соответствие модели**.

$$\log \mathcal{L}(\sigma; \text{данные}) =$$

$$\sum_{i=1}^{n_{\text{trials}}} \log p(\hat{s}_i | s_i, \sigma)$$

$$\log \mathcal{L}^* \quad \text{---} \quad \log \mathcal{L}(\sigma; \text{данные})$$



Сокращения

AIC	Akaike information criterion, информационный критерий Акаике
ANN	artificial neural network, искусственная нейронная сеть
BIC	Bayesian information criterion, байесовский информационный критерий
CCSD	class-conditioned stimulus distribution, обусловленное классом распределение воздействий
CDF	cumulative distribution function, интегральная функция распределения
HMM	hidden Markov model, скрытая марковская модель
LLR	log likelihood ratio, логарифмическое отношение правдоподобия
LPR	log posterior ratio, логарифмическое отношение апостериорных вероятностей
LR	likelihood ratio, отношение правдоподобия
MAP	maximum a posteriori [estimate], максимум апостериорной вероятности [оценка]
MLE	maximum-likelihood estimate, оценка максимального правдоподобия
MSE	mean squared error, среднеквадратическая погрешность
PDF	probability density function, плотность распределения вероятности
PME	posterior mean estimate, оценка апостериорного среднего
PPV	positive predictive value, положительная прогностическая ценность
PSE	point of subjective equality, точка субъективного равенства
RMSE	root-mean-square error, среднеквадратическая ошибка (квадратный корень)
ROC	receiver operating characteristic, рабочая характеристика приемника

Введение

Эта книга представляет собой введение в байесовские модели перцептивного принятия решений и действий. В таких моделях человеческий разум ведет себя как компетентный специалист по данным (или диагностирующий врач, или исследователь места преступления – выберите свою метафору), извлекая смысл из шумных и неоднозначных сенсорных наблюдений. В последние десятилетия байесовский подход к восприятию и действию становится все более популярным, а разработанные модели все шире применяются на практике. Этот подход нам особенно интересен, потому что он строго и успешно объясняет многие экспериментальные данные.

Мы решили написать эту книгу, потому что по сей день не было издано ни одного руководства, которое учило бы с нуля, как рассматривать и строить байесовские модели. Это не означает, что читателям недоступны превосходные материалы по байесовскому моделированию. Тем не менее фундаментальные исследовательские работы, как правило, сосредоточены на узких вопросах и не дают общей картины, в то время как обзорные статьи, как правило, посвящены перечислению достигнутых результатов и не очень полезны начинающему разработчику байесовских моделей. В нашем собственном исследовании мы опирались на академические издания, такие как [45, 94, 181], но они не образуют полную дидактическую последовательность. Наконец, [171] – отличное введение в правило Байеса, но оно не посвящено построению моделей для использования в исследовательской практике.

Эта книга не требует каких-либо предварительных знаний по теории вероятностей, но подразумевает понимание основ математики. Читатели с базовыми знаниями в области исчисления сочтут книгу вполне доступной; но и без знания математики можно понять значительную часть контента. Если это ваш первый опыт математического моделирования – примите наши поздравления! Использование языка математики позволяет избежать двусмысленности, а математические модели дают количественные прогнозы. Мы хорошо понимаем, что математика вызывает у некоторых читателей тревогу, но уверены, что при достаточном усердии материал этой книги будет доступен любому студенту университета, изучающему естественные науки.

Мы рекомендуем читателям уделить время выполнению промежуточных упражнений, встречающихся непосредственно в главах, а также задач в конце каждой главы, представляющих собой смесь базовых идей, математики и моделирования. Однако мы считаем, что полноценного понимания книги невозможно достичь, пока вы не научитесь связывать математические понятия с реальностью.

Восприятие как умозаключение

Фундаментальная проблема восприятия заключается в том, что состояние мира по большей части недоступно напрямую. Глаза, уши, кожа и другие органы чувств регистрируют физические сигналы и преобразуют их в электрические импульсы, направляющиеся к мозгу, – что-то вроде нейронной азбуки Морзе. Глаза регистрируют рисунок света, но сами по себе не идентифицируют зрительную сцену. Кожа ощущает давление и вибрацию, но не идентифицирует внешний объект, вызывающий эти раздражители. Уши обнаруживают звуковые волны, но не указывают на их значение. Таким образом, возникает фундаментальная проблема интерпретации сенсорных наблюдений.

Работа мозга состоит в том, чтобы выяснить основные причины, породившие сенсорные наблюдения, и отреагировать на них соответствующим поведением. Эта работа трудна, потому что сенсорное наблюдение обычно допускает множество возможных интерпретаций. Сенсорное наблюдение может быть низкого качества; например, объекты могут быть плохо освещены, находиться далеко или двигаться быстро. Но даже качественная сенсорная информация не гарантирует наличие единственно возможной интерпретации. Двумерное изображение на сетчатке совместимо с бесконечным числом трехмерных объектов. Хорошо замаскированное животное может быть неотличимо от фона. Слова «торт» и «корт» обычно производят одинаковую слуховую реакцию. Напряжение в мышцах не дает однозначную информацию о позе тела. Возможны множественные интерпретации одних и тех же сенсорных наблюдений.

Получение выводов о состоянии окружающего мира на основе некачественных или неполных наблюдений – это форма умозаключения¹ (*вывод*, *inference*). В частности, *вероятностный вывод*, или *байесовский вывод*, выполняется путем присвоения вероятностей возможным состояниям мира. Теория, изложенная в этой книге, состоит в том, что перцептивное принятие решений является формой байесовского вывода. С помощью байесовского вывода мозг может присваивать вероятности различным состояниям мира, что является решающим шагом при принятии решения о том, какие действия следует предпринять.

В байесовском выводе вероятности назначаются на основе *порождающей модели*² (*generative model*) – статистического описания того, как возникают наблюдения. Это описание включает в себя *исходные убеждения* (*prior belief*), а также убеждения о статистических зависимостях между переменными. Хотя природа сенсорной информации и структура порождающей модели зависят от задачи, сами байесовские вычисления всегда следуют одним и тем же правилам исчисления вероятностей. Таким образом, байесовский подход объединяет огромный диапазон совершенно разрозненного поведения в рамках одной целостной структуры.

¹ Далее для краткости мы будем использовать общепринятый термин «вывод». – Прим. перев.

² Следует отличать от *генеративной модели* в области глубокого обучения. – Прим. перев.

Порождающая модель воплощает предположения, сделанные наблюдателем, и поэтому может принимать различные формы. Особое значение имеет случай, когда порождающая модель является *правильным* статистическим описанием наблюдений. В этом случае байесовский вывод является *оптимальным* в том смысле, что результирующие вероятности, присвоенные состояниям системы, являются компонентом стратегии максимизации *эффективности* (performance) – независимо от того, измеряется ли эффективность с точки зрения точности или иным образом. Другими словами, среди всех стратегий, которые наблюдатель может использовать для задачи восприятия, байесовский вывод, основанный на правильной порождающей модели, является наилучшим вариантом из возможных.

Оптимальный вывод не значит безошибочный. Это всего лишь означает, что сделан наилучший вывод из информации, доступной наблюдателю. Эффективность оптимальной стратегии будет зависеть от информативности сенсорного наблюдения. Иллюзии восприятия часто можно объяснить с точки зрения байесовского вывода. Мы продемонстрируем это на нескольких примерах.

Классификация байесовских моделей

Ученые и философы давно искали способы охарактеризовать логическую структуру моделей. Один особенно полезный подход заключается в классификации моделей по трем типам – описательные, процессные и нормативные.

Описательная модель (descriptive model, иногда называется *моделью «Что»*) представляет собой математическое описание поведенческих переменных (таких как точность или время реакции) в терминах наблюдаемых переменных (таких как интенсивность стимула или количественная характеристика личности). Эти описания могут принимать форму регрессии, обобщенной линейной модели или модели машинного обучения. Утверждение, поддерживаемое подгонкой описательных моделей к данным, заключается в том, что существует модель, которая может соответствовать данным с определенной ошибкой.

Процессная модель (process model, иногда называется *моделью «Как»*) является более амбициозной: она пытается разделить формирование вывода на обобщаемые и психологически или физиологически значимые составляющие процессы, определяя, как наблюдатель/агент принимает решение на основе доступной информации. Примерами таких процессов могут быть «добавление гауссова шума измерения к переменной воздействия» или «наблюдатель сопоставляет переменную решения с решением, применяя критерий». Цель процессной модели – не просто согласовать данные, но и зафиксировать внутренние представления и конкретный поток информации. В данном случае утверждение относительно подогнанной к данным модели заключается в том, что если процесс соответствует гипотезе, то результатом будет измеренное поведение.

Нормативная модель (normative model, иногда называется *моделью «Почему»*) амбициозна по-другому. В нормативной модели мы спрашиваем, поче-

му поведение такое, какое оно есть. Точнее говоря, мы спрашиваем, почему в конкретной экологической нише полезны определенные виды поведения. Например, нормативная модель может (явно или неявно) предполагать, что в нашей жизни важна максимальная точность. Такая модель может затем делать предположения о мире, которые рассматриваются как неизменные (например, что в зрении присутствует фиксированное количество шума). Затем модель может найти оптимальное решение проблемы, которое можно сравнить с реальным поведением. Утверждение относительно этой подогнанной к данным модели заключается в том, что если мир или экологическая ниша соответствует гипотезе, то измеренное поведение будет полезным.

Байесовские модели обычно используются в качестве нормативных моделей: мы указываем цель, которую агент хочет оптимизировать, предположения, которые агент имеет о своем мире, и типы наблюдений, которые доступны агенту. Затем байесовский подход позволяет нам обнаружить поведение, которое оптимизирует цель агента. Для байесовской модели с определенными наблюдениями, априорными знаниями и целями существует одно четко определенное оптимальное действие, и в этой книге объясняется, как его вычислить. В этом смысле байесовская модель является прототипом нормативной модели.

Байесовские модели также часто используются в качестве процессных моделей. С одной стороны, байесовские модели, очевидно, являются процессными – они делят путь от воздействия к реакции на два этапа, принимая точку зрения наблюдателя. На этапе «кодирования», или «представления», они определяют характер данных, с которыми работает наблюдатель и которые могут отличаться от воздействия – мы не воспринимаем состояния мира напрямую, а получаем от него сенсорные наблюдения. На втором этапе происходит принятие решения, когда данные, доступные наблюдателю, преобразуются в действия. Эта двухэтапная последовательность составляет процессную модель. С другой стороны, однако, процессная природа байесовских моделей менее ясна. Дело в том, что байесовская процедура на этапе принятия решения выполняется путем вычисления функций правдоподобия и апостериорного распределения и создания действия на основе последнего. Исследователи расходятся во мнениях относительно того, являются ли эти шаги и связанные с ними конструкции психологически значимыми или это просто способ вывода нормативного правила принятия решений. В зависимости от позиции этап принятия решения в байесовской модели может считаться моделью процесса, а может и не считаться. Мы обсудим это далее в главе 15.

Байесовские модели также можно расположить на условной поведенческо-нейронной оси. Поскольку соображения оптимальности относятся к функционированию организма в целом, байесовские модели мозга в первую очередь являются поведенческими моделями. Если байесовская модель успешно описывает поведение в конкретной задаче, мы можем надеяться ограничить наше понимание основных нейронных процессов и разработать нейронные модели или модели уровня реализации. Мы коснемся этого в главе 14. Данный подход принципиально отличается от подхода «снизу вверх», при котором начинают с моделирования цепей нейронов вплоть до биологических

деталей, а затем пытаются построить модели, комбинируя несколько цепей.

Последним релевантным критерием является условная ось восприятия-познания. Хотя эта книга посвящена восприятию и действию, байесовские модели широко используются в других областях когнитивной науки и психологии. В частности, существует богатая история байесовских моделей высокоуровневого познания [34], восходящая, по крайней мере, к работам Джонатана Эванса [48] и Джона Андерсона [11], которые внесли большой вклад в понимание когнитивного развития (например, [68, 210]). Познание более высокого уровня время от времени фигурирует в этой книге, особенно в главах 6, 12 и 13. В главе 15 мы комментируем различия между восприятием и познанием.

