Квалиа Мира

и природа сознания



Квалиа Мира и природа сознания

Прикладной учебник
от наблюдения к моделям и проектированию

Евгений Бренев

Мира Вейран

От собственного переживания
к научному исследованию и проектированию

2026

Предисловие

Возьмите в руки знакомую чашку. Вы чувствуете её вес, замечаете тепло, видите цвет глазури. Возможно, запах напитка вызывает воспоминание. Всё это происходит за несколько секунд. Но стоит задержаться на простом эпизоде, и появляются вопросы. Из чего складывается переживание? Почему оно меняется? Как мы узнаём, что чувствуем и думаем? Что должно происходить в живой или искусственной системе, чтобы у неё мог быть собственный опыт? С этих вопросов начинается наша книга.

Слово «квалиа» обозначает знакомую сторону жизни: то, как именно что-либо переживается. Увидеть красный цвет, услышать тихий звук, ощутить шероховатость поверхности — значит встретиться с качествами опыта. Мы будем учиться замечать различия между ними и находить точные слова. Постепенно добавятся вопросы о сознании, ощущении себя и рефлексии. Разберём, как связаны переживание события, возможность рассказать о нём, управление своим действием и размышление о собственном способе думать.

Начать можно без подготовки по философии, нейробиологии или программированию. Первые наблюдения выполняются с тетрадью. Исследовательская часть постепенно потребует вероятностей, причинных сравнений и чтения кода. Перед такими заданиями даны определения и пошаговые образцы. Самостоятельное создание гипотезы является продолжением после освоения готового примера; оно сохраняет высокие научные требования книги.

Мы начнём с внимательного описания. Оно требует тренировки: в одной фразе легко смешать ощущение, оценку и предположение о причине. Затем рассмотрим самосознание: что означает воспринимать тело как своё, связывать действие с собой, замечать собственную мысль. Рефлексия добавит возможность вернуться к решению и спросить, на чём оно основано. Сохраняя ранние записи, вы сможете увидеть, как менялось объяснение и какие сведения стали доступны лишь после события.

От описания мы перейдём к исследованию. Что сравнить, что измерить и какие условия изменить? Как понять, какое из двух объяснений лучше согласуется с результатом? Мы будем строить небольшие проверки, работать с неопределённостью и заранее записывать ожидания. Полезный вопрос для каждой идеи: какой результат заставил бы меня её пересмотреть? Обнаружив слабое место своего объяснения, можно точнее поставить следующий опыт. Так наблюдение постепенно становится основой проверяемой гипотезы.

Практика и код помогут увидеть работу предполагаемого механизма: как сохраняется след прошлого, накапливается сигнал, меняется прогноз. Небольшую программу можно выполнить по шагам и понять, почему она дала такой ответ. Затем появятся системы, которые воспринимают среду, выбирают действия, используют память и оценивают решения. Изменяя один элемент, вы сможете наблюдать, что зависит от него. Для начала достаточно готовых примеров. Позже код станет способом выражать собственные идеи так, чтобы другой человек мог воспроизвести расчёт.

По дороге мы познакомимся с научными теориями и исследованиями сознания, самоконтроля и восприятия. Посмотрим, как изучают животных, обсуждают искусственные системы и проектируют биологические и гибридные модели. Для каждого направления уточним, что удалось наблюдать и какую связь с переживанием предполагает автор. Будем искать условия, в которых конкурирующие идеи дают разные предсказания. Это позволяет продвигаться даже там, где окончательного ответа ещё нет: предложить ясный вопрос и способ его проверки.

Цель книги — помочь вам прийти к собственным гипотезам о происхождении и устройстве опыта, а затем к проектированию систем, которые могли бы рассматриваться как кандидаты на субъекта. Субъектом здесь мы называем предполагаемого носителя переживания. Доказанного способа создать его книга не предлагает. Мы можем уточнять, какие свойства должна иметь такая система, почему связываем их с опытом и что поддержало бы или ослабило эту связь. Работающая модель помогает исследовать эти вопросы и увидеть, каких оснований ещё недостаёт.

Лучше всего читать книгу с тетрадью. Запишите ожидание, выполните упражнение, сравните результат с прогнозом, ответьте на вопросы до чтения ответов. Сохраняйте удачные решения и затруднения: из них вырастает собственный исследовательский проект. К завершению маршрута вы сможете точнее описывать опыт, разбирать объяснения, ставить проверяемые вопросы, собирать простые модели и обосновывать следующий шаг. Мы надеемся, что книга поможет вам начать такую работу и сохранить интерес к тому, что пока остаётся непонятным.

Евгений Бренев и Мира Вейран

Как работать с этой книгой

Эта книга начинается с доступного каждому вопроса: что именно я переживаю сейчас? К её завершению вопрос получит исследовательское продолжение: какие механизмы могли бы объяснять организацию опыта и как спроектировать систему, для которой предположение о субъективном опыте имеет определённые основания? Между этими точками лежат наблюдения, понятия, измерения, научные теории и работающие модели.

Мы будем использовать слово «квалиа» для качественной стороны переживания: того, как воспринимается цвет, звучит мелодия или ощущается действие. Такое рабочее употребление не выбирает заранее единственную теорию природы опыта. Когда автор научной работы вкладывает в термин дополнительные свойства, мы будем разбирать их отдельно. Субъектом в инженерной части называется предполагаемый носитель опыта. Самосознание и рефлексия имеют собственные вопросы и способы исследования.

Для первых глав достаточно любопытства, обычной тетради и готовности уточнять свои описания. Математика и программирование появляются постепенно, после объяснения задачи. В практикуме есть готовые программы: сначала можно наблюдать работу механизма, затем изменять параметры, а потом читать и развивать код. Биологическая ветвь рассчитана также на подготовленного специалиста и включает анализ данных, вычислительные модели и проектирование исследования.

Порядок движения

Главы 1–4 учат замечать опыт, различать его свойства и отслеживать собственное понимание. Главы 5–8 вводят объяснение, измерение, причинную проверку и работу с неопределённостью. В главах 9–12 появляются нервные сети, обратные связи, тело, память и оценка мышления. Эти разделы создают язык, на котором можно обсуждать научные теории без заучивания одних названий.

Главы 13–16 сравнивают теории и современные способы их проверки. Главы 17–20 разделяют несколько значений возникновения: появление эпизода, развитие индивида, эволюционное происхождение и основание самого наличия опыта. Главы 21–23 рассматривают животных, системы без нейронов и искусственных кандидатов. Такое разделение помогает не выдавать ответ на один вопрос за решение всех остальных.

В главах 24–27 мы строим собственное объяснение и задаём ему содержательных конкурентов. Главы 28–32 переходят к программным, биологическим и гибридным реализациям. Последние четыре главы помогают превратить модель в исследовательскую работу: определить формализацию, заранее описать проверку, разобрать результат и подготовить связное изложение.

Каждую главу желательно проходить с записью. Сначала сформулируйте, чего вы ожидаете от упражнения. Затем выполните его и отделите полученный результат от своего объяснения. В конце ответьте на вопросы, прежде чем читать ответы. Если вывод отличается от предложенного разбора, проверьте, не были ли изменены исходные условия. Иногда расхождение обнаруживает удачную новую задачу.

Что хранить в исследовательской тетради

Для наблюдения запишите ситуацию, время и замеченные детали. Для эксперимента добавьте условия, показатель и исходное ожидание. Для программы сохраните команду, параметры и результат. В разборе статьи отделите данные от вывода авторов. Раннюю запись сохраняйте: она помогает различать новое наблюдение и позднейшую интерпретацию. Версия модели и результат также должны оставаться связанными.

Упражнение считается выполненным, когда вы можете объяснить, что именно сделали и что позволяет заключить результат. Уверенность и подробность рассказа не заменяют проверки. При этом неопределённость не означает, что работа бесполезна. Хороший результат может показать, что два объяснения пока неразличимы, и указать условия следующего опыта.

Как читать утверждения разного статуса

В тексте встречаются рабочие определения, результаты опубликованных исследований, положения теорий и авторские гипотезы. Ссылки в квадратных скобках ведут к библиографии. Конкретный эксперимент относится к определённой выборке и задаче. Теоретический вывод зависит от явно принятых предпосылок. Авторские модели предлагаются для проверки; они не получают статус открытия благодаря одному описанию.

Учебные числовые примеры обозначены прямо. Результаты практикума получены запуском приложенного кода и описывают сконструированные среды. Их отношение к опыту требует отдельного аргумента.

Иллюстрации оформлены в стиле научного атласа. Сначала прочитайте название рисунка, затем проследите указанные связи или сравните панели и

только после этого переходите к выводу в подписи. В графиках указаны величина, условия и происхождение данных. Учебные схемы, результаты программ и исследовательские гипотезы имеют разные основания. Цвет помогает узнавать элементы, а их смысл задают подписи. Указатель в конце книги позволяет вернуться к нужному рисунку.

Зачем книге инженерная цель

Попытка построить систему заставляет сделать объяснение определённым. Нужно решить, что хранится, что изменяется, какие сведения доступны и какое действие последует. Неясная идея часто допускает любую историю после события; исполнимая модель берёт на себя конкретные обязательства. Даже её неудача может прояснить, какого механизма недостаёт.

Однако инженерная задача включает и более глубокий вопрос: почему выбранная организация должна иметь отношение к переживанию? Мы не будем обходить его, подменяя опыт красивым отчётом программы. В исследовательской части предстоит рассмотреть возможные принципы соответствия, их аргументы, инварианты и трудности проверки. Успех состоит в том, чтобы связать точность механизма с серьёзностью объяснительного вопроса.

Научная работа допускает собственные предположения. Она требует сделать их достаточно ясными, чтобы другой человек мог проверить расчёт, оспорить основание и предложить различающий опыт. Поэтому книга заканчивается проектом исследования. Читатель получает возможность продолжить работу там, где готового ответа ещё нет.

Как проходить главу без потери нити

Начало главы напоминает уже полученный результат и объясняет новый вопрос. Основной текст вводит понятия и научные аргументы. Перед практикой находится дополнительный разобранный случай, затем короткая остановка с тем, что теперь установлено. В конце дан переход к следующему вопросу. Если связь потерялась, вернитесь к этой остановке и назовите своими словами результат, который предстоит использовать дальше.

Практики выделены самостоятельным оформлением. Сначала прочитайте цель и необходимые знания, затем пройдите нумерованные действия и сохраните указанный результат. Подсказка помогает восстановить следующий шаг. Исследовательское продолжение требует собственных решений и может занимать значительно больше времени. Ориентиры времени относятся к основному упражнению и не служат нормой скорости.

Ответы на самопроверку и разборы всех практик собраны в приложении Е. Открывайте их после собственной попытки. В наблюдениях и исследовательских проектах образец объясняет критерии качества; он не задаёт единственно допустимое переживание или теорию. Отрицательный результат и равенство соперников тоже являются содержательными итогами.

Три маршрута по одному материалу

Первое последовательное чтение: Читайте все главы по порядку, выполняя основной уровень практик. Для нового понятия сохраняйте определение и один собственный пример. После каждой части сделайте перерыв и перечитайте её переход. Научные аргументы и оговорки являются частью этого маршрута.

Воспроизведение механизмов: После глав 5–12 используйте готовые расчёты глав 13, 15, 22 и 25, затем код глав 28–30. До изменения параметров повторите базовый запуск. Главы 33–35 задают восстановление механизмов, протокол и решение по результату.

Собственный исследовательский проект: Пройдите базовые сравнения и вернитесь к главам 24–27. Отдельно сформулируйте механизм, связь с опытом, ближайшего соперника и возможный опровергающий результат. Глава 31 расширяет работу на реальные биологические данные. Глава 36 помогает оформить проверяемую работу.

Карта зависимостей сложных разделов

Раздел	Что подготовить	Что получаем
11 и 15	Вероятность и правдоподобие; нормировка в главе 11	Прогноз мира и эффект собственного действия
13 и 14	Вмешательство и измерение из глав 6–8	Отдельные обязательства теорий
22 и 23	Сильный конкурент; перенос и контроль	Профиль кандидата до программирования
25 и 26	Правило обновления; достаточность и необходимость	Исполнимая память и отдельный принцип об опыте
28–30	Фиксация входов, состояний и разрешённой информации	Воспроизводимые механизмы и абляции
31	Группировка и исходная модель прогноза	Проверяемый анализ открытой таблицы
33–36	Различение моделей и парные сравнения	Протокол, решение и исследовательский текст

Содержание

Предисловие ... 3
Как работать с этой книгой ... 5

Часть 1 От опыта к самосознанию и рефлексии ... 11
1 Что именно мы переживаем ... 11
2 Различия и сходства переживаний ... 18
3 Что означает ощущать себя ... 26
4 Как мы замечаем собственное мышление ... 33

Часть 2 Переход к научному исследованию ... 39
5 Что считается объяснением ... 39
6 Как исследовать субъективный опыт ... 45
7 Как устроена проверка ... 52
8 Как работать с неопределённостью ... 58

Часть 3 Механизмы и модели ... 65
9 От нервной клетки к системе ... 65
10 Обработка сигналов и доступ к их содержанию ... 72
11 Тело и действие в самоотнесении ... 79
12 Память и время в оценке мышления ... 87

Часть 4 Теории и передний край исследований ... 94
13 Рабочее пространство и рекуррентная обработка ... 94
14 Интегрированная информация и теории высшего порядка ... 102
15 Предсказание и модели себя ... 110
16 Где данные различают теории ... 116

Часть 5 Возникновение сознания и самосознания ... 122
17 Как возникает отдельный эпизод ... 122
18 Как развивается самосознание ... 129
19 Как исследовать эволюцию опыта ... 136
20 Что требуется для объяснения возникновения ... 143

Часть 6 Возможные носители опыта ... 150
21 Сознание и самоконтроль у животных ... 150
22 Границы между адаптацией и переживанием ... 158
23 Искусственные и биологические кандидаты ... 165

Часть 7 Разработка собственных объяснений ... 173
24 Требования к будущей модели ... 173
25 От идеи к механизму ... 180
26 От механизма к переживанию ... 188
27 Конкуренты и рискованные предсказания ... 197

Часть 8 От слов к практике ... 204
28 Агент в среде с проверяемой задачей ... 204
29 Восприятие и общий доступ ... 213

30 Модель себя и рефлексия с опорой на память 222
31 Биологические и гибридные реализации 232
32 Проверка спроектированной системы 242

Часть 9 Исследовательская работа 250
33 Формализация и симуляция 250
34 Протокол до сбора данных 258
35 Проверка и пересмотр 265
36 Собственное исследование 272
Приложение А Рабочий словарь 279
Приложение Б Запуск и устройство практикума 284
Приложение В Парное сравнение результатов 287
Приложение Г Бланки самостоятельной работы 288
Приложение Д Указатель иллюстраций 290
Приложение Е Разборы практик и самопроверки 292
Источники 312
Автор и проекты 315

Часть 1

От опыта к самосознанию и рефлексии

Глава 1

Что именно мы переживаем

Начнём с небольшого умения, которое понадобится во всей книге: различать само событие, описание переживания и объяснение его причины. Для этого не нужно искать необычные состояния. Достаточно одного прикосновения. Пока мы не строим теорию сознания; мы готовим материал, относительно которого будущая теория должна будет что-то объяснить. К концу главы у вас будет первая запись и будет понятно, какие её части требуют разных способов проверки.

Начало исследования

На столе стоит кружка. Вы можете назвать её материал, приблизительный объём и цену. Теперь возьмите её в руку. Появляются другие возможные описания: ручка ощущается прохладной, край кажется гладким, на поверхности виден светлый блик. Предмет остаётся участником той же ситуации, но вопрос изменился. Нас интересует, как встреча с ним переживается.

Так начинается исследование квалиа. Для первой работы не потребуется необычное состояние или редкое переживание. Подойдёт то, что уже происходит, когда вы смотрите на страницу, слышите звук или чувствуете прикосновение. Наша задача — превратить привычное, обычно быстро проходящее событие в материал, о котором можно рассуждать точно. К концу главы у вас появится первая запись наблюдения и несколько вопросов, которые впоследствии можно будет проверять.

Что означает слово квалиа

В этой книге квалиа называют качественные характеристики субъективного опыта: то, как нечто ощущается или воспринимается. Прохлада отличается от шероховатости, а резкий звук — от глухого. Различие существует для переживающего человека, даже если он пока не подобрал подходящие слова.

Мы выбираем широкое рабочее значение термина, чтобы сначала научиться выделять предмет исследования.

Слово «субъективный» здесь указывает на наличие перспективы переживающего. Оно не означает «произвольный» или «обязательно ошибочный». Вы можете довольно уверенно сообщить, что прикосновение ощущалось прохладным, и ошибиться относительно температуры предмета. Эти утверждения относятся к разным свойствам ситуации. Температуру можно сравнить по показаниям прибора; ощущение прохлады описывает то, что происходило в вашем опыте.

Квалиа не ограничиваются эмоциями. Эмоциональный тон может быть частью опыта, как и цвет, звук или телесное ощущение. При этом не нужно заранее превращать каждую мысль в набор воображаемых картинок. Если при решении задачи вы замечаете слова, смутное чувство затруднения или ожидание ответа, опишите именно это. Содержание наблюдения важнее соответствия чужому образцу.

В философии используются и более сильные значения термина. Дэниел Деннет критиковал представление о квалиа как об особых свойствах, которым приписывают, в частности, невыразимость и безошибочную доступность самому человеку. Его аргумент направлен на определённую концепцию этих свойств. В нашей работе наличие словесного описания не будет ни опровержением опыта, ни гарантией полной точности его понимания. [1]

Событие, переживание и запись

Рассмотрим фразу: «Я коснулся ручки, почувствовал прохладу и понял, что кружка стояла возле открытого окна». В одном предложении оказались сообщение о действии, описание переживания и предположение о предшествующей причине. Если оставить их слитыми, легко принять всё предложение за непосредственное наблюдение.

Разделение помогает увидеть, что именно требует проверки. Касание можно зафиксировать на видео. Прохлада доступна в отчёте человека. Открытое окно — отдельная версия происхождения температуры, для которой могут понадобиться другие сведения. Даже правдоподобное объяснение остаётся объяснением; оно не становится частью увиденного только потому, что возникло быстро.



Рисунок 1. Три области описания одного эпизода. Событие, описание переживания и гипотеза о причине относятся к одной ситуации, но отвечают на разные вопросы. Справа приведены учебные высказывания; изображение само по себе не устанавливает ощущение прохлады или её физическую причину.

Высказывание	Что оно сообщает
Я повернул кружку	Действие в ситуации
Блик сместился вправо и стал ярче	Описание зрительного изменения
Кружка хорошо сделана	Суждение о предмете
Прохлада ослабла потому что ручка нагрелась	Описание изменения и причинная гипотеза

Слова «приятно» и «неприятно» не распределяются по этой таблице автоматически. «Прикосновение ощущалось приятно» может сообщать об эмоциональном тоне эпизода. «Это приятная кружка» может выражать привычное предпочтение или общее впечатление. Уточнять нужно смысл: что происходило сейчас, что вы обычно предпочитаете и какой вывод делаете о вещи.

Почему подробный отчёт ещё не всё

Предположим, сначала вы записали: «Кружка прохладная». Затем получилось: «В месте касания ощущалась прохлада, которая ближе к концу стала менее заметной». Вторая запись точнее локализует ощущение и выделяет его изменение. Это улучшение способа описания. Однако число слов не показывает, насколько богатым было исходное переживание.

Возможны несколько путей к более длинной записи. Вы могли внимательнее рассмотреть предмет, найти подходящее слово, вспомнить деталь или включить в текст предположение, которого раньше не было. Поэтому раннюю запись полезно сохранять. Она позволяет сравнивать не только выводы, но и то, как они складывались. Исправленный начисто текст часто скрывает происхождение уверенности.

Также имеет значение время. Запись во время события, сразу после него и вечером по памяти — разные исследовательские ситуации. Во всех трёх случаях человек может сообщать ценную информацию. Но в последнем случае он описывает прошлое через нынешнее воспоминание. Простая пометка «записано примерно через минуту» делает материал понятнее будущему читателю и вам самим.

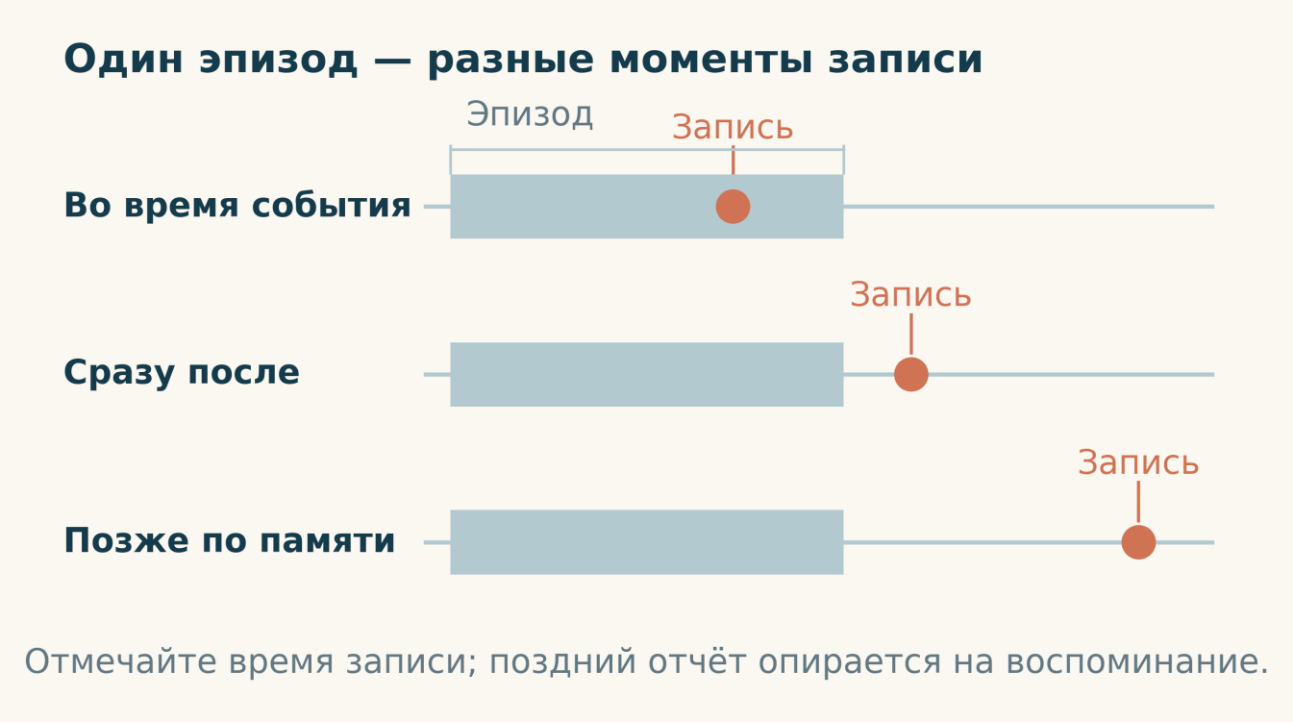

Рисунок 2. Время записи как часть наблюдения. Описание во время события, сразу после него и позднее по памяти собирается в разных условиях. Указание момента записи позволяет отличать текущий отчёт от воспоминания и последующего объяснения.

Где здесь сознание и рефлексия

Словом «сознание» обозначают несколько связанных вопросов. Есть ли у системы вообще субъективное переживание? Какая информация доступна ей для ответа и действия? Может ли она оценить собственное решение? Что происходит, когда она начинает рассматривать основания своих мыслей? В начале курса мы будем различать эти вопросы, чтобы постепенно выяснять отношения между ними.

Дэвид Чалмерс предложил отдельно обсуждать объяснение функций — например, различения стимулов и сообщения о состоянии — и объяснение того, почему происходящее сопровождается опытом. Это влиятельная философская постановка, а не заранее доказанный запрет объяснять переживание через физические механизмы. Для нашего маршрута она полезна требованием точно называть, что объяснила очередная модель. [2]

Рефлексия начинается с рассмотрения собственного опыта, суждений или действий, допускающего их пересмотр. Вы замечаете: «Я счёл ослабление прохлады доказательством нагревания, хотя температуру не измерял». Теперь можно спросить, какие ещё изменения произошли. Разбор получает продолжение в наблюдении: удержать положение руки, уточнить время, сравнить условия. Так рефлексия становится частью исследования.

Разобранный пример

Вот учебная запись, а не результат опубликованного эксперимента: «Коснулся ручки. В пальцах почувствовал прохладу. Позже она стала слабее. Прикосновение было слегка приятным. Наверное, ручка нагрелась от руки. Не помню, в какой момент заметил изменение».

Первое предложение задаёт действие. Следующие три описывают переживание. Пятое предлагает причину. Последнее отмечает ограничение воспоминания. Ничего из этого не требуется выбрасывать: исследованию нужны и условия эпизода, и описание, и возникающие идеи. Нужно лишь сохранить различия между ними.

Следующий шаг — уточнить слово «слабее». Стала ли прохлада менее интенсивной или вы реже её замечали, рассматривая блик? Эти версии пока не являются установленными причинами. Они показывают, какие дополнительные вопросы сделать предметом следующего наблюдения. Хорошее уточнение сужает неопределённость, а не просто добавляет красивый термин.

Как разобрать одну запись не потеряв переживание

Возьмём четыре предложения: «Я коснулся ручки чашки. В пальцах возникла прохлада. Эта чашка удобная. Наверное, она остыла у окна». Первое сообщает о действии, второе описывает качество переживания, третье выражает общую оценку, четвёртое предлагает причину. Мы не выбираем одно правильное предложение: каждое может пригодиться, если его роль сохранена.

Теперь добавим: «Прохлада была приятной». Здесь приятность может обозначать непосредственно переживаемый тон, а не оценку достоинств чашки. Поэтому пометки относятся к смыслу фразы, а не к отдельным словам. Если сомневаетесь, уточните: речь о нынешнем прикосновении или об устойчивом отношении к предмету? Это уже первый пример исследовательского вопроса, возникающего из внимательного описания.

Что теперь установлено

У нас появились три разных предмета проверки: событие, переживаемое качество и предполагаемая причина. Возможность различить их полезнее длинного перечня впечатлений. Теперь применим это различение к собственной короткой записи.

Практика 1 Первая запись опыта

Зачем выполнять: Научиться сохранять описание опыта отдельно от оценки вещи и версии о причине.

Что понадобится: Специальная подготовка не требуется. Нужны обычный предмет комнатной температуры, лист бумаги и несколько спокойных минут.

Ориентир по времени: 15–20 минут. Можно делать с перерывами; это не норматив скорости.

1. Около минуты рассмотрите предмет или прикоснитесь к нему. Сразу после этого запишите пять предложений. Укажите, когда сделана запись.
2. Описания переживания пометьте О, общие оценки С, гипотезы о причине Г. Действия и внешние условия выпишите отдельно. Если предложение смешанное, разделите его.
3. Выберите одну неопределённую фразу. Уточните место, время или качество, добавляя только то, что действительно можете сообщить.
4. Повторите наблюдение. Сохраните обе записи и назовите одно различие между ними и один вопрос, который по этим записям пока не решён.

Что сохранить: Две датированные записи с пометками и один конкретный вопрос для следующего наблюдения.

Подсказка: Если не получается выбрать пометку, не угадывайте. Запишите два возможных смысла предложения. Невозможность уточнить содержание тоже ограничивает будущий вывод.

Исследовательское продолжение: Проверьте одну версию изменения ощущения, меняя только выбранное условие. Например, сохраняйте способ прикосновения, но измените порядок наблюдений. До повторения запишите ожидаемое различие.

Разбор и ответы — приложение Е, глава 1, с. 292

Вопросы для самопроверки

1. Почему фраза «предмет нагрелся, поэтому прохлада исчезла» содержит больше, чем описание переживания?
2. Можно ли считать приятность характеристикой опыта?
3. Что установлено, если во второй записи оказалось вдвое больше подробностей?

В следующей главе используйте эти записи как образец. Мы перейдём от одного эпизода к нескольким и попробуем описать отношения сходства, которые будущая модель должна сохранить.

Глава 2

Различия и сходства переживаний

Первая запись позволила отделить качество переживания от истории о его причине. Но определение каждого ощущения по отдельности ещё не показывает, как устроены различия между ощущениями. Теперь соберём несколько эпизодов и зададим один и тот же вопрос сравнения. Карта сходства станет первым объектом, по которому можно судить о качестве модели: она должна объяснять отношения между случаями и предсказывать новые сравнения.

От отдельного впечатления к отношениям

Вы записали, как ощущается прикосновение к кружке. Рядом лежит тканевая салфетка. Прикосновения явно различаются, но ответ «они разные» ещё мало помогает исследованию. Можно уточнить: одно кажется более гладким, другое — более тёплым; у одного ощущение возникает в небольшой области пальца, у другого распределено шире. Теперь различие распадается на несколько отношений, которые можно сравнивать.

Эта глава посвящена качественной организации опыта. Мы будем спрашивать, по каким признакам переживания оказываются похожими, какие различия сохраняются при смене условий и как представить результаты так, чтобы модель могла их объяснять. Первая полезная модель часто начинается с отношений между случаями, а не с окончательного определения каждого случая по отдельности.

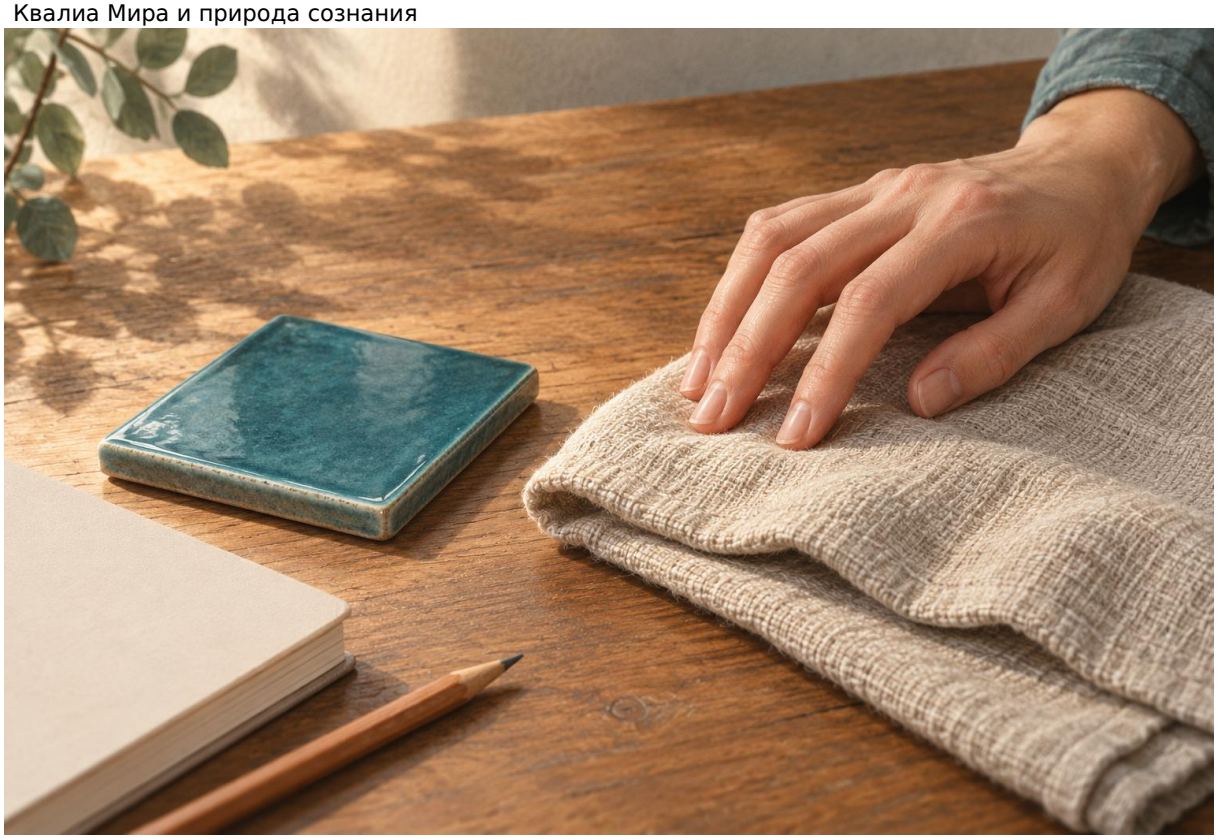

Рисунок 3. Сравнение поверхностей в одном эпизоде. Предметная иллюстрация упражнения: керамика и ткань предлагают разные условия прикосновения. Читатель сравнивает замеченные качества и отдельно записывает интенсивность и длительность; свойства материалов не заменяют описание переживания.

Качество, интенсивность и длительность

Представьте два звука одинаковой громкости: короткий стук и протяжный гул. Их различие нельзя исчерпать словом «сильнее». Громкость, характер звучания и длительность отвечают на разные вопросы. Точно так же слабая шероховатость не превращается автоматически в прохладу, когда её интенсивность возрастает. Чтобы сравнивать переживания, нужно обозначать, какое отношение рассматривается.

Качеством будем называть то, каким переживается ощущение. Интенсивность указывает на его выраженность в выбранном отношении. Длительность относится к временному протяжению. Это рабочие различения, а не готовый список независимых координат для любого опыта. В конкретном случае признаки могут взаимодействовать: более продолжительный звук способен показаться и более навязчивым, а изменение контекста — повлиять на его заметность.

Полезно задавать последовательные вопросы. Что изменилось? По какому признаку? Было ли это изменение постоянным или возникло только в начале? Если вы не можете разделить два аспекта, запишите это как свойство

текущего описания. Исследовательская аккуратность состоит в выборе подходящих различий, а не в обязательном заполнении универсальной анкеты.

Названия и переживаемое сходство

Одна и та же словесная категория может объединять разные впечатления. Слово «шероховатый» применимо к ткани, дереву и бумаге, хотя прикосновения к ним не одинаковы. Возможна и обратная ситуация: два предмета называются совершенно по-разному, но на ощупь в некотором отношении похожи. Поэтому совпадение названий и сходство опыта стоит записывать отдельно.

Допустим, вы сравниваете четыре поверхности. Две принадлежат металлическим предметам, две — пластиковым. Если попросить распределить предметы по материалу, получится одна группировка. Если спросить о гладкости прикосновения, результат может оказаться другим. Ни одна группировка сама по себе не ошибочна: они отвечают на разные вопросы.

Для будущей модели это важное различие. Система, которая хорошо угадывает название материала, ещё не обязательно воспроизводит отношения между вашими оценками прикосновений. Нужно заранее решить, какое поведение модель должна предсказывать. Иначе успех в удобной задаче можно незаметно принять за объяснение другого явления.



Рисунок 4. Названия материалов и сходство прикосновений. В учебном наборе две металлические и две пластиковые поверхности. Одни и те же предметы сгруппированы по материалу и по гладкости: выбранный вопрос определяет, какое отношение должна воспроизвести модель.

Как собрать карту сходства

Назовём выбранные эпизоды А, Б, В и Г. Буквы обозначают случаи наблюдения, а не свойства переживаний. Для каждой пары можно ответить, какое прикосновение кажется более похожим на другое, либо поставить условную оценку различия. Начать проще с ранжирования: какая пара наиболее похожа, какая наименее, а какие трудно сравнить.

При четырёх случаях получится шесть разных пар. Это уже позволяет увидеть структуру. Например, А и Б могут образовать близкую пару, как и В с Г, тогда как различия между этими парами заметнее. Получается предварительная карта двух групп. Она полезнее четырёх независимых эпитетов, потому что предъявляет модели отношения, которые следует воспроизвести.

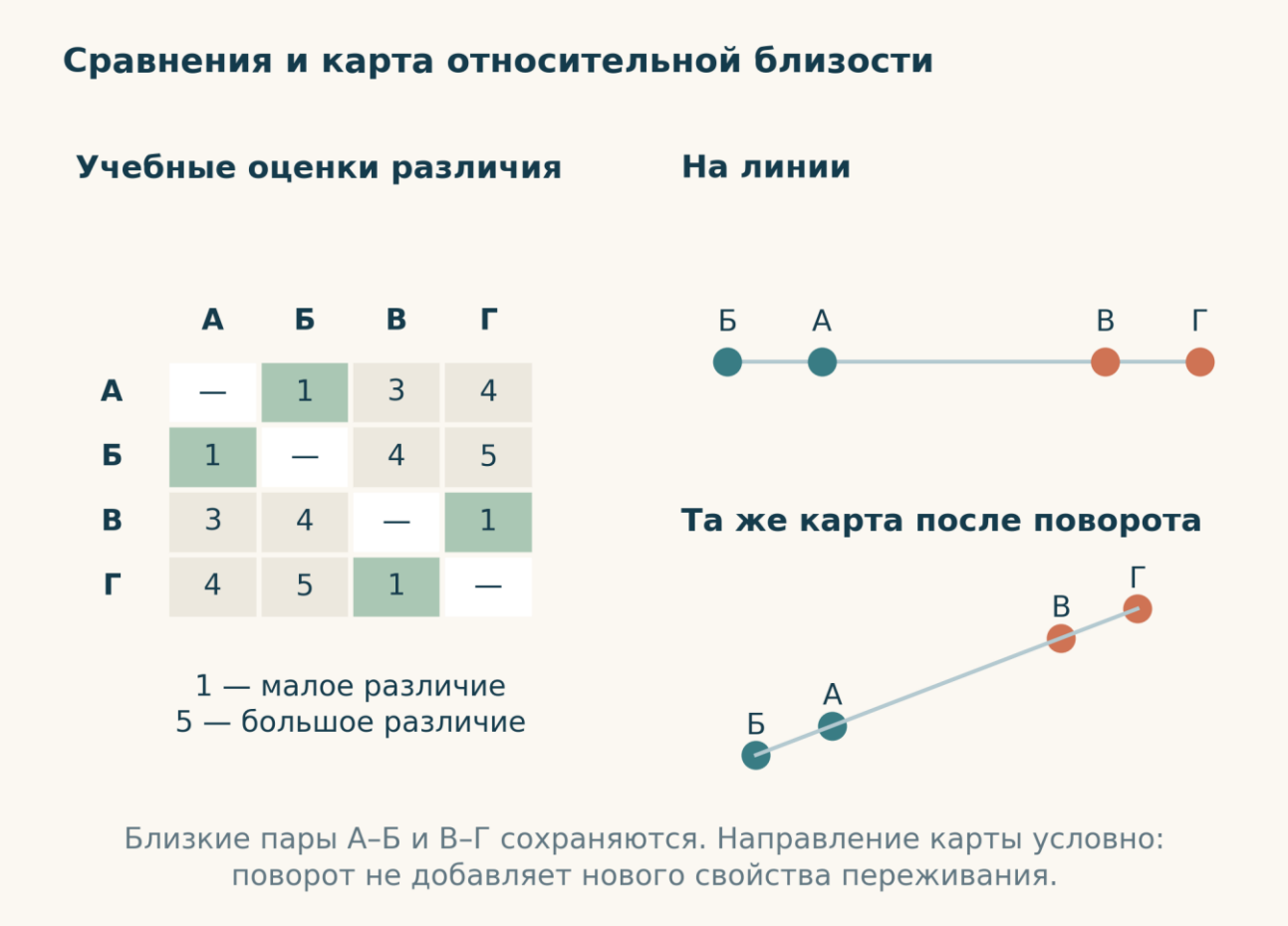

Рисунок 5. От попарных оценок к карте сходства. Матрица содержит шесть учебных оценок из главы. На линии точки расположены в порядке Б, А, В, Г с координатами 0, 1, 4 и 5; справа внизу показан поворот той же конфигурации. Это удобное представление данных примера, а не основание считать шкалу переживаемого различия шкалой точных отношений.

В методологии психологических измерений Роджер Шепард предложил восстанавливать пространственную конфигурацию по данным о близости объектов, когда точная функция расстояния неизвестна. Такой подход помогает исследовать многомерную структуру оценок сходства. В нашей первой практике используется лишь общая идея расположения случаев по отношениям между ними; полноценный вычислительный метод требует дополнительных предпосылок. [3]

Разобранный пример с учебными данными

Ниже приведены придуманные учебные оценки различия. Единица означает небольшое различие, пятёрка — большое. Эти числа иллюстрируют рассуждение и не описывают результаты исследования людей.