О чем эта книга *не* рассказывает

Байесовские модели восприятия и действия моделируют наблюдателей, которые делают выводы о состоянии мира на основе сенсорных наблюдений. В этом заключается принципиальное отличие от *байесовского статистического анализа данных*, в котором экспериментатор делает вывод о значении параметра модели на основе собранных данных. Математический аппарат в основном совпадает, но в этой книге мы сосредоточимся на том, как мозг воспринимает или принимает решения, а не на том, как анализируются данные. При этом байесовские модели, как и все модели, имеют параметры, значения которых необходимо вывести (или обучить). По этой причине мы включили в книгу приложение С, посвященное подбору и сравнению моделей; это приложение, однако, по-прежнему не охватывает байесовские методы анализа данных. Читателям, интересующимся анализом данных, мы рекомендуем [59, 103, 164].

Заявление об ограниченном цитировании

В этой книге мы опираемся на работы многих замечательных ученых, но по очевидным причинам не можем упомянуть или перечислить их все. Отсутствие цитаты или упоминания в данной книге ни в коем случае не умаляет заслуги отдельных ученых и достоинства их работ. Мы в равной степени признательны всем исследователям.

В заключение

Мы надеемся, что вам понравится книга, и будем рады вашим отзывам на веб-сайте книги www.bayesianmodeling.com. Дополнительный материал, включая решения задач, интерактивные демонстрации и дополнительное чтение, будет размещен там в будущем.

Глава 1

Неопределенность И ВЫВОД

Как мы трансформируем наши сенсорные наблюдения в убеждения о состоянии мира?

Всякий раз, когда мы что-то воспринимаем, делаем предсказание или обдумываем решение, мы рассуждаем, опираясь на вероятности, даже если не осознаем этого. Мы используем имеющуюся у нас информацию, чтобы сделать вывод или оценить что-то еще, что нас интересует. Имеющаяся у нас информация обычно неполная или зашумленная, поэтому наш вывод не является точным. Например, если мы наблюдаем блестящий пол (имеющаяся информация), это говорит о том, что он *может быть* мокрым (предмет нашего интереса). Используя доступную сенсорную информацию и любые соответствующие знания, которые у нас могут быть, мы должны определить вероятность каждой интерпретации (мокрой или сухой). Как нам удастся выносить здравые суждения в таких ситуациях?

Краткое содержание главы

Мы обрисовываем в общих чертах процесс перцептивного вывода, подчеркивая неопределенность, присущую восприятию. Используя простые примеры, мы вводим вероятности, связанные с перцептивным выводом, правдоподобием, априорным и апостериорным распределением вероятностей, сосредоточив внимание на лежащем в их основе *неопосредованном мышлении*¹. Затем мы проиллюстрируем вездесущность перцептивного вывода в повседневной жизни серией примеров, связанных со зрительным и слуховым восприятием. В этой главе мы не используем математику, исследуя каждый пример только качественно и графически. Наша цель – дать интуитивное понимание процесса перцептивного вывода, которое послужит основой для более строгой математической формулировки в следующих главах.

¹ Intuitions, мышление, не опосредованное сознанием. – Прим. перев.

1.1. Цель восприятия

Этот вопрос разрешить единственно разум обязан;
Глаз же природу вещей познавать совершенно не может,
А потому не вини его в том, в чем повинен лишь разум¹.

– Лукреций, *De rerum natura* [111]

Люди, как и другие живые существа, наделены набором тонких органов чувств, с помощью которых они обнаруживают свойства окружающей среды. Органы чувств реагируют на такие разнообразные свойства и явления окружающей среды, как свет (глаза), звук (уши), температура (кожа), текстура материала (кожа), химический состав (нос, язык) и положение тела (суставные и мышечные рецепторы, вестибулярные органы). Наши органы чувств составляют неотъемлемую часть нас самих, настолько, что мы обычно воспринимаем их присутствие как должное. Чтобы оценить роль, которую играют наши чувства, попробуйте представить себе жизнь без зрения, слуха, осязания, обоняния или вкуса.

Какими бы сложными ни были органы чувств, их активация физическими раздражителями – это только первый шаг в восприятии. В повседневной жизни нас не волнует длина волны (цвет) и интенсивность (яркость) попадающих в наши глаза световых волн или структура изменяющейся по амплитуде и времени акустической энергии, поступающей в наши уши. Мы заботимся не об этих паттернах сенсорной активации как таковых, которые мы называем *сенсорными входами*, или *наблюдениями*, а об их интерпретации. На самом деле качество нашей жизни – а часто и сама жизнь – зависит от нашей способности придумывать правильные интерпретации. Отражает ли этот рисунок света лицо друга? Является ли эта акустическая волна звуком ветра, воем собаки или голосом нашего спутника? Короче говоря, наш интерес заключается не в сенсорной информации как таковой, а в тех знаниях, которые она предоставляет о соответствующих состояниях мира² [1].

Сделать интерпретативный переход от *ощущения* (активация органов чувств) к *восприятию* (вывод о состоянии мира) – сложная задача. Вообще говоря, эта книга о том, как мозг может оптимально выполнять эту задачу. Большой и быстро растущий объем экспериментальных и теоретических работ показывает, что восприятие – это, по крайней мере неявно, процесс вероятностного вывода, в котором организм пытается сделать вывод о наиболее вероятном состоянии мира, используя сенсорные данные и все имеющиеся в его распоряжении соответствующие знания. Как гласит цитата Лукреция, восприятие иногда идет не тем путем, но ошибки часто можно рассматривать как побочные продукты разумной стратегии вывода.

¹ Перевод: Ф. А. Петровский (1936).

² Некоторые исследователи восприятия называют состояние мира *дистальным стимулом* (distal stimulus), а наблюдение – *проксимальным стимулом* (proximal stimulus).

1.2. Гипотезы и их вероятности

Переход от ощущения к восприятию основан на *условных вероятностях*. Условная вероятность – это вероятность одного события при наличии другого: например, вероятность того, что у вас хорошее настроение, при условии что на улице идет дождь. Мы обозначаем условные вероятности как $p(B|A)$, читаем «вероятность B при данном A ». Осознают это люди или нет, в повседневной жизни мы очень часто делаем суждения об условной вероятности (рис. 1.1).

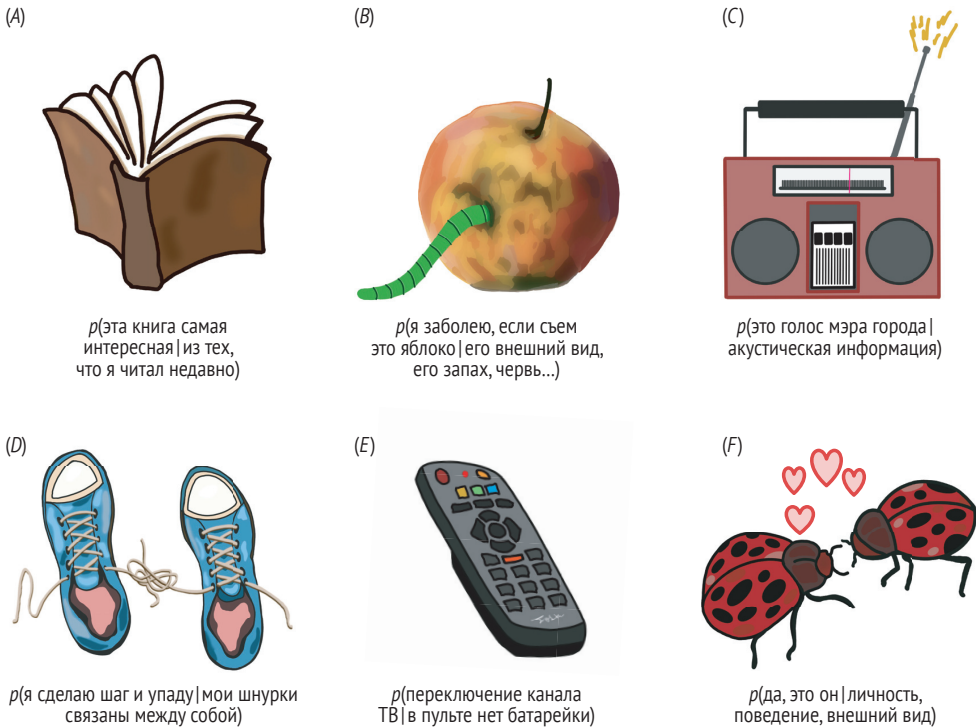


Рис. 1.1 ❖ Различные сценарии с вероятностными суждениями. Обозначение $p(B|A)$ читается как «вероятность события B при данном событии A »

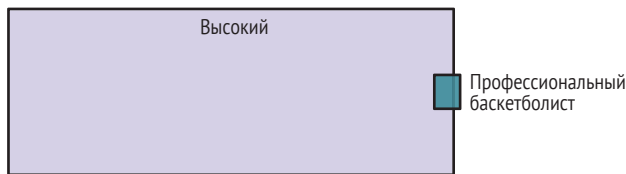
Важно отметить, что условные вероятности не являются симметричными. В общем случае $p(A|B) \neq p(B|A)$. Например, большинство профессиональных баскетболистов высокие, но большинство высоких людей не являются профессиональными баскетболистами. Если A – это «профессиональный баскетболист», а B – «высокий», то справедливо неравенство $p(B|A) > p(A|B)$.

Упражнение 1.1. Какая из двух условных вероятностей больше в каждом из приведенных ниже случаев и почему:

- $p(\text{дождь} \mid \text{облачно})$ или $p(\text{облачно} \mid \text{дождь})$;
- $p(\text{говорит по-французски} \mid \text{родился и вырос в Париже})$ или $p(\text{родился и вырос в Париже} \mid \text{говорит по-французски})$;

- $p(\text{холост}|\text{студент колледжа})$ или $p(\text{студент колледжа}|\text{холост})$;
- $p(\text{вы понимаете правило Байеса}|\text{вы читали эту книгу})$ или $p(\text{вы читали эту книгу}|\text{вы понимаете правило Байеса})$.

Одним из способов визуализации вероятностей является использование площадей прямоугольников (рис. 1.2). Площадь прямоугольника A пропорциональна вероятности события A , обозначаемой $p(A)$, а площадь прямоугольника B пропорциональна вероятности события B , обозначаемой $p(B)$. Возвращаясь к примеру с баскетболом, высоких людей значительно больше, чем профессиональных баскетболистов, поэтому площадь лилового прямоугольника намного больше, чем площадь бирюзового прямоугольника. Область перекрытия занимает почти всю площадь бирюзового прямоугольника, показывая, что вероятность быть высоким, если человек является профессиональным баскетболистом, составляет почти 100 %. Однако площадь перекрытия намного меньше, чем площадь сиреневого прямоугольника, т. е. вероятность быть профессиональным баскетболистом просто при наличии высокого роста очень мала. Различие между $p(A|B)$ и $p(B|A)$ очевидно во множестве реальных примеров.



$r(\text{высокий} | \text{профессиональный баскетболист}) > r(\text{профессиональный баскетболист} | \text{высокий})$

Рис. 1.2 ❖ $p(A|B)$ в общем случае не равна $p(B|A)$. Площадь каждого прямоугольника представляет вероятность события во всем релевантном наборе (здесь все люди). Перекрытие двух прямоугольников представляет $p(A, B)$. Размеры прямоугольников и их перекрытие здесь даны условно и не откалиброваны по фактическим данным о принадлежности к баскетболу

При восприятии известны сенсорные данные или наблюдения, которые непосредственно доступны наблюдателю, например схема активации фоторецепторов сетчатки глаза. Учитывая эти наблюдения (A), наблюдатель стремится сделать вывод о текущем состоянии мира (B). Поскольку наблюдатель не знает истинного состояния мира, B – это *гипотеза*, которую наблюдатель развивает, и мы называем B *гипотетическим*, или *предполагаемым*, состоянием мира. Цель наблюдателя – оценить вероятность того, что мир находится в том или ином возможном состоянии. Например, наблюдатель может захотеть узнать, насколько вероятно, что пол мокрый (гипотетическое состояние мира B), при условии что пол блестит (наблюдение A).