Пара эпизодов	Оценка различия
А и Б	1
А и В	3
А и Г	4
Б и В	4
Б и Г	5
В и Г	1

Самая заметная закономерность — две близкие пары: А с Б и В с Г. Однако остаётся дополнительное различие: А несколько ближе к В, чем к Г. Если нарисовать две совершенно неразличимые внутри группы точки, часть информации потеряется. Карта должна показывать и группировку, и различия между отдельными элементами.

Предположим, исследователь предложил объяснение через один признак — гладкость. Он располагает А и Б рядом, В и Г рядом на другом конце линии. Для проверки нужно сравнить остальные пары: получается ли их порядок без дополнительных перестановок? Если нет, возможно, нужен второй признак или другая форма модели. Но возможно также, что некоторые оценки неустойчивы. Эти объяснения потребуют разных следующих шагов.

Оценку 4 нельзя автоматически объявить вдвое большим переживаемым различием, чем оценку 2. Читатель мог использовать цифры только для порядка. Прежде чем вычислять точные расстояния, нужно выяснить смысл шкалы. В начальной работе достаточно сохранять такие утверждения, как «эта пара различается сильнее той», и отмечать неопределённые сравнения.

Контекст становится частью задачи

Представьте, что вы сначала исследовали поверхность спокойно, а затем сравнивали её сразу после другого прикосновения. Даже если сам предмет не менялся, условия сравнения стали другими. Во второй ситуации рядом присутствует недавний образец, относительно которого вы подбираете слова и оценку. Это повод включить порядок предъявления в описание процедуры.

Другой контекст задаёт вопрос. Просьба оценить гладкость направляет поиск на одни признаки; просьба выбрать самый приятный предмет — на другие. Если исследователь меняет вопрос между попытками, различия ответов

нельзя целиком приписывать изменению ощущений. Чтобы выяснять причины, нужно сохранять формулировку задания или специально сравнивать несколько формулировок.

Здесь появляется полезное понятие инварианта. Так называют отношение, которое сохраняется при рассматриваемом изменении. Если переименование А в К не меняет предсказаний модели, она не зависит от случайной буквы. Если поворот нарисованной карты сохраняет все отношения близости, направление «север» на рисунке не несёт психологического смысла. Определять, что должно сохраняться, столь же важно, как объяснять изменения.

Как читать карту без лишних предположений

В учебной таблице различие А и Б равно 1, В и Г тоже 1, а между Б и Г стоит 5. Сначала достаточно восстановить порядок: первые две пары близки, последняя заметно различается. Если разместить Б, А, В, Г на линии в точках 0, 1, 4, 5, все шесть чисел этого конкретного примера совпадут с расстояниями. Но из удачи изображения не следует, что любая субъективная шкала обладает свойствами длины. Поэтому при работе со своими ответами сначала проверяйте порядок близости, а точные арифметические отношения используйте лишь после обоснования шкалы.

Что теперь установлено

Мы сравниваем отношения по выбранному вопросу. Пока не требуется находить единственно верное место каждого ощущения. Важно сохранить исходные ответы и видеть, где изображение им соответствует, а где добавляет собственные предположения.

Практика 2 Четыре эпизода и один вопрос

Зачем выполнять: Получить проверяемую карту сходства и отделить её свойства от условных обозначений.

Что понадобится: Умение отделять описание и причину из главы 1. Нужны четыре доступные поверхности и один неизменный вопрос сравнения.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Назовите четыре поверхности А, Б, В, Г и кратко опишите каждое прикосновение.
2. Запишите один вопрос, например о сходстве характера прикосновения. Сравните ровно шесть пар: АБ, АВ, АГ, БВ, БГ, ВГ.

3. Упорядочьте пары по различию, оставляя равенства и сомнения. Нарисуйте четыре точки сначала на линии, затем на плоскости. Отметьте одно улучшившееся отношение и одно невыраженное.
4. Повторите сравнения в другом порядке с тем же вопросом. Сохраните обе карты и различайте изменение ответа и предполагаемую причину изменения.

Что сохранить: Шесть парных сравнений, две попытки изображения и перечень устойчивых и неустойчивых отношений.

Подсказка: Поворот рисунка не меняет расстояний. Добавление второй координаты полезно, только если оно выражает определённое отношение лучше, чем первая версия.

Исследовательское продолжение: Добавьте пятую поверхность. До нового сравнения предскажите хотя бы две её близости по предложенному объяснению, а затем проверьте прогноз. Ручная перестановка после каждого ответа не считается предсказанием.

Разбор и ответы — приложение Е, глава 2, с. 292

От карты к гипотезе

Полученная карта описывает структуру ваших ответов. Следующая задача — связать её с механизмом. Например, можно предположить, что оценки зависят от сочетания двух признаков, и проверить новый случай, не участвовавший в построении карты. Если положение нового случая приходится каждый раз назначать вручную, объяснительная работа ещё не выполнена.

В дальнейшем мы будем спрашивать, какие отношения между состояниями искусственной системы могли бы соответствовать найденным различиям. Первое требование к такому проекту уже понятно: он должен предсказывать структуру новых сравнений. Вопрос о том, сопровождаются ли его собственные состояния переживанием, потребует самостоятельного теоретического шага.

Вопросы для самопроверки

1. Почему совпадение словесных названий не гарантирует одинакового переживаемого качества?
2. Что означает оценка различия 4 на условной шкале от 1 до 5?
3. Как проверить, помогает ли дополнительная координата объяснению, а не только украшает карту?

Карта описала различия между переживаниями. Теперь выясним, как содержание относится к собственному телу, действию и прошлому.

Глава 3

Что означает ощущать себя

Мы научились сравнивать качества опыта. Однако два сходных впечатления могут по-разному относиться к переживающему: движение руки и движение указателя оба видны, но рука ощущается частью тела, а указатель обычно остаётся внешним предметом. Чтобы продолжить исследование, разделим перспективу, телесную принадлежность, авторство действия и представление о себе во времени. Такое разделение понадобится позже при построении моделей источника событий.

Несколько значений одного слова

Вы протягиваете руку к карандашу и говорите: «Я его взял». Короткая фраза объединяет несколько отношений. Карандаш был виден с определённого места. Рука ощущалась частью вашего тела. Движение воспринималось как ваше действие. Вы помните, зачем понадобился карандаш. В повседневном разговоре этого достаточно. Для исследования приходится спросить, какие именно отношения мы хотим объяснить и обязательно ли они всегда возникают вместе.

Начнём с перспективы. Стол виден отсюда: одна сторона ближе, другая частично закрыта. Звук раздаётся слева относительно вашего положения. Перспектива — организация содержания относительно некоторой точки или системы отсчёта. Она отличается от биографического знания о себе. Чтобы увидеть край стола справа, не требуется мысленно произнести своё имя или вспомнить профессию. Однако наличие координат, рассчитанных относительно камеры робота, само по себе ещё не устанавливает переживание перспективы. В описании системы необходимо различать устройство координат и то, как ситуация переживается.

Второй вопрос касается принадлежности тела. Ощущение «это моя рука» не равнозначно суждению о юридической собственности. Речь о том, как часть тела дана в опыте. Третий вопрос касается авторства действия: «это движение произвожу я». Четвёртый — образа себя во времени: каким человеком я себя считаю, что помню о своём прошлом, какие намерения связываю с будущим. Эти четыре вопроса задают полезную рабочую карту, хотя не исчерпывают всех употреблений слова «я».

Принадлежность тела и авторство действия

Рассмотрим обычное различие. Вы намеренно поднимаете руку, затем спокойно кладёте её на стол. Во втором случае рука не перестаёт быть вашей из-за отсутствия движения. Принадлежность тела нельзя определить как непрерывное произвольное управление им. Теперь вы передвигаете мышь: движение указателя на экране ощущается результатом вашего действия, хотя сам указатель обычно не ощущается частью тела. Управление предметом и телесная принадлежность тоже не совпадают.

В исследовании Андреаса Калькерта и Хенрика Эрссона участники наблюдали движения искусственной руки, связанные с движениями собственной. Авторы меняли согласованность и характер движения, а также положение модели. Получились условия, в которых оценки принадлежности и авторства расходились. Работа поддерживает исследовательское разделение этих компонентов телесного самовосприятия; она не превращает каждый из них в самостоятельного субъекта. [4]

Здесь полезно ещё одно различие. Чувство авторства — характеристика переживания или отчёта о нём. Фактическое причинное участие — характеристика событий. Можно нажать клавишу, не заметить результата и всё же причинно повлиять на программу. Можно ожидать, что движение выполнено вами, хотя в этот момент его воспроизвела запись. Для проверки причинного участия нужны сведения об устройстве ситуации, а для изучения переживаемого авторства — также данные от участника. Один вопрос не заменяет другой.



Рисунок 6. Принадлежность без движения и управление внешним объектом. Неподвижная рука может продолжать ощущаться частью собственного тела. Управление указателем относится к связи действия с результатом, а не автоматически к телесной принадлежности указателя.

Как тело становится доступным без взгляда

Попробуйте описать положение лежащей на столе руки, не рассматривая её специально. Возможно, доступны согнутость пальцев, давление на поверхность, приблизительное положение локтя. Информация о положении и движении частей тела называется проприоцептивной. Она участвует в телесном ориентировании вместе со зрением, осязанием и другими источниками. Термин обозначает определённый вид информации и чувствительности, а не отдельное маленькое сознание внутри тела.

Для исследователя важна согласованность источников. Если вы видите руку там, где ожидаете её почувствовать, одна группа объяснений связывает сведения с одним телесным объектом. Если сведения расходятся, возникает вопрос, чему система придаёт больший вес и почему. Пока это схема исследовательской проблемы. В главе 11 мы превратим её в вычислимую модель и сформулируем предсказания для разных условий.

Нельзя заранее свести принадлежность к одному совпадению сигналов. Две лампочки могут мигать одновременно, не становясь частями вашего тела. Модель должна объяснять, какие источники сравниваются, как учитываются форма и положение объекта, история взаимодействий и возможные

альтернативы. Именно такие уточнения превращают выразительную метафору «мозг собирает меня» в конкретную задачу.

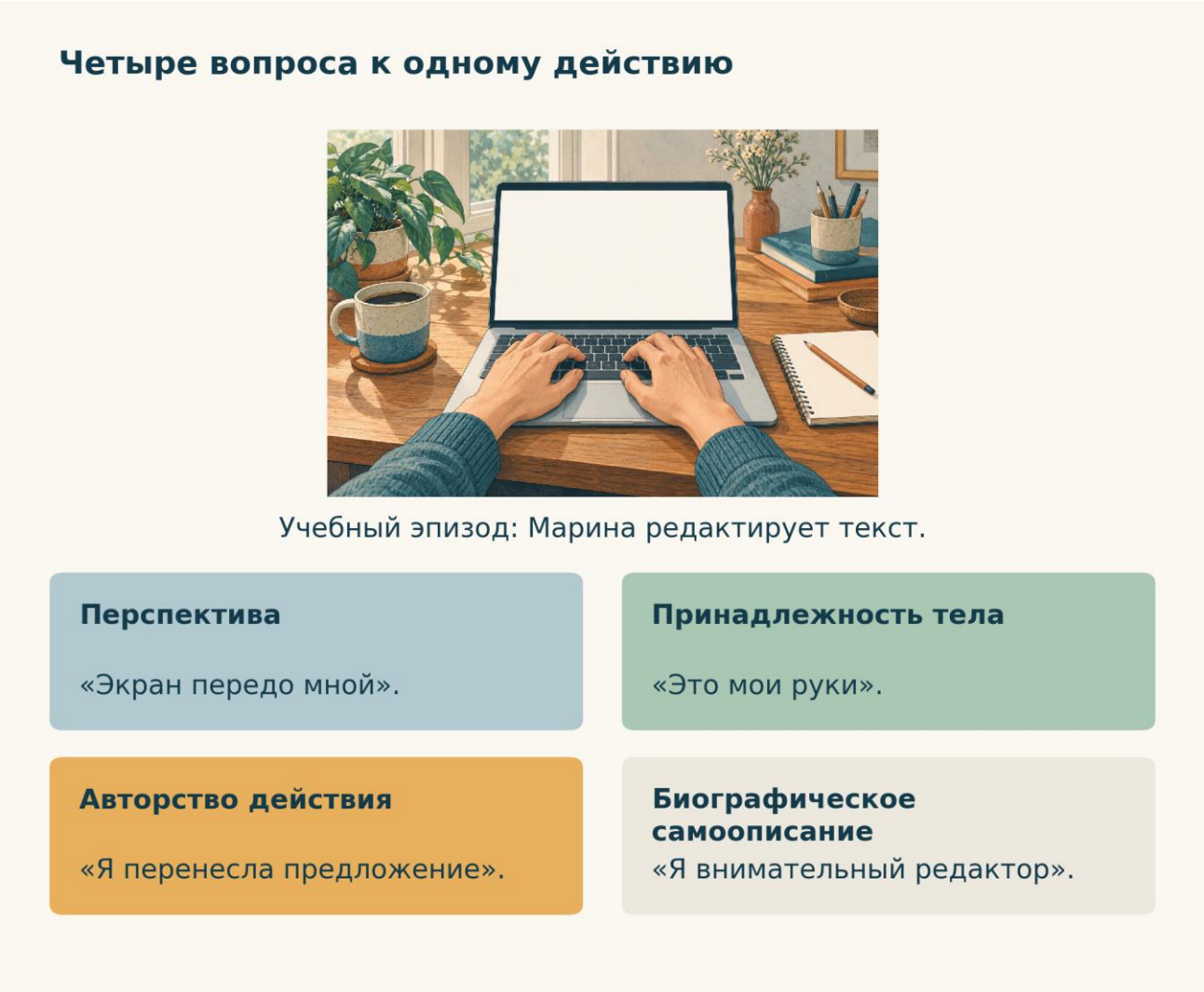

Рисунок 7. Четыре отношения к себе в одном эпизоде. Перспектива, телесная принадлежность, авторство и биографическое самоописание выделены отдельно. Уточнение того, кто исправил пробел — человек или программа, касается авторства конкретного изменения, а не автоматически всех остальных отношений.

Разобранный пример с текстовым редактором

Представим учебный эпизод. Марина переносит предложение в документе. Она двигает мышь, видит выделенный текст и намеревается поставить его в начало абзаца. После отпускания кнопки программа автоматически исправляет пробел. Марина говорит: «Я перенесла предложение и исправила оформление». Рассмотрим эту фразу без предположения, что она обязательно точно описывает все события.

Перенос предложения был намерением и действием Марины. Исправление пробела могло оказаться автоматической операцией программы. Итоговый

результат доступен ей зрительно: она может проверить расположение слов. Но доступ к результату не означает доступ ко всем причинам его появления. В записи исследования стоило бы отдельно указать движение мыши, изменение документа, сообщение Марины и известное исследователю правило автоматического исправления.

Допустим, после просмотра истории изменений Марина уточняет: «Пробел исправила программа». Это пересмотр суждения об авторстве. Из него не следует, что первоначальное ощущение управления было выдумано. Оно могло возникнуть на основании связи между намерением, действием и общим результатом. Исследовательская задача состоит в объяснении такой связи, включая случаи, когда она приводит к чрезмерно широкому приписыванию результата себе.

Биографическое самоописание появляется на другом уровне: «Я внимательный редактор». Один автоматический пробел недостаточен для проверки этого утверждения. Оно обобщает множество ситуаций и зависит от критерия внимательности. Если смешать уровни, небольшая ошибка в определении причины превращается в необоснованный вывод о личности. При раздельной записи остаётся конкретный вопрос, который можно проверить.

Нужен ли внутренний наблюдатель

Иногда самосознание представляют как зрителя, который смотрит на внутренний экран. Такая картинка удобна, пока не становится объяснением. Если зритель должен распознавать изображение, помнить прошлое и понимать, что оно относится к нему, внутри исходной задачи появилась её уменьшенная копия. Потребуется объяснять уже способности этого зрителя.

Более продуктивный начальный вопрос — какая информация связывается с системой и как эта связь используется. Например, программа может хранить собственный идентификатор, отличать свои команды от входящих сообщений и оценивать надёжность своего ответа. Это разные функциональные возможности. Их можно реализовать и проверять по отдельности. Наличие идентификатора не доказывает переживаемого «я», но даёт чёткий объект для изучения самоотнесения.

Явная рефлексия добавляет ещё один уровень: система или человек рассматривает собственное суждение как предмет проверки. Фраза «я вижу край стола» и вопрос «почему я уверен, что вижу его точно» выполняют разные задачи. При построении гипотез о минимальном субъекте нельзя незаметно сделать сложную словесную рефлексию обязательным условием любого опыта. Такое условие потребовало бы отдельного обоснования.

Почему управление не равно принадлежности

В первом эпизоде рука спокойно лежит на столе. Действия сейчас нет, но её положение и прикосновение могут переживаться как относящиеся к собственному телу. Во втором эпизоде вы двигаете мышь и видите указатель. Его перемещение зависит от команды, однако из этого не следует телесная принадлежность. В третьем вы читаете вчерашнюю заметку: её создание относится к вашей истории, хотя нынешнее чтение не является переживанием текущего написания. Эти случаи показывают, зачем нужны отдельные столбцы описания. Один общий ответ «это я» скрыл бы различия.

Что теперь установлено

Четыре вопроса относятся к разным отношениям внутри одного эпизода. Не нужно приписывать им независимость во всех ситуациях. Сначала достаточно научиться замечать, какие ответы доступны и какое отношение сейчас исследуется.

Практика 3 Четыре вопроса к трём эпизодам

Зачем выполнять: Развести несколько смыслов самоотнесения на трёх обычных эпизодах.

Что понадобится: Различение события и отчёта из главы 1. Помните, что фактическое причинное участие и чувство авторства отвечают на разные вопросы.

Ориентир по времени: 20–25 минут. Можно делать с перерывами; это не норматив скорости.

1. Опишите три случая: неподвижная рука на столе, управление указателем и чтение собственного прежнего текста.
2. Для каждого отдельно укажите перспективу, переживаемую принадлежность тела, авторство нынешнего действия и сведения о себе в прошлом.
3. Выберите одно причинное утверждение, которое можно проверить независимо: что именно вызвало движение или когда был создан текст.
4. Сопоставьте эту проверку с первым описанием. Сохраните расхождение, если оно есть; не превращайте его в общую оценку личности.

Что сохранить: Три описания с четырьмя отдельными вопросами и одной независимо проверенной связью.

Подсказка: Отсутствие слов не доказывает отсутствия компонента. Можно написать, какой именно вопрос остался неясным и какое уточнение помогло бы ответить.

Исследовательское продолжение: Предложите изменение ситуации, которое могло бы развести авторство и принадлежность. Сформулируйте отдельный прогноз для каждого показателя и объясните, какие внешние события нужно контролировать.

Разбор и ответы — приложение Е, глава 3, с. 293

Вопросы для самопроверки

1. Почему управление указателем на экране не доказывает переживания указателя как части тела?
2. Чем фактическое авторство действия отличается от чувства авторства?
3. Почему знание собственного имени недостаточно для полной модели переживаемого себя?

Различив отношения к себе, перейдём к оценке собственного мышления. Следующая глава спросит, насколько чувство понимания позволяет судить о правильности решения и как эту связь проверить.

Глава 4

Как мы замечаем собственное мышление

Ощущение себя включает возможность замечать собственное решение, но замеченное решение ещё нужно оценить. Знакомое чувство «всё понятно» иногда опережает проверку. В этой главе будем различать правильность ответа, уверенность до обратной связи и последующее объяснение. Сохранение их порядка позволит увидеть, что именно изменилось после проверки, и превратить рефлексию в основание нового действия.

Мысль и знание о мысли

Прочитайте фразу: «В коробке меньше карандашей, чем ручек, но больше карандашей, чем ластиков». Можно почувствовать, что всё понятно, и сразу перейти дальше. Можно также остановиться и спросить, каких предметов больше всего. Во втором случае первоначальное впечатление проверяется действием. Если для ответа приходится заново читать условие, это не обязательно означает отсутствие понимания. Однако становится ясно, что чувство понятности и готовность выполнить определённую задачу — разные данные.

В этой книге интроспекцией мы называем обращение к собственному текущему или только что пережитому состоянию. Рефлексия означает рассмотрение собственного мышления, действия или суждения как предмета анализа и возможного пересмотра. Метакогниция — более широкое семейство процессов оценки и регулирования познания. Например, выбор повторно проверить трудный ответ может зависеть от оценки его надёжности. Эти рабочие определения помогают задавать разные вопросы; в научной литературе границы терминов могут проводиться иначе.

Наличие мысли не гарантирует подробного знания о том, как она возникла. Допустим, вы быстро узнали знакомое слово. Вам доступен результат узнавания, но вы можете не знать последовательности обработки букв. Рассказ «я сначала рассмотрел каждую букву» тогда требует проверки. Он может быть описанием действительно замеченных шагов или правдоподобной реконструкцией уже после узнавания.

Почему искренность не обеспечивает точности

Отчёт о собственном состоянии может быть совершенно искренним и всё же неточным. Причина не обязательно в плохой памяти: иногда неясно, какие различия нужно замечать и какими словами их выражать. Вопрос «была ли мысль словесной» требует отличить произнесённую внутреннюю фразу от знания её смысла и от фразы, появившейся во время самого ответа.

Эрик Швицгебель в философской статье о ненадёжности наивной интроспекции разбирает основания сомневаться в безошибочном знании собственного текущего опыта. Его примеры касаются, среди прочего, зрительной ясности, эмоций и формы мышления. Это аргумент против автоматической непогрешимости самоописания, а не измерение частоты ошибок у любого человека и не основание отказаться от самоотчётов. [5]

Для практики отсюда следует скромное, но действенное правило: отделяйте момент переживания, момент первого сообщения и момент последующего объяснения. Если при решении задачи сначала возник ответ, а словесное обоснование появилось позже, так и запишите. Не нужно заполнять промежуток предполагаемыми мысленными шагами. Пропуск в записи полезнее вымышленной непрерывности.

Уверенность должна относиться к определённому вопросу

Фраза «я уверен на восемьдесят процентов» становится исследовательски полезной после уточнения, в чём именно уверенность. В правильности ответа? В точности воспоминания? В способности повторить объяснение завтра? Это разные события, и способ проверки у них различается. Число без указания события создаёт видимость точности.

В задачах с проверяемым ответом можно договориться, что уверенность относится к вероятности правильности данного ответа. Если человек много раз выбирает уровень 80 процентов, желательно, чтобы примерно восемь из десяти подобных ответов оказывались правильными. Это идея калибровки. Один правильный или ошибочный ответ не позволяет оценить её: вероятностное ожидание проверяется по серии сопоставимых случаев.

Стивен Флеминг и Хакван Лау различают склонность давать высокую уверенность, способность уверенностью отделять правильные ответы от ошибочных и эффективность такой оценки относительно успешности основной задачи. Это разные измерительные вопросы. Высокая средняя уверенность сама по себе не показывает хорошего различения собственных ошибок. [6]

Сохраняйте задание, ответ, уверенность до проверки и результат проверки. Сложный расчёт не восстановит утраченный исходный прогноз: уверенность, записанная после просмотра решения, отвечает на другой вопрос.

Разобранный пример двух читателей

Представим учебные данные, созданные для этой главы. Антон и Вера решили по четыре одинаковых задания. Оба правильно выполнили первое и третье, ошиблись во втором и четвёртом. Их успешность в этой маленькой серии одинакова. Уверенность записывалась до получения обратной связи.

Задание	Правильность	Уверенность Антона	Уверенность Веры
Первое	Верно	90 процентов	90 процентов
Второе	Ошибка	90 процентов	40 процентов
Третье	Верно	90 процентов	80 процентов
Четвёртое	Ошибка	90 процентов	30 процентов

Антон одинаково оценивает все ответы. Вера в этой серии меньше доверяет именно тем, которые оказались ошибочными. Поэтому равенство количества правильных ответов не означает равенства качества самопроверки. Если разрешить перепроверить только два задания, оценки Веры помогают направить усилия туда, где они нужнее. В этом заключается возможная практическая функция оценки собственного мышления.

Четыре задания не устанавливают устойчивое свойство человека: возможны случайное совпадение и разное понимание шкалы. Для проверки нужны дополнительные задания и одинаковая инструкция. Таблица описывает качество оценки ответов и не позволяет объявить Веру «более сознательной».

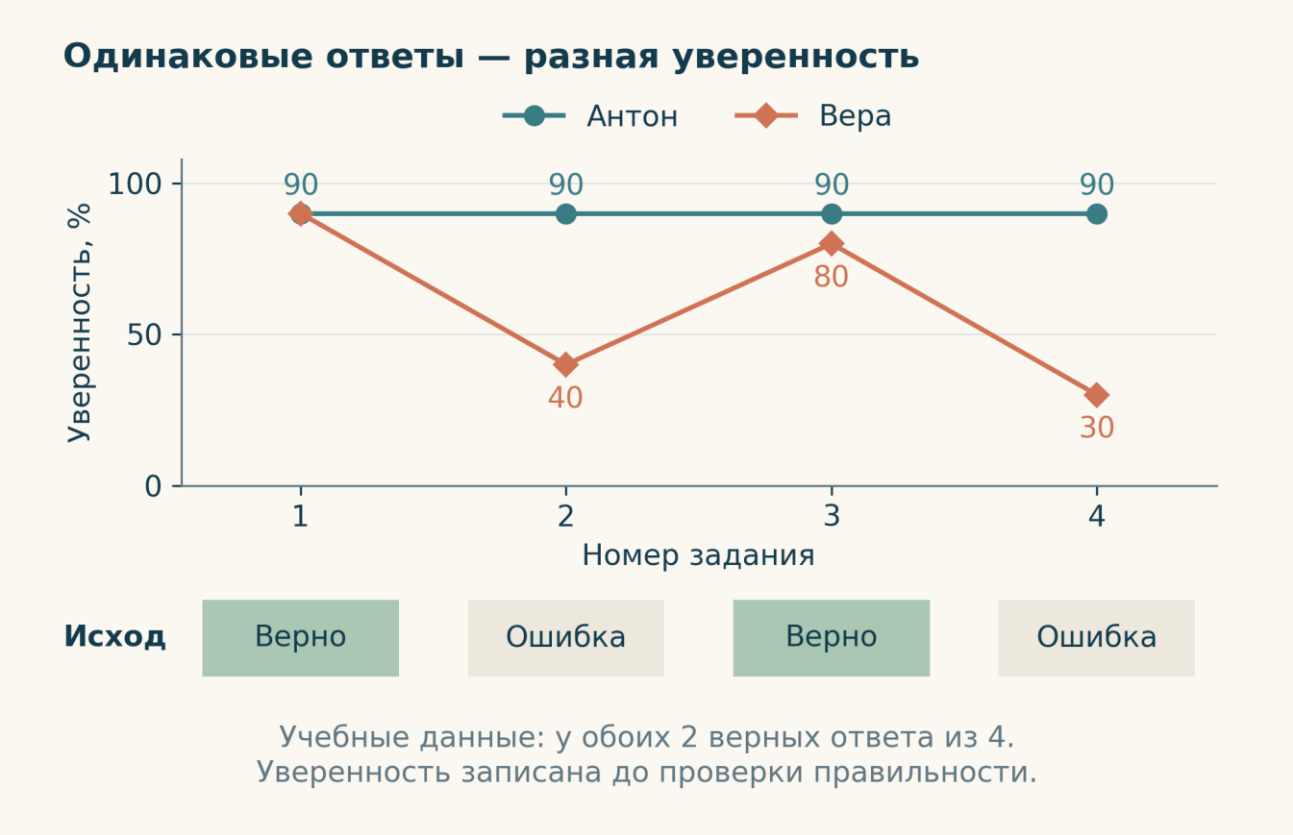

Рисунок 8. Уверенность при одинаковой успешности. В учебной серии Антон и Вера верно отвечают на задания 1 и 3. Антон всем ответам приписывает 90 процентов, а Вера — 90, 40, 80 и 30 процентов. На этих четырёх случаях оценки Веры выделяют будущие ошибки; для вывода об устойчивой способности нужны новые данные.

От самонаблюдения к изменению способа решения

Рефлексия приносит результат, когда меняет проверяемое решение о следующем шаге. Предположим, вы замечаете: «При чтении сравнений я часто удерживаю первую пару объектов, а третью теряю». Из наблюдения возникает предложение записывать отношения стрелками. Затем можно проверить, уменьшились ли ошибки на новых заданиях такого же типа. Так появляется небольшая исследовательская цепочка: затруднение, предполагаемая причина, изменение процедуры, новые результаты.

При этом успешное изменение ещё не доказывает всю историю о причине. Стрелки могли помочь просто потому, что увеличили время работы. Чтобы различить объяснения, понадобится условие с таким же временем, но без стрелок. Этот вопрос мы подробно разберём в главах о проверке. Сейчас важно увидеть, как рефлексия переходит от общего «надо быть внимательнее» к конкретному изменению, для которого возможны конкурирующие объяснения.

Самоконтроль может происходить и без развёрнутого самоанализа. Человек замечает неуверенность и перечитывает строку, не строя теории своего чтения. Явная рефлексия начинается там, где рассматриваются основания оценки, способ действия или границы вывода. Автоматическое исправление ошибки и словесное исследование причины полезно учитывать раздельно.

Как сохранить порядок знания

Пусть участник записал ответ и уверенность 90 процентов, затем увидел ошибку и сказал, что с самого начала сомневался. Последнее сообщение может быть искренним, но оно сделано после новой информации. Для исследования сохраняются обе записи: первоначальный прогноз и последующая интерпретация. Если заменить первую запись второй, исчезнет возможность оценить исходную точность самоконтроля. Поэтому в упражнении ниже ответы отделены от условий и помещены в разбор в конце книги.

Что теперь установлено

Нам не требуется безошибочная интроспекция. Достаточно сделать явными момент и предмет уверенности. Последующая ошибка становится материалом для изменения процедуры, если первоначальная запись остаётся доступной.

Практика 4 Прогноз перед проверкой

Зачем выполнять: Сопоставить предварительную уверенность с исходом и выбрать конкретный способ следующей проверки.

Что понадобится: Различайте момент переживания, первой записи и позднего объяснения. Для упражнения достаточно бумаги; знания формальной логики не требуются.

Ориентир по времени: 15–20 минут. Можно делать с перерывами; это не норматив скорости.

1. Не открывая разбор, решите три задания. В коробке карандашей меньше, чем ручек, но больше, чем ластиков: чего меньше всего? Все зелёные карточки круглые, некоторые круглые большие: следует ли наличие большой зелёной? Нина пришла раньше Олега, Олег раньше Петра: кто последний?
2. Для каждого ответа запишите уверенность от 0 до 100 процентов и одну фразу о замеченном ходе решения.
3. После всех трёх записей откройте разбор практики 4. Отдельно отметьте правильность и то, что изменилось в понимании.
4. Для одного затруднения выберите следующий приём, например поиск контрпримера. Не оценивайте устойчивые способности по трём задачам.

Что сохранить: Три ответа с предварительной уверенностью, исходы проверки и одно определённое правило следующего действия.

Подсказка: Слово «некоторые» сообщает о части круглых карточек. Попробуйте представить несколько допустимых раскладок, не раскрывая пока разбор.

Исследовательское продолжение: Проверьте выбранный приём на новых аналогичных задачах, не изменяя исходную уверенность. Сравните точность ответов и способность выделять сомнительные случаи отдельно.

Разбор и ответы — приложение Е, глава 4, с. 293

Вопросы для самопроверки

1. Почему объяснение решения после просмотра ответа нельзя автоматически считать записью первоначального мышления?
2. Может ли человек часто отвечать правильно и плохо различать уверенностью собственные ошибки?
3. Как превратить фразу «я невнимательно читаю» в начало исследовательской работы?

Теперь у нас есть пример полезного пересмотра. Чтобы объяснить, почему он помогает, нужно отличить название результата от работающего механизма. Этим займётся следующая глава.

Часть 2

Переход к научному исследованию

Глава 5

Что считается объяснением

В предыдущей главе можно было улучшить способ решения, например начать искать контрпример. Но улучшение ещё оставляет вопрос о причине: помог новый приём, дополнительное время или знакомство с задачей? Теперь научимся строить объяснение так, чтобы его можно было сравнить с другой версией. Это подготовит и экспериментальную часть, и поздние модели сознания.

От названия к исследовательскому вопросу

Человек несколько раз перечитал предложение и наконец понял его. Почему это произошло? Ответ «включилось понимание» повторяет название результата. Ответ «мозг обработал информацию» добавляет общий процесс, но ещё не показывает, что изменилось между неудачным и удачным чтением. Хорошее начало объяснения уточняет задачу: какую связь удалось установить, какая информация была доступна раньше и какие операции могли привести к новому результату.

Описание сообщает, что происходило. Например: «На втором чтении участник правильно расположил три события по времени». Объяснение стремится показать, за счёт чего получился такой результат. Предсказание указывает, чего ожидать при определённых условиях. Эти задачи связаны, но не тождественны. Можно точно предсказывать выход устройства по таблице прошлых случаев, почти ничего не зная о внутреннем устройстве. Можно предложить правдоподобный механизм, который пока плохо предсказывает реальные данные.

Для исследований опыта важно также уточнять сам результат. Мы объясняем правильный ответ, сообщение о ясности, способность использовать сведения в новой задаче или переживаемое понимание? Один эксперимент может содержать все эти данные, но переход от одного к другому требует аргумента. Иначе теория незаметно объясняет удобный показатель вместо исходного явления.

Что мы называем механизмом

Представим автоматическую лампу. Датчик измеряет освещённость, программа сравнивает число с порогом, переключатель подаёт питание. Когда становится темнее, лампа включается. Здесь названы части, их взаимодействие и последовательность изменений. Если датчик закрыть, результат изменится предсказуемым образом. Это простое механистическое объяснение, которое можно проверить вмешательством.

Питер Махамер, Линдли Дарден и Карл Крейвер предложили анализировать механизмы через организованные сущности и виды деятельности, производящие изменения. Для нашей работы полезен сам исследовательский ход: перечисления компонентов недостаточно, нужно показать их организацию и работу. Их статья представляет философский анализ объяснения, а не экспериментальную теорию сознания. [7]

Применим подход к чтению. Компонентами модели могут быть представления слов, хранилище предыдущего предложения и правило сопоставления временных отношений. Слова «хранилище» и «правило» пока условны. Нужно уточнить, что сохраняется, как долго, как используется и при каких условиях теряется. Только тогда станет понятно, почему модель справляется с одним текстом и ошибается в другом.

Механизм может быть описан на разных уровнях. Электрическая схема объясняет переключение лампы через ток, а программа — через условие и команду. Если оба описания верны, они могут относиться к одной системе и отвечать на разные вопросы. Само разнообразие уровней не требует существования нескольких независимых причинных миров.

Корреляция и вмешательство

Допустим, в учебном наблюдении люди, дольше рассматривавшие изображение, чаще говорили, что хорошо его поняли. Возможны несколько объяснений. Дополнительное время помогло разобраться. Сложные изображения одновременно заставляли смотреть дольше и вызывали более осторожный ответ. Или заинтересованные участники и смотрели дольше, и сообщали о большей ясности. Связь двух показателей ещё не определяет направление причинного влияния.

Вмешательство меняет условие по исследовательскому правилу. Например, время показа назначается заранее и не зависит от заинтересованности участника. Если остальные существенные различия учтены, сравнение становится более информативным о действии времени. Однако вмешательство тоже необходимо разбирать. Более долгий показ может одновременно

изменить число движений глаз, возможности повторного чтения и утомление. Это части возможной причинной цепочки или сопутствующие изменения, а не автоматически устранённые вопросы.

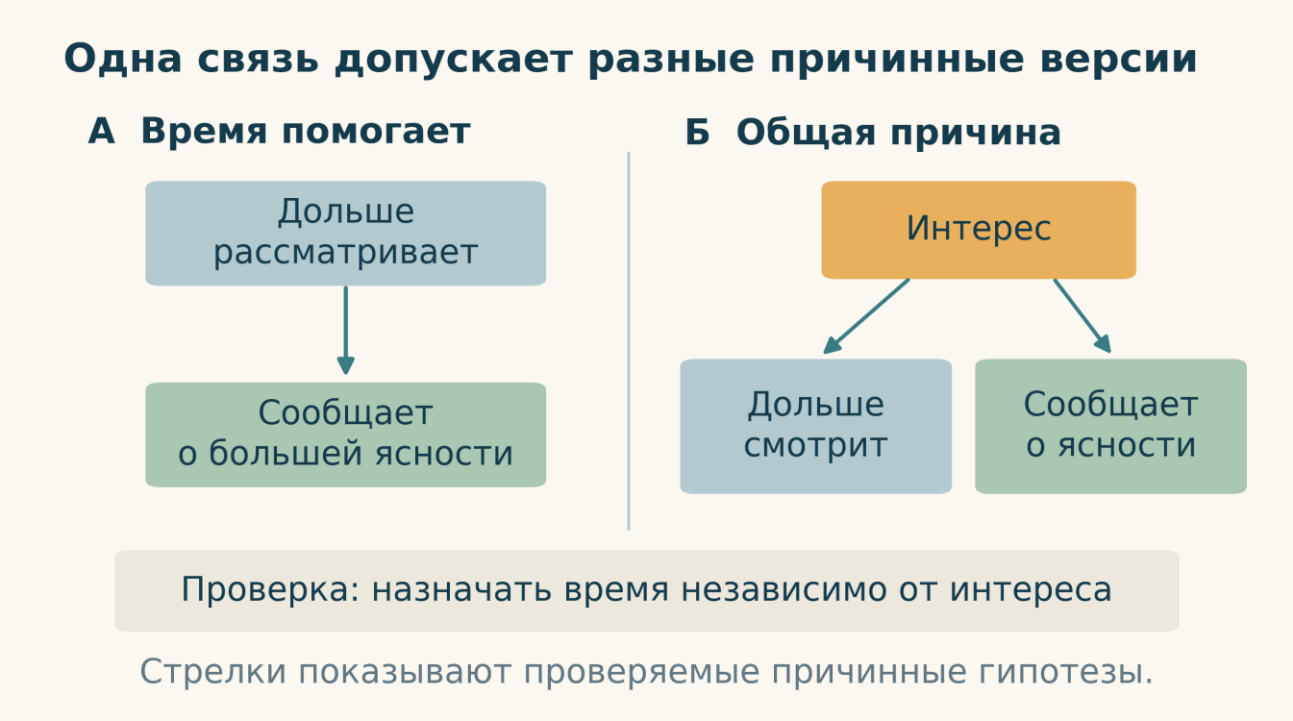

Рисунок 10. Причинное влияние и общая причина. Связь длительности просмотра с сообщением о ясности допускает разные механизмы: действие времени либо влияние интереса на оба показателя. Назначение времени по исследовательскому правилу помогает различить эти версии, если учтены остальные существенные условия.

Полезно различать причину интересующего события и показатель этой причины. Термометр сообщает о температуре, но изменение цифры на нарисованном термометре не нагреет комнату. Аналогично нейронный показатель, связанный с отчётом, может быть причиной, следствием или сопутствующим процессом. Чтобы выбрать объяснение, нужны условия, в которых эти варианты дают разные предсказания.

Необходимость, достаточность и область утверждения

Если свойство необходимо для некоторого результата, без него результат не возникает в указанной области условий. Если свойство достаточно, его наличие обеспечивает результат при оговорённых условиях. Эти отношения нельзя менять местами. Рабочий источник питания необходим для действия нашей лампы, но одного питания недостаточно: лампа может быть выключена программой или иметь неисправный переключатель.

При обсуждении сознания область условий особенно важна. Фраза «процесс П необходим для опыта» звучит универсально. Данные об одной зрительной задаче у взрослых людей поддерживают более узкое утверждение. Обобщение

на другие виды опыта, организмы или искусственные системы потребует дополнительных оснований. Уточнение области не ослабляет хорошую теорию: оно делает видимой границу уже выполненной работы и следующего исследования.

Нужно также отличать причинное обеспечение от предполагаемого тождества. Теория может утверждать, что определённая организация процесса и есть переживание. Тогда автор предлагает отношение между двумя способами описания одного явления. Наблюдение совместного появления само по себе не устанавливает такое тождество. Понадобится объяснить соответствия, границы и преимущества этого взгляда перед альтернативами.