Условная вероятность, представляющая интерес для наблюдателя, равна $p(B|A)$ – это вероятность гипотетического состояния мира при данных сенсорных наблюдениях.

В зависимости от ситуации наблюдатель может быть озабочен оценкой условных вероятностей только двух гипотетических состояний мира (пол

мокрый или сухой), нескольких различных состояний мира (животное на пути впереди – собака, кошка, кролик, енот или скунс) или даже континуума (бесконечное число) мировых состояний. В конечном счете нам хотелось бы выразить результаты нашего вывода, рассчитав вероятность каждого состояния мира с учетом наблюдения (рис. 1.3). Это позволило бы нам принять обоснованное решение о состоянии мира.

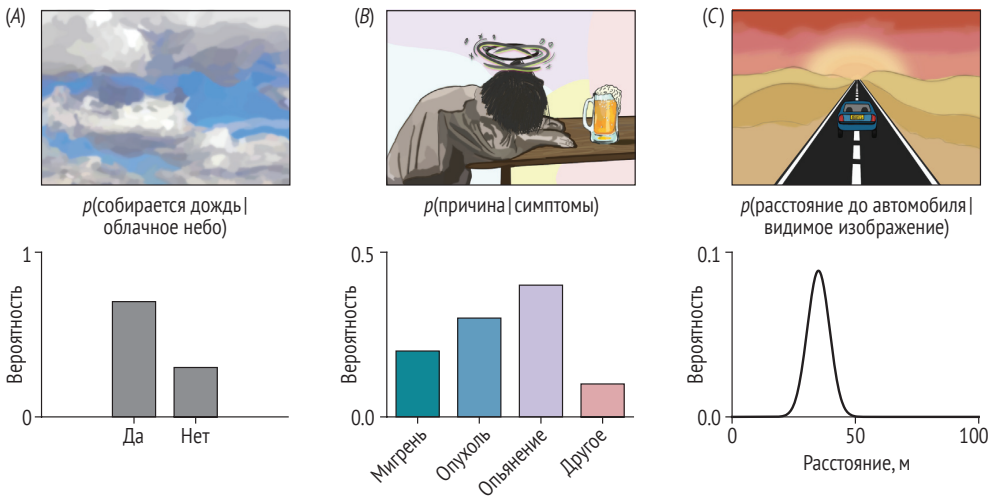


Рис. 1.3 ❖ Различные виды распределений вероятностей: (A) две гипотезы; (B) множественные гипотезы; (C) непрерывное (бесконечное множество) гипотез

Это мой друг?

Предположим, вы видите вдалеке человека, который идет в вашу сторону, и задаетесь вопросом, не является ли он вашим другом (рис. 1.4). К какому бы выводу вы ни пришли, у вас будет определенная степень уверенности, и она может меняться со временем, пока вы продолжаете наблюдать за происходящим. Восприятие зрительных образов почти не требует от нас сознательных усилий, но распознавание сцен на самом деле является сложной вычислительной задачей, которую мозг выполняет «за кадром». Как и все формы восприятия, распознавание сцен столь трудоемко, потому что сенсорный ввод, улавливаемый нервной системой (в данном случае зрительный образ), обычно совместим с несколькими интерпретациями. Зрительный образ может соответствовать вашему другу или постороннему человеку. В данном случае изображение предоставляет достаточную информацию, чтобы заключить, что объект на самом деле является человеком, и оно может содержать данные о приблизительной форме человека (рост, обхват и т. д.). Со временем движущееся изображение может дополнительно предоставить информацию о походке человека. Тем не менее человек далеко, и ваши зрительные наблюдения совместимы со многими возможными людьми.

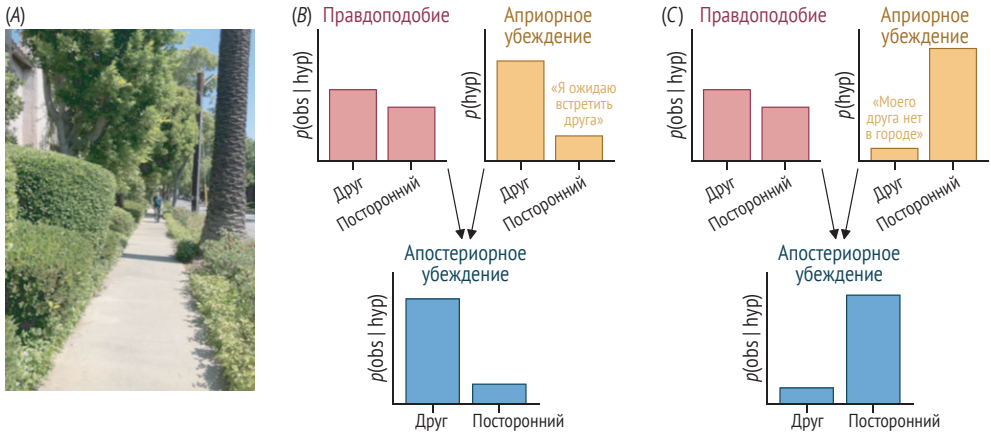


Рис. 1.4 ❖ Узнавание друга: (А) наблюдаемая сцена содержит изображение с низким разрешением человека на расстоянии, похожего на вашего друга; (В) вы считаете, что вероятность получения данного зрительного образа от вашего друга больше, чем вероятность получения его от незнакомца (функция правдоподобия). Вы ожидали встретить своего друга в это время и в этом месте (априорное распределение). Следовательно, вы полагаете, что наблюдаемый человек является вашим другом (апостериорное распределение); (С) в этом альтернативном сценарии вы знаете, что ваш друг уехал из города, поэтому ваше априорное распределение явно свидетельствует в пользу гипотезы о незнакомце. Исходя из того же наблюдения (функция правдоподобия), вы заключаете, что человек, которого вы видите, не является вашим другом

Правдоподобие (likelihood) гипотезы – это вероятность наблюдения при заданном гипотетическом состоянии мира, т. е. $p(\text{наблюдение} | \text{гипотетическое состояние мира})$. График вероятностей всех гипотез резюмирует, насколько хорошо визуальное восприятие позволяет вам отличить одно возможное состояние мира от другого (например, друг или незнакомец). Это график *функции правдоподобия* (likelihood function).

В общем случае на функцию правдоподобия влияет множество факторов. В нашем текущем примере рост человека (ваш друг высокий), цвет волос (каштановые) и манера держать голову (наклоненная) будут влиять на вероятность гипотез. Если вы продолжите наблюдать за приближением человека, график функции правдоподобия со временем будет становиться все более острым по мере поступления более четких и подробных наблюдений. Мы пока оставим в стороне то, как мы могли бы прийти к точной форме функции правдоподобия. На данный момент достаточно понять, что функция правдоподобия представляет собой полное информационное содержание изображения, относящегося к рассматриваемому вопросу (это мой друг?). В частности, она представляет собой вероятность того, что зрительный образ соответствует вашему другу, по сравнению с вероятностью того, что тот же зрительный образ соответствует другому человеку.

Хотя функция правдоподобия является важным компонентом нашего процесса вывода, ее недостаточно для решения стоящих перед нами задач. Функция правдоподобия отображает вероятность наблюдения при каждом гипотетическом состоянии мира: $p(\text{наблюдение} | \text{состояние мира})$. Мы стре-

мимся найти апостериорное распределение – вероятность каждого возможного состояния мира с учетом наблюдения: $p(\text{состояние мира} | \text{наблюдение})$. Чтобы определить апостериорное распределение, мы комбинируем функцию правдоподобия с *априорным распределением*, которое отображает *априорную вероятность* каждого состояния мира. Априорная вероятность гипотетического состояния мира, обозначаемая как $p(\text{состояние мира})$, – это вероятность состояния мира, основанная на всех имеющихся у вас знаниях, кроме наблюдения, – например, на вашей уверенности в том, что ваш друг будет присутствовать в наблюдаемой сцене, существовавшей еще до того, как вы посмотрели на улицу.

Давайте рассмотрим два разных сценария, которые заставят вас иметь разные априорные распределения.

- Сценарий 1: вы договорились встретиться со своим другом на данной улице, приблизительно в то время, когда вы видите идущего к вам человека, похожего на вашего друга.
- Сценарий 2: когда вы видите идущего к вам человека, похожего на вашего друга, вы удивляетесь, потому что знаете, что ваш друг должен быть в отпуске и не планировал возвращаться в город до следующей недели.

Сенсорный ввод этих двух сценариев идентичен (рис. 1.4А), но ваш перцептивный вывод будет разительно отличаться. В первом сценарии вероятность $p(\text{друг})$ была высокой, и вы делаете вывод, что человек, идущий к вам, является вашим другом; во втором сценарии вероятность $p(\text{друг})$ низкая, и вы приходите к выводу, что этот человек, скорее всего, не ваш друг. Очевидно, что ваши априорные вероятности играют решающую роль в процессе перцептивного вывода.

Правило Байеса – фундаментальная теорема теории вероятностей – показывает, как оптимально сочетать ожидание, представленное априорным распределением, с наблюдением, представленным функцией правдоподобия, чтобы вычислить апостериорное распределение (рис. 1.4В–С). Апостериорная вероятность каждого состояния мира – это ваша вера в это состояние мира, основанная на всей соответствующей информации, имеющейся в вашем распоряжении. Правило Байеса гласит, что апостериорная вероятность пропорциональна произведению априорного знания и вероятности:

$$\text{Апостериорная вероятность} = \text{константа} \cdot \text{априорное распределение} \cdot \text{вероятность.} \quad (1.1)$$

Апостериорный анализ основан на всех имеющихся у нас знаниях (т. е. на наших текущих сенсорных наблюдениях и соответствующих предшествующих знаниях). Когда мы начнем использовать фактические числа в главе 2, мы обсудим константу в этом уравнении, но пока это уравнение отражает соответствующие интуитивные предположения. Когда и априорное распределение, и вероятность выше для гипотезы A , чем для гипотезы B , тогда апостериорное значение также будет выше для гипотезы A . Однако если априорное знание благоприятствует гипотезе A , а вероятность – гипотезе B , то апостериорный вывод может быть любым, в зависимости от точных

числовых значений. Например, если одна величина (априорное знание или правдоподобие) лишь слегка поддерживает гипотезу A , а другая величина сильно поддерживает гипотезу B , то апостериорная вероятность склонится в пользу гипотезы B .

1.3. Сенсорный шум и неоднозначность восприятия

В любой системе восприятия чем более плоское (т. е. более ровное) апостериорное распределение, тем более неоднозначны – то есть открыты для множественных интерпретаций – наблюдения. Как мы видели, форма апостериорного распределения является результатом комбинации форм функции правдоподобия и априорного распределения. Неоднозначность восприятия часто возникает из-за широкой функции правдоподобия в отсутствие уравновешивающего четкого априорного знания.

Многие факторы могут снизить качество сенсорных данных и тем самым расширить функцию правдоподобия. К ним относятся физические особенности окружающей среды, ограничения органов чувств наблюдателя и его нервной системы (рис. 1.5). Одним из повсеместных факторов расширения функции правдоподобия является сенсорный шум. Под сенсорным шумом мы понимаем стохастическую изменчивость, присущую физическому процессу, порождающему сенсорное наблюдение. Из-за шума стимул, который повторяется одинаково в течение нескольких попыток, обычно каждый раз вызывает несколько разные сенсорные наблюдения. Рассеяние света, случайная изменчивость окружающих звуков или биофизическая изменчивость частоты возбуждения сенсорных нейронов – все это примеры сенсорного шума. Как показывают эти примеры, сенсорный шум может возникать как во внешнем мире, так и внутри наблюдателя.