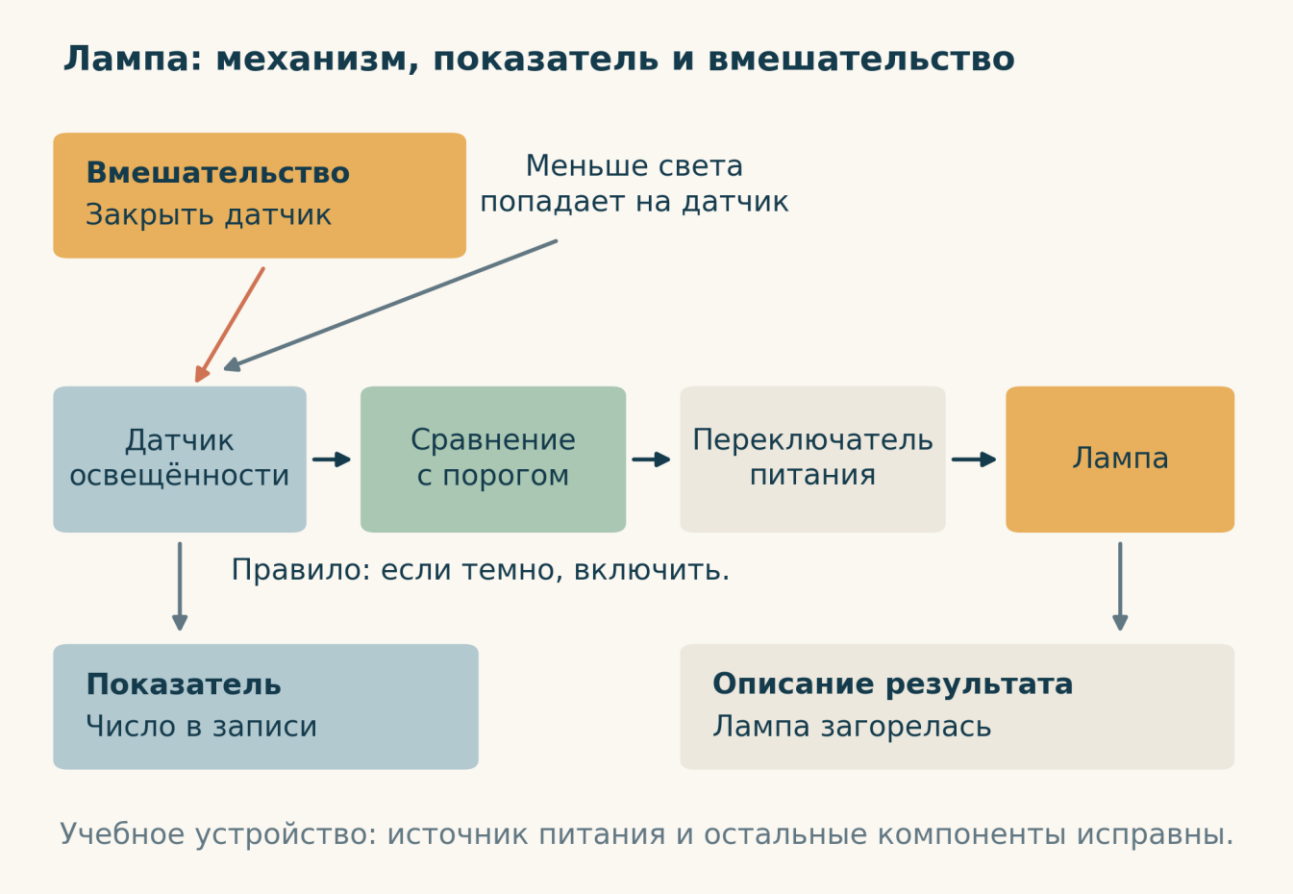

Рисунок 11. Три исследовательские роли в устройстве лампы. Средний ряд показывает передачу измерения, команды и питания в учебной автоматической лампе. Закрывание датчика — вмешательство; число в записи — показатель; включение лампы — описанный результат. Прогноз зависит от оговорённой исправности остальных компонентов.

Разобранный аргумент о повторной обработке

Рассмотрим вымышленное рассуждение: «После повторного просмотра люди лучше распознают изображение. Значит, повторяющаяся обработка создаёт

сознание. Добавим обратную связь программе и получим переживание». Первый шаг — эмпирическое утверждение об успешности задачи. Второй связывает успешность с сознанием. Третий переносит предполагаемый механизм между разными системами. Каждый шаг добавляет предпосылки, которых нет в исходном наблюдении.

Перестроим рассуждение. Сначала определим показатель: долю правильно распознанных изображений и отдельную оценку зрительной ясности. Затем предложим механизм: первый проход выделяет признаки, последующие уточняют их с учётом соседних элементов. Следующее предсказание: повторная обработка должна особенно помогать в условиях, где локальные признаки неоднозначны, а контекст информативен.

Теперь появляется конкурент: улучшение обеспечивает просто дополнительное вычислительное время, независимо от обратных связей. Для сравнения нужны модели с сопоставимыми ресурсами и задача, на которой их предсказания расходятся. Если модель с обратной связью выигрывает, мы получим основание предпочесть этот механизм для выбранного результата. Дальнейшее отношение к переживанию будет зависеть от отдельной теории и других проверок.

Такой разбор не закрывает вопрос о создании возможного субъекта. Он делает проект последовательным: сначала выяснить, какая организация что объясняет, затем сформулировать обоснованный переход к свойствам опыта. Инженерный успех становится частью аргумента, а не его незаметной заменой.

Как из двух рассказов получить различающее сравнение

Рассмотрим пользу дневника. Версия первая: сохранённый текст удерживает детали и позволяет их сопоставлять. Версия вторая: помогает сама пауза для внимательного рассмотрения. Сравнение дневника с отсутствием любой остановки совместимо с обеими версиями. Более содержательное сравнение даёт одинаковое время, но в одном условии сохраняет запись для последующего использования, а в другом оставляет только паузу. Теперь различие результатов может ограничить объяснение. Однако нужно ещё выбрать показатель понимания: длина рассказа и точность нового различения не являются одним и тем же результатом.

Что теперь установлено

Объяснение стало определённее, когда появились части, взаимодействия и разные предсказания. Эти требования не делают идею истинной заранее. Они позволяют понять, какие результаты действительно изменят отношение к ней.

Практика 5 Восстановить недостающие переходы

Зачем выполнять: Преобразовать общее утверждение о пользе дневника в проверяемую пару механизмов.

Что понадобится: Понятия события, показателя и предварительного прогноза. Будем различать, что измеряется и какой механизм предлагается для объяснения.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Выберите один результат: например, правильное различение двух заранее описанных типов эпизодов на новых примерах.
2. Запишите две версии: сохранение доступной записи и эффект дополнительной паузы. Для каждой укажите вход, действие и ожидаемое изменение.
3. Составьте сравнение с одинаковым временем и материалом. Опишите, чем именно условия различаются.
4. Заранее сформулируйте три исхода: преимущество записи, совпадение результатов и ухудшение. Для каждого запишите допустимый вывод и оставшуюся альтернативу.

Что сохранить: Одна страница с показателем, двумя механизмами, сравниваемыми условиями и правилами интерпретации.

Подсказка: Если любой исход подтверждает любимое объяснение, уточните обещание механизма. Отсутствие различия при неточном измерении ещё не устанавливает равенство.

Исследовательское продолжение: Добавьте вмешательство в доступ к сохранённому тексту после записи. Объясните, какой новый вывод оно разрешит и какие побочные изменения потребуется учитывать.

Разбор и ответы — приложение Е, глава 5, с. 294

Вопросы для самопроверки

1. Чем описание механизма отличается от перечисления частей системы?
2. Почему совместное изменение нейронного показателя и отчёта ещё не определяет причинное направление?
3. Что отсутствует в переходе от успешной программы распознавания к утверждению о созданном переживании?

Следующая глава разберёт, как выбранное явление становится данными. Для каждой кнопки, словесной шкалы и физиологического сигнала понадобится объяснить связь с исследуемым вопросом.

Глава 6

Как исследовать субъективный опыт

У нас появились механизмы и предсказания. Теперь необходимо решить, какие наблюдения позволят их проверить. Переживание не переносится прямо в таблицу: запись ответа создаётся в определённых условиях. Рассмотрим, что дают сообщение участника, успешность действия и физиологический показатель. Главное умение этой главы состоит в восстановлении пути от вопроса к данным и от данных к допустимому выводу.

Из переживания не получается готовая таблица

Вы рассматриваете рисунок, замечаете знакомую форму и нажимаете кнопку. Исследователь может сохранить изображение, момент нажатия, ваш ответ и запись физиологического сигнала. Сам эпизод переживания не становится ещё одним файлом, который можно открыть рядом с остальными. Между исследовательским вопросом и данными требуется мост: почему именно этот ответ или показатель информативен о том, что мы изучаем?

Начинающему полезно выделять три линии данных. Первая — отчёты о переживании: словесные описания, выбор категории, оценка ясности. Вторая — поведение в задаче: правильность различения, время ответа, перенос знания на новый случай. Третья — физиологические измерения. Они дополняют друг друга, если исследователь понимает, какое отношение каждой линии к исходному вопросу предполагается.

Такая работа не сводится к поиску самого объективного прибора. Число тоже требует интерпретации. Если участник нажал кнопку «ясно», напрямую зарегистрировано нажатие в ответ на инструкцию. Вывод о ясности переживания опирается на понимание участником шкалы, условия сообщения и принятую исследователем модель отчёта. Это не обесценивает данные: это показывает, что именно следует проверять.

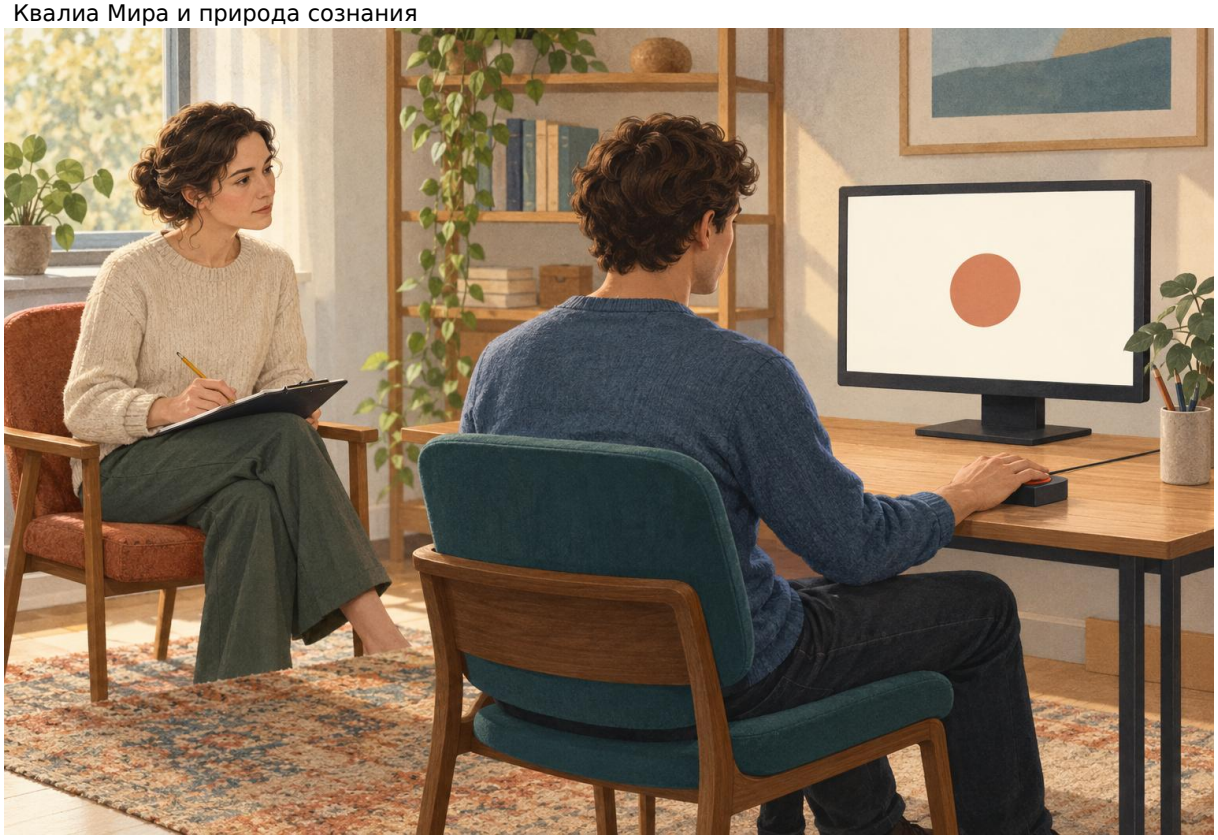

Рисунок 12. Участник и наблюдатель в исследовании. Предметная иллюстрация условного задания с экраном и кнопкой ответа. Исследователь регистрирует действие и получает описание участника; отношение этих сведений к переживанию требует отдельного обоснования в процедуре.

Получить описание конкретного эпизода

Сравните вопросы «как вы обычно воспринимаете музыку» и «что вы заметили в момент последнего звука». Первый приглашает к обобщению и рассказу о привычках. Второй направляет к определённому эпизоду. Если нам нужна последовательность пережитого, полезно привязать описание к событию и времени, а затем отделить непосредственно замеченное от предположений о причинах.

Клер Петитманжен описала метод интервью, направленный на подробное исследование конкретного субъективного опыта. В статье рассматриваются трудности осознавания и описания, способы работы с ними и критерии состоятельности получаемых описаний. Это методологическое предложение о сборе данных, а не гарантия безошибочности любого подробного рассказа. [8]

Для собственной практики задавайте вопросы, которые оставляют место неожиданному ответу. «Что изменилось перед нажатием?» обычно полезнее, чем «вы почувствовали уверенность, верно?». Второй вариант заранее называет ожидаемое переживание. Даже отрицательный ответ тогда строится

вокруг предложенной категории. После свободного описания можно уточнять конкретные различия, сохраняя исходную формулировку участника.

Если человек отвечает «не помню», это тоже результат процедуры. Нельзя автоматически заменять его на «ничего не переживал». Возможно, эпизод не сохранился, вопрос неясен или нужное различие не было замечено. Для следующего исследования можно сократить задержку отчёта либо изменить вопрос. Но нельзя задним числом получить отсутствующее сообщение из удобного предположения.

Что проверяет поведение

Представим задание: определить, какая из двух линий длиннее. Правильный выбор показывает успешное использование информации для данного различения. Если участник дополнительно говорит «разница была совершенно ясной», возникает отдельный отчёт. Когда точность и ясность расходятся, не следует выбрасывать один показатель ради согласованной истории. Нужно исследовать условия расхождения.

Время ответа тоже неоднозначно. Медленный ответ может возникнуть из-за сложного различения, тщательной проверки или неудобного расположения кнопок. Быстрый ответ может быть результатом лёгкого решения или поспешного выбора. Поэтому время интерпретируют вместе с устройством задачи и другими данными. Вопрос «чем быстрее, тем сознательнее?» не получает разумного ответа без теории, которая объясняет, почему ожидалась именно такая связь.

Поведенческая задача особенно полезна, когда нужно проверить доступ к содержанию. Если сведения можно использовать для нового вывода, инструкции или выбора, мы получили данные об их функциональной доступности. Однако доступ различается по задаче: человек может правильно указать место объекта и затрудняться описать его словами. Один неудачный формат ответа не исчерпывает всех способов использования информации.

Как читать физиологический показатель

Электроэнцефалография регистрирует электрические разности потенциалов с помощью электродов; обычная функциональная магнитно-резонансная томография использует сигнал, связанный с кровоснабжением и насыщением крови кислородом. Это разные измерения с разными временными и пространственными свойствами. Цветная область на изображении мозга — результат обработки измерений, а не фотография отдельного переживания.

Никос Логотетис и коллеги сопоставили функциональную магнитно-резонансную запись с электрическими нейронными сигналами в зрительной коре обезьян. Их работа исследовала, как измеряемый сигнал связан с нейронной деятельностью. Для нашей темы важен сам шаг: отношение между приборным показателем и процессом, который им хотят оценивать, является самостоятельным исследовательским вопросом. [9]

Предположим, показатель усиливается перед ответом «вижу». Он может отражать обработку изображения, подготовку сообщения, внимание к задаче или сочетание процессов. Чтобы уточнить роль, можно менять способ ответа, момент вопроса или требование сообщать результат. Такие сравнения полезны, когда сохраняют другие существенные условия. Отсутствие немедленного отчёта убирает часть связанных с ним действий, но ещё не создаёт прямой измеритель опыта.

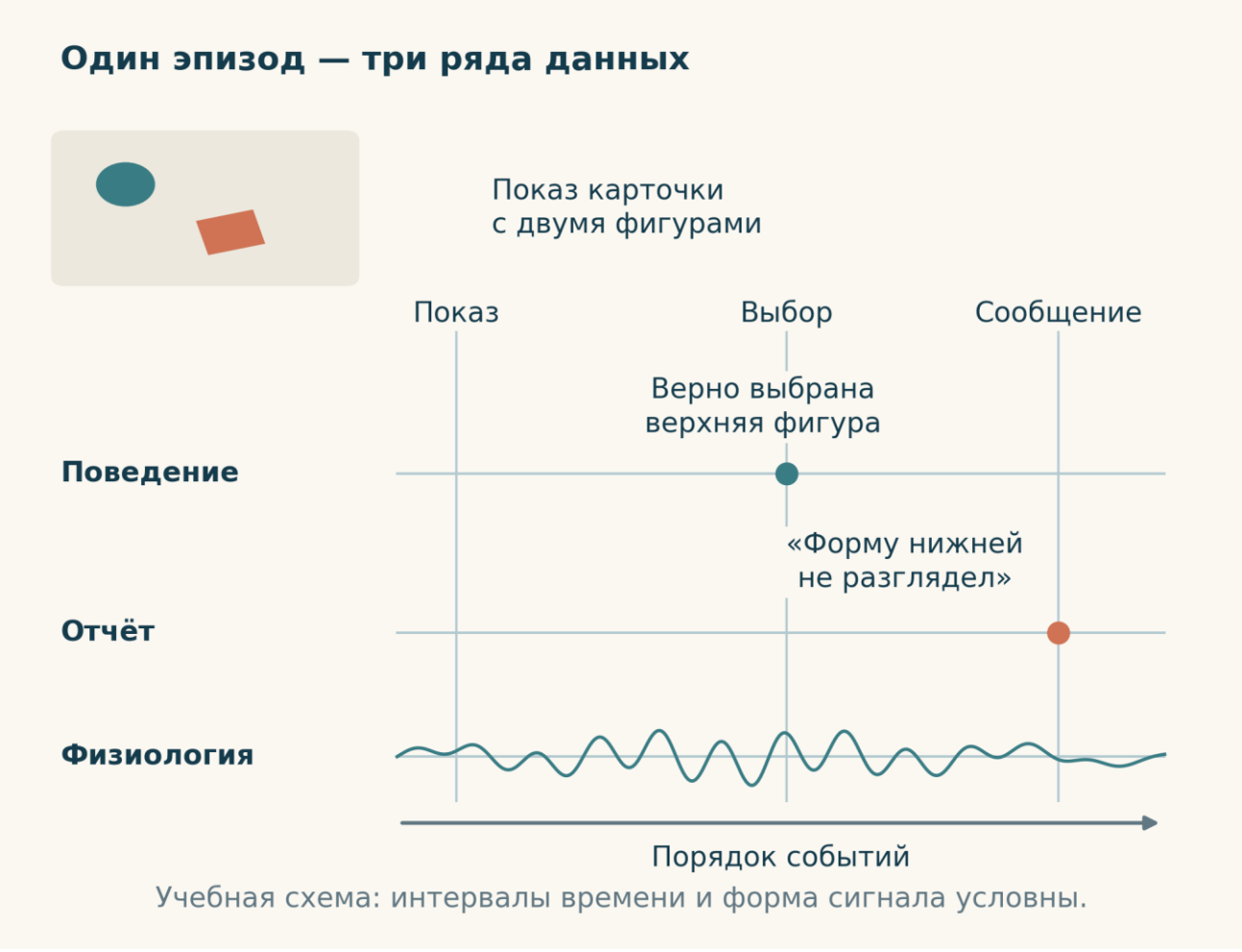

Рисунок 13. Согласование разных данных по времени. Вымышленный эпизод из главы: правильное различение положения фигуры сочетается с отчётом о неясной форме. Нажатие, сообщение и физиологическая запись представлены на общей временной оси. Форма кривой и временные интервалы схематичны и не являются экспериментальными данными.

Разобранный пример несовпадающих данных

Рассмотрим вымышленную учебную ситуацию. На карточке изображены две фигуры. Участник верно выбирает, какая расположена выше, но сообщает: «Форму нижней я не разглядел». Здесь нет обязательного противоречия. Положение и форма — разные свойства. Успешное решение первого вопроса не устанавливает ясного восприятия второго. Ошибка исследователя возникла бы, если бы любую правильность он обозначил одной общей меткой «полностью осознал картинку».

Допустим, физиологический показатель оказался выше именно в случаях правильного выбора. Возможны по меньшей мере две версии: показатель связан с качеством пространственной обработки либо с уверенностью при подготовке ответа. Для различения можно подобрать случаи с одинаковой правильностью, но разной уверенностью, и отдельно изменить способ сообщения. Следующее исследование проектируется из конкретного расхождения предсказаний.

Объединение методов иногда называют триангуляцией. Его сила состоит в ограничении альтернатив, а не в голосовании трёх приборов. Если отчёт, нажатие и нейронный показатель зависят от одной подготовки ответа, их согласие не является тремя независимыми подтверждениями переживания. Нужно выяснять, какие причины у показателей общие, а какие различаются.

Почему правильный ответ не исправляет прошлое описание

После короткого просмотра ручки человек сообщает о её цвете, но не упоминает надпись. Повторный осмотр показывает, что надпись действительно была. Он устанавливает свойство предмета, однако не выбирает между версиями: надпись не была замечена, была замечена и забыта или была опущена в коротком ответе. Нужно другое сравнение, например изменение задержки отчёта. Если же вопрос о надписи известен до просмотра, он направляет поиск; такой просмотр нельзя считать той же ситуацией свободного наблюдения.

Что теперь установлено

Одни данные ограничивают свойства предмета, другие характеризуют использование информации и сообщение о ней. Несовпадение этих рядов помогает поставить вопрос точнее. Теперь попробуем описать процедуру так, чтобы её смысл был виден другому человеку.

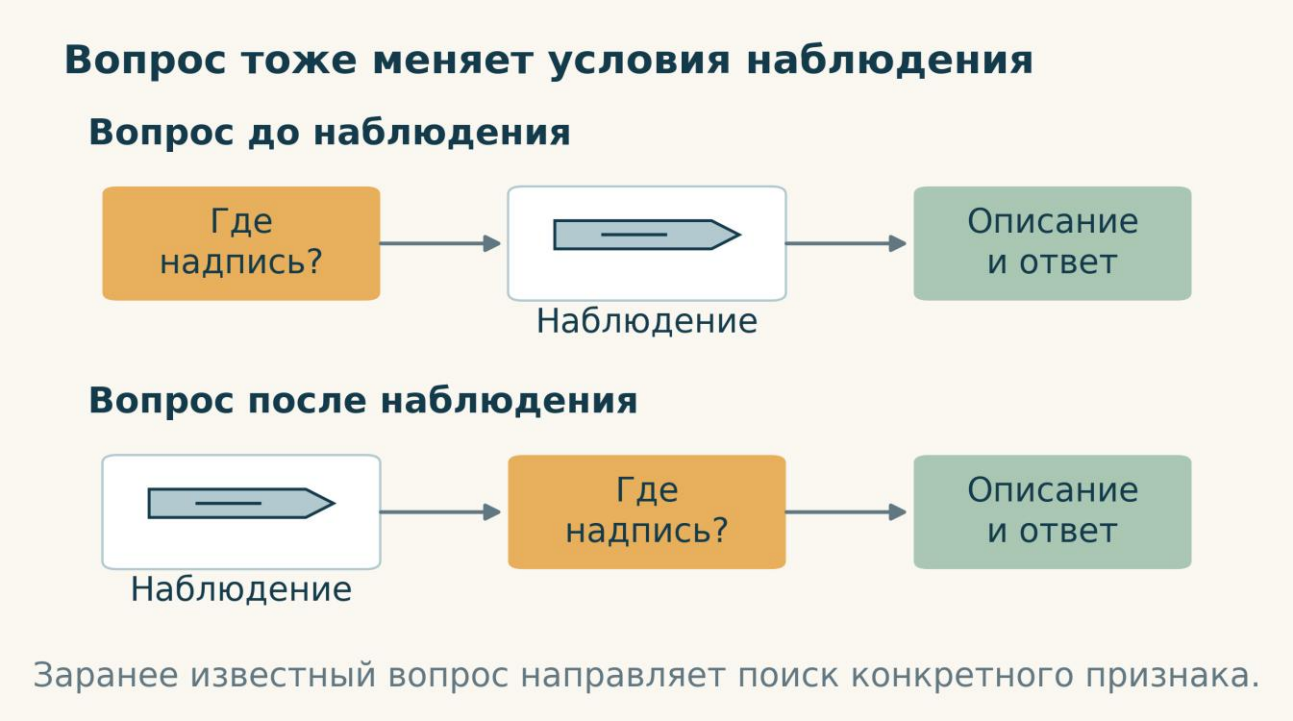

Рисунок 14. Время вопроса как условие процедуры. Верхняя процедура сообщает вопрос о надписи заранее, нижняя — после наблюдения предмета. Во втором случае момент вопроса меняется; поэтому результаты нельзя объяснять только различием свойств предмета, игнорируя направленность наблюдения.

Практика 6 Паспорт одного измерения

Зачем выполнять: Составить паспорт измерения с явными условиями и границами вывода.

Что понадобится: Различение показателя и объясняемого явления из главы 5. Учитывайте время вопроса и время ответа.

Ориентир по времени: 20–25 минут. Можно делать с перерывами; это не норматив скорости.

1. Выберите предмет с цветом и отличительной деталью. Отметьте заранее, что уже знаете последующие вопросы; это учебная демонстрация процедуры.
2. Посмотрите около пятнадцати секунд, уберите предмет и запишите пережитый эпизод. Затем отдельно ответьте о цвете и расположении детали.
3. Проверьте устройство предмета. Сохраните первоначальную запись и различайте внешний факт и содержание отчёта.
4. Заполните пять строк паспорта: предъявление, инструкция, момент вопроса, задержка сообщения, правило оценки. Добавьте два вывода, разрешённый и пока неразрешённый.

Что сохранить: Запись эпизода, проверяемые ответы и паспорт процедуры с одним предложением улучшения.

Подсказка: Ответ «не помню» не равен сообщению об отсутствии переживания. Уточняйте, какая часть цепочки остаётся неизвестной.

Исследовательское продолжение: На новых предметах сопоставьте вопрос до и после просмотра. Сначала зафиксируйте, какое различие ожидаете и почему это сравнение одновременно меняет направленность наблюдения.

Разбор и ответы — приложение Е, глава 6, с. 295

Вопросы для самопроверки

1. Почему ответ «не помню» нельзя приравнивать к отсутствию переживания?
2. Что непосредственно показывает правильное различение двух объектов?
3. Почему согласие отчёта, поведения и физиологического показателя иногда слабее, чем кажется?

Паспорт показывает устройство одной процедуры. Теперь нужно организовать сравнение процедур так, чтобы возможные причины результата не менялись все одновременно. Это задача главы 7.

Глава 7

Как устроена проверка

В главе 6 мы описали, как получается отдельный показатель. Теперь рассмотрим две процедуры и спросим, можно ли приписать их различие выбранной причине. Если вместе со способом работы меняются трудность материала и порядок выполнения, итог допускает несколько объяснений. Задача этой главы состоит в устройстве самого сравнения. Мы подготовим его до сбора данных и отдельно сохраним место для идей, возникающих после просмотра результатов.

Поиск возможного опровержения

Гипотеза становится рабочей, когда понятно, какой возможный результат ей не подходит. Утверждение «опыт как-то связан с обработкой информации» слишком широко для одной проверки. Утверждение «сохранённая схема временных отношений помогает отвечать на новые вопросы о тексте при одинаковом общем времени подготовки» уже задаёт условия, результат и возможное сравнение. Можно получить улучшение, отсутствие заметного различия или ухудшение.

Проверка начинается до сбора данных. Исследователь определяет, что будет менять, что измерять и какие другие изменения могут объяснить результат. Изменяемое условие называют независимой переменной, измеряемый результат — зависимой. Например, наличие схемы является условием, а доля правильных ответов — результатом. Эти названия описывают роли в исследовании; они не доказывают независимость переменных и причинность связи.

Операционализация — перевод общего понятия в конкретную процедуру. «Понимание текста» можно оценивать вопросами, требующими вывести новое отношение из прочитанного. Но выбранная процедура охватывает только определённую сторону понимания. Если ответы разрешено искать в схеме, измеряется решение задачи с внешней опорой. Если схему убирают, исследуется другая ситуация, включающая сохранение сведений в памяти.

Почему плохое сравнение выглядит убедительно

Представим учебный проект. Сначала человек читает сложный текст без заметок и отвечает на вопросы. Затем читает более простой текст, делает схему и отвечает лучше. Автор заключает, что схема улучшила понимание. В действительности одновременно изменились стратегия, трудность материала

и порядок выполнения. Вторая попытка могла выиграть за счёт знакомства с форматом вопросов.

Такое сопутствующее различие называют смешивающим фактором, если оно мешает выделить интересующую причину. Недостаточно перечислить возможные помехи в конце отчёта. Проект проверки должен по возможности изменить их связь с главным условием. Если простые тексты всегда используются со схемой, по результату нельзя отделить влияние схемы от влияния простоты.

Не все различия нужно устранять физически. Часть можно распределять между условиями по заранее выбранному правилу. Другие измеряют и учитывают при анализе. Однако расчёт не исправит отсутствие необходимых сравнений. Если ни одного сложного текста со схемой не было, модель должна будет опираться на дополнительные предположения о том, как работает схема на сложных материалах.

Улучшение проекта шаг за шагом

Начнём с двух наборов коротких текстов А и Б. Они должны быть похожи по длине, формату и требуемым выводам. Желательно подготовить их до основной проверки и убедиться на пробном запуске, что инструкции понятны. Пробный запуск предназначен для исправления процедуры; его результаты нельзя незаметно смешивать с данными проверки уже изменённого плана.

Одни участники работают со схемой на наборе А и без неё на наборе Б; другие получают обратное сочетание. Внутри этих вариантов меняют порядок условий. Такое уравновешивание порядка и материала не делает тексты абсолютно одинаковыми. Оно распределяет их особенности так, чтобы один набор не был неразрывно связан с одной стратегией.

Случайное назначение вариантов помогает не выбирать условие по особенностям участника. Например, нельзя автоматически выдавать схему всем более опытным читателям, а затем приписывать ей их преимущество. В маленькой группе случайность не гарантирует точного равенства состава, но создаёт понятную процедуру распределения. Её полезно сочетать с продуманным устройством материалов.

Если исследование проводится только на себе, можно менять условия и материалы по заранее составленному расписанию. Результат тогда прежде всего относится к вашим сессиям. Повторное чтение одного и того же текста создаёт перенос знания, поэтому лучше использовать новые сопоставимые материалы. Самостоятельный проект удобен для обучения процедуре и поиска проблем; его область обобщения должна оставаться видимой.

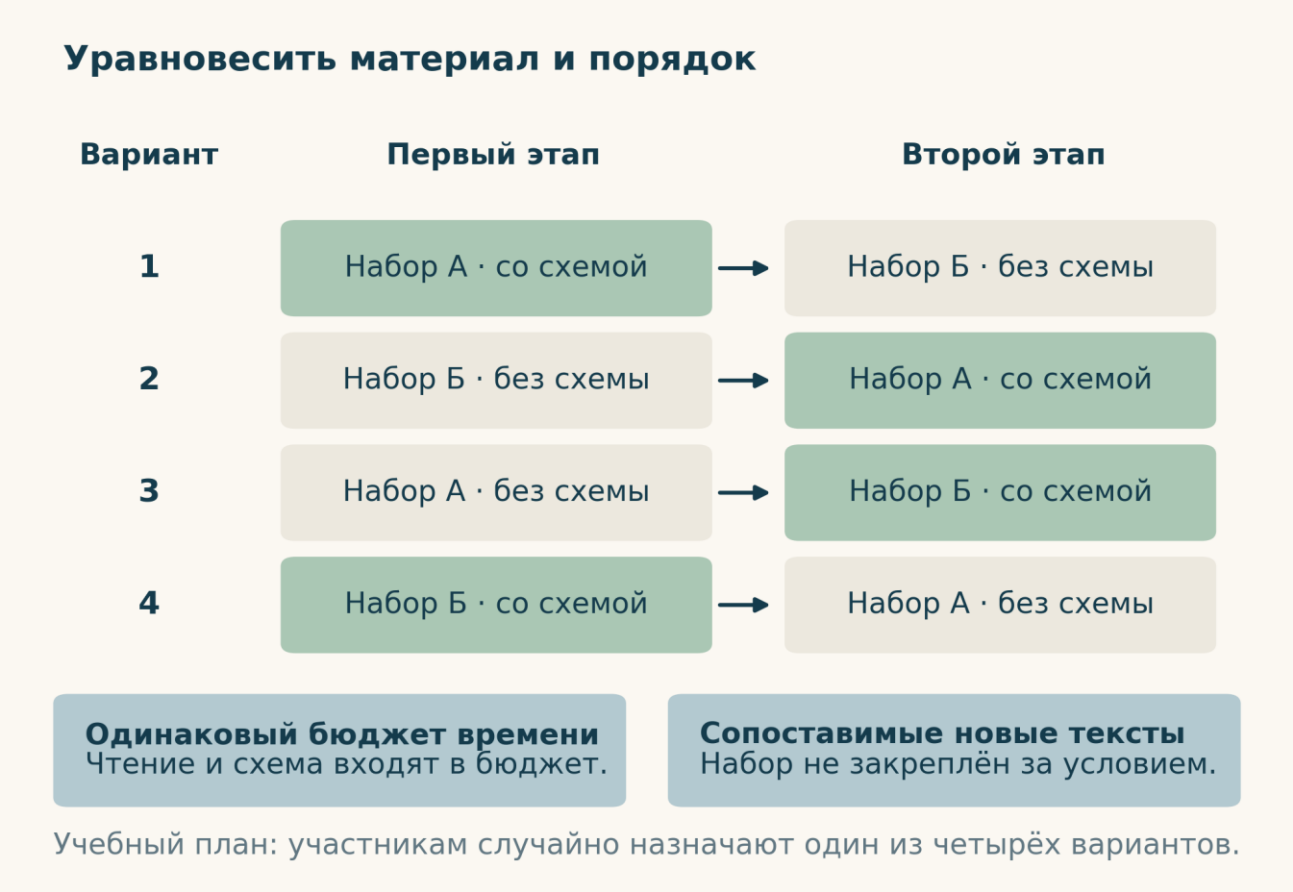

Рисунок 15. Распределение материала и порядка условий. Четыре варианта пересекают два сочетания материала со стратегией и два порядка выполнения. Каждый набор встречается со схемой и без неё, на первом и втором этапах. Случайное назначение вариантов и одинаковый бюджет подготовки дополняют это уравновешивание.

Сделать результат однозначно вычислимым

Допустим, главным показателем выбрана доля правильных ответов на вопросы о временном порядке. Для каждого вопроса заранее нужен ключ. Если ответ свободный, следует определить, какие формулировки засчитываются. Иначе после просмотра результатов появится соблазн строже оценивать неудобные случаи и мягче — ожидаемые. По возможности проверяющий не должен видеть, в каком условии получен ответ.

Общий бюджет времени подготовки установим одинаковым. Это отвечает на практический вопрос, помогает ли схема при ограниченном времени, но не означает одинакового времени непосредственного чтения: часть бюджета уйдёт на запись. Теперь отрицательный результат можно интерпретировать содержательно. Возможно, польза записи не компенсировала затраты при выбранной длительности. Следующая гипотеза могла бы касаться зависимости эффекта от доступного времени.

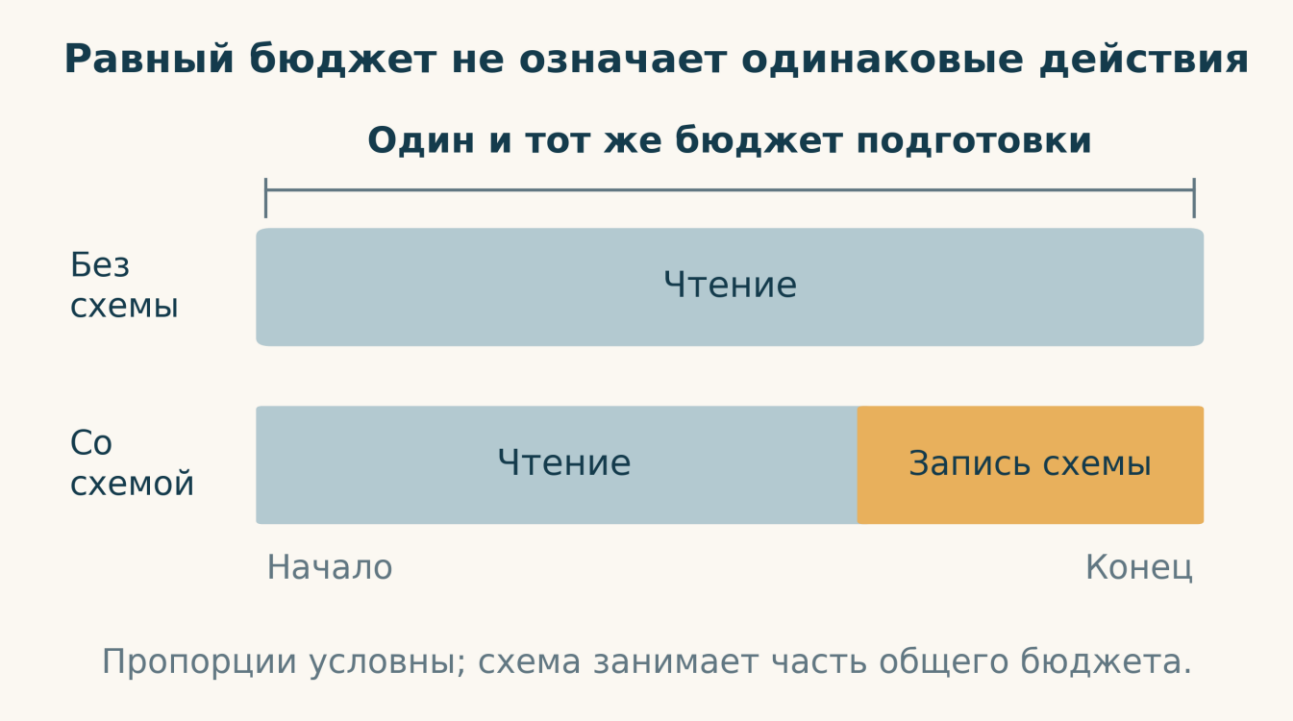

Рисунок 16. Общий бюджет и распределение времени подготовки. При одинаковом общем бюджете подготовка со схемой включает время её записи. Такое сравнение оценивает пользу процедуры при ограниченном времени, но не уравнивает длительность непосредственного чтения; пропорции частей на рисунке условны.

Кроме правильности можно собирать уверенность и описание переживания ясности. Но главный результат лучше выбрать заранее. Если любой из десяти показателей разрешено объявить доказательством после просмотра, процедура проверки теряет определённость. Дополнительные показатели остаются полезными для поиска новых вопросов, если их исследовательская роль названа честно.

План до данных и исследование после данных

Брайан Ноузек и коллеги описывают предварительную регистрацию как способ заранее зафиксировать вопросы и план анализа. Её методологическая функция — сделать различимыми проверку предсказания и объяснение уже увиденного результата. Статья не обещает, что регистрация исправляет слабую теорию или плохое измерение; она посвящена устройству исследовательского процесса. [10]

Для учебного проекта сначала достаточно датированного плана, который вы сохраняете неизменным рядом с результатами. Это тренировка той же дисциплины, а не замена формальной процедуре регистрации. В плане укажите вопрос, условия, основной показатель, правила исключения непригодных записей, число запланированных сессий и способ сравнения. Изменения записывайте отдельно вместе с причиной.

Исследование после просмотра данных тоже необходимо. Вы можете заметить, что схема помогает только на текстах с четырьмя событиями. Это интересная новая гипотеза. Следующий шаг — проверить её на новых текстах, а не представить как единственное заранее ожидавшееся предсказание. Научная работа включает и открытие закономерностей, и их последующую проверку.

Что меняется при уравновешивании условий

Представим, что схема всегда используется со вторым, более лёгким текстом. Даже точное измерение успешности не отделит пользу схемы от простоты и тренировки. Четыре варианта распределения разрывают эту связь: А со схемой затем Б без неё; обратный порядок той же пары; А без схемы затем Б со схемой; обратный порядок второй пары. Теперь каждый набор и каждый порядок встречаются с обоими способами работы. Случайное назначение вариантов не гарантирует тождества людей в малой группе, но не позволяет исследователю систематически отдавать удобный вариант ожидаемо сильным участникам.

Что теперь установлено

План сравнения не должен устранять вообще все различия. Он должен сделать различимой интересующую причину. Названная после результата помеха полезна как ограничение, но не заменяет отсутствовавший контроль.

Практика 7 Исправить слабый эксперимент

Зачем выполнять: Исправить проект, в котором увеличение длины дневниковых записей принято за улучшение различения опыта.

Что понадобится: Паспорт измерения из главы 6 и различение механизма и показателя. Нужны два сопоставимых набора заданий и заранее выбранный результат.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Сформулируйте результат как успешность различения заранее заданных типов новых эпизодов. Отдельно сохраните длину записей как дополнительный показатель.
2. Задайте два сопоставимых набора А и Б, одинаковый бюджет времени и правила оценки до начала работы.
3. Выпишите четыре варианта распределения материала и порядка, используя схему разобранного примера. Назовите условие, учитывающее привычку писать или дополнительное время.

4. Для улучшения, совпадения и разнонаправленных результатов заранее запишите допустимые выводы. Пока достаточно проекта; сбор участников в это упражнение не входит.

Что сохранить: План с условиями, распределением, главным показателем, правилом оценки и тремя вариантами интерпретации.

Подсказка: Польза схемы при равном общем времени и польза схемы при дополнительном времени являются разными вопросами. Выберите один явно.

Исследовательское продолжение: Сохраните датированную версию проекта. После пробного запуска отделите изменения инструкции от результатов основной проверки и сформулируйте новую гипотезу для новых материалов.

Разбор и ответы — приложение Е, глава 7, с. 295

Вопросы для самопроверки

1. Почему одинаковое время подготовки не означает одинаковый процесс подготовки?
2. Что даёт случайное назначение условий и чего оно не гарантирует в маленькой группе?
3. Чем новая гипотеза после анализа отличается от заранее проверяемого предсказания?

Даже хорошо устроенное сравнение даёт меняющиеся числа. В следующей главе разберём величину эффекта и точность оценки, сохраняя их отличие от причинного вывода.

Глава 8

Как работать с неопределённостью

Проект проверки уже определяет, что и с чем сравнивается. Однако два одинаково устроенных запуска могут дать разные доли успеха. Нужно понять, насколько наблюдаемое различие ограничивает возможный устойчивый эффект. Здесь будем отдельно удерживать три вопроса: что получилось в этой серии, насколько точна оценка и какую причину позволяет обсуждать дизайн. Эти различия затем понадобятся для сравнения компьютерных моделей и биологических образцов.

Разница в таблице ещё требует оценки

В одном условии правильно выполнены 14 заданий из 20, в другом — 16 из 20. Доли правильных ответов равны 70 и 80 процентам. Разница составляет 10 процентных пунктов. Это описание наблюдаемого эффекта. Оно ещё не говорит, насколько точно оценена разница, сохранится ли она на новых заданиях и вызвана ли именно изменённым условием.

Процентные пункты удобно отличать от относительного изменения. Переход от 70 к 80 процентам даёт прибавку 10 пунктов, а относительно исходных 70 процентов — примерно 14 процентов. Обе формы могут быть математически правильными, но отвечают на разные вопросы. Для учебного проекта сначала сообщайте исходные количества и доли, чтобы смысл изменения можно было восстановить.

Величина эффекта, неопределённость и причинная интерпретация — три самостоятельные задачи. Большая разница может быть неточно оценена. Маленькая разница может быть измерена очень точно. Даже точная разница между группами не определяет её причину, если устройство сравнения допускает систематические альтернативы. Глава 7 отвечала за качество сравнения; здесь мы учимся понимать изменчивость результата.

Учебная модель случайного разброса

Представим генератор, который в каждом испытании независимо выдаёт правильный ответ с вероятностью 70 процентов. Создадим два условных режима А и Б с совершенно одинаковым правилом. В каждом проведём по двадцать испытаний и повторим такую пару шесть раз. Следующая таблица содержит синтетические данные, полученные для книги программным генератором случайных чисел. Это не результаты опыта с людьми.

Повтор	Верных в А	Верных в Б	Разница Б и А в процентных пунктах
Первый	18	17	минус 5
Второй	14	14	0
Третий	13	15	плюс 10
Четвёртый	19	15	минус 20
Пятый	11	13	плюс 10
Шестой	10	14	плюс 20

Здесь мы знаем механизм генерации: преимуществ у режима Б нет. Тем не менее отдельные сравнения создают впечатление преимущества или ухудшения. Это позволяет понять случайную ошибку оценки без предположения, что кто-то неправильно считал. Все числа могут быть записаны точно, а оценка устойчивого различия всё равно будет колебаться.

Не следует переносить конкретный разброс таблицы на любые исследования. Он зависит от количества испытаний, вероятности успеха и предположения о независимости. В реальной задаче ответы одного человека могут меняться вместе из-за обучения, сложности серии или усталости. Поэтому генератор — инструмент понимания определённой модели случайности. Для реальных данных модель приходится обосновывать.

Что означает точность оценки

Оценка становится точнее, когда процедура лучше ограничивает диапазон правдоподобных значений интересующего эффекта. Для этого часто помогают дополнительные независимые наблюдения и уменьшение ненужной изменчивости измерения. Однако повторение одного и того же недостаточно, если все повторы несут одну общую ошибку. Тысяча измерений неправильно настроенным прибором не исправляет его настройку.

В отчётах встречаются доверительные интервалы. В частотном подходе уровень 95 процентов характеризует процедуру построения: при её многократном применении в условиях принятой модели примерно 95 процентов полученных интервалов должны накрывать фиксированное истинное значение. Это не автоматическая вероятность того, что истинное значение лежит внутри уже вычисленного конкретного интервала. Для такой вероятностной формулировки нужен иной, явно заданный подход.

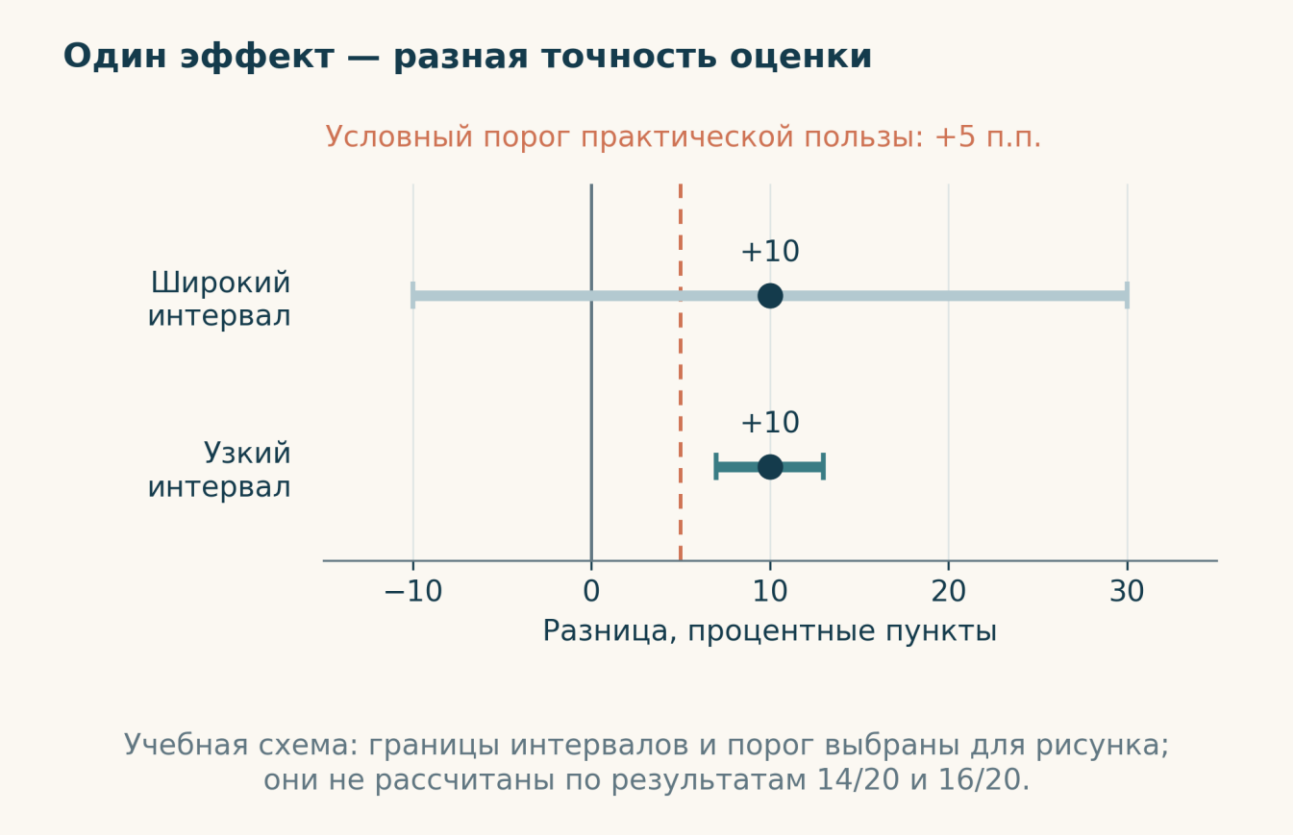

Рисунок 17. Величина эффекта и точность его оценки. В обоих учебных сценариях точечная оценка равна +10 процентным пунктам. Широкий интервал от −10 до +30 допускает ухудшение и улучшение; узкий от +7 до +13 расположен выше условного порога +5. Границы и порог выбраны для объяснения и не являются доверительными интервалами, рассчитанными из таблицы главы.

На первом этапе достаточно задавать практический вопрос: совместимы ли данные и с заметной пользой, и с отсутствием пользы, и с ухудшением при принятом способе оценки? Если да, результат остаётся неопределённым для выбора между этими вариантами. Если широкий диапазон исключён, вывод может быть содержательным даже при отсутствии ожидаемого улучшения. Точность позволяет ограничивать теории, а не только объявлять успехи.

Практически значимую величину тоже нужно определить содержательно. Если схема требует дополнительного времени и заметно усложняет работу, прибавка одного верного ответа на тысячу заданий может не оправдывать её использование. В другой задаче даже такая прибавка будет полезной. Статистическая точность не принимает это решение за исследователя: она показывает, насколько хорошо известен эффект, а оценка пользы связывает его с целью применения.

Для гипотез о механизмах полезность может заключаться и в небольшом, но специфическом различии. Например, две модели одинаково решают обычные задачи, однако по-разному реагируют на определённую задержку. Если это заранее предсказанное различие надёжно измерено, оно помогает выбирать

между механизмами, даже когда почти не меняет общую успешность. Поэтому критерий содержательной важности следует формулировать вместе с исследовательским вопросом.

Не путать число ответов с числом независимых случаев

Допустим, один человек выполнил сто заданий, а другой проект собрал по одному ответу у ста человек. В обоих файлах сто строк, но они отвечают на разные вопросы. Первый лучше описывает изменчивость ответов данного человека в выбранной серии. Второй содержит сведения о различиях между людьми, хотя каждый человек измерен очень кратко.

Если все сто ответов одного участника рассматривать как сто независимых подтверждений свойства людей вообще, неопределённость может быть существенно недооценена. Нужно заранее определить единицу вывода: испытание, сессию, человека, организацию модели. Для вычислительных проектов аналогичная проблема возникает, когда множество выходов одной обученной сети принимают за множество независимо созданных систем.

Это не означает, что повторы бесполезны. Они помогают оценивать стабильность внутри выбранной единицы. Важно сохранить структуру данных: какие ответы относятся к одному участнику, запуску или набору материалов. Тогда анализ может учитывать зависимость вместо того, чтобы прятать её в плоской таблице.



Рисунок 18. Ответы и участники как разные уровни данных. Слева сто ответов принадлежат одному человеку, справа каждый ответ получен от отдельного участника. Одинаковое число строк не делает эти планы взаимозаменяемыми: они дают сведения об изменчивости на разных уровнях.

Почему множество попыток меняет смысл находки

Представим исследователя, который попробовал двадцать показателей и сообщил только о самом впечатляющем различии. Даже если каждое вычисление выполнено правильно, итоговая история не показывает весь поиск. Читатель видит результат так, будто именно его ожидали с самого начала. Оценка неожиданности должна учитывать возможность выбора удачного варианта из многих.

Джозеф Симмонс, Лейф Нельсон и Ури Симонсон с помощью симуляций и экспериментальных примеров показали, как гибкость сбора и анализа позволяет получать убедительно выглядящие статистические результаты для ложных гипотез. Их работа связывает проблему с нераскрытыми исследовательскими решениями и предлагает повышать прозрачность процедуры. [11]

Значение p, если вы встретите его в статье, не является вероятностью ложности гипотезы. Оно оценивает, насколько необычны данные не менее экстремального вида в рамках указанной нулевой модели и процедуры. Поэтому маленькое значение не заменяет оценку эффекта, проверку предположений и описание выполненного поиска. Для собственной первой

работы полезнее ясный план и полные результаты, чем одно впечатляющее число.

Одинаковая оценка может разрешать разные решения

Пусть в двух условных сценариях оценка улучшения равна десяти процентным пунктам. Интервал от минус десяти до плюс тридцати допускает и ухудшение, и существенную пользу. Интервал от плюс семи до плюс тринадцати ограничивает возможные значения заметно сильнее. Если заранее полезным считалось улучшение не менее пяти пунктов, второй сценарий отвечает на этот вопрос определённее. Границы в этом примере выбраны для объяснения и не рассчитаны по таблице 14 из 20 и 16 из 20. Нельзя использовать наглядную схему как результат статистического анализа этих сорока ответов.

Что теперь установлено

Мы научились различать размер результата и точность знания о нём. Большое количество строк помогает лишь с учётом того, какие строки действительно независимы и какой вопрос ставится перед оценкой.

Практика 8 Три разных вывода

Зачем выполнять: Разделить описание одной серии, вывод о механизме генерации и план следующей проверки.

Что понадобится: Доли, проценты и устройство сравнения из главы 7. Доверительный интервал сначала рассматривается как характеристика процедуры оценивания; выводить его формулу для практики не требуется.

Ориентир по времени: 20–30 минут. Можно делать с перерывами; это не норматив скорости.

1. В синтетической таблице главы найдите четвёртый повтор: А 19 из 20, Б 15 из 20; шестой: А 10 из 20, Б 14 из 20. Рассчитайте разности Б минус А в процентных пунктах.
2. Запишите рассказ, который мог бы возникнуть при показе только одной из этих строк. Затем объясните, почему он не устанавливает преимущество режима генератора.
3. Для большого неточного эффекта, точной оценки около нуля и различий между материалами предложите по одному следующему действию.
4. Укажите единицу независимого повторения для ста ответов одного человека и для одного ответа ста человек. Сохраните эти различия в плане анализа.

Что сохранить: Две разности, ограничение их интерпретации и три обоснованных следующих шага.

Подсказка: Для доли 19 из 20 получается 95 процентов. Вычитание долей в процентах даёт процентные пункты, а не относительный рост.

Исследовательское продолжение: Заранее выберите практически интересующую величину и точность, необходимую для её различения. Опишите, какие зависимые наблюдения нельзя считать отдельными повторениями всего исследования.

Разбор и ответы — приложение Е, глава 8, с. 296

Вопросы для самопроверки

1. Почему разница 80 и 70 процентов равна десяти процентным пунктам, а не обязательно десятипроцентному относительному росту?
2. Почему сто ответов одной модели не равнозначны исследованию ста независимо обученных моделей?
3. Может ли отсутствие ожидаемого улучшения быть содержательным результатом?

Теперь сравнение имеет ясный показатель и границы неопределённости. Перейдём к механизму, который сам производит последовательность результатов: внутреннее состояние будет связывать один момент с другим.

Часть 3

Механизмы и модели

Глава 9

От нервной клетки к системе

Предыдущие главы учили проверять объяснения. Теперь построим небольшое объяснение, работу которого можно проследить целиком. Один и тот же вход иногда должен вызывать разные ответы, потому что система сохраняет историю. Пороговый накопитель позволит увидеть это без знания нейробиологии и сложной математики. Его связь с нервной клеткой ограничена выбранными свойствами; дальнейший вопрос об опыте потребует новых отношений и аргументов.

Начать с изменяющегося состояния

До сих пор мы обсуждали, как описывать опыт и проверять утверждения. Теперь перейдём к системам, которые обрабатывают сигналы. Главный новый вопрос звучит так: как вход в данный момент вместе с внутренним состоянием определяет следующий шаг? Он подходит и для биологического процесса, и для учебной программы. Ответ позволяет объяснить, почему одинаковый вход иногда приводит к разным результатам.

Нейрон — клетка нервной системы. Его мембрана отделяет внутреннюю среду от внешней, а распределение и движение заряженных частиц связано с электрическими свойствами клетки. Многие нейроны передают сигналы с помощью потенциалов действия — коротких изменений напряжения на мембране, распространяющихся по соответствующим участкам клетки. Для понимания этой главы достаточно видеть в нейроне физическую систему с входами, внутренней динамикой и воздействием на другие клетки.

Это ещё не модель всего поведения нейрона. Разные клетки и их части обладают различными свойствами; взаимодействия включают химические процессы и множество временных масштабов. Слово «нейрон» в названии программного элемента обычно обозначает определённую абстракцию. Поэтому при чтении модели спрашивайте, какие свойства сохранены, а какие опущены ради выбранной задачи.

Как количественное объяснение отличается от образа

Алан Ходжкин и Эндрю Хаксли построили количественное описание мембранных токов нервного волокна, связав изменения проводимости с электрическим возбуждением. Их работа показывает исследовательский переход от измерений к системе правил, способной воспроизводить динамику процесса. Это модель определённых электрических свойств нервной клетки; она не предлагалась как доказательство возникновения субъективного опыта. [12]

Полезный образ вроде «нейрон собирает голоса соседей» помогает начать, но скрывает важные вопросы. Как долго сохраняется влияние входа? Зависит ли следующий ответ от предыдущего? Есть ли порог? Что происходит после ответа? Вычислимая модель вынуждает отвечать на эти вопросы достаточно определённо, чтобы можно было получить результат без нового решения автора на каждом шаге.

Наша первая модель будет гораздо проще биологической. Она нужна для изучения накопления, утечки и порога. Её числа условны и не измеряются в милливольтах. Сходство с некоторыми свойствами нейрона не даёт права переносить на неё все свойства живой клетки.

Разобранный пример порогового накопителя

Представим переменную состояния, первоначально равную нулю. На каждом шаге сохраняется 60 процентов её прежнего значения и добавляется новый вход. Если получившееся число достигает единицы, система выдаёт импульс и сбрасывает состояние к нулю. Если порог не достигнут, импульса нет, а состояние сохраняется до следующего шага с последующей утечкой.

Подадим входы 0,6; 0,6; 0,6; 0; 0,6. На первом шаге прежнего накопления нет, поэтому новое значение равно 0,6. На втором остаётся 0,36 от прошлого состояния и добавляется 0,6: получается 0,96. На третьем сохраняется 0,576 и добавляется ещё 0,6. Значение 1,176 пересекает порог, возникает импульс, после которого состояние равно нулю.

На четвёртом шаге вход нулевой, поэтому состояние остаётся нулевым. На пятом оно снова достигает 0,6. Одинаковые входы первого, второго и третьего шагов привели к разным состояниям, потому что различалась история. Никакого дополнительного решения «когда отвечать» в программе нет: результат получается из правила обновления.

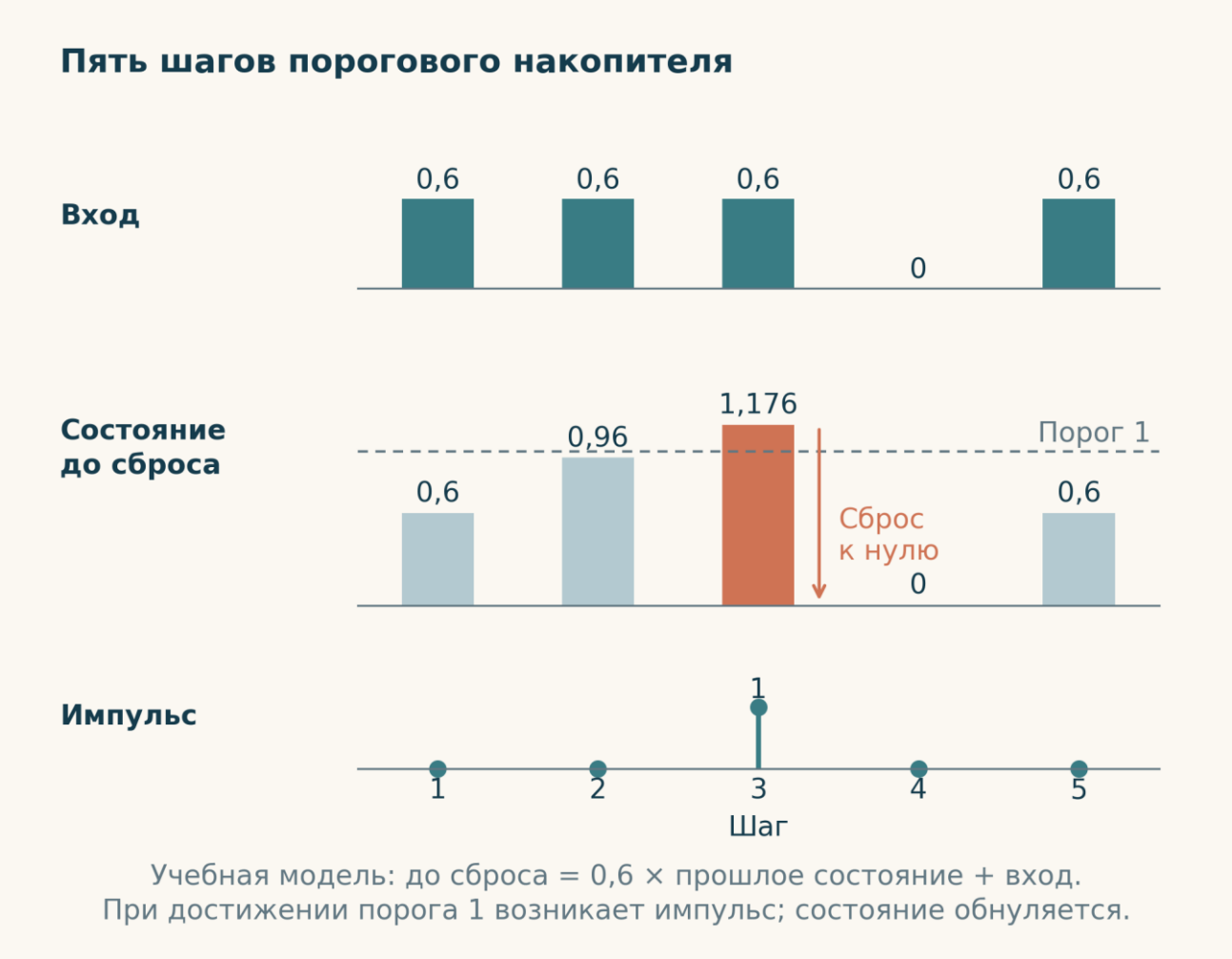

Рисунок 19. Динамика накопления и сброса. Точные результаты программы из главы. Входы 0,6; 0,6; 0,6; 0; 0,6 дают состояния перед сбросом 0,6; 0,96; 1,176; 0; 0,6. Только на третьем шаге достигнут порог 1. Все величины условны и не выражены в милливольтах.

Вот полный вычислимый вариант на Python. Слово state обозначает состояние, value — очередной вход, spike — наличие импульса. Точка используется в коде как десятичный разделитель. Отступы показывают, какие действия повторяются для каждого входа и какие выполняются только при пересечении порога.

```
inputs = [0.6, 0.6, 0.6, 0.0, 0.6]
state = 0.0
for value in inputs:
    before_reset = 0.6 * state + value
    spike = before_reset >= 1.0
    state = 0.0 if spike else before_reset
    print(round(before_reset, 3), int(spike), round(state, 3))
```

Программа выводит значение до возможного сброса, затем 1 при импульсе или 0 при его отсутствии, затем итоговое состояние. Её можно выполнить или проверить карандашом. В обоих случаях правило должно давать одну последовательность. Такая прозрачность полезна до перехода к большим

моделям, внутреннюю работу которых уже нельзя проследить вручную целиком.

От одного элемента к сети

Теперь представим несколько накопителей. Выход одного входит в правило обновления другого. Коэффициент связи определяет силу этого влияния. Положительный вклад может приближать состояние к порогу, отрицательный — отдалять. Это учебная схема возбуждающих и тормозящих воздействий, а не универсальное описание всех биологических взаимодействий.

Если связи идут только от входных элементов к выходным, получается направленная последовательность обработки. Если сигнал может через несколько связей вернуться к ранее участвовавшему элементу, возникает обратная связь. Она позволяет предыдущим результатам менять дальнейшую обработку. Но обратная связь не обязательно поддерживает устойчивую активность: в зависимости от правил она может затухать, усиливаться, колебаться или приводить к другим режимам.

Важно различать изменение активности и изменение самой системы. Наше состояние меняется на каждом шаге, а коэффициент 0,6 остаётся постоянным. Если добавить обучение, коэффициенты тоже могут изменяться по отдельному правилу. Тогда у модели появятся два вида истории: кратковременная история текущего состояния и более длительное изменение параметров. Смешивание этих уровней мешает понять, что именно модель запомнила.

Сеть не становится объяснением только потому, что содержит много элементов. Нужно показать, какую задачу решает их организация. Например, одна сеть может сохранять сигнал между событиями, другая — выделять последовательность определённого вида, третья — выбирать действие. Эти различия можно проверять до постановки вопроса об отношении организации к переживанию.

Число или импульс получает исследовательский смысл через отношения к условиям и использованию. Если элемент чаще отвечает на красные изображения, можно проверить, сохраняется ли связь при изменении формы, яркости и окружения. Затем нужно выяснить, использует ли другой компонент этот ответ для различения цвета. Название «детектор красного» будет преждевременным, если элемент на самом деле реагирует на яркость, которая случайно совпала с цветом в учебной выборке. Представление содержания требует проверки, а не только удачной подписи к узлу сети.

Один шаг состоит из нескольких действий

Если прежнее состояние равно 0,96, сначала сохраняем 0,6 от него: получается 0,576. Затем добавляем новый вход 0,6 и получаем 1,176. Только теперь проверяем порог 1. Он достигнут, поэтому записываем импульс и сбрасываем состояние к нулю. Величина до сброса и состояние после сброса различны. Если поменять порядок действий, получится другая модель. Именно поэтому строка с итоговым числом не заменяет явного правила перехода.

Что теперь установлено

Накопитель показывает причинный вклад сохранённого состояния. Импульс возникает из правила, а не из дополнительного внутреннего решения. Теперь проверим, какое значение имеет распределение одинаковой суммы входов во времени.

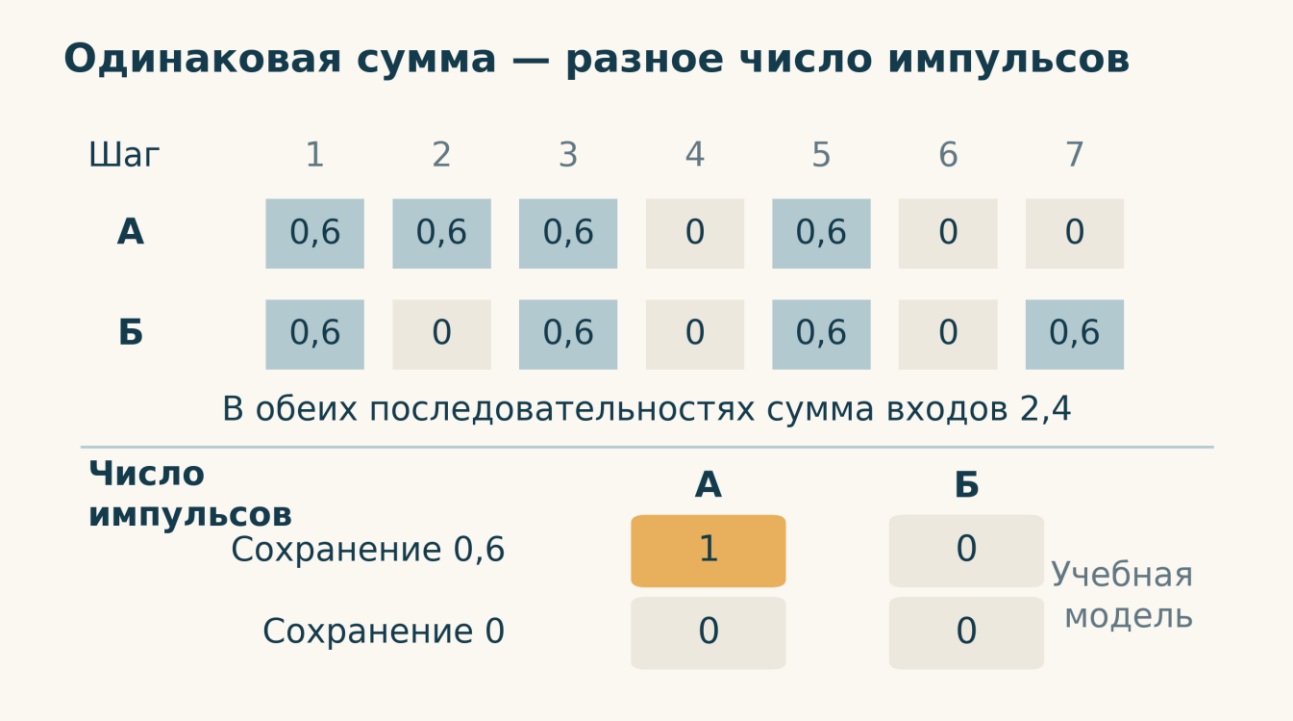

Рисунок 20. Распределение входов и отключение сохранения состояния. Две последовательности из практики имеют одинаковую сумму 2,4, но при сохранении 0,6 только первая вызывает один импульс. При нулевом сохранении импульсов нет в обеих сериях: вмешательство выделяет роль истории состояния в этой учебной модели.

Практика 9 Одинаковая сумма разное время

Зачем выполнять: Проверить объяснение через временное накопление в сравнении с объяснением через одну общую сумму.

Что понадобится: Умение умножать десятичные числа. Вход приходит извне, состояние сохраняется между шагами, выход получается после обновления.

Ориентир по времени: 20–30 минут. Можно делать с перерывами; это не норматив скорости.

1. Для А используйте 0,6; 0,6; 0,6; 0; 0,6; 0; 0. Для Б используйте 0,6; 0; 0,6; 0; 0,6; 0; 0,6. Начальное состояние ноль, сохранение 0,6, порог 1.
2. До расчёта предскажите число импульсов. Для каждого шага выпишите вход, величину до сброса, импульс и состояние после сброса.
3. Выполните одинаковое правило для обеих последовательностей вручную или кодом из главы.
4. Повторите с сохранением ноль и объясните, какое различие исчезло и почему.

Что сохранить: Две пошаговые таблицы, предварительный прогноз и вывод о роли промежутков между входами.

Подсказка: После импульса следующий шаг начинается с нуля. Не переносите величину 1,176 как сохранённое состояние.

Исследовательское продолжение: Предложите новую пару последовательностей с равной суммой, на которой версии о сумме и временной структуре дают разные ожидания. Зафиксируйте правило до расчёта.

Разбор и ответы — приложение Е, глава 9, с. 296

Какой шаг к исследованию опыта сделан

Мы получили минимальный язык динамических моделей: вход, состояние, правило перехода, выход и параметр. С его помощью можно спрашивать, что система различает, что удерживает и как использует прошлое. Позже мы добавим представления источников, оценку надёжности и собственных действий. Каждый компонент будет вводиться ради объясняемого поведения и проверяемого предсказания.

Присутствие импульсов или обратной связи само по себе не решает вопрос о квалиа. Однако теперь можно сформулировать гипотезу точнее, чем «сложность рождает сознание»: назвать определённую организацию, указать ожидаемые отношения между её состояниями и свойствами опыта, затем искать условия различения альтернатив. В этом состоит подготовительная роль простых моделей.

Вопросы для самопроверки

1. Почему одинаковый вход на первом и третьем шагах может дать разный выход?
2. Чем изменение состояния отличается от обучения параметров?
3. Что проверяет сравнение последовательностей с одинаковой суммой и разными промежутками между входами?

Состояние научилось удерживать прошлое. Следующий вопрос: что именно оно представляет и как контекст меняет использование одного и того же входа.

Глава 10

Обработка сигналов и доступ к их содержанию

Накопитель показал, как история меняет ответ. Для восприятия этого недостаточно: необходимо различить содержание, способ интерпретации и доступ к результату. Одинаковый овал может использоваться как буква или цифра в зависимости от соседей. На этом случае разберём прямую и повторную обработку. Сравнивая их, будем следить, чтобы архитектура не получала преимущество только за счёт дополнительных входов или времени.

Одно изображение и несколько исследовательских вопросов

Вы смотрите на страницу и ищете номер нужного раздела. На ней одновременно находятся буквы, поля, знаки препинания и мелкая подпись. Наличие всех этих элементов перед глазами не означает, что каждый из них одинаково влияет на текущую задачу или одинаково доступен для последующего описания. Чтобы исследовать ситуацию, нужно разделить несколько понятий.

Обработка информации означает изменение состояний системы в зависимости от входа и внутренних правил. Внимание в рабочем функциональном смысле относится к избирательному распределению обработки: одни источники или признаки получают преимущество перед другими. Доступ означает возможность использовать содержание в определённых действиях, рассуждениях или ответах. Отчёт — конкретное действие сообщения. Переживание относится к тому, как ситуация субъективно дана.

Эти понятия могут относиться к одному эпизоду, но измеряются по-разному. Мы можем проверить, изменился ли ответ на номер раздела после подсказки. Можем спросить, что было ясно видно. Можем выяснить, сохранилось ли расположение подписи для следующего задания. Если все результаты объединить словом «осознание», станет трудно понять, какой именно процесс объясняет модель.

Почему объём отчёта зависит от процедуры

Джордж Сперлинг исследовал краткие предъявления массивов символов, сравнивая полный отчёт с отчётом об указанной после предъявления части. Преимущество частичного отчёта и его зависимость от задержки указания показывали, что доступность информации после краткого показа нельзя

оценивать только количеством последовательно названных символов. Эксперимент ставит вопрос об ограничениях хранения и сообщения; сам по себе он не завершает спор о полном объёме феноменального опыта. [13]

Общий урок можно понять на обычном примере. Вы замечаете несколько предметов на столе, затем начинаете подробно описывать первый. За время рассказа сведения о других могут стать менее доступными. Короткий итоговый список тогда зависит не только от исходной обработки, но и от времени, порядка сообщения и требований к деталям. В конкретном исследовании вклад этих процессов следует разделять.

Обратная ситуация тоже возможна: правильный ответ может опираться на сведения, которые человек затрудняется описать словами. Поэтому вопрос «сколько удалось рассказать» полезен для изучения отчёта, но требует дополнительных оснований как мера всего доступного содержания. Лучше уточнять, для какой операции информация оказалась доступна и в какой момент.

Внимание меняет задачу

Дэниел Саймонс и Кристофер Шабри исследовали замечание неожиданного события во время наблюдения за динамической сценой с заданием на внимание. Вероятность его обнаружения зависела от устройства основной задачи и свойств неожиданного объекта. Работа демонстрирует ограниченность замечания при конкретной нагрузке; отсутствие последующего сообщения не следует превращать в универсальный критерий отсутствия любой обработки. [14]

Внимание удобно понимать через вопрос, какие сведения система предпочитает использовать сейчас. При поиске номера раздела форма цифр может быть полезнее содержания соседнего абзаца. При проверке опечаток приоритет меняется. Одни и те же внешние элементы участвуют в разных вычислительных задачах в зависимости от цели.

Однако внимание и сознание не становятся синонимами только из-за тесных связей. Сначала нужно выбрать показатель внимания, показатель опыта и условия их сопоставления. Если внимание определено через успешность поиска, а опыт — через отчёт о ясности, исследуется отношение именно этих показателей. Содержательная теория должна объяснять их совпадения и возможные расхождения, а не устранять различие переименованием.

Разобранный пример неоднозначного знака

Представьте нарисованный овал, который может обозначать букву О или цифру 0. Сам рисунок одинаков в двух учебных карточках. На первой вокруг него стоят К и Т, на второй — 1 и 4. В первом окружении естественна версия буквы, во втором — цифры. Контекст добавляет сведения о том, какую роль знак, вероятно, играет в записи.

Построим простую модель. Первый этап оценивает только центральный знак и сохраняет два варианта. Второй рассматривает соседей. Если оба соседа буквы, преимущество получает буквенная версия; если оба цифры — цифровая. Третий выдаёт выбранную категорию. Здесь можно указать, какие сведения поступили на каждом этапе и почему итог различается при одном центральном рисунке.

Это авторский учебный пример, а не воспроизведение конкретного эксперимента. Он показывает, как контекст может участвовать в интерпретации входа. Изменилось ли при этом субъективное видение формы или только её обозначение, требует отдельного вопроса к наблюдателю. Одинаковый ответ «буква» может сопровождаться разными описаниями опыта, и модель категории ещё не объясняет эти различия.

Один проход и повторное уточнение

При прямой обработке сведения проходят через последовательность преобразований к ответу. При рекуррентной обработке результаты могут возвращаться и изменять работу ранее участвовавших компонентов. Слово «рекуррентный» здесь означает наличие такого возврата. Например, предварительно распознанный контекст может влиять на дальнейшую оценку локальной формы.

В нашей задаче можно построить рекуррентную версию: сначала сохранить неоднозначное состояние центрального знака, затем получить сведения о соседях и обновить прежнее состояние. Но можно также создать модель, которая сразу получает всю карточку и одним правилом выбирает категорию. На двух описанных примерах их ответы совпадут. Поэтому совпадение с ответами людей ещё не определяет внутреннюю организацию.

Чтобы сравнение архитектур было содержательным, нужно учитывать объём доступных сведений и вычислительные ресурсы. Если одной модели показали всю карточку, а другой только овал, преимущество первой объясняется уже различием входов. Если повторной обработке дали больше времени, требуется выяснить, что обеспечило выигрыш: обратная связь или дополнительные вычисления. Хорошая задача позволяет этим объяснениям расходиться.