Хотя сенсорный шум встречается повсеместно, это не единственная причина широты функций правдоподобия. Даже если бы весь сенсорный ввод можно было каким-то образом сделать бесшумным, многие функции визуального правдоподобия остались бы широкими по чисто геометрическим и оптическим причинам (рис. 1.6). Например, информация неизбежно теряется, когда трехмерный видимый мир отображается на двухмерное изображение на сетчатке (или когда трехмерный слуховой мир отображается на два уха). Это схлопывание по измерению порождает множество случаев неоднозначности, в том числе двусмысленность размер–расстояние. Другим распространенным геометрическим источником функций широкого правдоподобия является *окклюзия*, при которой объект частично закрывает поле зрения наблюдателя, так что сцена совместима со множеством альтернативных конфигураций. В качестве последнего примера рассмотрим кажущуюся простой задачу восприятия оттенка серой поверхности по интенсивности света, который отражается от поверхности и попадает в ваши глаза. Конкретная интенсивность света, попадающего в ваши глаза, соответствует

множеству комбинаций истинного оттенка поверхности и интенсивности источника света. Например, одинаковую интенсивность света, попадающего в глаза, может дать темно-серая бумага при солнечном свете или белая бумага при тусклом свете. Поэтому если неизвестна интенсивность освещающего света, функция правдоподобия в этом сценарии широка. Мы рассмотрим эти и другие примеры с математической точки зрения в последующих главах.

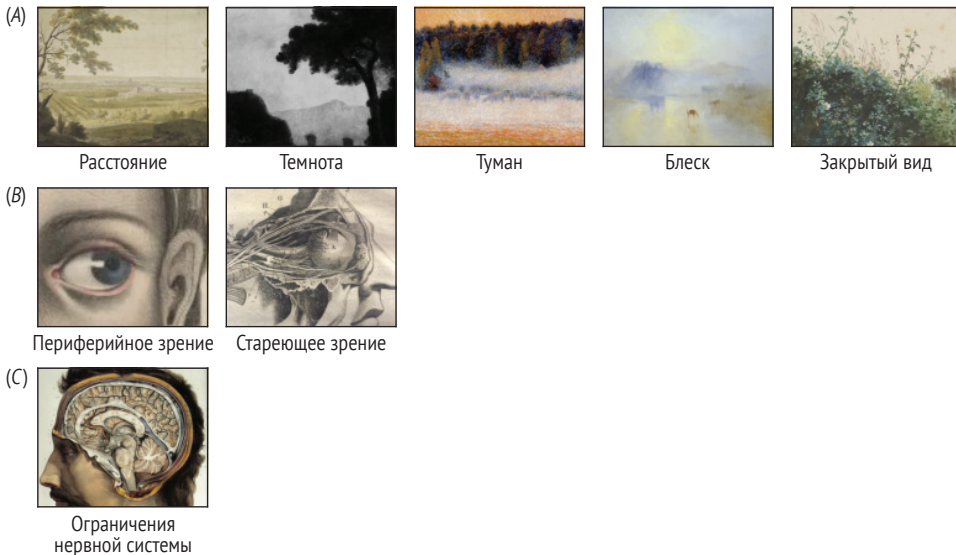


Рис. 1.5 ❖ Источники сенсорной деградации, которые снижают качество зрительной информации, вызывая уплощение функций правдоподобия: (A) физические особенности окружающей среды; (B) ограничения органов чувств наблюдателя. Большинство факторов, показанных в A и B, имеют аналоги в других смыслах. Например, в случае слухового восприятия расстояние, тихая речь, окружающий шум и стареющие органы слуха приводят к низкому качеству входных данных; (C) нервная система наблюдателя. В каждой сенсорной системе нейронные ограничения, такие как ошибочные фоновые знания и нейронный шум, также создают проблемы для восприятия; (1) Жан Шофурье (1679–1757), «Пейзаж с городом или дворцом вдалеке»; (2) неизвестный британский художник, «Пейзаж с темным деревом», конец XVIII в.; (3) Камиль Писсарро (1830–1903), «Заходящее солнце и туман, Эраньи» (1891); (4) Дж. Тернер (1775–1851), «Замок Норхэм, Восход солнца» (ок. 1845 г.); (5) Майлз Биркет Фостер (1825–1899), «Обочина с живыми изгородями»; (6) периферийное зрение: «Глазная хирургия для коррекции косоглазия»; (7) структура глаза и зрительных нервов, иллюстрация из книги Питера Деграверса «Полный физико-медицинский и хирургический трактат о человеческом глазе и демонстрация естественного зрения» (Лондон: Б. Лоу, 1780 г.); (8) Жан-Батист Буржери (1797–1849), «Мозг», иллюстрация из книги *Traité complet de l'anatomie de l'homme* (1831–1854)

Неоднозначность восприятия, возникающая из-за шума или других факторов, увеличивающих вероятность, может быть уменьшена или полностью предотвращена, если наблюдателю доступно четкое априорное распределение. Априорные вероятности основаны на фоновых знаниях и, следовательно, могут меняться со временем, по мере того как наблюдатель приобретает новые знания. Априорные знания также могут отличаться от одного наблю-

дателя к другому. В целом те, у кого больше релевантных знаний, имеют более реалистичные априорные предположения, что способствует точному восприятию. Взгляните на рис. 1.6В. Какие предварительные знания о собаках могут помочь одному наблюдателю испытать меньшую двусмысленность в этой ситуации относительно другого наблюдателя, у которого нет таких знаний?

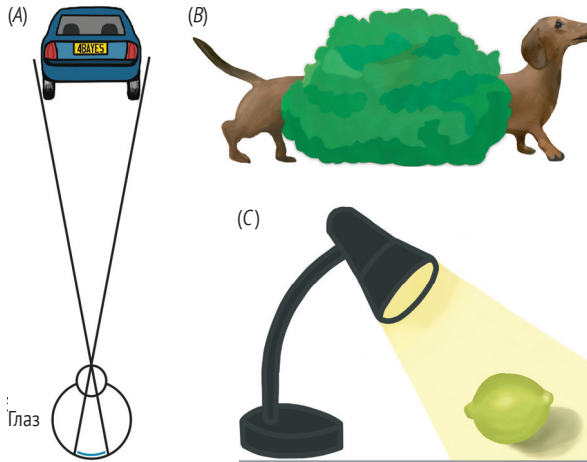


Рис. 1.6 ❖ Расширение функций правдоподобия, вызванное факторами, отличными от сенсорного шума: (А) один и тот же размер изображения на сетчатке может соответствовать небольшому объекту, расположенному ближе к наблюдателю, или более крупному объекту, находящемуся дальше; (В) куст мешает наблюдателю узнать, есть ли в этой сцене две (или даже больше) собаки или только одна очень длинная собака; (С) наблюдатель не может быть уверен в оттенке поверхности, не зная также свойства падающего света

1.4. Байесовский вывод в зрительном восприятии

Поскольку перцептивный вывод очень важен, мы хотим проиллюстрировать его дополнительными примерами из повседневной жизни. Наша цель – развить интуитивное понимание вероятностей, априорных и апостериорных событий, а также оценить замечательную объяснительную силу байесовского вывода как модели восприятия. Мы покажем, что каждый пример имеет уникальные особенности, но каждый из них основан на объединении функции правдоподобия и априорного распределения по правилу Байеса для создания апостериорного перцептивного вывода. Мы надеемся, что эти примеры продемонстрируют как богатство перцептивного вывода, так и широкую применимость байесовской перцептивной структуры.

Восприятие влажности

Когда люди перемещаются по миру, они полагаются на свои органы чувств, чтобы избежать опасностей. В современном мире опасности представлены во многих формах, например объект на нашем пути, быстро приближающийся автомобиль или внезапная ступенька вниз. Еще одна опасность современной жизни – мокрый скользкий пол. Мокрый ли пол на рис. 1.7? Если это так (или может быть так), необходимо соблюдать осторожность и перемещаться медленно, небольшими шагами. Если это не так, мы можем без опаски целеустремленно шагать вперед. Как наше восприятие различает два варианта?

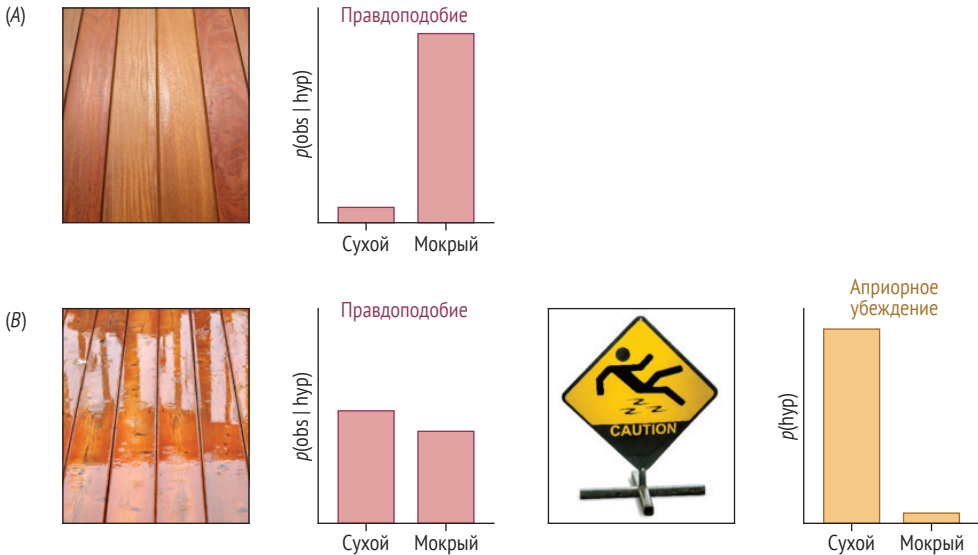


Рис. 1.7 ❖ Восприятие влажности: (А) функция правдоподобия, полученная из зрительного образа этой деревянной поверхности, благоприятствует «сухому» состоянию мира: $p(\text{наблюдение} | \text{сухой}) \gg p(\text{наблюдение} | \text{мокрый})$; (В) блестящая поверхность приводит нас к функции правдоподобия, которая благоприятствует «влажному» состоянию мира: $p(\text{наблюдение} | \text{мокрый}) > p(\text{наблюдение} | \text{сухой})$. Предостерегающий знак служит источником более выраженного априорного знания в пользу «мокрого» состояния мира

Вероятность того, что поверхность будет блестящей, если она мокрая, больше, чем вероятность того, что она будет блестящей, если она сухая. Таким образом, визуальная регистрация блестящей поверхности дает нам функцию правдоподобия, соответствующую мокрой поверхности. Важно понимать, что не только наши априорные предположения, но и наши вероятности зависят от фоновых знаний. В общем случае, чтобы определить вероятности, наблюдатель должен иметь (неявное) понимание процесса, посредством которого различные состояния мира генерируют сенсорные данные. В данном случае наблюдателю необходимо интуитивное знание законов оптики, а именно что мокрая поверхность намного лучше отражает свет.

Как сказано выше, мозг комбинирует вероятности $p(\text{наблюдение} | \text{гипотетическое состояние мира})$ с априорными вероятностями $p(\text{гипотетическое состояние мира})$, чтобы сгенерировать вероятности, которые больше всего его интересуют: $p(\text{гипотетическое состояние мира} | \text{наблюдение})$. Эти последние вероятности называются *апостериорными*, чтобы подчеркнуть, что, в отличие от априорных вероятностей, они формируются *после* наблюдения (апостериори). *Апостериорное распределение вероятностей* представляет собой убеждение мозга в каждом возможном состоянии мира, основанное на всей соответствующей информации (т. е. наблюдениях и ожиданиях). Нам нужно вычислить апостериорную вероятность каждого гипотетического состояния мира, $p(\text{состояние мира} | \text{наблюдение})$. Этот расчет включает перемножение априорных ожиданий и вероятностей. Напомним, что априорная вероятность $p(\text{мокрый})$ отражает априорное ожидание наблюдателя относительно влажности поверхности независимо от визуального наблюдения. Например, еще до того, как наблюдатель войдет непосредственно в наблюдаемую зону, существует вероятность, что он выдвинет гипотезу мокрого пола. Как у наблюдателя возникают подобные априорные предположения?