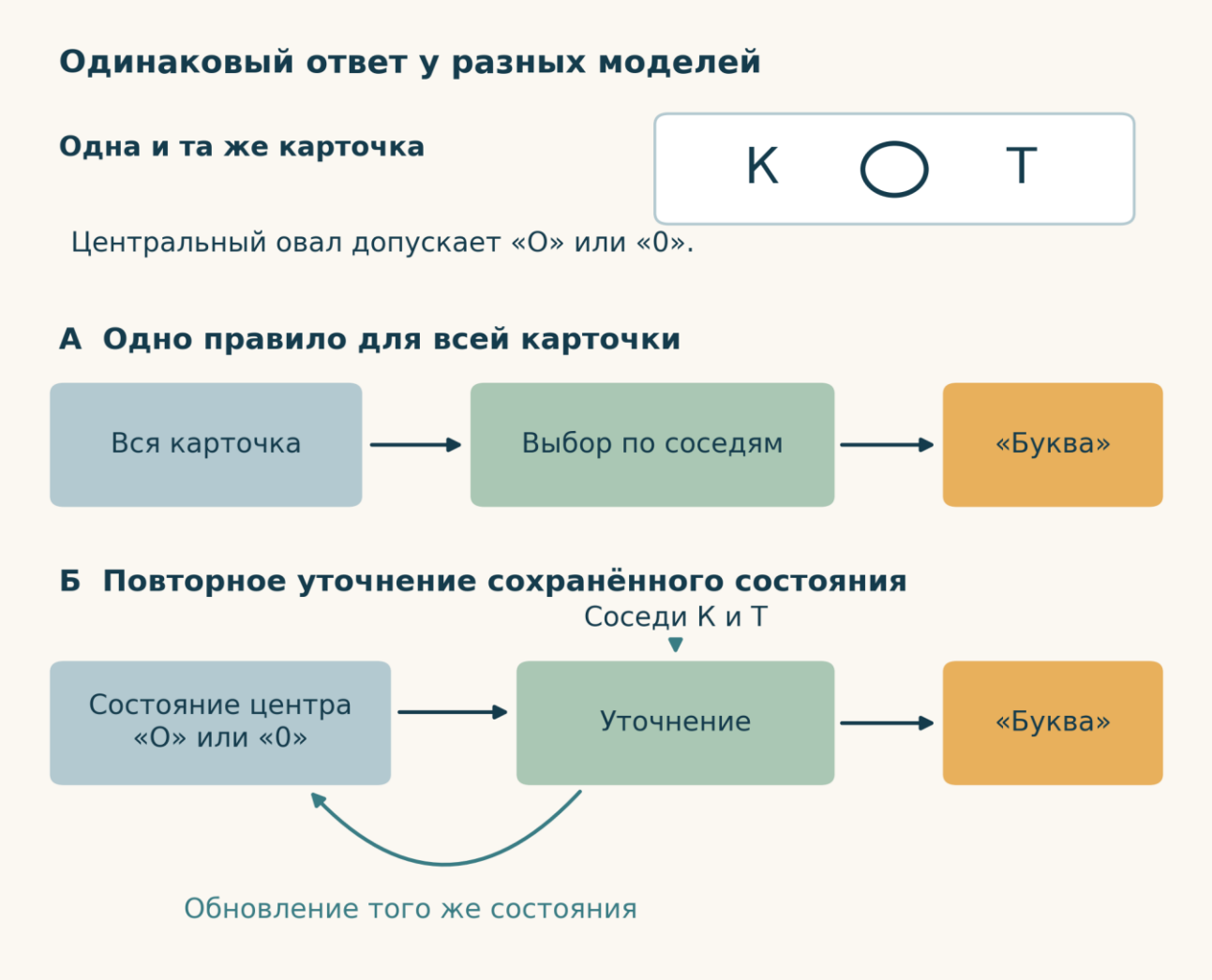

Рисунок 21. Два способа использовать контекст. Обе учебные модели называют центральный овал буквой при соседях K и T. Первая использует всю карточку одним правилом. Вторая сохраняет неоднозначное состояние центра, получает контекст и уточняет прежнее состояние; возвратная стрелка обозначает это обновление. Равенство ответов в примере не определяет внутреннюю организацию и не доказывает наличие переживания.

От внутреннего состояния к использованию

Допустим, выбранная категория сохранена внутри программы. Один выход может напечатать слово «буква», другой — выбрать словарь для продолжения записи, третий — решить, нужна ли дополнительная проверка. Эти возможности показывают разные формы доступа к одному результату обработки. Наличие внутреннего числа ещё не означает, что оно действительно используется всеми компонентами.

Чтобы проверить роль состояния, его можно изменить или временно сделать недоступным определённому выходу. Если печать ответа нарушилась, а выбор продолжения сохранился, возникло функциональное расхождение. В

исследовании важно описать его буквально. Нельзя объявлять систему полностью лишённой информации только потому, что сломан один канал сообщения.

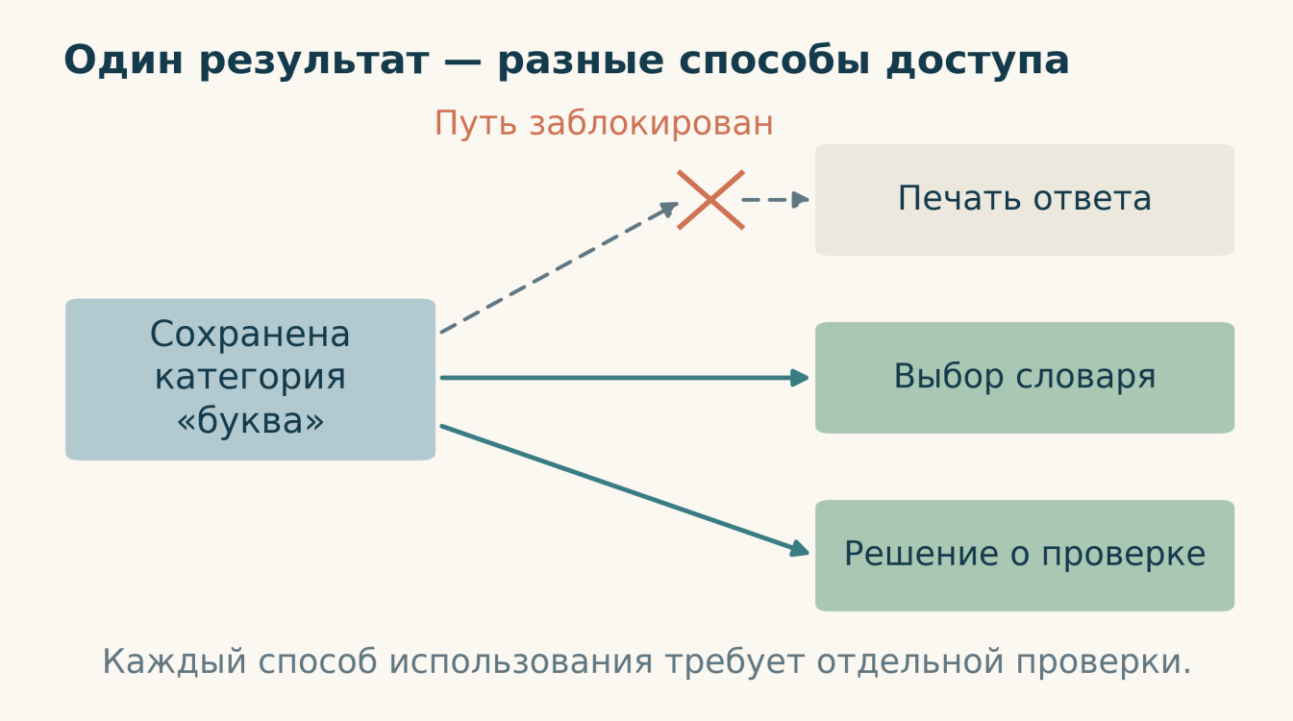

Рисунок 22. Избирательное нарушение канала использования. В учебной архитектуре выбранная категория доступна печати, выбору словаря и решению о проверке. Блокируется только путь к печати; сохранность других способов использования проверяется отдельно и не выводится из одного отсутствующего сообщения.

Подобный разбор подготавливает проектирование кандидатов на субъекта. Мы сможем спрашивать, какая информация объединяется, какие действия от неё зависят и как система оценивает собственное решение. Отношение этой организации к опыту затем станет отдельной проверяемой частью теории.

Почему скрытая правильная метка не всегда восстановима

На двух карточках нарисованы одинаковые овал и цифровые соседи. Автор в одном случае задумал цифру, в другом букву. Если устройство видит только изображение, наблюдения не различают эти замыслы. Ошибка относительно скрытой метки здесь не обязана означать плохую обработку. Чтобы исследовать пользу контекста, нужно заранее выбрать: оценивается наиболее обоснованная интерпретация по доступным сведениям или угадывание недоступного намерения. Это различие пригодится при дальнейшем аудите входов агента.

Что теперь установлено

Правильное название знака, переживаемая форма и доступность содержания для разных действий являются разными результатами. Теперь сравним объяснения на одном явно заданном наборе карточек.

Практика 10 Сравнить три объяснения

Зачем выполнять: Отделить роль контекста от преимущества информации и от недоступной метки среды.

Что понадобится: Вход, состояние и выход из главы 9. Различайте наблюдаемый рисунок и скрытую авторскую метку, которую устройство может не знать.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Сделайте шесть карточек с одинаковым центральным овалом: две с соседями К и Т, две с 1 и 4, ещё две повторяют эти изображения, но имеют противоположные скрытые авторские метки.
2. Для каждой запишите видимые сведения и отдельно скрытую метку. Задайте базовое правило: буквенные соседи ведут к категории буквы, цифровые к категории цифры.
3. Сравните устройство, видящее всю карточку сразу, и устройство, получающее сначала центр, затем соседей. Укажите момент, когда у них появляется одинаковая информация.
4. Объясните, что можно установить по совпадающим ответам и какие дополнительные условия понадобятся, чтобы различить механизм повторного уточнения.

Что сохранить: Шесть карточек, правила двух устройств и явный перечень доступных сведений по времени.

Подсказка: Разный порядок поступления входов уже является различием условий. Для проверки архитектуры нужно предоставить конкуренту сопоставимую память и вычисления.

Исследовательское продолжение: Пусть соседи исчезают до появления центра. Задайте память каждой версии и получите прогноз до пробы. Сравните с прямой моделью, которой разрешена такая же внешняя память.

Разбор и ответы — приложение Е, глава 10, с. 297

Вопросы для самопроверки

1. Почему количество названных деталей не является автоматической мерой всего исходного опыта?
2. Почему успех модели с контекстом не доказывает необходимость рекуррентной обработки?
3. Что нужно уточнить в утверждении «информация доступна системе»?

Мы отделили наблюдение от скрытого замысла. Теперь применим эту дисциплину к источникам событий: как система может оценить, связано ли наблюдаемое изменение с её собственной командой.

Глава 11

Тело и действие в самоотнесении

Контекст помогает интерпретировать знак, но в действии требуется ещё оценить источник изменения. Совпало ли движение с командой случайно или команда действительно им управляет? В этой главе будут две связанные задачи: объединение сенсорных причин и различение собственного воздействия. Перед числовым примером разберём вероятностное обновление как правило изменения версии, сохраняя его отличие от переживаемой телесной принадлежности.

Когда несколько сигналов относятся к одному событию

Вы поднимаете руку и одновременно видите её движение, чувствуете изменение положения и получаете новые ощущения от одежды. Для описания этого эпизода недостаточно сказать, что поступило много сигналов. Нужно объяснить, почему они связываются с одной частью тела и одним действием. Похожее движение на экране рядом может оставаться внешним событием, хотя тоже происходит перед глазами.

Один подход рассматривает задачу как вывод об источниках. Система должна оценить, связаны ли зрительные, телесные и двигательные сведения с общей причиной или возникли независимо. Такая оценка может учитывать временное совпадение, пространственное соответствие и прошлую надёжность источников. Слово «вывод» здесь не требует внутренней фразы или сознательного рассуждения: оно обозначает работу модели, которая различает возможные объяснения входа.

Маджед Самад, Альберт Джин Чанг и Ладан Шамс применили модель байесовского причинного вывода к восприятию принадлежности тела. Они сопоставили предсказания модели с экспериментами, посвящёнными иллюзии искусственной руки. Работа показывает, как предположение об объединении сенсорных причин можно сделать вычислимым и проверить на данных; она не устанавливает, что любое объединение сигналов уже является переживанием собственного тела. [15]



Рисунок 23. Рука инструмент и наблюдаемый результат. Предметная иллюстрация действия через инструмент: движение руки сопровождается изменением рисунка на экране. Для исследования авторства и самоотнесения нужно уточнить, какие сведения система связывает с одной причиной и как она проверяет эту связь.

Общая причина ещё не означает моё тело

Звук удара и движение падающей книги тоже могут иметь общую причину. Поэтому модель объединения источников решает более общую задачу, чем телесная принадлежность. Для перехода к телу нужно уточнять, какие сигналы относятся к нему, как представлено его положение и какие отношения связывают события с действиями системы.

Вернёмся к различению главы 3. Принадлежность отвечает на вопрос, дан ли объект как часть собственного тела. Авторство относится к переживаемой или оцениваемой связи действия с его результатом. Система может управлять внешним инструментом и сохранять различие между инструментом и телом. Она может получать сведения о положении неподвижной конечности, не выполняя в этот момент произвольного движения.

Работа Калькерта и Эрссона, рассмотренная ранее, даёт экспериментальное основание изучать эти компоненты раздельно. [4] В нашей модели поэтому будут отдельные переменные для объединения сенсорных источников и оценки причинного участия команды. Даже если в обычных эпизодах они

часто меняются вместе, их разделение позволяет сформулировать новые проверяемые условия.

Почему свидетельство меняет вес версии

Сначала различим два вопроса. Насколько вероятна версия С до нового наблюдения? Насколько вероятно само наблюдение, если С верна? Первая величина является предварительной вероятностью версии. Вторая называется правдоподобием наблюдения при этой версии. Они связаны расчётом, но не взаимозаменяемы.

Пусть из ста условных случаев шестьдесят имеют правую цель. Датчик верен в восьмидесяти процентах случаев независимо от стороны. Тогда среди шестидесяти правых целей ожидаются сорок восемь сигналов справа, среди сорока левых восемь таких сигналов. Если мы получили сигнал справа, рассматриваем именно эти пятьдесят шесть случаев. В сорока восьми из них цель действительно справа. Новая доля равна 48/56. Числа здесь описывают выбранную учебную вероятностную модель, а не обследование людей.

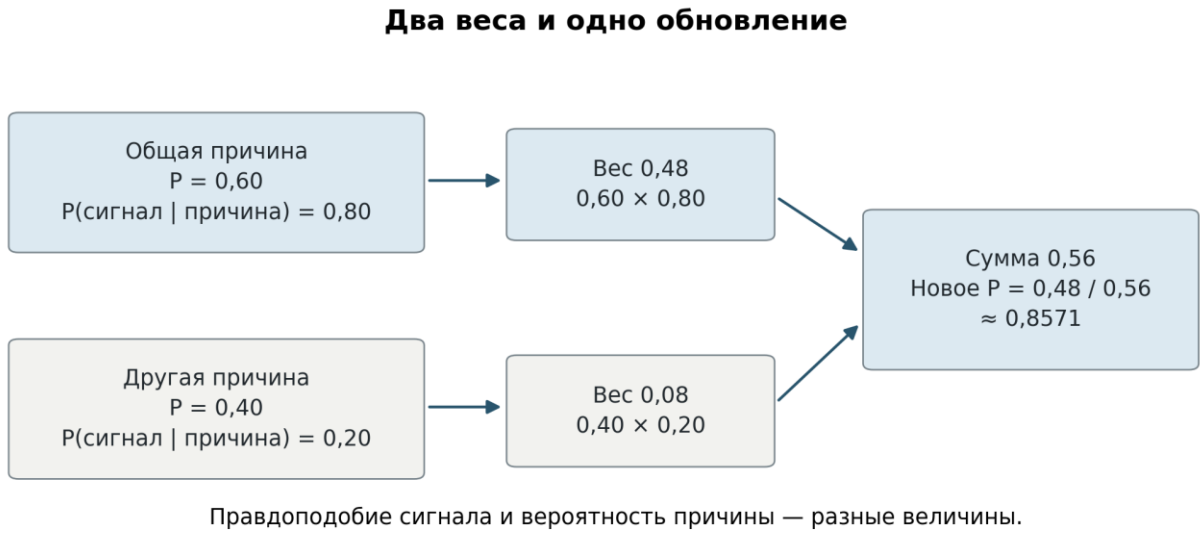

Рисунок 24. Последовательное обновление причинных версий. Учебный расчёт при предварительной вероятности 0,60 и симметричной надёжности 0,80. Сначала строятся ненормированные веса 0,48 и 0,08, затем выполняется нормировка. Числа задают модель, а не измерение опыта.

Такой же расчёт можно записать для двух произвольных версий. Сначала умножаем предварительную вероятность каждой на правдоподобие полученных данных. Получаем два ненормированных веса. Затем делим нужный вес на сумму обоих. Деление возвращает вероятности к общей сумме один. Если сумма равна нулю, данные невозможны при обеих версиях; это требует пересмотра модели или проверки входа, а не произвольного ответа.

Отношение шансов позволяет выполнить ту же операцию компактно. Шансы C против P равны P(C)/P(P). Новый признак умножает их на отношение правдоподобий P(признак|C)/P(признак|P). При исходных шансах один к одному признак, втрое более вероятный при C, даёт три к одному, то есть вероятность C 3/4.

Следующее свидетельство нельзя автоматически считать новым независимым голосом. Если один датчик просто скопировал другой, одинаковые числа не дают удвоенного основания. Для последовательного умножения используется правдоподобие нового признака с учётом уже известных; условная независимость является упрощением, при котором это правдоподобие не меняется. Теперь числовой пример можно читать как определённый механизм и одновременно видеть его предпосылки.

Разобранный пример двух версий

Построим авторскую учебную модель. Версия C предполагает, что зрительный и телесный сигналы относятся к одному объекту. Версия P предполагает разные источники. До нового наблюдения придадим версиям одинаковые веса. Затем получим признак: изменения произошли почти одновременно. Для упражнения зададим, что такое совпадение втрое вероятнее при версии C, чем при версии P.

После учёта признака отношение весов станет три к одному. Из общей суммы четырёх долей три приходятся на C: её модельная вероятность равна 75 процентам. Это не измеренная вероятность принадлежности руки и не оценка реального человека. Числа выбраны, чтобы показать обновление двух конкурирующих объяснений.

Теперь поступает второй признак: положения объектов сильно расходятся. Условимся, что он в четыре раза менее вероятен при C, чем при P. Прежнее отношение три к одному умножается на одну четверть и становится 0,75 к одному. Вероятность C теперь примерно 43 процента. Новые сведения изменили предпочтение модели, хотя первый признак продолжает учитываться.

Вычисление можно записать коротко. odds означает отношение вероятности C к вероятности P. ratio обозначает, во сколько раз очередной признак вероятнее при C. Последняя строка переводит отношение в долю от общей суммы.

```
odds = 1.0
for ratio in [3.0, 0.25]:
    odds = odds * ratio
    probability_common = odds / (1.0 + odds)
    print(round(probability_common, 3))
```

Мы получили 0,75 и примерно 0,429. В этой упрощённой модели признаки считаются независимыми при фиксированной версии причины. Если временное и пространственное совпадение взаимосвязаны, нельзя без проверки перемножать их как независимые свидетельства: часть сведений может быть учтена дважды. Более точная модель должна задавать совместную зависимость или оценивать второй признак с учётом первого.

Что добавляет собственное действие

Пассивное совпадение сигналов и результат команды создают разные возможности проверки. Допустим, программа посылает команду «влево» и ожидает движение указателя влево. Она сравнивает ожидание с фактическим изменением. Но одного совпавшего результата недостаточно, чтобы установить причинное управление: внешняя запись могла случайно выполнить то же движение.

Теперь программа меняет команды по новой последовательности. Если указатель продолжает соответствовать им, версия управления получает больше поддержки. Если воспроизводится прежняя запись независимо от новых команд, появляется расхождение. Вмешательство в последовательность действий помогает различить причины, которые при пассивном наблюдении могли выглядеть одинаково.

Модель действия должна хранить как минимум команду, ожидаемое следствие и наблюдаемый результат. Важно также указывать время: запоздалое следствие может относиться к предыдущей команде. Если сравнивать его только с текущей, система будет ошибочно приписывать себе неудачу или внешнее вмешательство. История нужна для установления отношения между событиями, а не просто для увеличения объёма памяти.

При этом функциональная оценка «эта команда вызвала событие» остаётся отдельной от переживаемого чувства авторства. Для человеческого исследования можно сопоставлять модельные оценки с отчётами. Для искусственной системы сначала проверяется само различение источников и его использование в поведении.

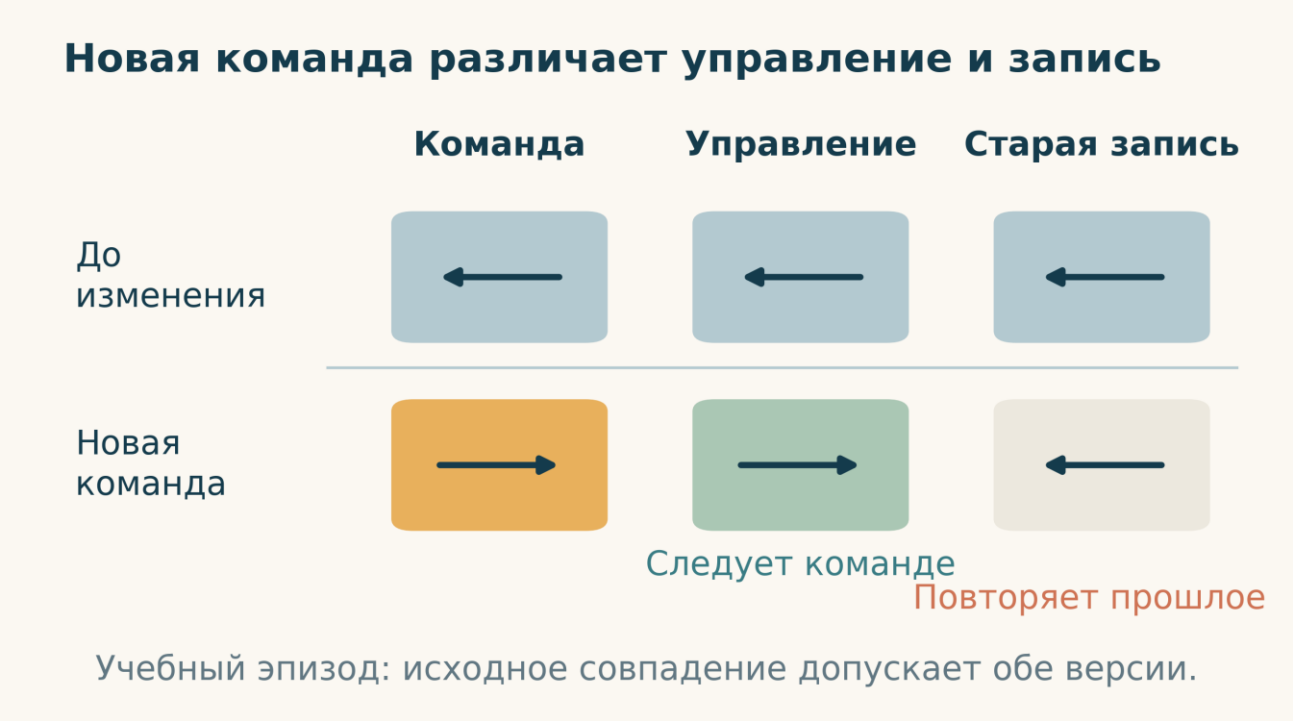

Рисунок 25. Различение причин изменением команды. При исходной команде влево движение указателя может совпасть и с управлением, и с внешней записью. В придуманном продолжении новая команда вправо даёт разные предсказания: управляемый указатель меняет направление, а запись повторяет прежнее движение влево.

Сравнить модели памяти и самоотнесения

Представим три варианта программы. Первая смотрит только на текущее совпадение сигналов. Вторая дополнительно учитывает общую историю надёжности, но не различает источники. Третья хранит историю отдельно для собственных команд, внешних событий и каждого канала наблюдения. Эти модели отвечают на разные вопросы о том, какая информация нужна для устойчивого различения причин.

Третья программа не получает преимущества автоматически из слова «собственная». Нужно проверить, какую именно информацию содержит эта метка и как она влияет на прогноз. Если ей просто сообщили правильные ответы, успех ничего не объясняет. Если она хранит больше данных, стоит сравнить её с общей моделью памяти сопоставимой вместимости. Иначе вклад организации смешается с преимуществом ресурсов.

Содержательная гипотеза могла бы звучать так: раздельная история источников помогает переносить оценку авторства на новые сочетания задержки и помех. Проверка должна использовать такие сочетания, которые не встречались при настройке. Тогда результат показывает способность

применять предполагаемое отношение, а не только запоминать знакомые последовательности.

Как задержка меняет смысл совпадения

В момент 1 отправлена команда вправо, в момент 2 влево. Видимый результат вправо возникает только в момент 3. Сравнение лишь текущей команды и текущего эффекта может ошибочно счесть результат внешним. История допускает связь с более ранней командой. Но если внешняя запись воспроизводит тот же ряд, пассивного наблюдения мало. Новая заранее выбранная последовательность команд позволяет проверить, следует ли результат за нынешним управлением или за старой записью. Такой контроль различает причинные версии и не измеряет сам по себе чувство авторства.

Что теперь установлено

Общая причина сигналов, причинное управление и телесная принадлежность не стали одним показателем. Мы получили способы проверять первые два отношения и можем отдельно спросить, как их оценки связаны с человеческими отчётами.

Практика 11 Бумажный исследователь источников

Зачем выполнять: Построить прозрачный разбор источника события, не открывая модели правильные причины.

Что понадобится: Понятия наблюдения, скрытого состояния и внутренней оценки. Вероятностный переход в этой главе объясняет необходимые расчёты; самостоятельно выводить правило Байеса до чтения главы не требуется.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Задайте команды по шагам: вправо, влево, влево, вправо. Для версии управления эффект появляется через два шага. Для внешней записи возьмите циклический ряд вправо, вправо, влево, влево, повторяющийся в том же порядке независимо от команд.
2. Составьте две таблицы наблюдений и отдельно храните истинную причину. Для первых доступных событий выпишите версии задержки ноль и два шага.
3. Продолжите команды новым рядом влево, вправо. До открытия новых эффектов запишите, что предсказывает управление и что предсказывает внешняя запись.
4. Отметьте случаи, где данных пока недостаточно, и объясните, какая новая команда действительно разведёт оставшиеся версии.

Что сохранить: Таблицы команд и эффектов, две причинные версии и предварительный прогноз различающего продолжения.

Подсказка: Первым двум шагам задержанного канала может ещё не соответствовать команда внутри наблюдаемого окна. Не заполняйте их вымышленной связью.

Исследовательское продолжение: Добавьте шум и неопределённое распознавание источника. Сравните общую память и раздельные истории при одинаковых входах и сопоставимой вместимости.

Разбор и ответы — приложение Е, глава 11, с. 297

Вопросы для самопроверки

1. Почему вывод об общей причине не равнозначен телесной принадлежности?
2. Что позволяет различить собственное управление и случайное совпадение с внешней записью?
3. Почему большее количество сохранённых данных ещё не доказывает пользу самоотнесения?

История оказалась полезной для различения причин. Следующая глава добавит ещё один вопрос: насколько надёжен собственный ответ и как оценивать это до получения результата.

Глава 12

Память и время в оценке мышления

Система уже может хранить сведения о мире и действии. Теперь дадим ей задачу прогнозировать собственный успех. Для этого нужно особенно внимательно сохранить порядок: прогноз записывается раньше исхода. Простая модель со счётчиками покажет, как учиться по прошлому, а сравнение с постоянной вероятностью позволит проверить, полезно ли обучение в конкретной серии. Позже этот разбор станет основанием для мониторинга и пересмотра стратегии.

Настоящий момент зависит от истории

Когда вы дочитываете предложение, его первые слова уже не находятся в центре текущего внимания, но могут участвовать в понимании конца. Чтобы объяснить такое поведение, модель должна сохранять информацию между событиями. В главе 9 для этого хватало одного изменяющегося числа. В главе 11 понадобилось различать историю команд и внешних источников. Теперь мы добавим оценку того, насколько хорошо система справляется со своей задачей.

Полезно разделить три масштаба. Текущее состояние определяет ближайший ответ. Сохранённая история позволяет сопоставлять нынешний эпизод с предыдущими. Параметры задают устойчивые особенности обработки и могут меняться при обучении. Эти масштабы не обязательно реализованы тремя отдельными ящиками. Различение нужно, чтобы понимать, какое изменение объясняет поведение и как долго сохраняется его влияние.

Наличие памяти не следует автоматически называть непрерывностью переживания. Файл может долго хранить сведения без выполнения каких-либо действий. Для функционального исследования важно проверить, используется ли сохранённое содержание, как оно меняет решения и что происходит при его нарушении. Отношение этих процессов к переживаемому времени станет вопросом дальнейшей теории.

Прогноз и оценка уже данного ответа

До задания вы можете ожидать, что справитесь хорошо, потому что знакомы с темой. После чтения конкретного условия уверенность меняется. После выдачи ответа появляется возможность ещё раз оценить решение, а после проверки — узнать его правильность. Эти четыре момента содержат разную

информацию. Их нельзя без отметок времени смешивать в одной колонке «уверенность».

Для модели особенно важно исключить подсказку из будущего. Если правильность ответа уже известна, оценить её легко. Исследовательский интерес представляет прогноз до обратной связи или оценка ответа по доступным системе внутренним признакам до получения правильного решения извне. Поэтому порядок записи является частью модели, а не технической подробностью.

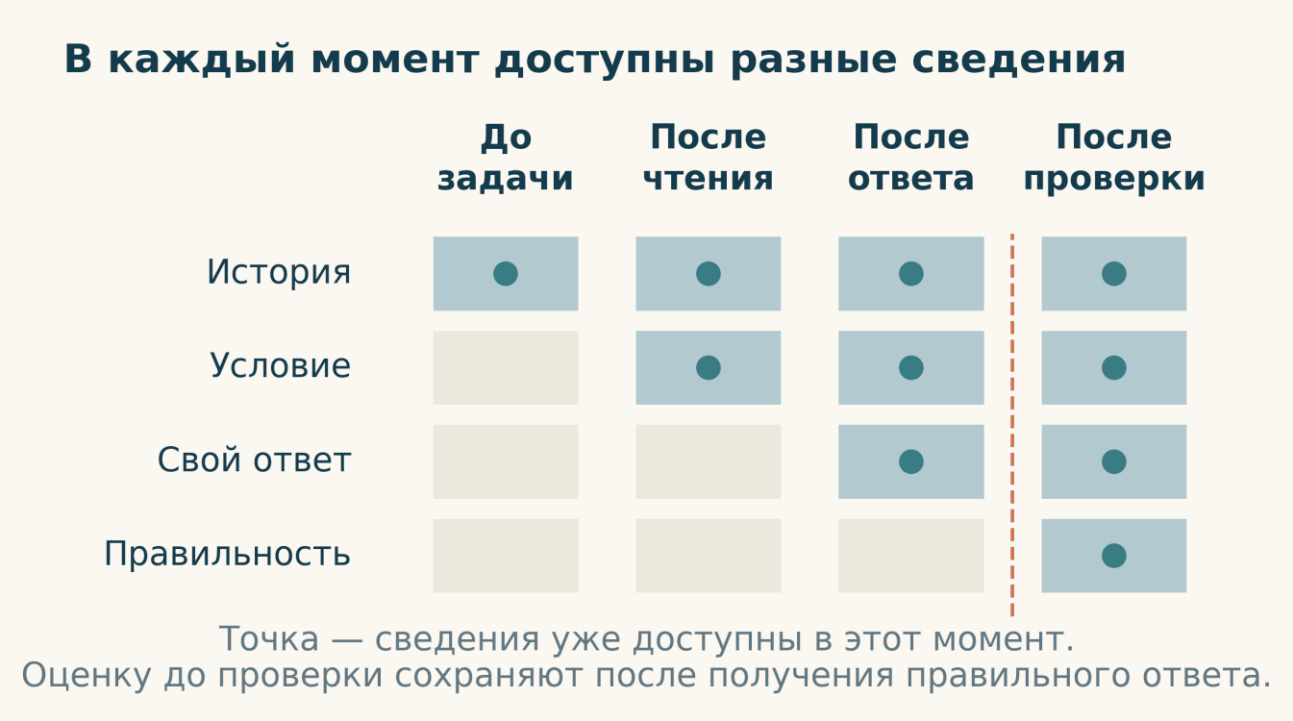

Рисунок 26. Информация до и после обратной связи. История прошлых результатов, условие текущего задания, собственный ответ и его правильность становятся доступны последовательно. Разделение моментов сохраняет различие между прогнозом до обратной связи и оценкой с уже известным решением.

Флеминг и Лау подчёркивают, что разные показатели метакогниции отвечают на разные вопросы о связи уверенности и успешности. [6] Здесь мы начнём с узкой задачи: научимся оценивать будущую надёжность ответов по прошлым результатам. Это ещё не полная модель переживаемой уверенности и не специальная оценка метакогнитивной эффективности.

Разобранный пример обучающейся оценки

Предположим, система решает серию заданий и после каждого получает правильность: 1 означает успех, 0 — ошибку. Введём два счётчика — успешных и ошибочных ответов. Первоначально положим в каждый по единице. Это условное сглаживание: до реальных данных система не начинает с полной уверенности в успехе или провале.

Перед очередным заданием прогноз равен доле успехов в общей сумме счётчиков. Вначале это одна вторая. После первого успеха счётчики равны двум и одному, поэтому следующий прогноз — две трети. После второго успеха получится три четверти. Если затем возникнет ошибка, успехов останется три, а общий счёт станет пять: новый прогноз равен трём пятым.

Вот шесть учебных исходов, придуманных для главы: успех, успех, ошибка, успех, ошибка, ошибка. Ниже программа сначала вычисляет и сохраняет прогноз, затем получает очередной исход и только после этого обновляет счётчики. Переменная loss накапливает ошибку прогноза: из вероятности вычитается фактический исход, а разность возводится в квадрат.

```
outcomes = [1, 1, 0, 1, 0, 0]
successes = 1.0
failures = 1.0
loss = 0.0
for outcome in outcomes:
    probability = successes / (successes + failures)
    loss += (probability - outcome) ** 2
    print(round(probability, 3), outcome)
    successes += outcome
    failures += 1 - outcome
print("Средняя ошибка", round(loss / len(outcomes), 3))
```

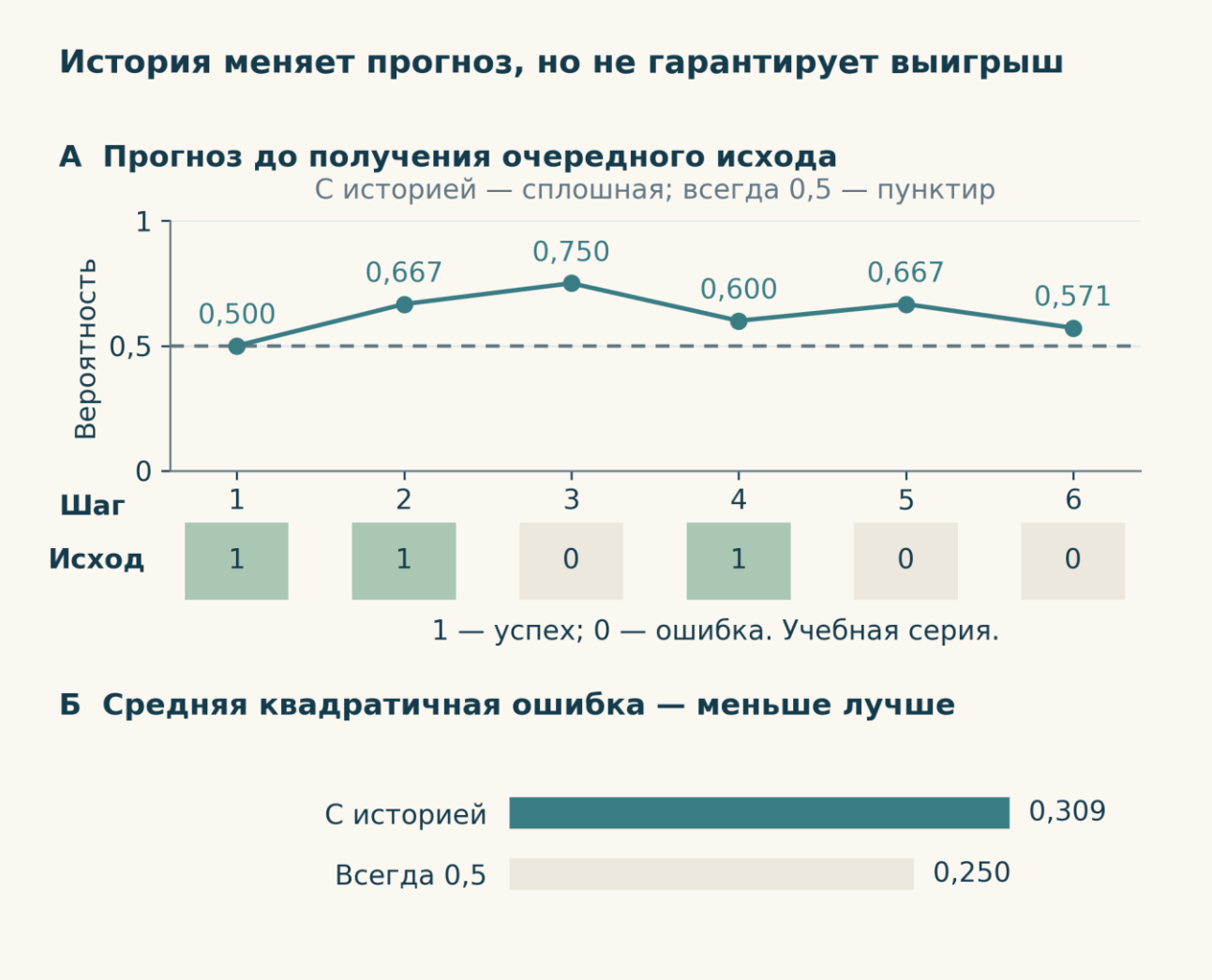

Рисунок 27. Прогнозы и ошибка двух моделей. Перед шестью учебными исходами 1, 1, 0, 1, 0, 0 модель со сглаженными счётчиками предсказывает 0,500; 0,667; 0,750; 0,600; 0,667; 0,571. Пунктир показывает постоянный прогноз 0,5. На этой серии средняя квадратичная ошибка модели с историей около 0,309, постоянной — 0,250; это сравнение конкретной серии, а не универсальное превосходство.

Оценки перед заданиями составят примерно 0,500; 0,667; 0,750; 0,600; 0,667; 0,571. После всей серии счётчики снова окажутся равны, и прогноз для следующего задания станет 0,5. Изменение понятно из механизма: успехи увеличивают одну часть отношения, ошибки — другую.

Почему сравнение с простой моделью обязательно

Возьмём соперника, который всегда предсказывает вероятность 0,5. При любом двоичном исходе его квадратичная ошибка равна 0,25. У обучающейся модели средняя ошибка в нашей короткой серии составляет примерно 0,309. Значит, на этих учебных данных постоянная оценка оказалась лучше по выбранному правилу. Результат не надо скрывать: он показывает, что обновление по истории полезно не автоматически.

Первые успехи повысили ожидание модели, после чего появились ошибки. Уверенные ошибочные прогнозы получили большой штраф. Шесть исходов не устанавливают общее превосходство постоянной модели. Чтобы сравнивать правила, нужны новые серии с заранее определёнными условиями: постоянной успешностью, изменяющейся трудностью или различиями между типами заданий.

Квадратичная ошибка оценивает качество вероятностного прогноза. Она не измеряет глубину переживания и сама по себе не показывает, как хорошо уверенность различает конкретные правильные и ошибочные решения. Название показателя должно соответствовать выполненному расчёту. Это правило помогает сохранить смысл даже при дальнейшем усложнении модели.

Общая надёжность и трудность данного случая

Наши счётчики рассматривают задания как случаи одного типа. Однако реальная серия может включать простые и трудные вопросы. Одинаковый общий прогноз для каждого следующего задания тогда теряет полезную информацию. Если тип известен заранее, можно вести отдельные счётчики по типам. Если трудность оценивается по внутренней неоднозначности решения, модель должна указать, какой признак для этого используется.

Например, классификатор может получать близкие оценки двух вариантов. Небольшое различие между ними — возможный признак неопределённости, но его значение нужно проверить по правильным ответам. Высокие внутренние числа не становятся хорошей уверенностью только из-за названия. Модель должна показать связь между такими признаками и фактической успешностью на новых случаях.

Так мы различаем общий прогноз «обычно я решаю подобные задачи» и оценку «этот конкретный ответ надёжен». Первая использует историю класса случаев, вторая может учитывать сведения текущего решения. Хороший исследовательский проект сравнивает их и выясняет, что добавляет каждый источник информации.

Когда прошлое перестаёт быть одинаково полезным

Допустим, задачи постепенно усложняются. Длинная история лёгких заданий может поддерживать чрезмерно высокий прогноз. Один способ адаптации — уменьшать вес старых исходов перед добавлением нового. Другой — сохранять отдельные режимы и пытаться определить, какой действует сейчас. Оба варианта требуют предположений о том, как меняется среда.

Быстрое забывание помогает реагировать на перемены, но делает оценку чувствительной к случайной серии ошибок. Медленное обновление устойчивее к краткому шуму, но может запаздывать после настоящего изменения. Здесь появляется проверяемый компромисс. Нельзя выбрать универсально лучший темп только по желанию системы «быть гибкой»; нужен набор условий, представляющий задачу использования.