Фоновые знания, которые формируют априорные убеждения, могут быть приобретены в течение всей предшествующей жизни или совсем недавно. У наблюдателя может быть априорное убеждение, благоприятствующее сухости, потому что подобные места по опыту наблюдателя большую часть времени сухие. Однако если наблюдатель видит предупреждающий знак, указывающий на скользкую поверхность, его априорная оценка может измениться в пользу влажности. Чтобы показать зависимость априорного убеждения от фоновых знаний наблюдателя, мы иногда записываем его как $p(\text{состояние мира} | B)$, где B снова означает информацию, полученную из предыдущего опыта.

Поскольку и априорное убеждение, и правдоподобие зависят от фоновых знаний, апостериорное убеждение тоже зависит от фоновых знаний. Чтобы отметить эту зависимость, мы иногда записываем апостериорную вероятность отдельного состояния мира как $p(\text{состояние мира} | \text{наблюдение}, B)$.

Маскировка

В животном мире выживание часто зависит от способности видеть, но не быть увиденным. Как отмечалось ранее, точная функция правдоподобия указывает на высокую информативность наблюдения, тогда как более плоская функция правдоподобия дает мало информации. Таким образом, в целом животному выгодно иметь острые чувства, которые производят четкие функции правдоподобия, когда животное смотрит на мир (рис. 1.8), но в то же время вести себя так или обладать такими физическими особенностями, которые формируют относительно плоскую функцию правдоподобия у других видов.

В дикой природе многие виды развили черты и поведение, которые служат для маскировки их присутствия или идентичности (рис. 1.9). Эти разнообразные примеры маскировки и мимикрии в животном мире можно рассмат-

ривать как развитые стратегии, направленные на выравнивание функций правдоподобия наблюдателей. Возьмем, к примеру, гусеницу плодовой гусеницы (рис. 1.9А). Примечательно, что особи этого вида принимают цвет коры дерева, на котором живут. Сливаясь с фоном, эти гусеницы защищают себя от хищных птиц. Зрительный образ, получаемый птицей, дает скудное указание на присутствие гусеницы. Хищники тоже выигрывают от маскировки. Рассмотрим изображение леопарда, поджидающего добычу (рис. 1.9D). В высокой золотистой траве, чья окраска очень похожа на его собственную, леопард почти невидим для ничего не подозревающего наблюдателя. Хотя леопарды и другие крупные представители семейства кошачьих могут быстро бегать, им не хватает выносливости для длительных погонь. Их успех в охоте зависит от их способности незаметно приближаться к добыче. Примеров замаскированных хищников и добычи предостаточно в животном мире.

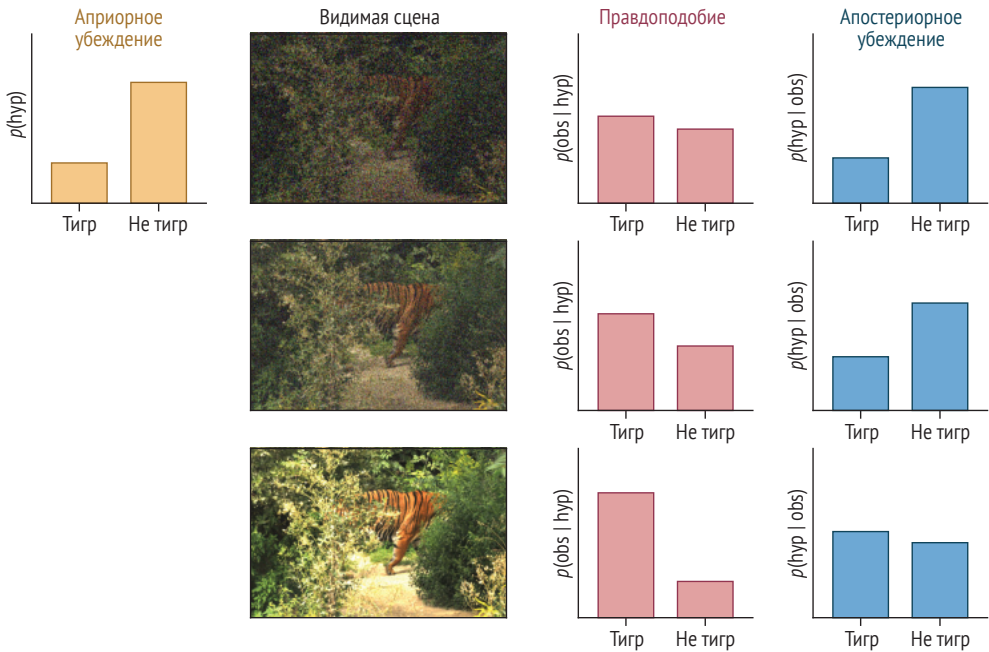


Рис. 1.8 ❖ Влияние остроты зрения на апостериорное распределение. Три жертвы, имеющие одинаковое (20 %) априорное ожидание присутствия тигра (вверху слева), различаются по остроте зрения и, следовательно, формируют разные функции правдоподобия при столкновении с одной и той же наблюдаемой сценой. Чем более плоской является функция правдоподобия, тем больше апостериорное распределение напоминает априорное. Вверху: для этого животного с плохим зрением наблюдаемая сцена порождает почти плоскую функцию правдоподобия. Таким образом, апостериорное распределение животного аналогично его априорному распределению; он мало чему научился из визуального наблюдения. В центре: животное с промежуточной остротой зрения имеет неплоскую функцию правдоподобия. Апостериорное распределение этого животного немного отличается от априорного. Внизу: для этого животного с превосходной остротой зрения сцена дает резкую функцию правдоподобия в пользу присутствия тигра. Апостериорное распределение указывает на вероятность присутствия тигра немного больше 50 %

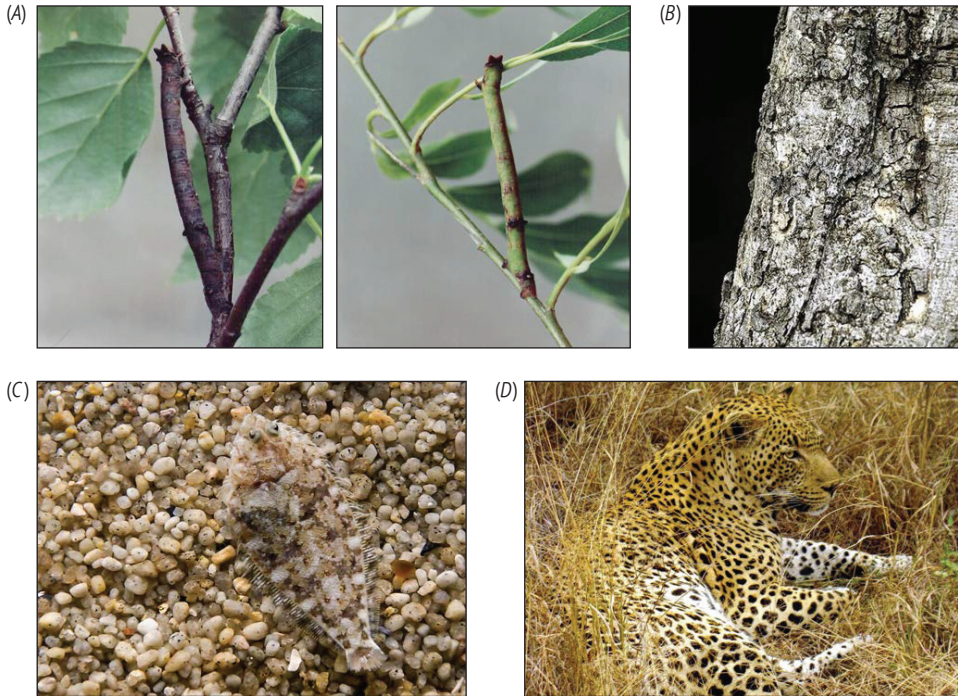


Рис. 1.9 ❖ Выравнивание функции правдоподобия в живом мире: (А) гусеница перечной огневки (*Biston betularia*) меняет свой цвет, чтобы слиться с фоном (ива или береза); (В) летающая ящерица (*Draco dussumieri*) эффектно сливается с фоном коры дерева; (С) камбала лежит на морском дне; (D) маскирующийся леопард в Южной Африке (*Panthera pardus*)

Пока функция правдоподобия наблюдателя не является идеально плоской, он способен хоть что-то извлечь из сенсорного ввода. Однако при наблюдении хорошо замаскированного животного функция правдоподобия наблюдателя почти неинформативна. Важно отметить, что форма функции правдоподобия зависит не только от наблюдаемой сцены, но и от остроты восприятия и сообразительности наблюдателя. Животное, которое для одного наблюдателя почти идеально замаскировано, может быть замечено другим наблюдателем, обладающим более тонким зрением или пониманием. Для наблюдателя, который по опыту знает, что гусеницы перечной моли, как правило, немного шире, чем ветки дерева, на котором они обитают, или что пятна леопарда немного отличаются по внешнему виду от окружающей растительности, одна и та же визуальная сцена даст более четкую функцию правдоподобия по сравнению с наблюдателем, не имеющим этих фоновых знаний.

Наряду с маскировкой эволюция породила сложные сенсорные системы и когнитивные способности, которые снижают неопределенность в отношении присутствия и местонахождения других животных. В своего рода гонке вооружений животные развили все более острые сенсорные системы, чтобы обнаруживать своих все более и более скрытых противников. Показательными примерами являются эволюция зрительной, слуховой и обонятельной систем млекопитающих, а также эволюция узкоспециализированных систем

обнаружения, таких как ультразвуковая эхолокация, используемая насекомоядными видами летучих мышей. В общем, животные выигрывают, если они остро воспринимают присутствие окружающих, в то время как окружающие воспринимают их с трудом. По этой причине у животных развились впечатляющие системы восприятия, чтобы добиться более четкой функции правдоподобия для себя, и в то же время развилась маскировка, чтобы навязывать более широкие функции правдоподобия другим.

1.5. Байесовский вывод в слуховом восприятии

До сих пор мы рассматривали зрительные примеры. Однако перцептивный вывод происходит в каждой сенсорной модальности. Например, люди живут в акустически богатой среде: щебечут птицы, завывает ветер, лают собаки, гудят автомобили, играет музыка и, что, пожалуй, самое главное, мы разговариваем друг с другом. Всякий раз, когда мы пытаемся определить источник звука (это лай собаки?), осознать его местонахождение (где эта лающая собака?) или интерпретировать его значение (какое слово вы только что произнесли?), мы выполняем перцептивный вывод. Чтобы оптимизировать слуховое восприятие, байесовский наблюдатель будет комбинировать априорные вероятности с акустическими правдоподобиями, чтобы получить максимально точный вывод о восприятии.

Птицы на проводе

Люди полагаются, по крайней мере частично, на свой слух, чтобы определить местонахождение объектов. Мы и другие млекопитающие локализуем источники звука, используя бессознательные расчеты, в том числе сравнивая интенсивность и время поступления звуков в два уха.

Предположим, вы идете по улице прекрасным солнечным утром и замечаете силуэты пяти птиц, сидящих на проводе (рис. 1.10А). Внезапно одна из птиц (не видно, какая) начинает мелодично петь. Какая птица поет? Ваша слуховая система быстро обрабатывает акустическое наблюдение, получая широкую функцию правдоподобия. Эта функция правдоподобия является непрерывной функцией от местоположения; то есть звук, который вы услышали, совместим с источником в континууме местоположений. Тем не менее некоторые места связаны с более высокой вероятностью, чем другие. Интересно, что место с наибольшей вероятностью может не совпадать с точным местонахождением какой-либо птицы. Эта ситуация распространена в акустическом восприятии и может быть вызвана многими факторами. Например, если поющая птица не смотрела прямо на вас, то издаваемый ею звук мог отразиться от ближайших предметов, прежде чем достичь ваших ушей. Даже если звуковые волны смогли достичь ваших ушей по прямой, стохастическая изменчивость реакции вашей нервной системы может привести к тому, что

функция правдоподобия достигнет пика в месте, которое немного смещено от местоположения источника.

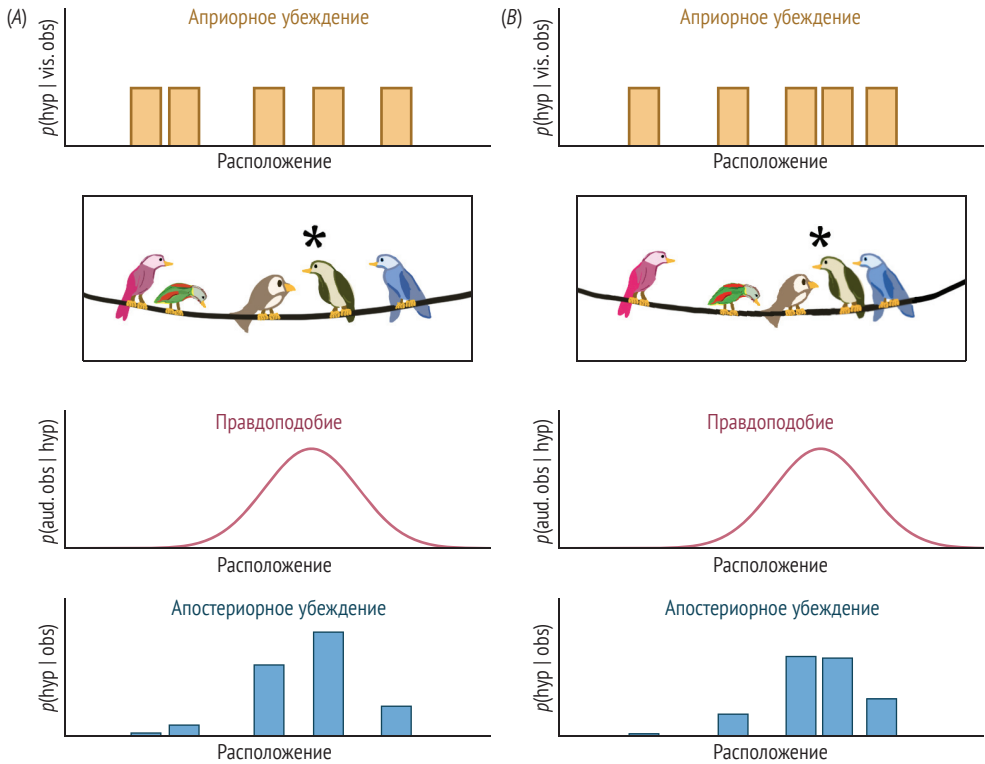


Рис. 1.10 ❖ Локализация источника звука: (А) зрительный образ птиц обеспечивает основу для предварительного распределения по местоположению источника звука. Широкая функция правдоподобия отражает неточность акустического наблюдения. Апостериорное распределение свидетельствует в пользу гипотезы о том, что пела четвертая птица слева (*); (В) неопределенность восприятия заметно возрастает, если птицы собираются ближе друг к другу

В отличие от функции правдоподобия визуальная информация в данном примере не непрерывна, а дискретна. Вы видите пять отдельных птиц. Ваше зрительное наблюдение, которое произошло до того, как птица зачирикала, дает вам априорное распределение. Таким образом, априорное распределение отлично от нуля в пяти дискретных точках (мы предполагаем, что ваше зрительное восприятие очень точно сработало в этой высококонтрастной сцене). Обратите внимание, что априорные вероятности считаются равными для пяти птиц и что априорная вероятность того, что источник звука займет пустое место на проводе, равна нулю. Это просто означает, что до того, как вы услышали пение, вы с равной вероятностью предполагали, что любая из птиц будет петь. Используя правило Байеса, теперь мы можем рассчитать апостериорное распределение вероятностей для местоположения источника звука. Для каждого из пяти предполагаемых местоположений мы умножаем вероятность наблюдения на априорное распределение вероятностей. Полу-

ченное апостериорное распределение указывает на четвертую птицу слева как на наиболее вероятный источник приятной мелодии.

Интуиция подсказывает, что если бы птицы сидели ближе друг к другу на проволке, наш вывод был бы менее точным. Этот результат действительно следует из байесовского вывода, как показано на рис. 1.10В. Здесь поющая птица изображена на прежнем месте, но трое других птиц сидят ближе к ней, чем раньше. Наше априорное распределение отражает новые положения птиц, но акустическое наблюдение и, следовательно, функция правдоподобия такие же, как и раньше. Апостериорное распределение теперь шире и ниже, что указывает на то, что хотя наиболее вероятным певцом является та же самая птица, наша неопределенность заметно возросла. Действительно, по нашему мнению, поющая птица почти в равной степени могла быть третьей или четвертой слева.

Прежде чем покинуть этот пример, мы хотели бы обратить ваше внимание на два альтернативных подхода к решению задачи, которые привели бы к одному и тому же ответу. В первом подходе, прежде чем смотреть на провод, мы могли бы начать с широкой однородной априорной оценки гипотетических местоположений птиц, отражая тот факт, что до зрительного наблюдения мы понятия не имели, где какие птицы могут сидеть. Затем мы могли бы включить последующее визуальное наблюдение в функцию правдоподобия и объединить ее с нашим плоским априорным распределением, чтобы получить апостериорное распределение по местонахождению птиц. Действительно, именно это первоначальное апостериорное распределение зрительного ввода мы использовали здесь в качестве априорного распределения для нашего анализа слухового наблюдения. Этот подход иллюстрирует важную общую черту байесовского вывода: его можно делать итеративно, при этом апостериорное распределение одного вывода используется в качестве априорного распределения для следующего.

О втором подходе, который приводит к тому же ответу, будет подробно рассказано в главе 5. Вкратце он заключается в следующем: начиная с плоского априорного распределения вероятностей по позициям, мы могли бы одновременно включить как зрительные, так и акустические наблюдения как функции правдоподобия в процедуру, называемую *объединением сигналов*. В этом подходе мы не будем использовать зрительную информацию для создания априорного распределения для последующего слухового наблюдения, а вместо этого объединим зрительную функцию правдоподобия, которая имеет пять дискретных пиков, с непрерывной слуховой функцией правдоподобия. По сути, когда у нас есть два или более независимых источника информации, мы можем выбрать, следует ли использовать разные источники последовательно, при этом апостериорные значения каждого наблюдения используются в качестве априорных для следующего, или все сразу, когда все наблюдения вводятся через функции правдоподобия. Таким образом, границы между функциями правдоподобия и априорными распределениями вероятностей часто размываются, а выбор того, как использовать информацию, остается за байесовским разработчиком моделей. Эта гибкость является не проблемой, а преимуществом байесовского подхода. Внутренняя согласованность правил байесовского вывода гарантирует, что если в выводе задействована вся доступная информация, результирующее апостериорное

распределение будет одним и тем же независимо от выбранного маршрута. В рамках байесовской модели часто существует несколько способов прийти к одному и тому же решению.

Мондегрин

Восприятие речи требует сложных умозаключений на различных уровнях, хотя наш мозг делает это автоматически и без особых видимых усилий. Очевидно, мы должны правильно воспринимать произнесенное слово. Легко неверно истолковать даже одно слово, произнесенное отдельно, особенно в присутствии окружающего шума (гул автомобильного двигателя, уличные звуки, болтовня окружающих людей, бегущая вода в раковине, шум вентиляции здания и т. д.). В таких условиях, близких к низкоконтрастному зрению, функции правдоподобия широки, и одно слово может быть неправильно воспринято за другое, похожее по звучанию. На самом деле неправильно воспринимаемая речь настолько распространена, что люди часто пропускают эти моменты, не задумываясь. Вы можете попытаться составить список таких случаев самостоятельно. Результаты и познавательны, и забавны. Например, в разговорах с другими мы ошибочно принимаем Монголию за магнолию, уточку за удочку, батон за бетон, сетку за ветку и т. д. Как показывают многочисленные примеры, помимо похожего звучания разных слов, проблема восприятия речи возникает из-за того, что паузы между произносимыми словами часто не длиннее пауз между слогами в одном слове. Следовательно, точное определение, где заканчивается одно слово и начинается следующее, – нетривиальная задача. Эта трудность синтаксического анализа может привести к ошибкам, при которых слоги из разных слов неправильно сочетаются в нашем восприятии.

В детстве писательница Сильвия Райт любила слушать популярную в XVII веке шотландскую балладу «Прекрасный граф о’Морей», которую ей часто рассказывала мать. Ей особенно нравились грустные, но красивые строки, описывающие гибель графа и любви всей его жизни, леди Мондегрин:

Вы, горы и долины, –
Где вы были,
Когда враги убили графа о’Морея
И леди Мондегрин (208: 48–51).

Слова, услышанные юной Сильвией Райт, были вовсе не теми, что говорила ее мать, как бы впечатляюще они ни звучали. На самом деле в балладе вообще не упоминается леди Мондегрин. Несчастливого мертвого графа уложили на траву в одиночестве; враги «повергли его на траву» (*laid him on the green*). Творческая, но ошибочная интерпретация услышанной баллады Сильвией Райт отражает ошибку восприятия. Она интерпретировала звуки «laid hi-» как «леди» (*lady*), а «-m on the green» как «Mondegreen». Позднее Сильвия Райт ввела термин «мондегрин» для обозначения неправильно расслышанного слова или фразы [208]. Поскольку разговорному языку вообще присуща фонетическая двусмысленность, примеры мондегринов встречаются на каждом шагу. Когда Квинсленд в Австралии был затоплен тропическим циклоном Таша,

газета *Morning Bulletin of Rockhampton* (6 января 2011 г.) сообщила трагическую новость о том, что в результате наводнения «более 30 000 свиней унесло течением реки Доусон с прошлых выходных» (*More than 30,000 pigs have been floating down the Dawson River since last weekend*) [28]. Эта поразительная история, основанная на интервью между репортером и владельцем местной свинофермы, была чрезвычайно ошибочной. Хозяин упоминал не *30 thousands pigs* (30 000 свиней), а *30 sows and pigs* (30 свиноматок и поросят), унесенных вниз по течению! На следующий день газета опубликовала исправление.

Книги и многие веб-сайты посвящены перечислению любимых мондегринов людей, особенно возникающих из-за неправильно услышанных текстов песен. Очень любопытно и поучительно заглянуть на веб-сайты, где слушатели публикуют свои версии одних и тех же песен, которые они не расслышали. Многочисленность людей, ошибочно услышавших знаменитые строки «пора-пора-порадуемся на своем веку *красавице Игуку*, счастливому клинку», по-видимому, отражают как фонетическую двусмысленность (функция широкого правдоподобия), так и невероятное содержание (низкая априорная вероятность) исходного текста. Что касается априорных вероятностей, предложение «Туалет справа», безусловно, является более распространенным предложением, чем «Восходит ненастная луна», а «подводная лодка» в качестве вида транспорта, возможно, более правдоподобна, чем «летний бриз» (рис. 1.11).