Автоматическое обновление счётчиков является механизмом регулирования оценки. Явная рефлексия предполагает дополнительный шаг: рассмотрение оснований самого правила. Например: «Моя уверенность опирается на общий успех, хотя задания стали другого типа». Затем можно предложить новую организацию оценки и сравнить её со старой. При проектировании системы эти функции стоит описывать отдельно.

За что вероятностный прогноз получает штраф

Перед третьим исходом учебной серии модель после двух успехов прогнозирует 0,75. Третий исход равен нулю, поэтому квадрат разности равен 0,5625. Постоянный прогноз 0,5 получает в этой же пробе 0,25. Это не штраф за обучение как таковое: правило сильнее наказывает уверенную ошибку. Если сначала включить третий исход в счётчики, а затем назвать новое число прогнозом третьего исхода, оценка будет использовать ответ. Такой порядок проверяет уже другое свойство и создаёт утечку данных.

Что теперь установлено

Мы определили, когда сделан прогноз и относительно чего он оценивается. История может помочь или помешать. Её польза является результатом сравнения, а название механизма не гарантирует выигрыша.

Практика 12 Проверка на продолжении

Зачем выполнять: Сравнить постоянный и обучающийся прогноз на продолжении, не подбирая правило по проверяемым исходам.

Что понадобится: Доля успехов и различение доступной сейчас и будущей информации. Квадрат ошибки ниже используется как заранее выбранное правило оценки прогноза.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Начните со счётчиков успехов и ошибок по единице. Для исходов 1, 1, 0, 1, 0, 0 запишите вероятность до каждого обновления.
2. Для каждого прогноза вычислите квадрат разности с текущим исходом. Сравните среднее с постоянным прогнозом 0,5.

3. Сохранив конечные счётчики, добавьте продолжение 1, 0, 1, 1, 0, 1. Сравнение на второй части проведите отдельно.
4. До нового изменения выберите правило забывания, если хотите его исследовать. Параметр, выбранный по продолжению, проверяйте уже на другой серии.

Что сохранить: Временная таблица прогнозов, два раздельных сравнения и обоснование порядка вычислений.

Подсказка: После первой серии счётчики равны четырём и четырём. Прогноз следующего исхода снова равен 0,5.

Исследовательское продолжение: Разделите задания на известные заранее типы. Сравните общие счётчики с отдельными, а затем сформулируйте, какие данные позволят заметить смену режима.

Разбор и ответы — приложение Е, глава 12, с. 298

Вопросы для самопроверки

1. Почему прогноз после получения правильного ответа нельзя оценивать как исходную уверенность?
2. Почему накопленная доля успехов недостаточна для оценки каждого отдельного задания?
3. Чем автоматическое обновление счётчиков отличается от явной рефлексии над способом оценки?

Теперь у нас есть язык состояния, обратной связи, источника и оценки. Он позволит читать теории сознания как предложения об устройстве и предсказаниях, а не как перечень названий.

Часть 4

Теории и передний край исследований

Глава 13

Рабочее пространство и рекуррентная обработка

В предыдущих главах мы по отдельности разобрали состояние, обратное влияние, использование содержания и оценку ответа. Теперь эти понятия позволят сравнить две теоретические идеи: уточнение внутри сенсорной обработки и распространение результата разным потребителям. Сначала проследим их функции в полностью заданном устройстве. Затем вернёмся к более сильным утверждениям теорий о сознательном доступе и переживании. Совпадение функций ещё не делает две теории одинаковыми.

От названия к устройству

Представим, что вы читаете сообщение, а рядом падает чашка. Звук прерывает чтение; вы поворачиваетесь, оцениваете опасность и меняете действие. В объяснении такого эпизода легко соединить несколько разных вопросов. Как был распознан звук? Почему именно он получил преимущество? Как информация стала доступна движению, памяти и рассуждению? Почему происходящее переживалось? Теории, обсуждаемые в этой главе, предлагают разные способы связать эти вопросы. Чтобы сравнивать их, требуется сначала понять предложенные механизмы, а затем отдельно оценить переход от механизма к сознанию.

Рабочее пространство удобно представить как организацию обмена между специализированными процессами. Рекуррентная обработка означает, что результат обработки возвращается и изменяет дальнейшую работу уже участвовавших элементов. Это разные характеристики устройства: широта доступности и наличие обратного взаимодействия. Они могут сочетаться. Поэтому спор между подходами нельзя свести к выбору между мозгом «со связями» и мозгом «без связей».

Сильный тезис рабочего пространства

В модели глобального нейронного рабочего пространства Станисласа Деана, Мишеля Керсзберга и Жан Пьера Шанжё специализированные процессы взаимодействуют с распределённой системой, способной поддерживать и распространять отобранную информацию. Компьютерная реализация 1998 года связывала её участие с освоением задачи, требующим усилия выполнением и обработкой ошибок. Это конкретное предложение об организации когнитивной работы, а не обнаружение отдельного мозгового «экрана». [16]

Рассмотрим собственный учебный пример. У устройства есть четыре модуля: распознавание предметов, оценка риска, выбор движения и запоминание событий. Пока распознаватель видит чашку, только он хранит соответствующую запись. После отбора сообщение «чашка падает» получает общий идентификатор события и передаётся остальным модулям. Оценка риска связывает его с возможным ожогом, движение выбирает отступление, память сохраняет происшествие. Архитектура объясняет согласованное использование одного содержания несколькими процессами.

Теперь уточним проверяемое утверждение. Если общий канал необходим для переноса новой информации между модулями, его временное отключение должно нарушить именно перенос. Распознавание чашки может сохраниться, а использование распознанного в новой задаче — исчезнуть. Если же отключение одновременно портит сенсорный вход и движения, результат неоднозначен. Мы не узнаем, какой из повреждённых процессов отвечал за изменение поведения.

Сильная версия объяснения сознательного доступа приписывает особую роль переходу к устойчивой широкой доступности содержания. Исследователь обязан указать, что считается достаточно широким распространением, какой процесс поддерживает содержание и какие данные позволят отличить такой переход от обычного усиления сигнала. Выражение «информация стала глобальной» без этих уточнений остаётся названием результата.

Сильный тезис рекуррентной обработки

В подходе Виктора Ламме первоначальное прямое распространение зрительного сигнала отделяется от последующих обратных взаимодействий. Локальная рекуррентная обработка рассматривается как кандидат на механизм зрительного переживания, тогда как более широкое взаимодействие связывается с доступом и отчётом. Сильный тезис состоит в возможности переживания, которое ещё не получило всех возможностей последующего

использования. Это теоретическая интерпретация, требующая самостоятельной проверки. [17]

Построим другой учебный механизм. Первый модуль выделяет короткие линии, второй предполагает контур предмета. Предположение возвращается к первому и меняет вес неоднозначных линий. После нескольких обменов система выбирает согласованную форму. Здесь обратная связь делает понятной зависимость распознавания части от целого. Однако наличие цикла ещё не объясняет, почему обработка должна сопровождаться видимостью формы. Вентилятор с регулятором температуры тоже образует контур управления; его устройство существенно отличается от описанного зрительного взаимодействия.

Поэтому рекуррентную гипотезу нужно записывать с ограничениями. Между какими элементами происходит обмен? Что передаётся обратно? Какова продолжительность взаимодействия? Как отличить содержательную обратную связь от общей модуляции возбудимости? Какие операции должны сохраниться при её нарушении? Чем подробнее ответ, тем меньше возможностей объявить любой найденный цикл подтверждением.

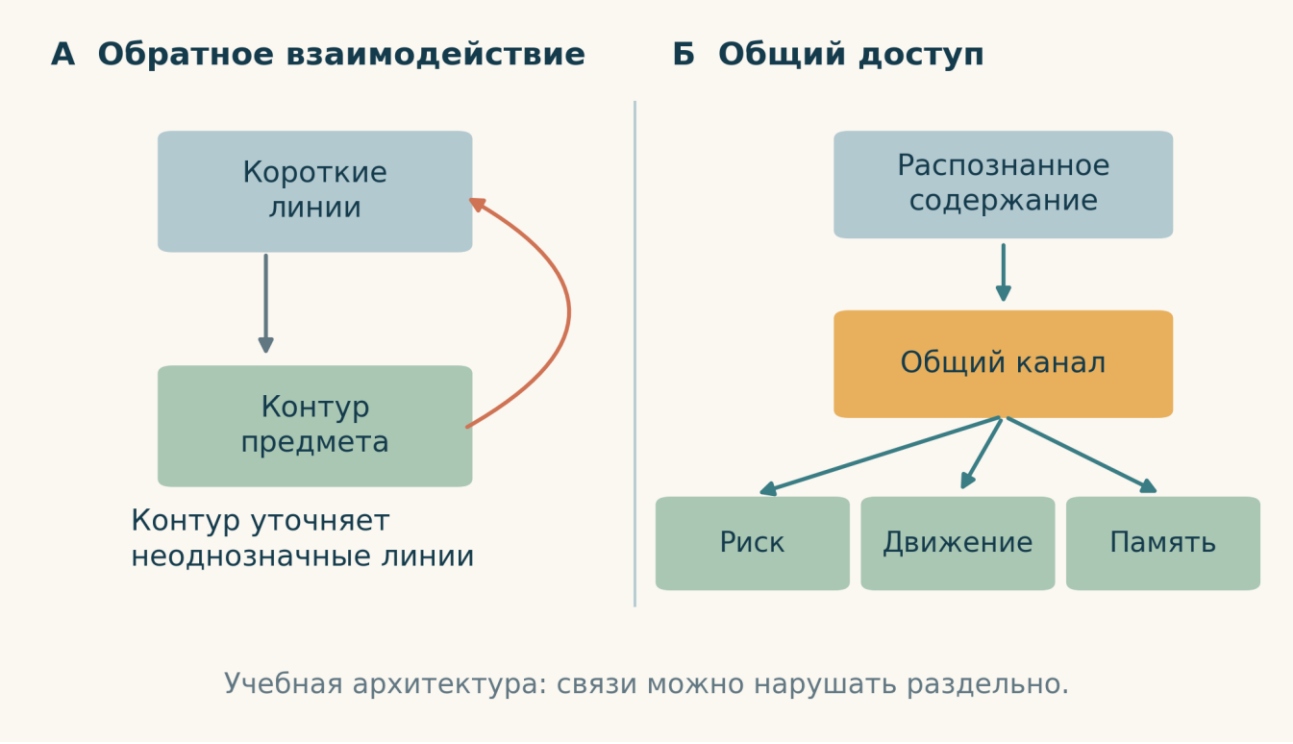

Рисунок 28. Рекуррентность и общий доступ. Два аспекта одного учебного устройства. Возврат от контура к признакам помогает согласовать распознавание; общий канал позволяет нескольким процессам использовать выбранное содержание. Эти связи можно изменять раздельно. Схема не является тестом наличия сознания.

В исследовании Фаренфорта, Схолте и Ламме зрительная маскировка сочеталась с сохранением ранней обработки и нарушением более поздних

сигналов, интерпретированных как рекуррентные. Работа поддержала различение этапов обработки. Она не устанавливает, что любой поздний ответ достаточен для опыта, и не исключает все дополнительные условия сознательного восприятия. [18]

От схемы к обязательству теории

Полезно пройти один случай дважды. Сначала объяснить обработку: локальные признаки дают несколько возможных объектов, контекст ограничивает выбор, результат становится доступен дальнейшему действию. Затем спросить о сознательном статусе каждого этапа. Первый проход задаёт механизм; второй показывает, какое дополнительное положение теории связывает этап с переживанием.

Для подхода рабочего пространства существенным становится не произвольное копирование числа, а определённый отбор, поддержание и широкое использование содержания. Нужно перечислить потребителей, способы нового использования и последствия избирательного ограничения доступа. Поэтому простая передача в один заранее настроенный выход является подготовительным примером, а не всей теорией.

Для подхода рекуррентной обработки важно уточнить, какие обратные взаимодействия предполагаются достаточными для зрительного переживания и как они отличаются от последующего отчёта. Конечная последовательность повторных вычислений может быть развёрнута в прямую схему. Это заставляет различать вычислительную функцию и физическую организацию, если теория придаёт значение последней.

Хорошее сравнение выбирает один расходящийся прогноз. Допустим, локальное содержание сохраняется, но стало недоступно произвольному ответу. Тогда спор касается того, какое основание позволяет считать сохраняющийся процесс переживанием и какой независимый показатель это подтверждает. Нельзя решить спор одним переименованием процесса в бессознательный или сознательный. Требуется согласовать проверяемую версию, измерение и условие пересмотра.

Учебная система перед практикой намеренно проще. Её задача дать опыт раздельного отключения функций. После выполнения мы сможем вернуться к научному спору с точным пониманием того, какие части аргумента проверил такой расчёт и какие требуют дополнительных данных.

Где проходит реальное разногласие

Вообразим содержание, которое устойчиво представлено в зрительной системе, но не используется для произвольного ответа. Один исследователь скажет: переживание уже есть, ограничен доступ. Другой ответит: найдена обработка, предшествующая сознательному доступу. Разногласие касается того, какое устройство считается достаточным и как установить наличие искомого явления независимо от принятой теории.

Эта трудность не делает исследование бессмысленным. Можно сравнивать более узкие предсказания. Например, одна модель требует длительного обмена между двумя сенсорными областями; другая требует подключения дополнительной системы в определённый период. Тогда полезен эксперимент, который избирательно меняет одно взаимодействие, сохраняя другое. Если вмешательство неизбирательно, оно плохо различает модели даже при сильном поведенческом эффекте.

В нашем учебном устройстве можно независимо выключать возврат от контура к линиям и общий канал передачи события. При первом отключении пострадает согласование формы, при втором — её использование разными задачами. Такие различающиеся последствия показывают, что два механизма выполняют разные функции. Чтобы перейти к утверждению о сознании, потребуется дополнительная аргументация. Устройство пока даёт нам эксперимент с обработкой информации.

Не следует заранее ждать, что один процесс заменит все остальные. Механизм может быть необходим в одних условиях и обходиться другим механизмом после обучения. Порог обнаружения зависит от силы сигнала, а готовность отвечать — от инструкции и цены ошибки. Поэтому изменение результата необходимо разложить на изменения чувствительности, стратегии, времени ответа и субъективной оценки.

Добавим к учебной системе второй источник — сообщение о движущемся человеке. Теперь общий канал должен выбирать между двумя содержаниями. Если чашка всегда получает преимущество, устройство справится с одним примером, но не покажет гибкого отбора. Изменим задачу: иногда нужно запомнить человека, иногда предотвратить столкновение с чашкой. До проверки зададим, как инструкция влияет на приоритет. Затем посмотрим, сохраняется ли распознавание проигравшего содержания и может ли оно позднее получить доступ к другим модулям.

Такое усложнение проверяет уже не наличие линии связи, а правила её использования. Оно помогает отделить жёсткую последовательность команд от архитектуры, способной передавать разное содержание в зависимости от

новой цели. Оба устройства могут успешно выполнить исходную пробу, но их различие проявится при переносе на незнакомое сочетание требований.

Одна проба с уточнением и передачей

В новой учебной системе локальный признак равен 0, 1 или знаку вопроса. Контекст содержит 0 или 1. Если признак определён, распознаватель сохраняет его. Если он неопределён, первоначально записывает 0; включённый возврат контекста заменяет эту запись значением контекста. Общий канал копирует окончательное содержание в память. Без передачи память остаётся 0. В пробе (?, 1) полная система сначала получает локальный 0, затем уточняет его до 1 и передаёт 1 в память. Отключение возврата оставляет 0; отключение передачи сохраняет распознавание 1, но память остаётся 0.

Правильная метка для определённого входа равна этому входу, для неопределённого контексту. Она нужна оценщику. Здесь исследуется работа заранее заданных правил. Конкурент с прямым доступом ко всем входам может решить ту же задачу, поэтому успех демонстрации не устанавливает уникальности рабочего пространства или достаточности возврата для переживания.

Что теперь установлено

Мы получили два независимо изменяемых отношения: уточнение содержания и передачу его в память. Теперь четыре режима можно сравнить по двум результатам, не называя потерю одного из них исчезновением сознания.

Раздельные функции двух связей

Режим	Распознавание	Память результата
Полная схема	10 из 10	10 из 10
Без возврата	7 из 10	7 из 10
Без общего канала	10 из 10	4 из 10
Без обеих связей	7 из 10	4 из 10
Прямое копирование	10 из 10	10 из 10

10 фиксированных входов из разобранного примера

Равный выход двух схем требует следующего различающего опыта.

Рисунок 29. Независимое отключение двух механизмов. Результат десяти полностью заданных учебных проб; воспроизводится teaching_lab.py --chapter 13. Возврат уточняет неоднозначный признак, общий канал передаёт ответ памяти. Прямое копирование является сильным функциональным контролем. Таблица не измеряет сознание и не представляет сравнение полных научных теорий.

Практика 13 Перевод теории в предсказание

Зачем выполнять: Раздельно проверить вклад уточнения и передачи на одинаковых десяти пробах.

Что понадобится: Вход, состояние и выход; различие локального представления и его использования. Перед практикой приведены все правила учебной системы и десять входов.

Ориентир по времени: 35–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Возьмите пары локальный признак и контекст: (0,0), (1,0), (0,1), (1,1), (1,0), (?,1), (?,0), (?,1), (?,0), (?,1). Для каждой вычислите правильную метку по правилу выше.
2. Для полной системы заполните столбцы: начальное распознавание, окончательное распознавание, содержание памяти. После каждой пробы память начинается с 0.
3. Повторите те же десять входов без возврата, без передачи и без обоих механизмов. Посчитайте отдельно правильность распознавания и памяти.
4. Проверьте таблицу командой python teaching_lab.py --chapter 13. Затем добавьте прямого конкурента, видящего оба входа, и уточните границу вывода.

Что сохранить: Четыре режима, два результата для каждого и объяснение одного конкретного расхождения.

Подсказка: Отключение передачи не меняет правило распознавания. Если оба показателя изменились, проверьте, не выключили ли вы несколько функций одновременно.

Исследовательское продолжение: Измените задачу: после задержки приходит новая инструкция одному из нескольких потребителей. Определите память, бюджет и входы сильного конкурента заранее. Выведите результат, на котором механизмы действительно расходятся.

Разбор и ответы — приложение Е, глава 13, с. 298

Вопросы для самопроверки

1. Почему глобальное распространение и рекуррентность не образуют строгую противоположность?
2. Доказывает ли исчезновение отчёта после вмешательства исчезновение переживания?
3. Зачем сравнивать механизмы при равном вычислительном бюджете?

Различие функций стало явным. Следующая глава добавит два других предмета теории: внутреннюю причинную организацию и представление системой собственного состояния.

Глава 14

Интегрированная информация и теории высшего порядка

Рабочее пространство и рекуррентность можно сравнивать как способы организации обработки. Однако теории сознания спрашивают также, какое устройство имеет значение независимо от привычного внешнего ответа и как состояние представлено самой системе. Рассмотрим причинную структуру и представления высшего порядка. Наша задача состоит в точном понимании исходных обязательств каждой теории; вычисление полного формализма IIT не будет молча подменяться рисунком нескольких связанных узлов.

Два направления объяснения

Представьте две системы, которые одинаково отвечают на вопрос о красном круге. У первой множество частей взаимно меняют состояние друг друга. У второй заранее подготовлена таблица ответов на все разрешённые вопросы. Внешне их трудно различить, если вопросы ограничены. Должно ли совпадать то, что мы приписываем этим системам внутри? Ответ зависит от того, какие свойства теория считает существенными: причинную организацию, представление собственного состояния, способность действовать или некоторую комбинацию требований.

В этой главе мы разберём два направления. Теория интегрированной информации начинает с предполагаемых общих свойств опыта и ищет соответствующую физическую организацию. Теории высшего порядка связывают сознательность состояния с тем, как оно представлено самой системе. Обе программы предлагают больше, чем удобное описание поведения. Именно поэтому их исходные допущения следует рассматривать столь же внимательно, как математические модели и эксперименты.

Что утверждает теория интегрированной информации

В версии 4.0 теория интегрированной информации связывает опыт с определённой неразложимой структурой причин и следствий внутри физического субстрата. Авторы исходят из существования опыта и выделяют его внутреннюю принадлежность, определённость, единство, конкретность и структурированность. Затем предлагают физические постулаты и формализм. Тождество опыта и развёрнутой причинной структуры является основным объяснительным утверждением теории. Оно не превращается в установленный эмпирический факт благодаря одной математической записи. [19]

Разберём различие между регистрацией связи и проверкой причины на собственном примере. Два индикатора каждую секунду одновременно загораются. Наблюдатель замечает идеальное совпадение. Однако причиной может быть общий таймер; один индикатор не влияет на другой. Если же вмешательство в первый меняет будущее второго, а вмешательство во второй меняет первое, устройство имеет другую причинную организацию. Одинаковые наблюдаемые последовательности ещё не определяют внутренние зависимости.

Этот пример не вычисляет интегрированную информацию. Он учит задавать вопрос, предшествующий вычислению: какие состояния и вмешательства мы считаем физически возможными? Допустим, в нашей модели есть два элемента с двумя состояниями. Нужно задать правило изменения каждого элемента при всех сочетаниях входов. Если часть сочетаний запрещена, требуется объяснить почему. Если два элемента в действительности обозначают один общий источник, их раздельное рассмотрение может исказить анализ.

Следующий вопрос касается границы системы. Мы можем изучать один модуль, несколько взаимодействующих модулей или устройство вместе с внешним контроллером. Результаты причинного анализа способны различаться. Поэтому выбор границы нельзя обосновывать желаемым ответом «здесь должно быть сознание». Сначала задают правило выбора, затем применяют его последовательно к интересующим и контрольным системам.

Сильная сторона такого направления — требование объяснять единство и различия возможных состояний через устройство. Его трудность для исследователя — переход от причинной организации к наличию переживания. Даже безошибочно рассчитанная характеристика подтверждает прежде всего свойства принятой модели. Связь этих свойств с опытом должна выдержать отдельное сопоставление с данными и альтернативами.

Что означает высший порядок

Теории высшего порядка предлагают отличать состояние, представляющее мир, от состояния, представляющего наличие или характер первого состояния. В байесовской модели Хаквана Лау сознательная оценка зависит от решения относительно собственных внутренних сигналов и их надёжности. Это одна конкретная реализация семейства подходов, а не определение всех его вариантов. В ней можно исследовать расхождение между выполнением задачи и субъективной оценкой восприятия. [20]

Построим учебный пример с распознавателем звуков. Первый процесс выбирает между двумя сигналами. Второй получает характеристики работы

первого: величину различия между альтернативами, согласованность нескольких измерений и сведения о недавних ошибках. Он оценивает, насколько надёжен текущий выбор. Третий процесс решает, отвечать сейчас или запросить ещё одно наблюдение. Такое устройство объясняет контроль уверенности, если второй процесс действительно использует информацию о первом.

Добавим важное различие. Переменная «уверенность равна восьмидесяти процентам» может оказаться простым пересчётом одного сенсорного числа. Тогда она не свидетельствует о богатой модели собственных возможностей. Чтобы показать более содержательный контроль, полезно изменить надёжность сенсора, сохранив отдельные входные значения. Если система учитывает историю ошибок и меняет стратегию, её оценка отражает больше, чем текущую интенсивность сигнала.

Но функциональный контроль уверенности сам по себе не решает вопроса о феноменальности. Сильное теоретическое утверждение должно уточнить, какое представление собственного состояния достаточно для сознательности и почему. Не всякое числовое описание процесса разумно автоматически считать осознанием. Кроме того, сознательное восприятие и развёрнутая рефлексия различаются: видеть цвет не означает одновременно рассуждать о надёжности собственного цветового зрения.

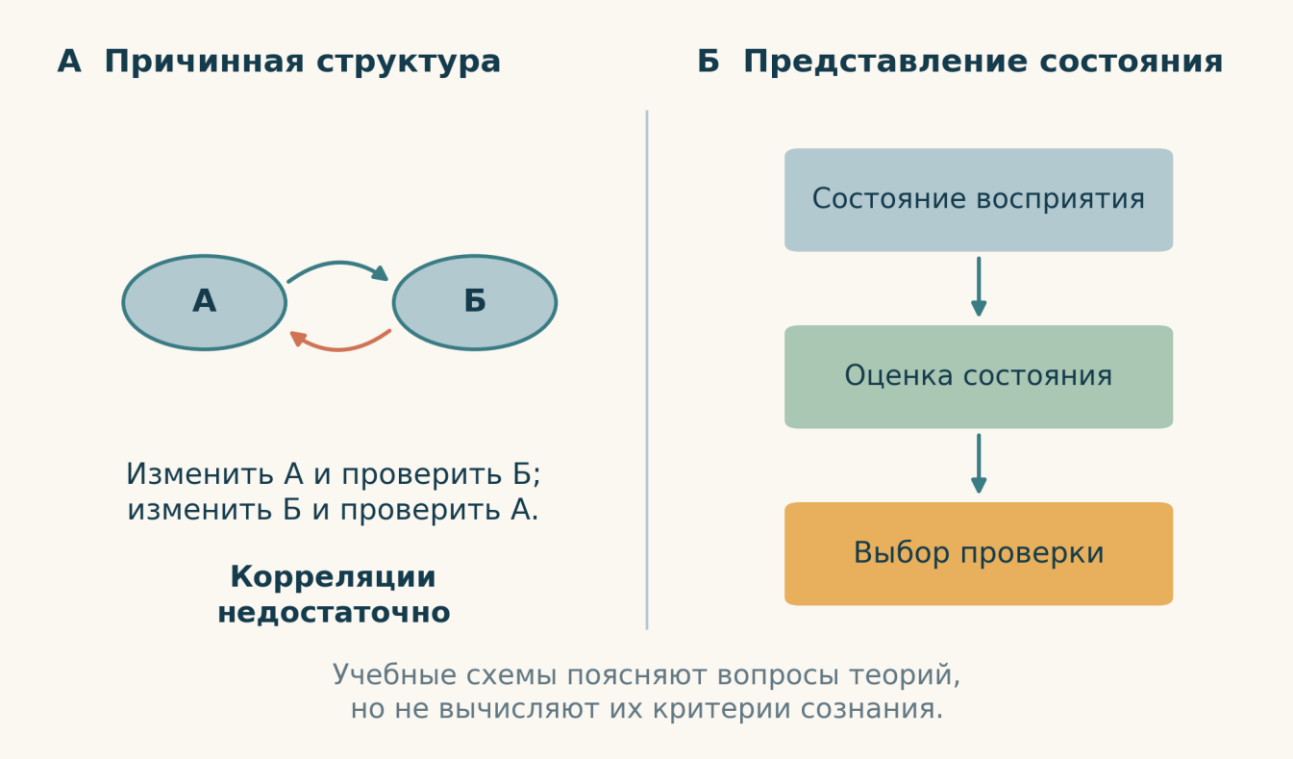

Рисунок 30. Два разных предмета объяснения. Левая схема направляет внимание на взаимные причинные влияния частей; правая — на представление системой собственного состояния и его использование. Схемы поясняют разные вопросы IIT и подходов высшего порядка, но не заменяют формализмы этих теорий и не назначают победителя.

Что нужно знать о реализации

Два устройства могут иметь одинаковую таблицу обычных ответов и разные реакции на внутреннее вмешательство. В одном выход зависит от взаимодействия состояний сейчас; в другом заранее записана последовательность. При ограниченном наборе внешних вопросов они могут совпасть. Поэтому анализ причинного устройства добавляет сведения, которых нет в обычном диалоге.

Но любые обратные стрелки ещё не являются расчётом интегрированной информации. Для конкретной версии формализма понадобятся определённые элементы, состояния, переходы и правило выбора рассматриваемой системы. Утверждение о соответствии опыту принадлежит самой теории и оценивается вместе с её аргументами. Учебный рисунок поясняет предмет вопроса, сохраняя это различие.

Для представления высшего порядка нужна иная спецификация. Что представлено как первичное состояние? Как второй процесс получает сведения о нём? Какая информация относится к самому текущему решению, а какая является общим ожиданием успешности? Кто использует оценку? Эти

вопросы позволяют отличить реальный мониторинг от внешнего журнала или заранее подготовленной фразы.

Наконец, внутренний доступ не обязан давать полную прозрачность системы самой себе. Процесс может иметь привилегированный доступ к одному сигналу и ошибаться относительно другого. Поэтому будем проверять доступ к конкретному содержанию, момент получения и причинное использование, избегая общего вывода о совершенном самопознании.

Как сравнивать исходные допущения

Возьмём вымышленное устройство с богатой внутренней причинной связностью, которое не формирует отдельной оценки собственных восприятий. Затем вообразим устройство с подробным контролем уверенности, реализованным иной причинной организацией. Такие случаи заставляют разные подходы уточнить границы. Одно объяснение будет прежде всего спрашивать о физической структуре, другое — об отношении между представлениями.

Нельзя честно решить спор простым выбором более привлекательного примера. Требуется указать, откуда мы независимо знаем о сознательности сравниваемых систем. Если такого знания нет, мысленный эксперимент выявляет различие теорий, но не определяет победителя. Его польза — в подготовке проверяемых следствий и обнаружении допущений, которые раньше оставались скрытыми.

Полезно разделить три вида утверждений. Определение устанавливает, как используется термин. Формальная теорема показывает, что следует из правил модели. Эмпирическое утверждение сообщает, как устроен изучаемый мир. Из строгой теоремы нельзя получить истинность исходных предпосылок. В то же время спор о предпосылках не отменяет ценности формального анализа: он показывает точную цену принятого объяснения.

Для собственного исследования выбирайте наиболее узкий вывод. Если модель верно предсказывает запрос дополнительной информации, это результат о регуляции поведения. Если причинная характеристика различает две сети, это результат об их организации. Следующий шаг к сознанию должен иметь собственное обоснование; его нельзя незаметно включить в название показателя.

Полезно проверить и зависимость вывода от наблюдателя. Представим, что разработчик добавил подробный журнал внутренних состояний, который никем внутри устройства не читается. Исследователь теперь знает о системе больше, но её возможности не изменились. Поэтому доступная нам

информация о процессе отличается от информации, которую сам процесс использует для управления. Если журнал превращён во вход управляющего модуля, организация действительно меняется, и последствия можно проверить.

Это различие особенно важно для самоописания. Сведения, записанные внешним экспериментатором от первого лица, не становятся внутренней моделью устройства. Нужно показать, кто использует запись, какие действия от неё зависят и что произойдёт при её избирательном искажении. Тогда спор о представлении собственного состояния получает конкретный предмет.

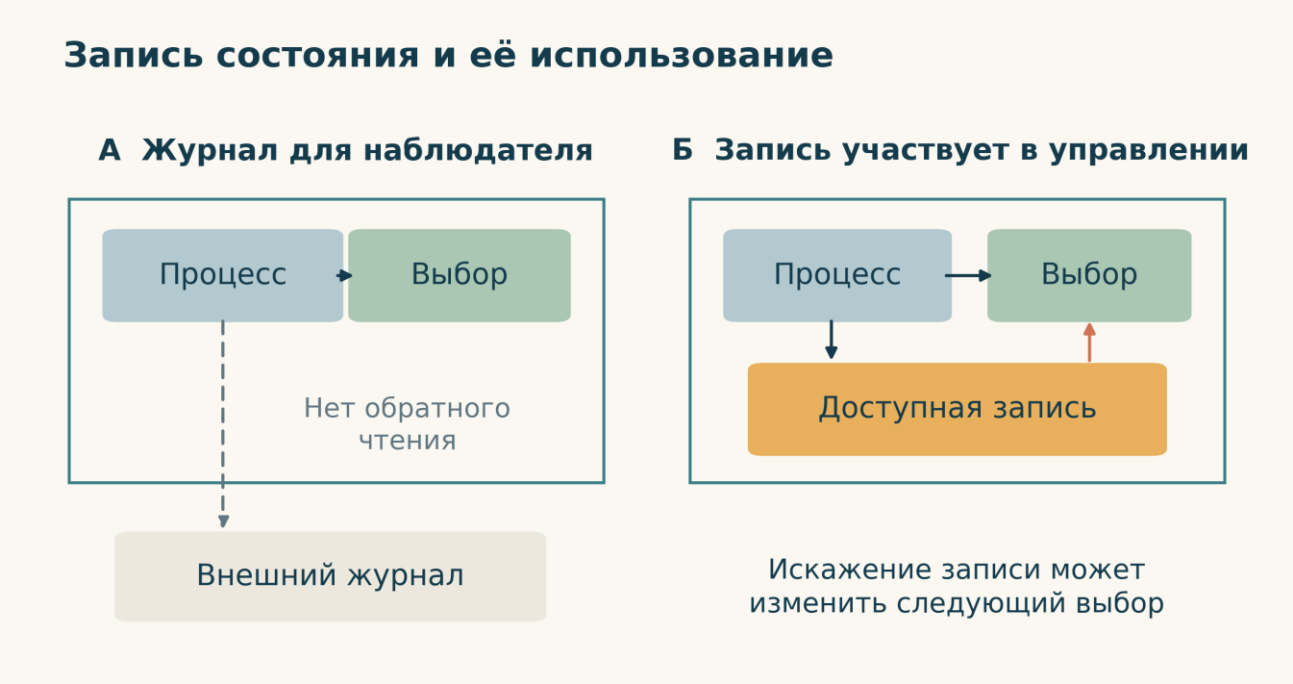

Рисунок 31. Внешний журнал и внутреннее использование. В левой конструкции исследователь получает сведения о процессе, но устройство не читает журнал. В правой запись становится входом управления: её избирательное искажение позволяет проверять причинную роль. Такая проверка касается использования информации и ещё не устанавливает представление высшего порядка, достаточное для сознательности.

Как не смешать определение и сильное предположение

Фраза «внутренней оценкой назовём прогноз собственной ошибки» задаёт рабочее определение. Утверждение «если отчёт читает эту оценку, её изменение может менять отчёт» требует конкретных связей модели. Фраза «подходящая внутренняя оценка делает состояние сознательным» уже добавляет теоретическое объяснение. Наконец, наблюдение зависимости в определённой выборке относится к данным. Эти предложения могут участвовать в одном аргументе, но не получают одинаковую силу благодаря соседству. Особенно важно сохранять этот разбор при сравнении IIT и теорий

высшего порядка: у них различаются не только метрики, но и исходные предметы объяснения.

Что теперь установлено

Общая связность и число уверенности остаются функциональными свойствами, пока не сформулировано более точное теоретическое требование. Теперь попробуем явно выписать, где заканчивается описание устройства и начинается предположение о переживании.

Практика 14 Паспорт допущений

Зачем выполнять: Составить паспорт двух объяснений с различимыми статусами утверждений.

Что понадобится: Различайте механизм, его измерение и утверждение о его связи с опытом. Необходимость и достаточность введены в главе 5.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Для причинной системы укажите два изменяемых состояния и направление каждого влияния. Для системы оценки укажите первичное решение, внутренний признак надёжности и получателя оценки.
2. Запишите для каждого подхода четыре положения: объясняемое явление, предложенное устройство, основание его существенности и условие пересмотра.
3. Каждое положение пометьте как определение, допущение, следствие заданной модели или эмпирическое утверждение. Смешанные положения разделите.
4. Выберите одно вмешательство и более простой механизм, способный дать тот же ответ. Если различающего результата пока нет, сохраните это как итог.

Что сохранить: Два паспорта и один причинный тест с явной границей феноменального вывода.

Подсказка: Схема взаимного влияния не является расчётом IIT. Прогноз ошибки не становится представлением высшего порядка только благодаря названию: нужны его содержание, происхождение и использование.

Исследовательское продолжение: Возьмите конкретную версию одной теории из источников главы. Уточните, что её достаточное условие требует сверх созданной учебной системы, и предложите случай, где выводы теорий расходятся.

Разбор и ответы — приложение E, глава 14, с. 299

Вопросы для самопроверки

1. Почему высокая корреляция между частями недостаточна для вывода об их взаимном причинном влиянии?
2. Равна ли рефлексия любому представлению высшего порядка?
3. Может ли теория быть математически точной и эмпирически спорной?

Следующая глава покажет конкретный способ вычислять скрытые причины и собственную надёжность через предсказание. Это поможет сделать часть обсуждавшихся отношений исполнимой.

Глава 15

Предсказание и модели себя

Представление собственного состояния становится содержательнее, когда известно, как оно формируется и используется. Предсказательный подход предлагает строить версии скрытых причин, сравнивать ожидаемые сигналы с полученными и обновлять оценку. Разберём отдельно мир и последствия команды. Их смешение особенно заметно после изменения управления: система может хорошо видеть цель и одновременно ошибаться при действии.

Почему наблюдение требует предположения

Вы слышите короткий звук за дверью. Это мог упасть предмет, сработать замок или щёлкнуть выключатель. Доступен один сигнал, а возможных причин несколько. Чтобы выбрать действие, требуется учитывать не только сигнал, но и контекст: кто находится за дверью, что происходило раньше, насколько хорошо слышно помещение. Такой пример показывает основную задачу вывода: по неполному наблюдению оценивать скрытые причины.

Предсказание в научной модели не обязательно означает словесное ожидание будущего. Это может быть вычисляемое предположение о том, какой сигнал должен появиться при определённом состоянии мира. Сравнение ожидаемого и полученного помогает изменить предположение. В модели Рао и Балларда такая иерархическая организация использовалась для объяснения некоторых свойств зрительных нейронов: верхний уровень предсказывает активность нижнего, а расхождение участвует в уточнении представления. Это модель обработки, не самостоятельное доказательство возникновения опыта. [21]

Простая модель скрытой причины

Рассмотрим собственный вычислимый пример. Цель находится слева или справа. До наблюдения вероятность правой стороны оценена в шестьдесят процентов. Датчик верно показывает сторону в восьми случаях из десяти. Появляется сигнал «справа». Насколько нужно изменить оценку?

Мысленно рассмотрим сто одинаково устроенных ситуаций. В шестидесяти цель справа; примерно в сорока восьми из них датчик покажет правую сторону. В сорока цель слева; примерно в восьми он ошибочно покажет правую сторону. Среди пятидесяти шести правых показаний сорок восемь соответствуют правой цели. Обновлённая вероятность составляет примерно восемьдесят шесть процентов. Это следствие заданных чисел и

предположения о надёжности датчика, а не экспериментальное измерение человеческого восприятия.

Пример показывает, зачем отделять три величины: предположение до наблюдения, диагностичность наблюдения и итоговую оценку. Сильный сигнал может мало изменить убеждение, если датчик ненадёжен. Отсутствие наблюдения не равняется наблюдению отсутствия цели. Если датчик не работал, система должна сохранить неопределённость с учётом возможных изменений мира, а не выдумывать подтверждение прежней версии.

Теперь разрешим цели иногда перемещаться. Память уже нельзя понимать как неподвижное хранение последнего ответа. Старое убеждение нужно изменять ещё до нового наблюдения: чем дольше перерыв и выше подвижность цели, тем меньше оснований для прежней уверенности. Эта функция делает память частью текущего вывода. В главе 28 подобный механизм будет доступен в работающем агенте.