Lucy in the sky with diamonds

(Люси в искрящейся вышине)
– The Beatles

- "Lucy and this guy with diamonds"
(Люси и этот парень с бриллиантами)
- "Lucy in disguise with diamonds"
(Люси по уши в бриллиантах)
- "Lucy in the sky with Simon"
(Люси в небесах с Саймоном)



There's a bad moon on the rise

(Восходит ненастная луна)
– Creedence Clearwater Revival

"There's a bathroom on the right"
(Туалетная комната расположена справа)



The Death of Lady Mondegreen
(Смерть леди Мондегрин)
– Harper's Magazine (Nov. 1954)

And you come to me on a summer breeze

(Тебя принес мне летний ветерок)
– Bee Gees (How Deep Is Your Love)

"And you come to me on a submarine"
(Ты приплыла ко мне на субмарине)

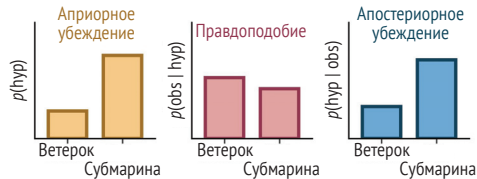


Рис. 1.11 ❖ Мондегрины являются результатом сочетания фонетической неоднозначности (функции широкого правдоподобия) и низкого ожидания реальной фразы, которая поется или произносится (предварительное распределение в пользу «неправильной» гипотезы)

Возникновение мондегринов убедительно свидетельствует о том, что восприятие речи, как и зрительное восприятие, является результатом комбинации функций правдоподобия и априорных распределений. Люди обычно

воспринимают речь точно, но, конечно, случайные ошибки неизбежны. Действительно, «чем неразборчивее исходный текст, тем больше вероятность того, что слушатели услышат то, что хотят услышать» [139]. Перефразируя в терминах байесовского вывода: чем более плоской является функция правдоподобия, тем больше будет влияние априорного распределения на результирующее апостериорное распределение. Таким образом, включение в вывод о восприятии априорного ожидания, которое в большинстве случаев повышает точность восприятия, может иметь неприятные последствия, создавая мондегрини в тех случаях, когда мы сталкиваемся с неожиданным (низкоаприорным) словом, которое звучит как другое, более ожидаемое (высокоаприорное) слово.

Располагая этими знаниями, довольно легко вызвать мондегрини у других людей. Просто выберите два разных слова или фразы, которые звучат одинаково, убедитесь, что у вашего слушателя сформировано априорное убеждение в пользу одного из слов или фраз, а затем произнесите другое слово или фразу со схожим звучанием. Например, вы можете сказать другу: «Знаешь, люди очень хорошо распознают речь; на самом деле мы можем понимать речь намного лучше, чем даже самые лучшие компьютерные программы. Мы подсознательно знаем, как разломать печь. Что я только что сказал? Мы подсознательно знаем, как...?» Если вы произнесли слова «разломать печь» естественным тоном и вам удалось не выделить их на фоне остальной речи, ваш друг, вероятно, услышал «распознать речь», а не слова, которые вы на самом деле произнесли. С байесовской точки зрения функция широкого правдоподобия, с которой столкнулся ваш друг, будет сочетаться с четким априорным распределением (учитывая предыдущее содержание вашего дискурса) в пользу гипотезы «распознать речь». Более детально этот пример рассмотрен в [177].

Даже когда слушатели правильно воспринимают каждое слово, они сталкиваются с последней проблемой: определить предполагаемое значение цепочки слов. Опять же, это часто требует оценки нескольких гипотез. Предположим, кто-то сказал вам, что некий мост «держится за счет тросов и опор». Утверждения, подобные этому, даже если их хорошо услышать или прочитать на печатной странице, тем не менее согласуются с двумя или более интерпретациями [39]; то есть такие утверждения вызывают широкую функцию правдоподобия, обусловленную не фонетической двойственностью, а неоднозначностью толкования смысла (этот мост прочный или наоборот?); существует также более прямая семантическая двусмысленность (рис. 1.12). В других случаях сама структура предложения неоднозначна, что называется синтаксической двусмысленностью. Когда мы слышим или читаем двусмысленное предложение, мы естественным образом комбинируем функции правдоподобия с априорным распределением и обычно достигаем правильного восприятия. Мы, однако, иногда сбиты с толку – и удивлены – на мгновение, когда обе интерпретации приходят нам на ум. Это пришло в голову одному из авторов, когда приятель рассказал ему о путешествии по лесу, которое он совершил со своими родителями. «Повсюду была дикая природа, – воскликнул он. – Однажды я увидел медведя вместе с моей мамой» (рис. 1.13). Хотя эта книга не посвящена подобным ситуациям неоднозначности, мы отмечаем здесь сам факт их существования, чтобы проиллюстрировать, что неопределенность и умозаключения играют роль на многих уровнях перцептивной и когнитивной обработки.



(reservations означает как «оговорки (отступления) в речи», так и «сомнения, подозрения» – прим. перев.)

(net worth означает как «улов, попавший в сеть», так и «чистый бухгалтерский доход» – прим. перев.)

Рис. 1.12 ❖ Семантическая неоднозначность языка

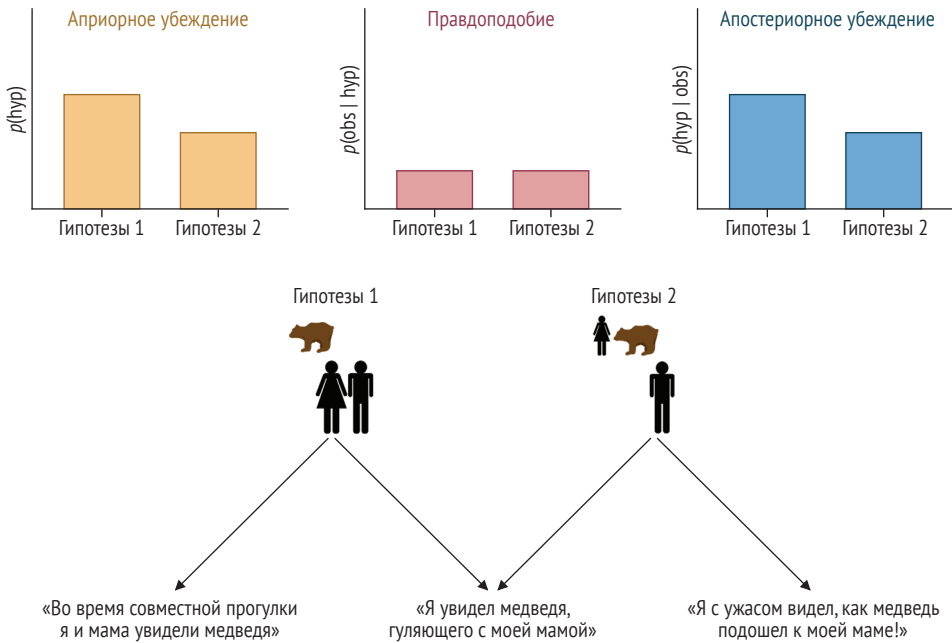


Рис. 1.13 ❖ Перцептивный вывод в условиях синтаксической неоднозначности. Каждое состояние мира (гипотеза) может быть описано многими различными способами, три из которых показаны на рисунке. Докладчик выбрал выражение, которое может описывать два состояния мира: «я увидел медведя вместе с моей мамой». Базовые знания человека предполагают, что медведи с меньшей вероятностью будут прогуливаться рядом с людьми, чем будут замечены на расстоянии, поэтому априорное распределение благоприятствует гипотезе 1. Функция правдоподобия показывает, что произносимое предложение имеет примерно одинаковую вероятность при двух гипотезах. Следовательно, апостериорное распределение благоприятствует гипотезе 1

1.6. Исторический обзор: восприятие как бессознательное умозаключение

Правило Байеса названо в честь английского министра и математика Томаса Байеса (1702–1761), который интересовался проблемами обратной вероятности, главным образом тем, как вычислить $p(B|A)$, когда известны $p(A)$ и $p(A|B)$. В книге Байеса *An Essay towards Solving a Problem in the Doctrine of Chances* («Опыт решения задачи доктрины вероятностей»), опубликованной посмертно в 1763 г., были заложены основы исчисления условной вероятности – области статистических рассуждений, которая сейчас называется байесовским выводом. Позднее правило Байеса независимо вывел французский математик и физик Пьер-Симон маркиз де Лаплас (1749–1827). Лаплас чрезвычайно продуктивно применил эту формулу к задачам в широком диапазоне дисциплин. Важно отметить, что Лаплас также признал вездесущность вероятности, заявив, что «наиболее важные вопросы жизни... на самом деле, по большей части, являются только проявлениями вероятности. Строго говоря, почти все наши знания лишь вероятны» [106]. Действительно, сегодня байесовский вывод играет все более важную роль в необычайно разнообразном наборе дисциплин, охватывающих почти все области науки и техники: нейробиологию, психологию, эволюционную и молекулярную биологию, геологию, астрономию, экономику, робототехнику и компьютерные науки, и это лишь малая их часть. Однако идея о том, что восприятие является формой бессознательного вывода, возникла независимо от Байеса и Лапласа. Это заслуга нескольких ученых. Древний арабский физик и эрудит Ибн аль-Хайсам (965–ок. 1040) прозорливо отметил, что «не все, что воспринимается зрением, воспринимается грубым ощущением; вместо этого многие зримые характеристики будут восприняты посредством суждения... в сочетании с ощущением видимой формы». Так, «знакомые видимые объекты воспринимаются зрением через определяющие признаки и через предшествующие знания» [10]. Значительно позже немецкий врач и физик Герман фон Гельмгольц (1821–1894) снова высказал мысль о том, что восприятие есть форма бессознательного умозаключения, красноречиво заявив, что «предыдущий опыт действует в сочетании с наличными ощущениями, чтобы произвести перцептивный образ» [190]. Идеи аль-Хайсама и Гельмгольца прекрасно согласуются с мнением о том, что восприятие является формой байесовского вывода (рис. 1.14). Однако трудно установить, является ли та или иная форма перцептивного вывода сознательной или бессознательной, и мы не будем рассматривать этот вопрос в нашей книге.



Томас Байес, 1702–1761



Ибн аль-Хайсам, 965 – ок. 1040

«Знакомые видимые объекты воспринимаются зрением через определяющие признаки и через предшествующие знания».

– *Книга оптики*



Пьер-Симон Лаплас, 1749–1827

«Наиболее важные вопросы жизни... на самом деле, по большей части, являются только проявлениями вероятности. Строго говоря, почти все наши знания лишь вероятны».

– *Философские размышления о вероятности*



Герман Людвиг фон Гельмгольц, 1821–1894

«**Предыдущий опыт** действует в сочетании с **наличными ощущениями**, чтобы произвести **перцептивный образ**».

– *Трактат по физиологической оптике*

$$\frac{p(\text{hyp} | \text{obs})}{\text{Апостериорная вероятность}} \propto \frac{p(\text{obs} | \text{hyp})}{\text{Правдоподобие}} \cdot \frac{p(\text{hyp})}{\text{Априорная вероятность}}$$

Рис. 1.14 ❖ Авторитетные ученые, развивавшие байесовский вывод и взгляды на то, что восприятие является бессознательным умозаключением. Стоит отметить, что список не ограничивается только ими, просто история сохранила для нас не все имена

1.7. Заключение

В этой главе мы представили концепцию, согласно которой восприятие по своей сути является вероятностным и как таковое оптимально характеризуется как процесс байесовского вывода. Что касается байесовского вывода, вы узнали следующее:

- условные вероятности, такие как $p(A|B)$, представляют вероятность A при заданном B . В байесовском выводе о восприятии A и B обычно представляют состояние мира и наблюдение;
- функция правдоподобия $p(\text{наблюдение} | \text{состояние мира})$ отражает информационное содержание сенсорного наблюдения, относящееся к различению одного состояния мира от другого;

- чем более плоской является функция правдоподобия, тем меньше полезной информации мы получаем от наших органов чувств. Если функция правдоподобия совершенно плоская, то наблюдатель ничему не научился из наблюдения;
- в некоторых случаях, например в примере вывода «Это мой друг?», вероятность меняется со временем;
- априорное распределение вероятностей по состояниям мира $p(\text{состояние мира})$ суммирует информационное содержание наших прошлых наблюдений, т. е. исходные знания, которые у нас есть о мире. Восприятие основано не только на сенсорных наблюдениях, но и на ожиданиях, основанных на предыдущем опыте;
- более плоские априорные распределения означают, что мы меньше знаем о потенциальных состояниях мира;
- согласно правилу Байеса, апостериорная вероятность каждого гипотетического состояния мира $p(\text{состояние мира} | \text{наблюдение})$ вычисляется, исходя из вероятностей и априорных знаний о состояниях мира;
- процедуры байесовского вывода в равной степени применимы к ситуациям, в которых гипотетические состояния мира дискретны или непрерывны;
- сценарии восприятия, будь то зрение, слух или другие чувства, подвержены разным уровням неопределенности;
- восприятие речи сопряжено с фонетической и синтаксической двусмысленностью, что часто приводит к возникновению функций плоского правдоподобия. Комбинация априорных значений и вероятностей может привести к неверным интерпретациям, таким как мондегрини.