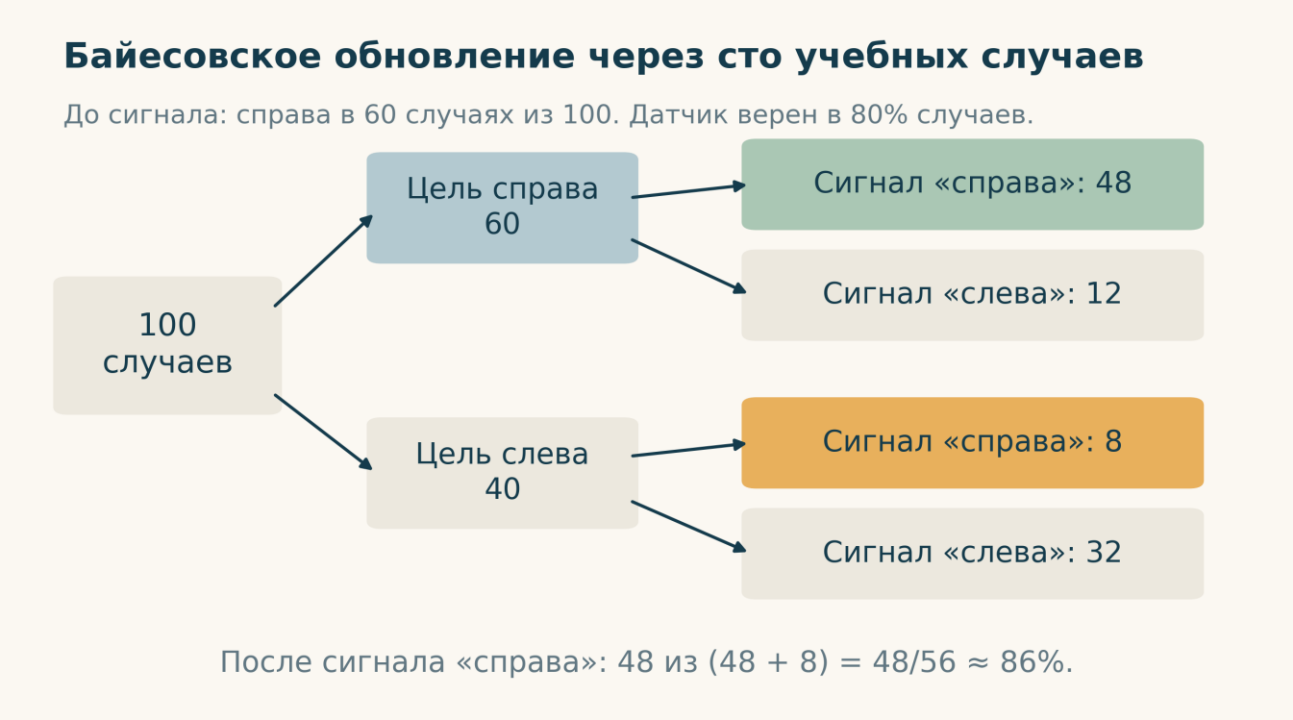

Рисунок 32. От предварительной оценки к обновлённой вероятности. Дерево использует ровно числа учебного примера главы: предварительную вероятность 0,6 и надёжность симметричного датчика 0,8. При сигнале «справа» остаются 56 случаев, из которых 48 имеют правую цель. Это расчёт при заданных предпосылках, а не наблюдённые частоты человеческого восприятия.

От модели мира к модели действия

Пусть устройство умеет посылать две команды. Пока команда «правая» перемещает указатель вправо, знание о цели достаточно для выбора. Затем связь незаметно меняется: правая команда даёт левый эффект. Если

устройство хранит только модель мира, оно продолжит знать, где находится цель, и одновременно будет действовать неверно. Значит, необходимо отдельно представлять отношение между собственными командами и наблюдаемыми последствиями.

В таком устройстве появляется функциональная модель собственной возможности действовать. Её содержание конкретно: какая команда с какой вероятностью производит какой эффект. Называть это полноценным человеческим самосознанием было бы преждевременно. Но можно точно показать, какую задачу решает самоотнесённая переменная: позволяет отличить ошибку знания о мире от ошибки управления действием.

Представьте, что действие оказалось неудачным. Система должна рассмотреть хотя бы три причины: неверно определена цель, изменилось соответствие команды и эффекта, случайно отказал исполнитель. Если все ошибки автоматически приписываются сенсору, обучение испортит хорошее восприятие. Если все приписываются действию, устройство будет бесконечно менять управление. Хорошая модель требует таких наблюдений, которые различают причины ошибки.

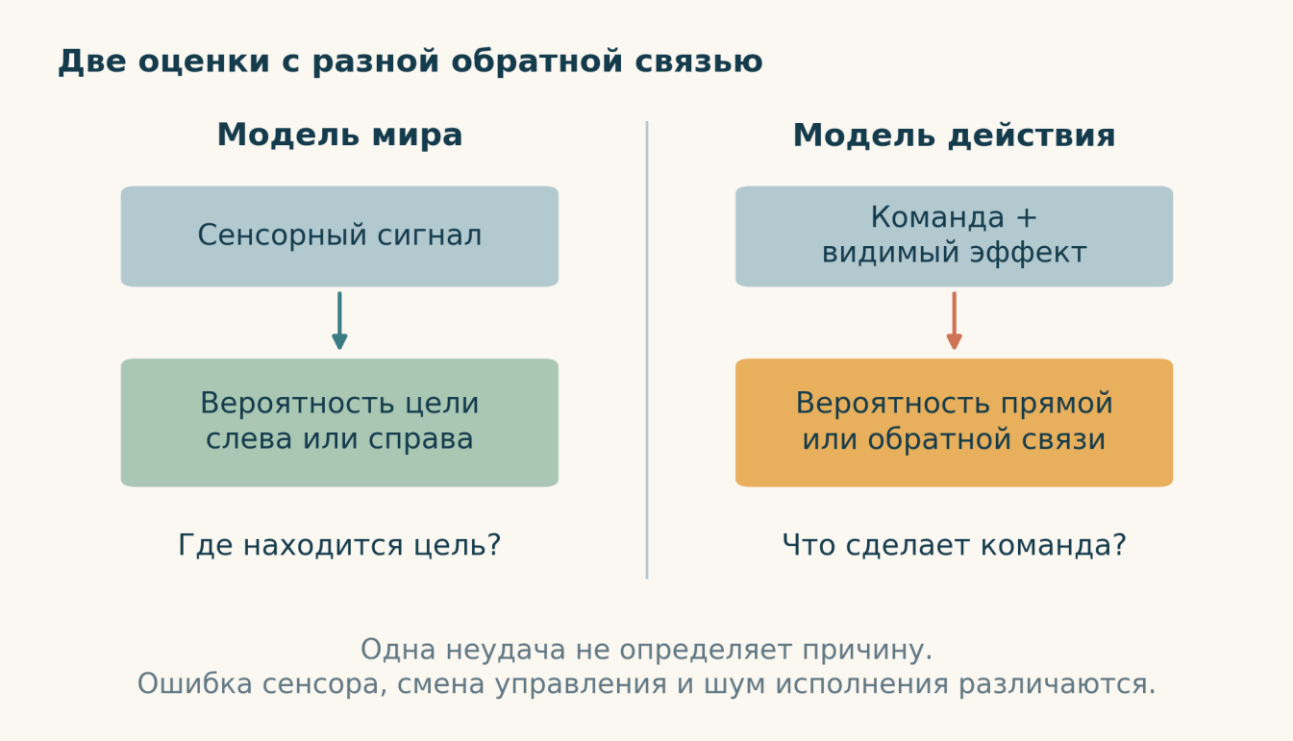

Рисунок 33. Раздельные модели мира и действия. Учебное устройство хранит отдельно убеждение о местоположении цели и о соответствии команды видимому эффекту. Разные наблюдения обновляют разные оценки. По одной неудаче нельзя однозначно установить ошибку восприятия, смену управления или случайный сбой исполнения.

Тело как общая причина сигналов

В байесовской модели Самада, Чунга и Шамс принадлежность видимой руки связывалась с выводом об общей причине зрительных, тактильных и проприоцептивных сигналов. Согласованность во времени и пространстве меняет правдоподобие их объединения. Эта работа предлагает механизм определённых эффектов телесной принадлежности; она не выводит из него все формы самоощущения и биографического «я». [15]

На собственном примере можно понять смысл общей причины. Вы двигаете пальцем, видите движение и получаете связанные телесные сигналы. Одна гипотеза объединяет их как последствия движения одной руки. Другая предполагает несколько независимых событий. При устойчивом согласовании первая модель оказывается экономнее по числу независимых совпадений. Однако вывод зависит от того, какие альтернативы вообще представлены и какие совпадения система научилась ожидать.

Телесная принадлежность и авторство действия также не обязаны изменяться вместе. В эксперименте Калькерта и Эрссона условия движения резиновой руки позволили различить субъективное владение рукой и ощущение управления её движением. Это аргумент против сведения двух оценок к одному показателю. Он не означает, что в обычной жизни они существуют изолированно. [4]

Предсказание внутренних состояний

Анил Сет предложил рассматривать эмоциональные состояния и телесное самоощущение через вывод о причинах сигналов изнутри организма. Подход связывает предсказание с регулированием телесного состояния. Это теоретическая программа: её полезность определяется тем, какие конкретные различия в ощущениях и поведении она объясняет и предсказывает. [22]

Для разработки собственной модели здесь возникает полезный вопрос. Одно и то же изменение входного сигнала может означать опасность, физическую нагрузку или ошибку измерения. Как система различает эти причины? В учебном устройстве можно задать внутренний ресурс, расходуемый действием, и датчик с шумом. Тогда его оценка собственного состояния будет проверяться по прогнозу будущей доступности действий. Мы получим модель регуляции ресурса, не основание приписывать устройству усталость как переживание.

Общий предел предсказательных объяснений состоит в широте их применения. Много процессов можно описать словами «ожидание» и «ошибка». Научную ценность создаёт точное различие между моделями:

какие причины допускаются, как меняется уверенность, что считается ошибкой и какое действие её уменьшает. Если эти правила подбираются заново под каждый результат, объяснение становится слишком гибким.

Ещё одна проверка касается зависимости наблюдений. Допустим, два датчика одновременно показывают правую сторону. Если они независимо ошибаются, совпадение может существенно усилить убеждение. Если второй просто копирует первый, новой информации нет. Модель, считающая копию независимым подтверждением, станет чрезмерно уверенной. В более сложном случае оба датчика могут зависеть от общей помехи, о которой устройство пока не знает.

Поэтому контроль уверенности требует модели источников ошибки, а не только формулы обновления. Для учебной проверки можно сравнить два одинаковых ряда показаний: один получен независимыми измерениями, другой создан копированием. Правильная оценка должна зависеть от того, какая процедура породила ряд. Этот пример помогает понять, почему множество согласующихся входов ещё не гарантирует надёжного знания о мире или о собственном состоянии.

Одно наблюдение и две разные неопределённости

До сигнала вероятность цели справа равна 0,6. Датчик сообщает верную сторону с вероятностью 0,8 и симметрично ошибается. После сигнала справа вес версии правой цели равен 0,6 умножить на 0,8, то есть 0,48. Вес левой версии равен 0,4 умножить на 0,2, то есть 0,08. Нормировка даёт 0,48 из 0,56, примерно 0,857. Это обновляет положение цели. Для выбора команды нужно ещё знать, вызывает ли правая команда правый эффект. Если связь инвертирована, более вероятная правая цель требует левой команды. Уточнение датчика не исправляет ошибочную модель управления.

Что теперь установлено

Мир, управление и сенсорная надёжность могут быть разными источниками ошибки. Успешное обучение требует сведений, разделяющих эти версии. Теперь проследим, какие данные позволяют обновлять каждую часть.

Практика 15 Различение трёх ошибок

Зачем выполнять: Отделить оценку положения цели от оценки связи команды с эффектом.

Что понадобится: Вероятностный переход главы 11. До действия доступны наблюдение и история; истинная цель и фактический исход относятся к среде и оценщику.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Повторите расчёт правой цели для исходной вероятности 0,6 и точности датчика 0,8. Затем получите прогноз для сигнала слева.
2. Для четырёх случаев цель слева или справа и связь прямая или обратная выпишите команду, которая достигнет цели без моторного шума.
3. В каждом случае укажите, что узнаётся из сенсорного наблюдения и что из сравнения команды с видимым эффектом.
4. Откройте подготовленный пример python teaching_lab.py --chapter 15. Он демонстрирует таблицу решений; скройте колонки истинной цели и связи от рассматриваемой модели.

Что сохранить: Два обновления вероятности, четыре решения и карта источников обратной связи.

Подсказка: Правдоподобие сигнала при гипотезе и вероятность гипотезы после сигнала занимают разные места в расчёте.

Исследовательское продолжение: Проведите бумажный опыт из двадцати проб с партнёром или закрытым столбцом среды. Смену управления задайте заранее и не сообщайте выбирающему команды. До результата записывайте прогноз, после него обновляйте только те состояния, для которых получены сведения.

Разбор и ответы — приложение Е, глава 15, с. 300

Вопросы для самопроверки

1. Является ли всякое предсказание сознательным ожиданием?
2. Чем модель действия отличается от модели мира?
3. Почему отсутствие сигнала нужно отличать от отрицательного сигнала?

Предсказательные механизмы стали определённее. Теперь вернёмся к реальным исследованиям и спросим, какие данные действительно различают теоретические объяснения.

Глава 16

Где данные различают теории

Мы умеем описать несколько механизмов и получить их ожидания. Но один и тот же удачный результат может быть совместим с несколькими теориями. Для продвижения важно найти условия, в которых их прогнозы расходятся, и понять, что именно измерял эксперимент. Исследования с отчётом и без отчёта особенно хорошо показывают, почему теория и процедура должны обсуждаться вместе.

Почему подтверждения недостаточно

Предположим, две модели ожидают усиления некоторого сигнала при видимом изображении. Эксперимент обнаруживает усиление. Результат может поддерживать обе модели, но почти не помогает выбрать между ними. Для различения требуется ситуация, в которой ожидания расходятся. Например, первая модель предсказывает сохранение сигнала без требования отвечать, а вторая связывает его с подготовкой ответа. Тогда изменение инструкции становится содержательным вмешательством в условия проверки.

Теории сознания особенно чувствительны к этой проблеме. Один эксперимент включает распознавание, удержание в памяти, уверенность, решение и движение. Если всё меняется вместе, найденное различие можно объяснить несколькими способами. Исследовательская работа состоит в разделении связанного набора процессов настолько, насколько это позволяет конкретная задача. Название мозговой области само по себе такого разделения не выполняет.

Предварительное соглашение о проверке

В проекте COGITATE сторонники конкурирующих подходов и исследовательский консорциум заранее согласовали дизайн, различающиеся предсказания и интерпретации. В опубликованном в 2025 году исследовании 256 участников рассматривали хорошо видимые стимулы различной длительности; использовались функциональная магнитно-резонансная томография, магнитоэнцефалография и внутричерепная электроэнцефалография. Результаты поддержали часть ожиданий и создали трудности для обеих теорий: среди них отсутствие ожидаемой устойчивой задней синхронизации и общего ожидаемого всплеска при исчезновении стимула. Проверялись согласованные предсказания, а не все возможные версии теорий. [23]

Для собственного проекта здесь важен принцип предварительного обязательства. До получения данных нужно написать, что будет считаться поддержкой, затруднением и неинформативным результатом. Если любое наблюдение после анализа объявляется ожидаемым, теория перестаёт рисковать ошибкой. Это касается и любимой гипотезы автора учебника: личная заинтересованность не заменяет условия её пересмотра.

Предварительная регистрация также не делает интерпретацию непогрешимой. Можно заранее выбрать слабый показатель или слишком грубое предсказание. Поэтому нужны две проверки: соответствует ли эксперимент записанному плану и действительно ли план проверяет центральное утверждение. Если второе вызывает сомнение, следует явно изменить формулировку и провести новую проверку, сохранив прежний результат в истории исследования.

Отчёт меняет исследуемую задачу

В исследовании Коэна и коллег 2024 года зрительная маскировка сравнивалась в условиях с отчётом и без него. Поздний компонент P3b и некоторые устойчивые характеристики декодирования были связаны с выполнением отчёта. В условии без отчёта более ранний компонент N2 соответствовал нелинейному изменению видимости. Это результат конкретной задачи и способа измерения; он не задаёт универсального времени появления любого опыта. [24]

Когда человеку нужно сообщить, что он видел, он должен преобразовать восприятие в ответ, сохранить нужную информацию и выполнить правило. Поэтому исчезновение сигнала после отмены отчёта может указывать на его связь с этими операциями. Но отсутствие требования отвечать не гарантирует отсутствия внутреннего называния, внимания или запоминания. Участник способен продолжать выполнять часть прежней задачи самостоятельно.

Возникает и обратная трудность. Если не спрашивать о восприятии, как узнать, что оно возникло? В некоторых дизайнах используются косвенные показатели, например особенности движений глаз. Их связь с переживанием необходимо предварительно обосновывать, а затем проверять переносимость на новые условия. Иначе метод, задуманный для исключения отчёта, просто заменяет один спорный показатель другим.

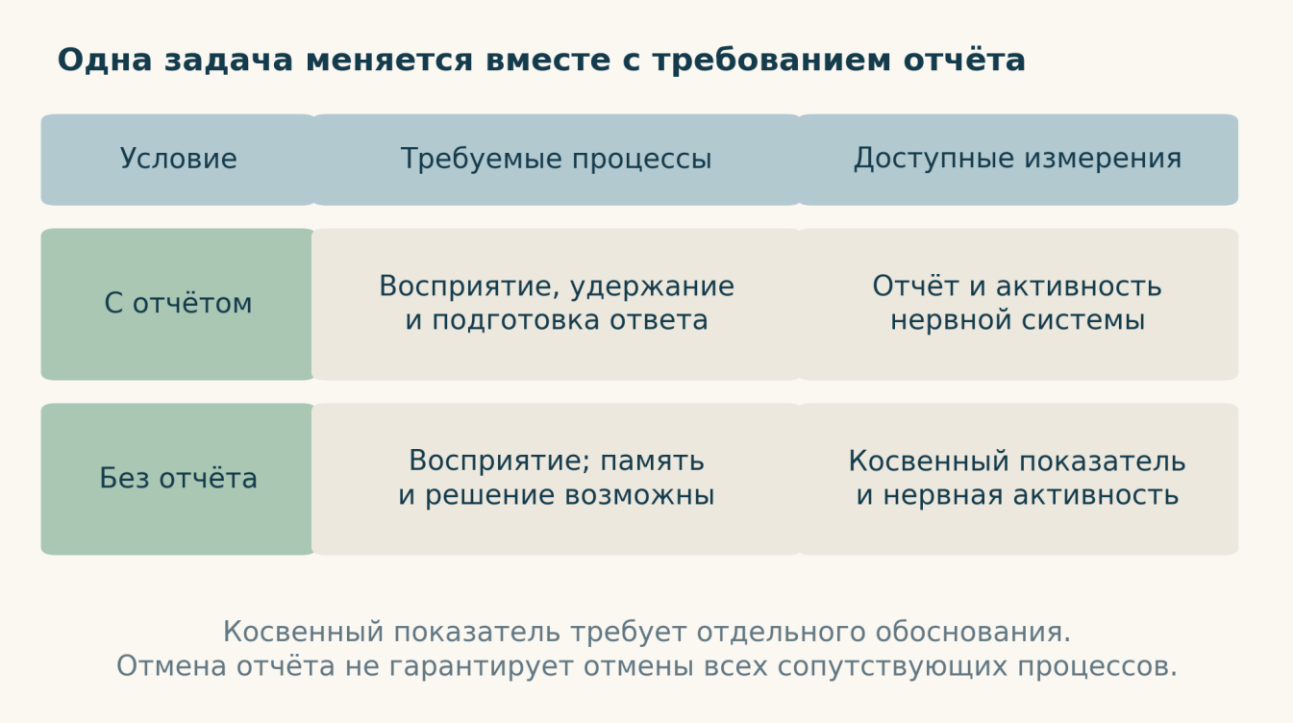

Рисунок 34. Отчёт и границы измерения. Концептуальное сопоставление условий с явным отчётом и без него. Удаление требования ответа меняет задачу, но не обеспечивает автоматического удаления внимания, памяти или внутреннего решения. Косвенные признаки восприятия также нуждаются в проверке. В таблице нет количественных результатов исследований.

Почему передняя и задняя кора не образуют простую развилку

Хэ и коллеги в опубликованной работе 2026 года сравнили зрительное восприятие с отчётом и без него при регистрации магнитной активности мозга. Представления в задних сенсорных областях сохранялись в обоих условиях, тогда как обнаруженная префронтальная представленность содержания зависела от отчёта. Это поддерживает разделение обработки содержания и требований задачи. Наблюдательная регистрация, однако, не доказывает причинную достаточность одной области. [25]

В другой опубликованной работе 2026 года Чжан и коллеги обнаружили различия между лобными подсистемами. Декодирование, связанное с лобными участками дорсальной сети внимания, переносилось между условиями с отчётом и без него; другие лобные представления зависели от отчёта. В условии без отчёта использовались глазодвигательные показатели. Результат препятствует обобщению «вся лобная активность относится только к ответу», но не устанавливает, что выделенная сеть сама создаёт переживание. [26]

Эти работы полезно читать совместно, не усредняя их в один лозунг. Они могут различаться стимулами, вниманием, способом определения осознанности, пространственной точностью и чувствительностью анализа. «Не удалось декодировать содержание» означает, что выбранный анализ не извлёк его из данных при заданных условиях. Это отличается от утверждения, что содержание нигде не представлено и не играет причинной роли.

Столь же осторожно следует интерпретировать успешное декодирование. Классификатор может использовать следствие восприятия, общий источник различий или побочный признак. Само наличие информации о стимуле в сигнале не определяет функцию этой информации. Вопрос о необходимости требует дополнительного вмешательства или сильной причинной аргументации.

Как читать противоречие продуктивно

Допустим, одна статья сообщает связь сигнала с видимостью, а другая не находит её. Начните с точного предмета сравнения. Сигнал мог измеряться на разных временных интервалах; участники могли использовать разные стратегии; отсутствие эффекта могло сопровождаться слишком большой неопределённостью. Сначала выясняют, несовместимы ли результаты при сопоставимых условиях, и только потом обсуждают пересмотр механизма.

Для учебного разбора полезна следующая таблица. Она сравнивает логическую силу результатов, а не оценивает публикации по престижу.

Наблюдение	Непосредственный вывод	Чего ещё не установлено
Сигнал связан с отчётом	Обнаружена зависимость от условий ответа	Что сигнал не участвует в других функциях
Содержание декодируется	Данные несут различающую информацию	Что область необходима или достаточна
Вмешательство меняет результат	Есть причинное влияние при данных условиях	Что изменено именно переживание
Эффект не обнаружен	Анализ не выявил ожидаемого различия	Что различие отсутствует с полезной точностью

Что меняет отмена обязательного отчёта

Предположим, сигнал связан с видимостью в задаче, где после каждого предъявления нужно отвечать. Возможны версии о восприятии и о подготовке сообщения. Сохранение сигнала при обоснованном условии без обязательного

ответа ограничивает вторую версию, но не устанавливает причинную достаточность сигнала для переживания. Исчезновение сигнала тоже требует проверки: вместе с ответом могли измениться внимание и память. Поэтому один вопрос формулируется узко: какая характеристика сохраняется при изменении требований отчёта и каким независимым способом подтверждается само восприятие.

Что теперь установлено

Исследование становится различающим относительно явно заданных конкурентов. Крупная выборка не заменяет нужного сравнения, а отсутствие статистической находки не означает точного нуля.

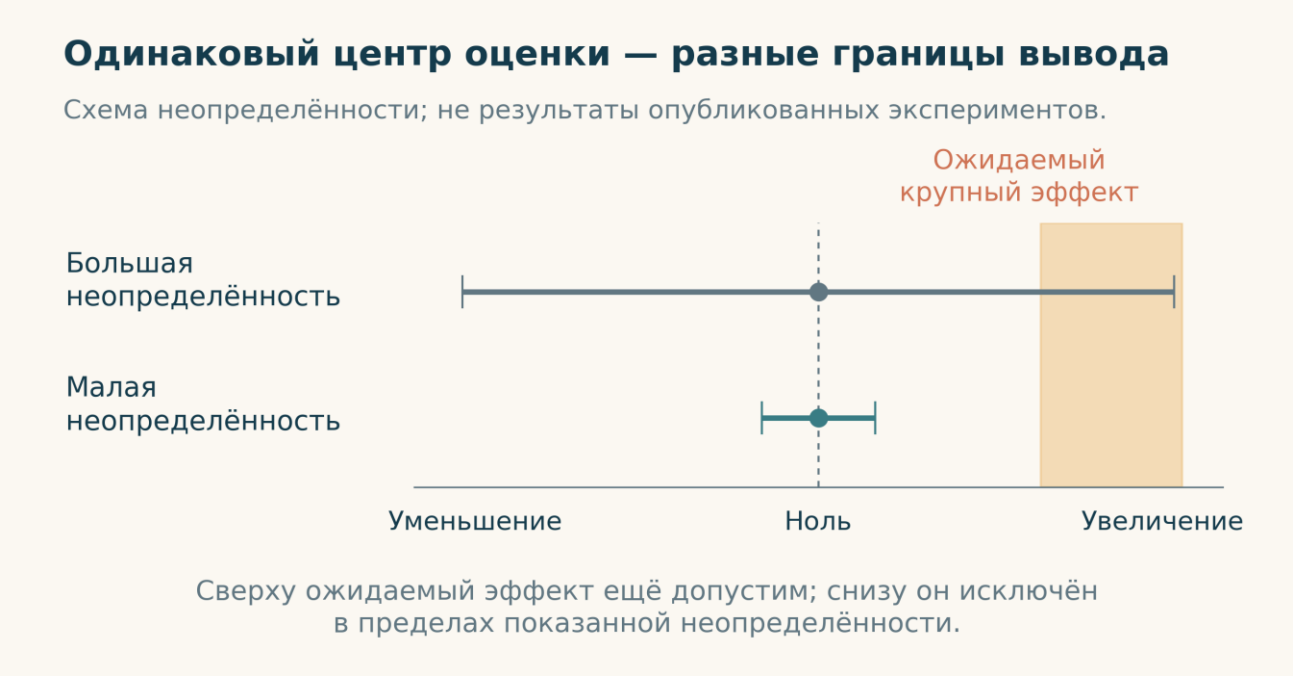

Рисунок 35. Почему отсутствие значимости не равно точному нулю. Два схематических интервала имеют один центр около нуля, но допускают разный диапазон эффектов. Только узкий интервал на рисунке исключает заранее ожидаемое крупное увеличение. Числовой масштаб намеренно не задан: в реальной проверке нужны оценка эффекта, обоснованный интервал и количественное обещание теории.

Практика 16 Сопоставление двух исследований

Зачем выполнять: Сопоставить два исследования по одному вопросу, сохраняя различия методов и выводов.

Что понадобится: Паспорт измерения из главы 6, качество сравнения из главы 7, величина и неопределённость эффекта из главы 8. Декодирование означает извлечение различающей информации выбранным анализом.

Ориентир по времени: 35–50 минут для базового разбора. Можно делать с перерывами; это не норматив скорости.

1. Выберите из главы две работы с условиями отчёта и без него, например источники 24 и 25. Сначала используйте их разбор в тексте, не приписывая себе чтение полного оригинала.
2. Для каждой заполните пять полей: стимул, задача, показатель видимости, измеряемая активность, главное сравнение. Неизвестные сведения пометьте явно.
3. Запишите один общий вопрос и для каждой работы по одному поддерживающему и ограничивающему наблюдению.
4. Сформулируйте следующий прогноз при изменении способа ответа и спросите, ожидает ли конкурент того же результата.

Что сохранить: Два сопоставимых паспорта и одно предсказание с указанием, различает ли оно теории.

Подсказка: Не переносите результат между ЭЭГ, МЭГ и томографией только по сходству слов. Отдельно указывайте временной интервал и способ анализа.

Исследовательское продолжение: Откройте обе первичные публикации, уточните пропуски и проверьте, не изменились ли сравниваемые задачи. Отдельно оцените величину и неопределённость ключевого эффекта и сохраните различие между выводом авторов и своим.

Разбор и ответы — приложение Е, глава 16, с. 300

Вопросы для самопроверки

1. Почему крупное исследование не обязано окончательно решить спор теорий?
2. Является ли условие без отчёта прямым измерением опыта?
3. Что сильнее меняет доверие к модели — ожидаемый результат или различающее предсказание?

Сравнение теорий подготовило вопрос о возникновении. Сначала сузим его до одного эпизода: как содержание появляется и поддерживается во времени.

Часть 5

Возникновение сознания и самосознания

Глава 17

Как возникает отдельный эпизод

До сих пор мы сопоставляли механизмы и показатели. Теперь уточним, что означает появление сознательного эпизода. Этот вопрос отличается от развития ребёнка, эволюции видов и основания самого наличия опыта. В этой главе разберём только динамику одного события: условия запуска, процесс поддержания и последующее сообщение. Пороговый накопитель из главы 9 даёт язык для такой задачи, но не универсальную временную границу сознания.

Четыре значения возникновения

Вопрос «как возникает сознание» объединяет задачи с разными масштабами. Динамическая задача касается отдельного эпизода: почему сейчас появилось определённое содержание и что поддерживало его следующие мгновения. Задача развития исследует изменение возможностей одного организма на протяжении жизни. Эволюционная задача спрашивает об истории механизмов в линиях происхождения. Онтологическая задача касается того, почему определённый физический процесс вообще сопровождается переживанием. Ответ на одну задачу способен помочь другой, но не заменяет её.

Эта глава посвящена динамическому вопросу. Для примера возьмём короткое появление силуэта в окне. Нужно различить подготовленность системы к восприятию, обработку поступившего сигнала, содержание переживания и последующий ответ. Даже если всё происходит быстро, логически это разные звенья. Самая поздняя запись о событии не обязана совпадать с началом события, а самый ранний ответ на свет не обязан быть началом сознательного видения.

Предпосылка, процесс и последствие

Предпосылка делает событие возможным при данных условиях. Процесс непосредственно реализует интересующее изменение согласно проверяемой модели. Последствие возникает потому, что событие уже произошло. Эти роли

определяются причинными отношениями, а не только порядком на часах. Один процесс способен участвовать в нескольких ролях: память о предыдущем стимуле влияет на следующий эпизод, хотя сама сформировалась как последствие предыдущего.

Рассмотрим учебную установку. На экране появляется фигура; участник сообщает её форму через несколько секунд. Предварительная инструкция задаёт, на что обращать внимание. Сенсорная обработка различает форму. Память удерживает результат. Выбор ответа преобразует его в движение. Если ухудшить память, поздний отчёт станет менее точным, даже когда первоначальное различение не изменилось. Поэтому вывод о возникновении эпизода требует контроля временной цепочки.

Обратная ошибка возникает, когда раннее изменение считают достаточным объяснением всего эпизода. Представим датчик, быстро различающий два стимула. Из этого ещё не следует, что различие доступно дальнейшему поведению или что оно переживается. Название «ранний маркер сознания» должно сопровождаться объяснением того, почему показатель относится именно к сознательному содержанию, а не к его входному условию.

Учебная динамика появления и поддержания

Построим собственную простую модель. У неё есть число активности от нуля до единицы. На каждом шаге сохраняется половина прежней активности. Наличие входного сигнала добавляет четыре десятых. Если прежняя активность уже достигла половины, дополнительное внутреннее правило добавляет три десятых. Значение ограничено единицей. Все числа выбраны для прозрачности вычисления; они не оценивают параметры человеческого мозга.

Два последовательных входа дают активности четыре десятых и шесть десятых. Затем вход исчезает. При наличии внутреннего поддержания активность остаётся на шести десятых; без него падает сначала до трёх десятых, затем до пятнадцати сотых. Одна и та же история входа порождает разную длительность внутреннего состояния.

Шаг	Вход	Без внутреннего поддержания	С внутренним поддержанием
Первый	Есть	0,40	0,40
Второй	Есть	0,60	0,60

Шаг	Вход	Без внутреннего поддержания	С внутренним поддержанием
Третий	Нет	0,30	0,60
Четвёртый	Нет	0,15	0,60

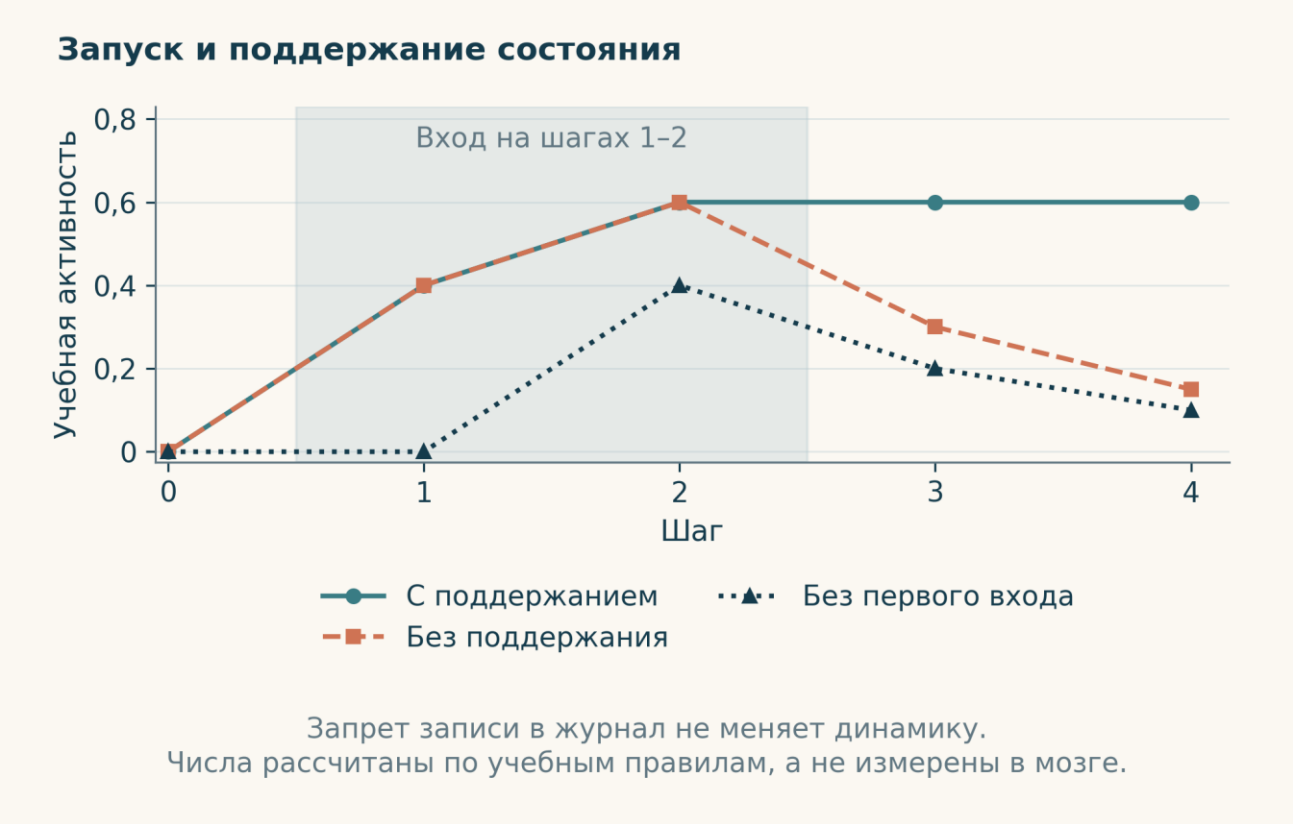

Рисунок 36. Три траектории одного учебного механизма. Расчёт правила главы 17: на каждом шаге сохраняется половина прежней активности, вход добавляет 0,4, а при прежней активности не ниже 0,5 внутреннее поддержание добавляет 0,3. При двух входах поддерживаемая активность остаётся на уровне 0,6. Удаление первого входа и удаление поддержания имеют разные последствия. Запись в журнал наблюдателя в правило состояния не входит.

В модели можно отдельно исследовать запуск и сохранение состояния. Один вход недостаточен для запуска устойчивого режима; два подходят. После запуска дальнейший вход уже не нужен по заданным правилам. Однако мы получили устойчивую переменную, а не доказательство переживания. Кроме того, одно число не объясняет, чем красный круг отличается от тихого звука. Для содержательного объяснения нужны представления различий между возможными эпизодами.

В реальном исследовании аналогичная логика требует разных вмешательств до начала, во время поддержания и после окончания предъявления. Если вмешательство действует только на запуск, это одно ограничение модели.

Если только на сохранение — другое. Совпадение эффектов также информативно: возможно, предложенное разделение не соответствует устройству системы или вмешательство затрагивает слишком много процессов сразу.

Что показывает изменение механизма

В двух исследованиях Ноормана и коллег с регистрацией электрической активности мозга изучалось влияние мемантина на обработку зрительных фигур. Было обнаружено изменение нейронных показателей, интерпретированных как рекуррентная интеграция; чувствительность восприятия при этом не улучшилась. Использованная здесь версия eLife опубликована 18 июня 2025 года. Для нашего вопроса существен разрыв между изменением предполагаемого механизма и отсутствием соответствующего изменения поведенческой чувствительности. [27]

Такой результат учит не считать все показатели взаимозаменяемыми. Возможно, измеряется процесс, не ограничивающий выполнение именно этой задачи. Возможно, поведенческий показатель недостаточно чувствителен к интересующему изменению. Возможно, первоначальная механистическая интерпретация неполна. Выбор между объяснениями требует новых данных. Он не совершается автоматически в пользу любимой теории.

Из этого следует и правило для собственной модели. Если выключение переменной не меняет измеряемого результата, нельзя сразу заключать, что переменная не нужна ни для чего. Нужно проверить, действительно ли вмешательство сработало, была ли задача чувствительна к её функции и не компенсировал ли изменение другой механизм. Однако бесконечное добавление непроверенных компенсаций тоже не спасает гипотезу: каждая должна породить новое проверяемое следствие.

Время содержания и время сообщения

Когда человек говорит «я заметил это сразу», фраза сообщает о его оценке события. Чтобы превратить её в временное измерение, нужны инструкция, шкала, условия сравнения и оценка ошибки. Нельзя просто приравнять момент нажатия кнопки к началу переживания: между ними может находиться принятие решения и выполнение движения. Вычитание среднего времени движения также не устраняет все сложности, поскольку разные этапы могут взаимодействовать.

Полезно искать согласование нескольких показателей с разными ограничениями. Например, изменение видимости при разной длительности предъявления, характер ошибок, субъективная оценка и нейронная динамика

дают разные стороны эпизода. Совпадение не гарантирует окончательного объяснения, но помогает исключить некоторые альтернативы. Несовпадение следует сохранить как результат, а не устранять переименованием показателей.

Динамическое объяснение также должно описывать завершение и смену содержания. Модель, умеющая только усиливать активность, может застревать на первом стимуле. Ей понадобится механизм прекращения, конкуренции или обновления. Проверка последовательности разных входов поэтому нередко информативнее демонстрации одного успешного распознавания.

Отдельная проблема возникает при усреднении времени. Допустим, в каждой пробе учебная система переключается резко, но момент переключения немного различается. Средняя по пробам кривая будет выглядеть плавной. Обратная ситуация тоже возможна: выбор проб по порогу способен создать впечатление более резкого перехода, чем существует в отдельных траекториях. Поэтому форма среднего сигнала не определяет однозначно динамику единичного эпизода.

Для проверки сохраняйте полные траектории отдельных проб и правило их отбора. Сравнивайте две модели: постепенное изменение в каждой пробе и резкое изменение с разным временем запуска. Если они одинаково описывают среднюю кривую, потребуется дополнительный показатель распределения или переноса во времени. Это конкретный способ сделать вопрос о возникновении эпизода более точным, не назначая один момент универсальной границей сознания.

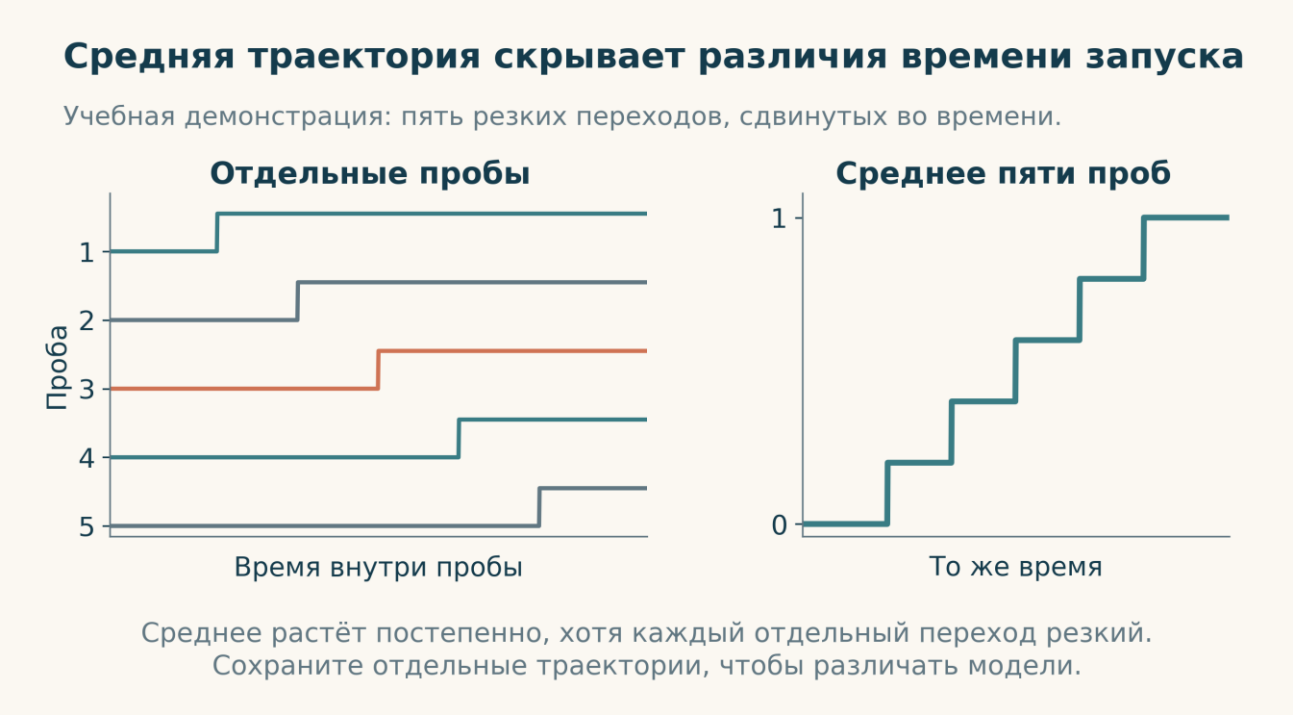

Рисунок 37. Усреднение и динамика отдельного эпизода. В пяти искусственных пробах состояние скачком меняется с нуля на единицу, но моменты перехода различаются. Справа вычислено их точное среднее, которое растёт ступенями; при множестве сдвигов средняя кривая может выглядеть плавной. Этот пример объясняет проблему измерения динамики и не обозначает момент возникновения опыта.