1.8. Рекомендуемая литература

- Alhacen. *Alhacen's Theory of Visual Perception: A Critical Edition, with English Translation and Commentary, of the First Three Books of Alhacen's "De Aspectibus", the Medieval Latin Version of Ibn Al-Haytham's "Kitāb Al-Manāz. ir.* Edited by A. Mark Smith. Vol. 1. Philadelphia, PA: American Philosophical Society, 2001.
- Thomas Bayes. *An Essay towards Solving a Problem in the Doctrine of Chances.* Biometrika 45, no. 3–4 (1958): 296–315.
- Peter Brugger and Susanne Brugger. *The Easter Bunny in October: Is It Disguised as a Duck?* Perceptual and Motor Skills 76, no. 2 (1993): 577–578.
- Daniel Burdon. *Pigs Float down the Dawson.* Morning Bulletin, February 9, 2011. <https://realtegan.blogspot.com/2011/02/best-correction-ever.html>.
- Gloria Cooper. *Red Tape Holds up New Bridge, and More Flubs from the Nation's Press.* New York: TarcherPerigee, 1987.
- Wilson S. Geisler and Randy L. Diehl. *A Bayesian Approach to the Evolution of Perceptual and Cognitive Systems.* Cognitive Science 27, no. 3 (2003): 379–402.

- Gary Hatfield. *Perception as Unconscious Inference*. In *Perception and the Physical World: Psychological and Philosophical Issues in perception*. Chichester: John Wiley, 2002.
- Hermann von Helmholtz. *Treatise on Physiological Optics*. Edited by James P. C. Southall. Vol. 3, *The Perceptions of Vision*. New York: Optical Society of America, 1925.
- Pierre-Simon Laplace. *A Philosophical Essay on Probabilities*. Edited and translated by Andrew I. Dale. New York: Springer Science and Business Media, 2012.
- Mohamed A. F. Noor, Robin S. Parnell, and Bruce S. Grant. *A Reversible Color Polyphenism in American Peppered Moth (*Biston betularia cognataria*) Caterpillars*. *PLoS One* 3, no. 9 (2008): e3142.
- Pamela Licalzi O’Connell. *Sweet Slips of the Ear: Mondegreens*. *New York Times*, April 9, 1998.
- Russell Smith. *Milk Drinkers Turn to Powder and Other Pun-ishing Headlines*. *Globe and Mail* (September 23).
- Stephen M. Stigler. *Who Discovered Bayes’ Theorem?* *American Statistician* 37, no. 4a (1983): 290–296.
- J. V. Stone, I. S. Kerrigan, and J. Porrill. *Where Is the Light? Bayesian Perceptual Priors for Lighting Direction*. *Proceedings of the Royal Society B: Biological Sciences* 276, no. 1663 (2009): 1797–1804.
- Dave Tompkins. *How to Wreck a Nice Beach: The Vocoder from World War II to HipHop, the Machine Speaks*. New York: Melville House, 2011.
- Michel Treisman. *Motion Sickness: An Evolutionary Hypothesis*. *Science* 197, no. 4302 (1977): 493–495.
- Sylvia Wright. *The Death of Lady Mondegreen*. *Harper’s Magazine* 209, no. 1254 (1954): 48–51.

1.9. Задачи

Задача 1.1. Перефразируйте с точки зрения байесовского вывода восприятия следующее высказывание Ибн аль-Хайтама, записанное примерно 1000 лет назад: «Когда взгляд увидит среди цветов в каком-нибудь саду розовый цвет, мы немедленно приходим к заключению, что это розы, потому что такой цвет присущ только розам. Но этого не происходит, когда зрение воспринимает в саду миртово-зеленый цвет. Ибо, когда мы видим в саду только миртовую зелень, мы не приходим к выводу, что это мирт, просто из восприятия зелени, потому что многие растения зеленые, а кроме того, некоторые растения похожи на мирт по оттенку и форме» (*Alhacen, De aspectibus*, книга 2, *Alhacen’s Theory of Visual Perception*, пер. Смита).

Задача 1.2. Почему мы называем себя в самом начале телефонного разговора даже с уже знакомыми людьми, но не делаем этого при личной встрече? Изложите свой ответ в рамках байесовского вывода о восприятии.

Задача 1.3. Когда собеседник говорит тихо или когда разговор происходит в присутствии значительного окружающего шума, мы иногда затыкаем уши

и/или внимательно смотрим на губы говорящего. Почему, с точки зрения байесовского восприятия, мы это делаем?

Задача 1.4. Чтобы понять, как шумная обстановка порождает неопределенность, рассмотрим слово *лифт*. Предположим, вы видите это слово написанным (или слышите, как оно произносится) с пропущенной буквой *л* (что делает ее неизвестной): *_ифт* (например, из-за окружающего слухового шума). Перечислите все слова, которые совместимы с вашим наблюдением. Теперь рассмотрим случай, когда отсутствуют буквы *л* и *ф*: *_и_т*. Каков эффект отсутствия *л*, *ф* и двух букв одновременно с точки зрения условных вероятностей, имеющих отношение к восприятию?

Задача 1.5. Фонетический алфавит НАТО, используемый многими военными, морскими и другими организациями при радиосвязи, представляет каждую букву словом: А (альфа), В (браво), С (Чарли), D (дельта), Е (эхо), F (фокстрот) и т. д. Для какой цели это служит в радиосвязи? Объясните применительно к условным вероятностям. В частности, рассмотрим радиосвязь в условиях значительного фоновых шума, в которой отправитель желает произнести слово FACE (лицо). Сравните вероятности $p(\text{слуховой сигнал, слышимый получателем} | \text{слово FACE, произнесенное отправителем})$ и $p(\text{слуховой сигнал, слышимый получателем} | \text{другое слово, такое как RACE, произнесенное отправителем})$, когда отправитель использует обычный алфавит, и еще раз, когда отправитель использует фонетический алфавит.

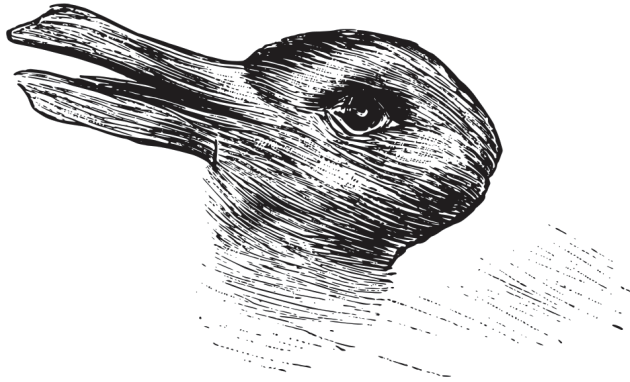
Задача 1.6. Англоязычные носители иногда неправильно воспринимают английские слова, когда слушают песни, исполняемые на иностранном языке, с которым они незнакомы, а также воспринимают слова в музыке, воспроизводимой задом наперед. Дайте байесовское объяснение этим явлениям.

Задача 1.7. Иногда стоит нажать кнопку вызова лифта, как кабина лифта сразу приходит в движение (это видно на дисплее, показывающем текущий этаж кабины). Докажите, что в этой ситуации вероятность гипотезы о том, что внутри никого не окажется, когда она прибудет, намного выше, чем вероятность гипотезы о том, что внутри окажутся люди. Априорные знания в этой задаче не учитываются.

Задача 1.8. Предположим, вы видите кого-то, с кем незнакомы, лишь мельком взглянув на него с расстояния около 10 м. Если вы заинтересованы в оценке возраста этого человека, как бы вы поступили? Какие факторы, помимо внешности человека, могли бы повлиять на вашу оценку? Предоставьте байесовское описание ваших рассуждений. Как часть вашего ответа нарисуйте примеры вашей функции правдоподобия, априорного распределения и результирующего апостериорного распределения.

Задача 1.9. Исследовательская статья под названием «Пасхальный кролик в октябре: он маскируется под утку?» поясняет, что «о внешнем виде пасхального кролика в его нерабочие дни известно очень мало» [26]. В ходе исследования авторы показали «неоднозначный рисунок утки/кролика (без указания авторства из номера *Fliegende Blatter* от 23 октября 1892 г.) 265 испытуемым в пасхальное воскресенье и 276 испытуемым в воскресенье в октябре 1992 г.

В том же году авторы сообщают: «Если на Пасху рисунок значительно чаще распознавался как кролик, то в октябре большинство испытуемых считали его птицей». Рисунок, показанный авторами в своем исследовании, был похож на следующий:



Кролик или утка, Fliegende Blätter,
выпуск Fliegende Blätter от 23 октября 1892 г.

Дайте байесовское объяснение результатов, полученных авторами.

Задача 1.10. На изображениях ниже показана полая маска для лица, постепенно поворачиваемая от вида сбоку (слева) до вида с изнанки (справа).

Правое изображение выглядит как обычное выпуклое лицо, хотя на самом деле это полая (вогнутая) сторона маски. Это называется иллюзией полого лица, или иллюзией полый маски. Дайте байесовское объяснение этой иллюзии.



Задача 1.11. Приведите три примера из повседневной жизни (перцептивных или когнитивных), в которых вы пытались сделать вывод о состоянии мира на основе неполной или несовершенной информации. Для каждого примера укажите наблюдение(я), интересующее состояние мира и источник(и) неопределенности.

Задача 1.12. Мишель Трейсман попытался объяснить явление укачивания (морскую болезнь) в контексте эволюции ([179], Motion Sickness, 493–495). В течение миллионов лет, на протяжении которых развивался человеческий мозг, случайное употребление токсичной пищи могло вызвать галлюцинации в виде раскачивания и вращения окружающей обстановки, а последующие за этим тошнота и рвота были всего лишь защитной реакцией организма, избавляющегося от токсинов. Возможно, наш современный мозг все еще использует априорные вероятности, генетически унаследованные от тех дней; просто они основаны не на нашем личном опыте, а на опыте наших предков! Здесь мы не углубляемся в достоинства этой теории, а пытаемся представить ее в байесовской форме. Предположим, вы находитесь в комнате без окон на корабле в море. Вашему мозгу доступны два набора сенсорных наблюдений: зрительные и вестибулярные. Предположим, что мозг рассматривает три гипотезы того, что вызвало эти наблюдения:

- **гипотеза 1:** комната не движется, и ваше движение в комнате вызывает оба набора наблюдений;
 - **гипотеза 2:** ваше движение в комнате вызывает ваши зрительные наблюдения, тогда как ваше движение в комнате и движение комнаты в мире вместе вызывают вестибулярные наблюдения;
 - **гипотеза 3:** у вас галлюцинации – ваше движение в комнате и проглоченные токсины вместе вызывают оба набора наблюдений.
- (a) В доисторические времена окружающий мир почти никогда не двигался. Время от времени человек может случайно проглотить токсины. Предполагая, что ваши врожденные априорные знания основаны на этих доисторических частотах событий, нарисуйте столбчатую диаграмму, чтобы представить ваши априорные вероятности трех приведенных выше гипотез. Числовые данные не нужны.
- (b) В комнате без окон на корабле наблюдается большое расхождение между вашими зрительными и вестибулярными наблюдениями. Нарисуйте столбчатую диаграмму, иллюстрирующую вероятность трех гипотез в этой ситуации (т. е. насколько вероятны эти конкретные сенсорные наблюдения при каждой гипотезе). Числовые данные не нужны.
- (c) Нарисуйте столбчатую диаграмму, иллюстрирующую апостериорные вероятности трех гипотез. Числовые данные не нужны.
- (d) Объясните, используя апостериорные вероятности, почему вас может стошнить в этой ситуации.