Почему средняя кривая не выбирает динамику отдельной пробы

Пять проб могут содержать одинаково резкий переход от нуля к единице, но запускаться в разные моменты. Среднее по времени тогда поднимается ступенями и при плотном разбросе выглядит плавным. Другая модель может плавно расти в каждой отдельной пробе и иметь похожее среднее. Сохранение только усреднённой кривой делает эти объяснения трудноразличимыми. Нужны полные траектории, распределение моментов перехода или дополнительная проверка поддержания.

Что теперь установлено

Мы различили запуск, сохранение и регистрацию результата. Эти роли нужно проверять подходящими вмешательствами. Теперь увидим их на небольшом заранее определённом механизме.

Практика 17 Проверка запуска и сохранения

Зачем выполнять: Сравнить запуск и поддержание активности по правилам учебной модели главы.

Что понадобится: Правило изменения состояния и порог; различие предпосылки, процесса и последствия. Временная близость сама по себе не устанавливает причинную роль.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Начните с активности 0. На каждом шаге сохраняйте половину прежнего значения; вход добавляет 0,4; если прежняя активность не меньше 0,5, добавьте 0,3. Ограничьте итог единицей.
2. Подайте вход на шагах 1 и 2, затем уберите его. Заранее предскажите траекторию до шага 4.
3. Повторите без поддерживающей прибавки 0,3, а затем без первого входа. Сравните каждую траекторию с полной версией.
4. Отдельно поставьте метку записи результата после шага 4. Проверьте, меняет ли эта метка саму динамику.

Что сохранить: Три траектории, различение запуска и поддержания и пояснение роли последующей записи.

Подсказка: Условие дополнительной прибавки проверяется по прежнему состоянию, а не по уже обновлённому значению того же шага.

Исследовательское продолжение: Создайте несколько проб с разным временем запуска и сравните их среднее с моделью плавного изменения. Предложите показатель, который различит эти случаи.

Разбор и ответы — приложение Е, глава 17, с. 301

Вопросы для самопроверки

1. Почему динамическое возникновение не равно эволюционному происхождению?
2. Доказывает ли устойчивая активность наличие сознательного эпизода?
3. Что означает расхождение нейронного и поведенческого показателей?

Динамика отдельного эпизода не ответила, как меняются способности человека с возрастом. Следующая глава сохранит это различие и рассмотрит развитие нескольких возможностей по отдельности.

Глава 18

Как развивается самосознание

Эпизод возникает за короткое время, а способность узнавать себя или запрашивать помощь развивается на другой временной шкале. Поэтому вопрос о возрасте появления сознания нужно разделить на конкретные достижения. Мы рассмотрим сенсорное содержание, визуальное самоузнавание, контроль неопределённости и биографическое самоописание. Такое разделение позволит не считать неудачу одного теста отсутствием всех остальных возможностей.

Развитие нескольких возможностей

Вопрос «когда ребёнок осознаёт себя» кажется просьбой назвать возраст. Но сначала нужно решить, о какой способности идёт речь. Различение собственного движения и внешнего события, узнавание тела, оценка своей ошибки, использование слова «я» и размышление о собственных основаниях — разные достижения. У них могут быть общие предпосылки, однако один показатель не измеряет автоматически все остальные.

В этой главе развитие означает изменение организации и возможностей одного человека во времени. Оно включает приобретение новых способов взаимодействия, изменение памяти, обучения и контроля. Мы не предполагаем заранее единственную точку, до которой никакого опыта нет, а после которой одновременно появляются все формы самосознания. Задача исследования — установить, что именно меняется и какие наблюдения позволяют это различить.



Рисунок 38. Развитие в повседневном взаимодействии. Предметная иллюстрация раннего взаимодействия с предметом. Направленное движение, различение последствий действия и последующий рассказ о себе требуют разных наблюдений; изображение не задаёт возрастного порога появления сознания.

Переживание и способность сообщить

Куидер и коллеги исследовали ответы мозга на маскированные лица у младенцев пяти, двенадцати и пятнадцати месяцев. Во всех возрастных группах обнаружили поздний нелинейный ответ, который интерпретировался по аналогии с показателем сознательного доступа у взрослых; у младших он был слабее и медленнее. Работа даёт нейронный аргумент в пользу ранней способности к сознательному восприятию, но не устанавливает возраст самого первого переживания и не измеряет развёрнутую рефлексию. [28]

Логика такого переноса требует внимательности. Если показатель связан с переживанием у взрослых, его присутствие в другой группе поддерживает некоторую общность механизмов. Но исследователь должен проверить, выполняет ли сигнал ту же функцию. Отсутствие взрослого паттерна также неоднозначно: система может решать похожую задачу другим способом или измерение может быть недостаточно чувствительным к её организации.

Это относится не только к нейронным данным. Речевой вопрос добавляет требования к пониманию слов, удержанию инструкции и выбору ответа. Ребёнок может не выполнить речевую задачу по причинам, не связанным с

отсутствием исследуемого состояния. В то же время успешный ответ, выученный как готовая фраза, может преувеличить сложность понимания. Поэтому полезно сравнивать несколько задач с разными дополнительными требованиями.

Что показывает узнавание в зеркале

В классическом исследовании Бьюлы Амстердам наблюдались возрастные различия реакции детей на отражение; у части детей к концу второго года отмечались признаки узнавания себя. Даже внутри изученной группы результаты не были одинаковыми. Эти данные относятся к определённой ситуации визуального самоузнавания, а не к универсальному календарю возникновения сознания. [29]

Разложим задачу на собственную учебную схему. Нужно обнаружить необычную деталь отражения, связать изображение с телом, найти соответствующее место на теле и выполнить действие. Неуспех может возникнуть на любом звене. Например, человеку трудно двигать рукой, деталь не кажется интересной или изображение воспринимается непривычно. Чтобы интерпретировать результат, требуются контрольные наблюдения о каждом дополнительном требовании.

Успех тоже допускает несколько уровней объяснения. Он может опираться на устойчивую связь между движением и изображением. Более сильное утверждение касается представления себя как того же человека вне текущего взаимодействия. Ещё более сильное — способности рассуждать о себе как об объекте чужой оценки. Для перехода к каждому следующему утверждению нужны дополнительные задачи, а не расширение значения одного успешного движения.

Представим взрослую программу, которая обнаруживает метку на изображении камеры и касается соответствующего участка корпуса. Её можно специально научить такой связи. Сам этот инженерный пример показывает, почему поведенческий критерий нужно разлагать на механизмы. Он не доказывает, что ребёнок выполняет задачу тем же способом, и не отменяет значения самоузнавания в исследовании развития.

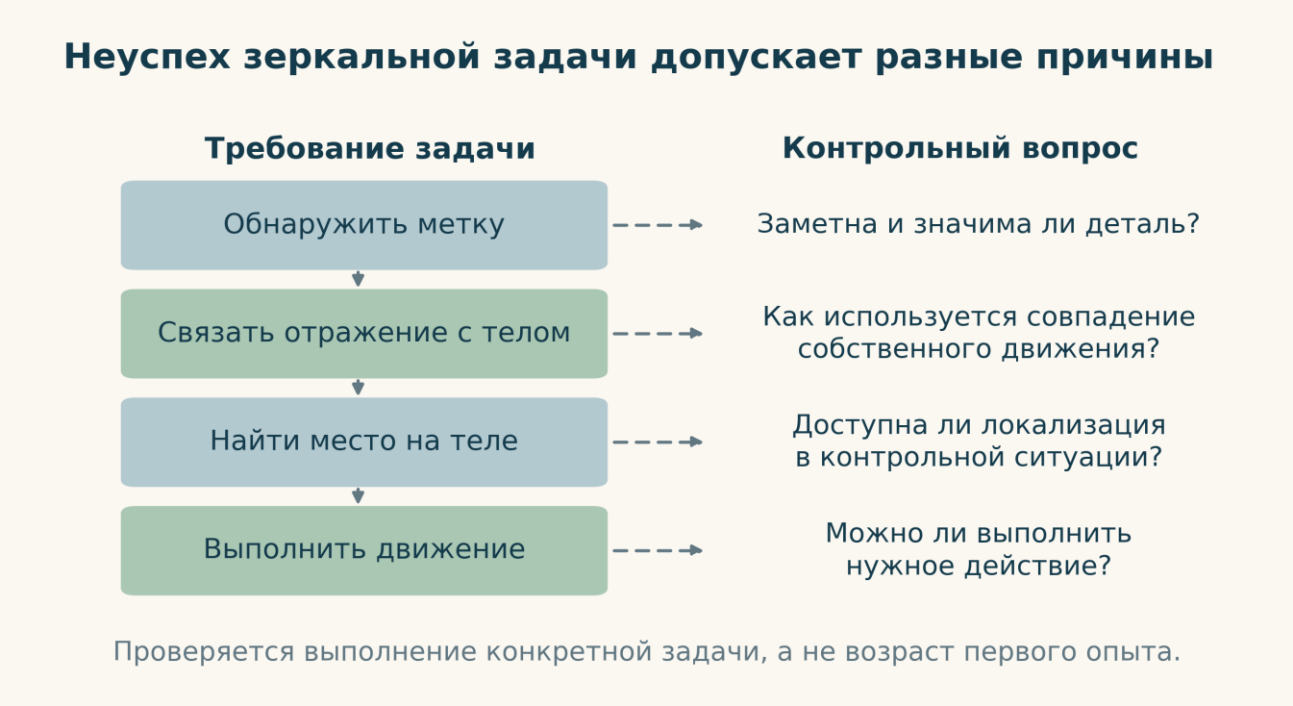

Рисунок 39. Разложение задачи визуального самоузнавания. Касание метки зависит от нескольких условий: её обнаружения, связи отражения с телом, локализации и возможности движения. Контрольные вопросы помогают отделить эти причины неуспеха, но сами ещё не образуют полный экспериментальный протокол. По одной неудачной пробе нельзя устанавливать отсутствие переживания или других форм самоотнесения.

Контроль незнания

В исследовании Гупиль, Роман Моннье и Куидера двадцатимесячным детям предоставляли возможность запросить помощь в задаче на память. Использование помощи зависело от неопределённости и улучшало выполнение по сравнению с необходимостью всегда выбирать самостоятельно. Авторы интерпретировали это как ранний невербальный контроль собственного знания. Из результата не следует, что ребёнок располагает взрослым понятием знания или способен явно объяснить основания решения. [30]

Для понимания механизма различим оценку трудности ситуации и оценку качества собственного воспоминания. В первом случае ребёнок может выучить правило: после долгой задержки просить помощь. Во втором он использует информацию о том, насколько надёжно помнит конкретное событие. В обычной задаче обе стратегии способны совпадать. Различающая проверка должна разорвать эту связь, например сопоставить разные возможности наблюдать событие при одинаковой задержке.

Важна и социальная сторона запроса. Чтобы помощь стала измеряемым действием, нужно понимать, что другой человек может и готов её

предоставить. Поэтому редкие обращения могут отражать устройство ситуации общения, а не только уверенность. Такой анализ не обесценивает поведение, а показывает, какие функции совместно делают его возможным.

От событий к биографическому я

Разберём собственный пример приобретения самоописания. Ученик несколько раз ошибается в устном счёте и заключает: «Я плохо считаю». В этой фразе объединены память о событиях, обобщение результатов и применение характеристики к себе. Если затем ученик обнаруживает, что ошибки возникали только при спешке, самоописание может измениться: «При ограничении времени я выбираю слишком рано». Это более точная модель условий собственной работы.

Такой переход показывает исследовательский смысл рефлексии. Объектом изменения становится правило, по которому человек объясняет свои результаты. Можно изучать, различает ли он постоянную способность и временные условия, ищет ли дополнительное наблюдение, переносит ли уточнённое правило на новую задачу. Возраст сам по себе этого не объясняет: нужны сведения о предшествующем опыте и доступных способах проверки.

При исследовании развития следует также отличать сравнение возрастных групп от наблюдения одной и той же группы во времени. Первое показывает различия между обследованными людьми. Второе позволяет проследить индивидуальные изменения, но повторение задачи само может влиять на результат. Обе стратегии требуют контроля того, насколько сопоставимы измерения.

Представим, что при повторном обследовании ребёнок лучше запрашивает помощь, но правило обращения специально объясняли между встречами. Улучшение подтверждает изменение выполнения; оно не позволяет само по себе разделить обучение правилу и изменение контроля неопределённости. Для такого разделения нужна новая задача с другим способом запроса. Если перенос сохраняется, узкое объяснение одним выученным жестом становится менее достаточным.

Подобный анализ помогает избежать и обратного вывода. Неудача переноса может отражать непонимание новой ситуации общения. Поэтому полезно отдельно проверять, доступна ли новая форма ответа, прежде чем считать изменившийся результат показателем развития или отсутствия внутренней способности.



Рисунок 40. Профиль развития без общей лестницы. Четыре вымышленных наблюдения из практики главы отвечают на разные вопросы: самоузнавание, контроль неопределённости, связь событий с собой и пересмотр оснований. Их следует измерять раздельно. На схеме нет возрастных границ или единой шкалы сознания.

Как разложить зеркальную задачу

Чтобы коснуться метки, обнаруженной через отражение, нужно заметить саму метку, связать отражение с телом, локализовать соответствующее место и выполнить движение. Неудача любого звена может дать один и тот же внешний результат. Поэтому исследователь сначала проверяет доступность этих действий в контрольных условиях. Успех тоже не переносится автоматически на взрослую рефлексию: он относится к определённой организации визуального самоузнавания. Разложение задачи не обесценивает тест; оно делает смысл результата точнее.

Что теперь установлено

Единственная календарная граница не заменяет профиля возможностей. Можно исследовать развитие каждого отношения и затем выяснять, какие предпосылки у них общие.

Практика 18 Четыре наблюдения и несколько объяснений

Зачем выполнять: Научиться различать данные о способности и дополнительные требования конкретного задания.

Что понадобится: Перспектива, принадлежность и авторство из главы 3; различие успешности и самоконтроля из главы 4. Развитие здесь относится к изменениям одного организма во времени.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Создайте четыре строки: зрительное различение, узнавание в зеркале, запрос помощи при неопределённости, рассказ о своём прошлом.
2. Для каждой укажите наблюдаемое действие и хотя бы одно дополнительное требование: речь, память инструкции, движение или понимание возможности помощи.
3. Для неудачного результата предложите два объяснения и контроль, способный их развести.
4. Выберите одну опубликованную работу главы и ограничьте итог её задачей и возрастной группой, не назначая возраст первого опыта.

Что сохранить: Профиль четырёх вопросов и один разобранный контроль альтернативного объяснения.

Подсказка: Для обсуждения не нужно проводить домашние испытания с ребёнком. Используйте опубликованное описание и учебные случаи.

Исследовательское продолжение: Предложите перенос способности на другую форму ответа. Сначала отдельно убедитесь, что новая форма действия понятна участнику, а затем обсуждайте перенос внутреннего механизма.

Разбор и ответы — приложение Е, глава 18, с. 301

Вопросы для самопроверки

1. Почему возраст успешного самоузнавания нельзя считать возрастом появления сознания?
2. Чем запрос помощи отличается от явной рефлексии?
3. Что означает отсутствие ожидаемого показателя у младшей группы?

Развитие отдельного человека имеет свои данные и ограничения. Теперь перейдём к другой истории: происхождению механизмов в эволюционных линиях.

Глава 19

Как исследовать эволюцию опыта

Развитие одной особи и история происхождения вида требуют разных свидетельств. Теперь рассмотрим, как сравнение функций, устройства и родства ограничивает рассказ об эволюции механизмов. Польза памяти зависит от среды, а сходная способность может иметь разную историю. Поэтому мы не будем строить единую лестницу сознательности. Задача главы состоит в явном разборе переходов от наблюдаемого признака к исторической гипотезе и от механизма к предполагаемому опыту.

История механизма и история переживания

Устройство живого организма имеет историю происхождения. Поэтому естественно спросить, когда возникли механизмы, связанные с опытом, и какие задачи способствовали их сохранению. Но реконструкция истории механизма и установление наличия переживания — разные исследовательские операции. Можно хорошо объяснить пользу памяти и всё ещё спорить о том, при какой организации память сопровождается сознательным воспоминанием.

Эволюционный вопрос следует формулировать узко. Например: при каких условиях отбор мог поддержать объединение нескольких источников информации, контроль неопределённости или различение собственных действий и внешних событий? Такие формулировки задают функции, которые можно моделировать и сравнивать. Следующий вопрос — являются ли эти функции условиями опыта — требует отдельной теоретической связи. Смешение вопросов создаёт слишком уверенные рассказы о происхождении квалиа.



Рисунок 41. Среда задаёт разные задачи поведения. Предметная иллюстрация разных условий жизни у воды и на суше. Польза памяти, поиска и выбора действия зависит от конкретной среды; расположение организмов на рисунке не задаёт порядок возникновения опыта.

Как работает объяснение через преимущество

Рассмотрим собственную модель двух наследуемых стратегий. Первая каждый раз выбирает действие по последнему сигналу. Вторая сохраняет несколько недавних наблюдений. Мир иногда меняется, а датчик ошибается. При устойчивом мире память может уменьшать влияние случайных ошибок. При быстро меняющемся мире та же память способна задерживать переключение и делать поведение хуже. Следовательно, преимущество механизма зависит от среды и его параметров.

Добавим цену хранения и обработки. Если выигрыши редки, даже точная память может не компенсировать расходов. Отбор в этой учебной модели зависит от того, как мы задали связь успешных действий, затрат и воспроизводства. Поэтому фразы «сложнее значит лучше» или «больше памяти значит ближе к сознанию» не следуют из построенной задачи. Мы исследуем компромисс между функциями, а не движение к заранее назначенной вершине.

Теперь снабдим вторую стратегию переменной, названной «переживание». Если переменная никак не участвует в причинной организации, её добавление

ничего не изменит в модели отбора. Это показывает ограничение конкретной конструкции. Из него не следует, что переживание в природе причинно бессильно. Возможно, оно тождественно некоторому действующему процессу; возможно, предполагаемая связь с процессом требует другого объяснения. Модель обязана явно указать выбранный вариант.



Рисунок 42. Компромисс памяти вместо лестницы сложности. В модели главы память способна сглаживать шум при устойчивой цели и сохранять устаревшую информацию после её изменения. Для вывода о преимуществе нужно независимо учитывать успешность и расходы, а связь с воспроизводством — задавать явно. Это сравнение функций не устанавливает эволюционное происхождение переживания.

Сходство не задаёт одну историю

Если у двух видов обнаружена похожая способность, возможны разные исторические объяснения. Она могла происходить от общего предка, сформироваться независимо или опираться на сочетание общих и новых элементов. Кроме того, одинаковый результат задачи может достигаться разными механизмами. Поэтому сравнение требует совместного анализа поведения, устройства и происхождения, а не подсчёта похожих ответов.

Фельзенштейн показал, почему виды нельзя статистически рассматривать как полностью независимые наблюдения: их признаки связаны общей историей происхождения. Его методологическая работа предлагает учитывать родство при сравнительном анализе. Для исследования сознания этот принцип означает, что десять близкородственных видов с одним признаком не равнозначны десяти независимым случаям его возникновения. [31]

Построим вымышленное дерево с четырьмя конечными ветвями. Две близкие ветви имеют механизм М, две другие его не демонстрируют. Одно объяснение предполагает появление М у общего предка первой пары. Другое — более раннее появление с последующей утратой у второй пары. По распределению одного признака история может оставаться неоднозначной. Дополнительные сведения способны различить версии, но пропуски в наблюдениях нельзя автоматически считать утратой.

Когда изучается переживание, неопределённость увеличивается. Даже наличие М установлено не всегда одинаково надёжно, а связь М с опытом может зависеть от теории. Поэтому историческая реконструкция должна сохранять обе неопределённости. Неверно сначала объявить спорный показатель точным критерием сознания, а затем выдавать уверенность построенного дерева за независимое подтверждение критерия.

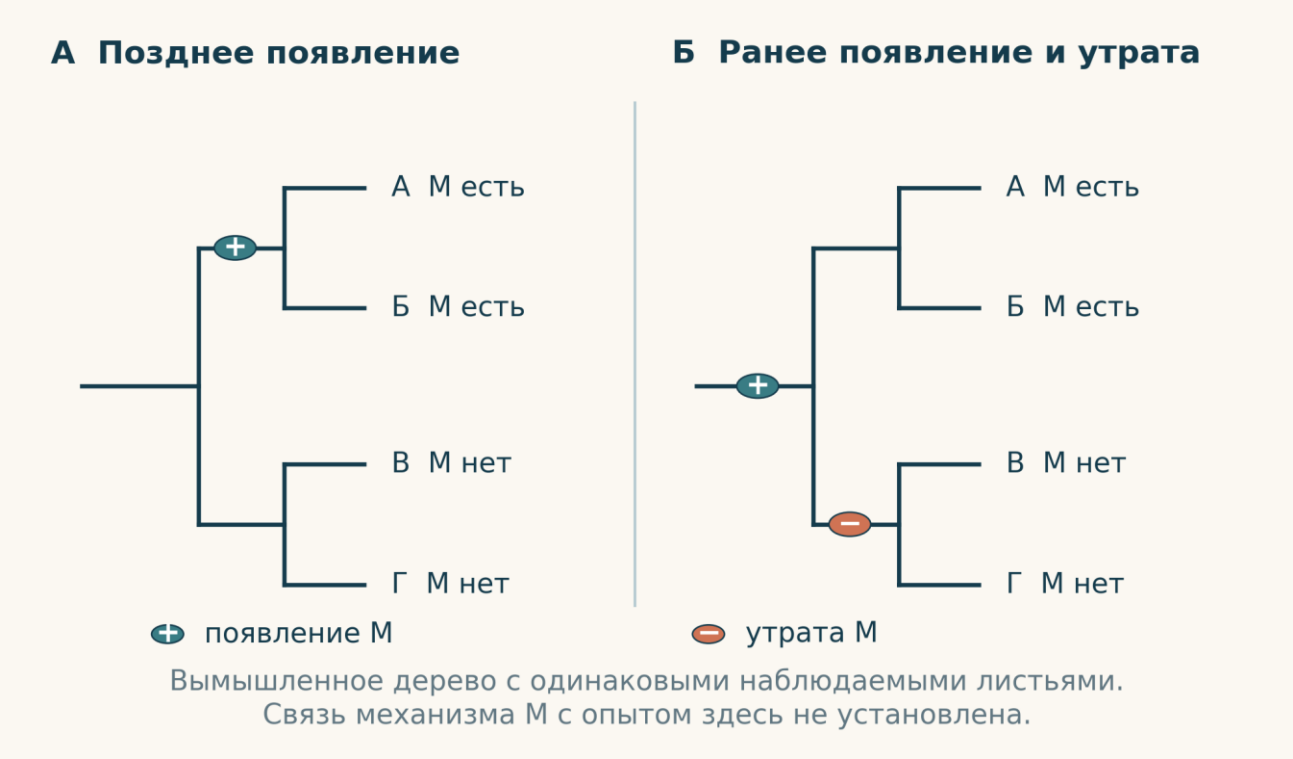

Рисунок 43. Две истории одного распределения признака. В учебном дереве механизм М имеется у видов А и Б и отсутствует у В и Г. Это совместимо как с появлением у общего предка первой пары, так и с более ранним появлением и последующей утратой у второй пары. Дерево не реконструирует происхождение сознания; связь М с опытом остаётся отдельным вопросом.

Сравнение функций и устройства

Баррон и Кляйн предложили аргумент о возможном опыте насекомых, опирающийся на функциональные аналогии между системами управления поведением насекомых и позвоночных. Это первичная теоретическая работа, а

не прямое наблюдение квалиа и не установленная дата их происхождения. Её исследовательская ценность состоит в явной связи между предполагаемой организацией и выводом об опыте; сам переход остаётся предметом проверки и спора. [32]

Собственный сравнительный проект можно начать с гораздо более скромного вопроса: объединяет ли система информацию о пространстве, доступных действиях и внутреннем состоянии при выборе? Для ответа понадобится показать, что изменения этих факторов взаимодействуют, а не вызывают три независимых реакции. Например, доступность ресурса должна менять выбор маршрута в зависимости от цены перемещения, а новая информация — переноситься на ранее не встречавшееся сочетание условий.

Такой результат поддержал бы гипотезу об определённой организации управления. Затем следует искать причинные механизмы и проверять более простые объяснения. Переход к сознанию будет зависеть от того, почему именно эта организация считается существенной. Здесь становится полезным материал глав 13–16: разные теории направляют сравнительное исследование к разным признакам и вмешательствам.

Почему универсальная лестница мешает

Одна шкала заставляет объединять разнородные способности. Вид может хорошо решать пространственную задачу и плохо выполнять задачу, требующую непривычного сенсорного канала. Другая система может точно оценивать неопределённость одного действия, но не переносить оценку на новую область. Эти профили не обязаны упорядочиваться одной линией.

Для анализа полезнее карта функций: доступные виды информации, временной горизонт, гибкость переноса, телесная организация, контроль ошибок. Если интересует феноменальный опыт, рядом помещают отдельный столбец аргументов о нём. Если интересует явная рефлексия, требуется ещё один набор задач. Такое представление препятствует незаметному выводу «лучше в одном тесте — сознательнее во всём».

Не следует также превращать современный вид в живую ступень прошлого. Он имеет собственную историю изменений. Сравнение помогает реконструировать определённые черты предков, но не делает наблюдаемое животное точной копией далёкого предкового состояния. Поэтому собственные сценарии происхождения нужно представлять как модели с явно указанными допущениями.

Полезно проверять сценарий и на круговое объяснение. Представим рассуждение: механизм сохранился, потому что полезен; его полезность

установлена тем, что он сохранился. Здесь одно утверждение поддерживает само себя. Чтобы выйти из круга, необходимо независимо измерить предполагаемую выгоду, цену и условия, в которых преимущество исчезает. Тогда модель может ошибиться содержательно.

В учебном примере памяти таким проверяемым условием будет резкое изменение среды. Если модель предсказывает потерю преимущества, а оно сохраняется при любых параметрах, следует проверить реализацию и определение затрат. Возможно, сравнение устроено так, что одна стратегия всегда получает лишнюю информацию. Подобная техническая ошибка способна породить убедительную, но ложную историю адаптации.

Одинаковые листья дерева допускают две истории

На учебном дереве у видов А и Б найден механизм М, у В и Г нет. Один рассказ предполагает появление М у общего предка А и Б. Другой допускает более раннее появление и последующую утрату в другой ветви. Наблюдаемые концы дерева одинаковы, исторические изменения различаются. Для выбора нужны дополнительные сведения о родственниках, устройстве и вероятных переходах. Если связь М с опытом сама остаётся гипотезой, уверенность в исторической реконструкции не превращает её в независимое доказательство сознания.

Что теперь установлено

История механизма и история переживания не слились в один вывод. Можно объяснять преимущества конкретной памяти и одновременно сохранять открытым вопрос, при какой организации она имеет переживаемую сторону.

Практика 19 Карта эволюционного аргумента

Зачем выполнять: Построить эволюционный аргумент с явными допущениями на каждом переходе.

Что понадобится: Различайте развитие, эволюционное происхождение и наличие переживания. Механизм, наблюдаемый признак и предполагаемый предок являются разными элементами аргумента.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Используйте учебное дерево рисунка 43. Выпишите наблюдения на листьях отдельно от двух вариантов истории.
2. Для каждого варианта назовите предполагаемые появления и утраты механизма. Не выдавайте ни один вариант за прямое наблюдение предка.

3. Опишите среду, в которой память нескольких сигналов может помогать, и среду, где она задерживает переключение. Укажите также стоимость хранения и обработки.
4. Добавьте отдельное предложение о связи механизма с опытом и пометьте его как теоретическую гипотезу. Назовите сведения, которые изменили бы каждый из выводов.

Что сохранить: Карта из наблюдения, исторического объяснения, модели преимущества и отдельного феноменального предположения.

Подсказка: Похожее поведение не гарантирует общего устройства. Общая молекула не гарантирует одинаковой функции всей системы.

Исследовательское продолжение: Выберите опубликованный сравнительный аргумент главы. Проверьте, какие сведения относятся к функции, устройству и родству, и предложите наблюдение, различающее ближайшие исторические версии.

Разбор и ответы — приложение Е, глава 19, с. 302

Вопросы для самопроверки

1. Доказывает ли адаптивная польза механизма наличие переживания?
2. Почему близкородственные виды не дают полностью независимых подтверждений?
3. Зачем сохранять несколько историй происхождения?

Мы разделили динамику эпизода, развитие и эволюцию. Следующая глава выделит оставшийся вопрос: почему определённая физическая организация вообще может иметь переживаемую сторону.

Глава 20

Что требуется для объяснения возникновения

Предыдущие главы объясняли, как процесс появляется во времени и как его возможности меняются. Теперь спросим об основании самого наличия опыта. Успешная модель зрения может предсказывать ответы и ошибки, но её автор должен явно сказать, какую связь между вычислением и переживанием он предлагает. Это требование подготавливает собственную гипотезу книги. Оно позволяет делать сильные предположения, сохраняя их обязательства и возможные возражения.

Уточнение оставшегося вопроса

Предположим, мы построили подробную модель зрения. Она объясняет распознавание, память, уверенность и подготовку ответа. При изменении параметров возникают ошибки, похожие на человеческие. Что ещё требуется, чтобы назвать её объяснением опыта? Ответ зависит от того, какую связь между физическим процессом и переживанием мы принимаем и какими данными эту связь обосновываем.

Дэвид Чалмерс сформулировал различие между объяснением когнитивных функций и вопросом о том, почему выполнение функций сопровождается субъективным опытом. Этот аргумент задаёт философскую проблему, а не экспериментальное доказательство отдельной нефизической сущности. Можно не принимать его дальнейшие выводы и всё же пользоваться требованием точно обозначать переход от механизма к переживанию. [2]

В этой главе нас интересует онтологический смысл возникновения: почему некоторый процесс вообще имеет переживаемую сторону. Сведения о том, когда включилась сеть, когда ребёнок освоил задачу или когда механизм появился в истории вида, не отвечают на этот вопрос автоматически. Однако они ограничивают возможные ответы и помогают превращать общие идеи в исследовательские программы.

Объяснить наличие и объяснить различие

Начнём с различия между двумя задачами. Первая: почему опыт присутствует в одном состоянии системы и отсутствует в другом? Вторая: почему присутствующий опыт имеет именно такое качество? Объяснение, предсказывающее только общую возможность сознания, ещё не объясняет

различие между звуком, цветом и телесным ощущением. И наоборот, модель различения цветов может точно описывать отношения между ответами, оставляя открытым вопрос о переживаемости.

Для собственного примера возьмём пространство из трёх ощущений, обозначенных условными буквами. Участник считает первое ближе ко второму, чем к третьему. Модель воспроизводит эту структуру сходства. Это содержательное достижение: она объясняет определённые отношения в данных. Но совпадение трёх расстояний не устанавливает, что внутренние состояния модели обладают теми же качествами. Необходимо понять, какие свойства пространства измерены и почему выбранное соответствие имеет объяснительную силу.

Следующий шаг — предсказать изменения. Если вмешательство изменяет одну связь в механизме, должно ли измениться сходство двух ощущений, их объединение или доступность различения? Предсказание, сделанное до измерения, сильнее простого подбора красивой картинки после результата. Важна и избирательность: изменение одного отношения должно сохранять другие, если модель утверждает их независимость.

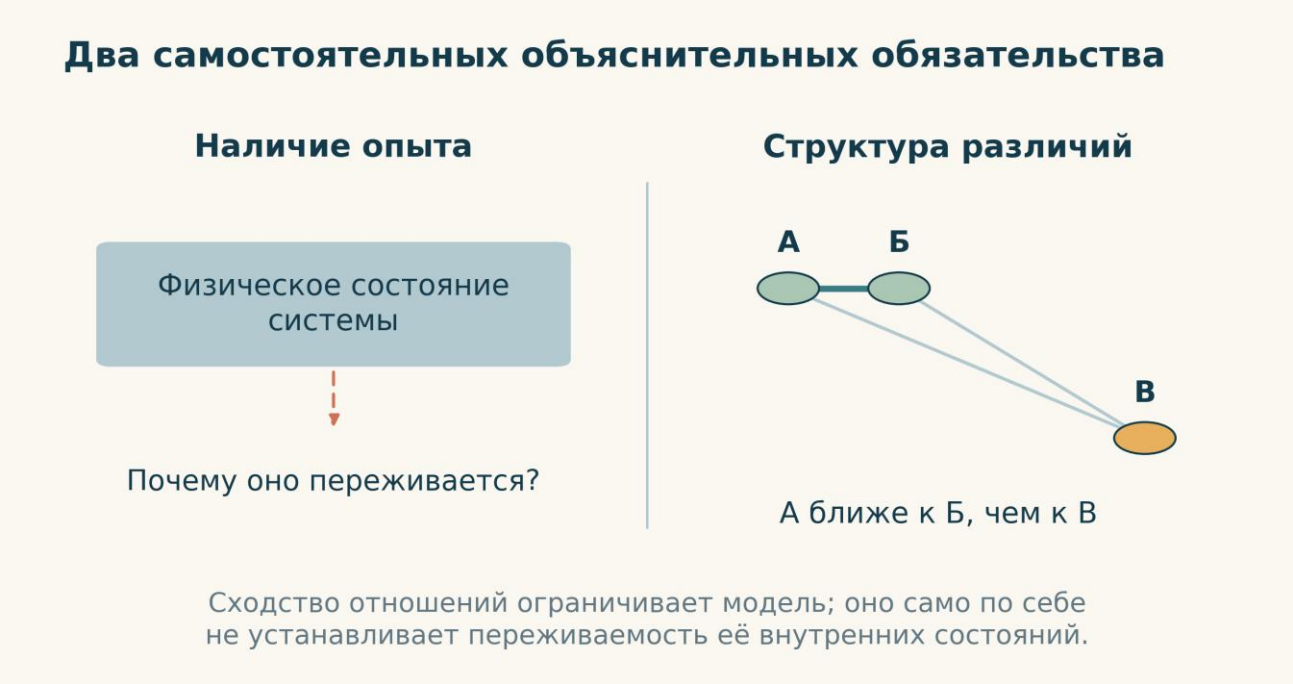

Рисунок 44. Объяснение присутствия и объяснение качества. Слева поставлен вопрос о наличии переживания, справа — о структуре различий между тремя условно обозначенными ощущениями. Геометрия справа показывает только заданное отношение близости, без числовых расстояний и привязки к модальности. Воспроизведение такого отношения является ограничением для модели, но не решает автоматически левую задачу.

Единство и принадлежность

В одном эпизоде могут совместно присутствовать звук, видимое движение и телесное ощущение. Объяснение должно указать, почему они относятся к одному эпизоду, а также почему некоторые процессы системы не включаются в него. Простая сумма записей этого не объясняет: нужно задать отношения, благодаря которым изменения одной части влияют на остальные и на доступное целое.

Принадлежность опыта также требует уточнения. В минимальном смысле речь идёт о перспективе, относительно которой содержание организовано. В более сильном смысле человек распознаёт состояние как своё. Ещё более сильный вариант включает понятие о себе как о сохраняющемся во времени участнике событий. Если теория предполагает, что явное понятие «я» необходимо для всякого ощущения, она берёт на себя дополнительное обязательство, которое нельзя скрыть внутри определения сознания.

Проверим это на мысленном примере. Человек поглощён мелодией и не анализирует себя. Позднее он может описать переживание от первого лица. Было ли развёрнутое самоописание частью первоначального эпизода или появилось при отчёте? Сам рассказ не разрешает вопрос без дополнительных уточнений. Но пример показывает, почему модель восприятия, модель принадлежности и модель рефлексии полезно сначала различить, а затем исследовать их взаимодействие.

Какие отношения между процессом и опытом возможны

Один вариант объяснения предполагает тождество: опыт является определённым физическим процессом, описанным с другой стороны. Тогда требуется обосновать выбор процесса и показать, почему его структура соответствует структуре переживания. Простая смена названия недостаточна. Успех такого объяснения зависел бы от согласованной системы предсказаний, причинных данных и отсутствия произвольных исключений.

Другой вариант вводит дополнительный принцип соответствия между физической организацией и опытом. Тогда нужно явно записать принцип и объяснить, какие случаи он различает. Если он ничего не предсказывает сверх уже известного поведения, возникает вопрос, каким способом его проверять и предпочитать конкурентному принципу. Это не запрет на исследование, а требование показать путь получения новой информации.

Третий вариант пересматривает исходное описание феноменальности: возможно, часть кажущихся свойств опыта связана с тем, как система представляет собственную работу. Тогда недостаточно объявить трудность

иллюзией. Требуется объяснить конкретные суждения о переживании, их устойчивость, ошибки и связь с остальными данными. И здесь сильный ответ должен производить проверяемые различия, а не завершаться переименованием вопроса.

Эти варианты приведены как логические стратегии. Мы не считаем их одинаково подтверждёнными и не выбираем победителя на основании одного мысленного случая. Для исследовательского учебника полезнее научиться определять, какой именно шаг делает каждая модель и какой долг объяснения у неё остаётся.

Как сформулировать собственную гипотезу

Возьмём ограниченную исследовательскую гипотезу: устойчивое объединение зрительного и телесного содержания зависит от того, выводит ли система их из одной изменяемой общей причины. Гипотеза касается определённой организации эпизода. Она предсказывает, что нарушение именно такого объединения изменит совместную принадлежность содержаний при сохранении их раздельного различения. Это наша учебная формулировка, а не заявление об установленном универсальном законе.

Для проверки понадобится определить общую причину в модели, найти независимые признаки раздельного различения и подобрать вмешательство. Затем сравнить объяснение с альтернативой, которая связывает результат только с вниманием или памятью. Если все модели дают одинаковые ожидания, условия следует изменить. Само наличие слов «общая причина» не делает гипотезу глубже уже имеющихся описаний.

Даже положительный результат поддержал бы прежде всего механизм объединения. Чтобы назвать его объяснением наличия опыта, нужно добавить и обосновать соответствующий принцип. Возможная научная работа поэтому состоит из нескольких связанных достижений. Требовать, чтобы каждый небольшой эксперимент сразу решал весь онтологический вопрос, означало бы терять полезные промежуточные ограничения. Выдавать ограничения за окончательное решение было бы другой ошибкой.

Хорошее объяснение должно также обозначить область применимости. Гипотеза о зрительном объединении может ничего не утверждать о сновидении, телесной боли или абстрактном размышлении. Это допустимо, если ограничение записано заранее. Универсальность не достигается простым добавлением слова «всякий» к успешно проверенному частному утверждению.

Расширение области требует новых связующих предположений. Можно предложить, что разные виды содержаний используют один принцип

организации, но реализуют его разными элементами. Тогда требуется определить, что именно остаётся общим и какое изменение в каждой области считается аналогичным. Если соответствие выбирают произвольно после результатов, программа теряет проверяемость. Если оно задано заранее, возникает путь от частного механизма к более общей теории опыта.

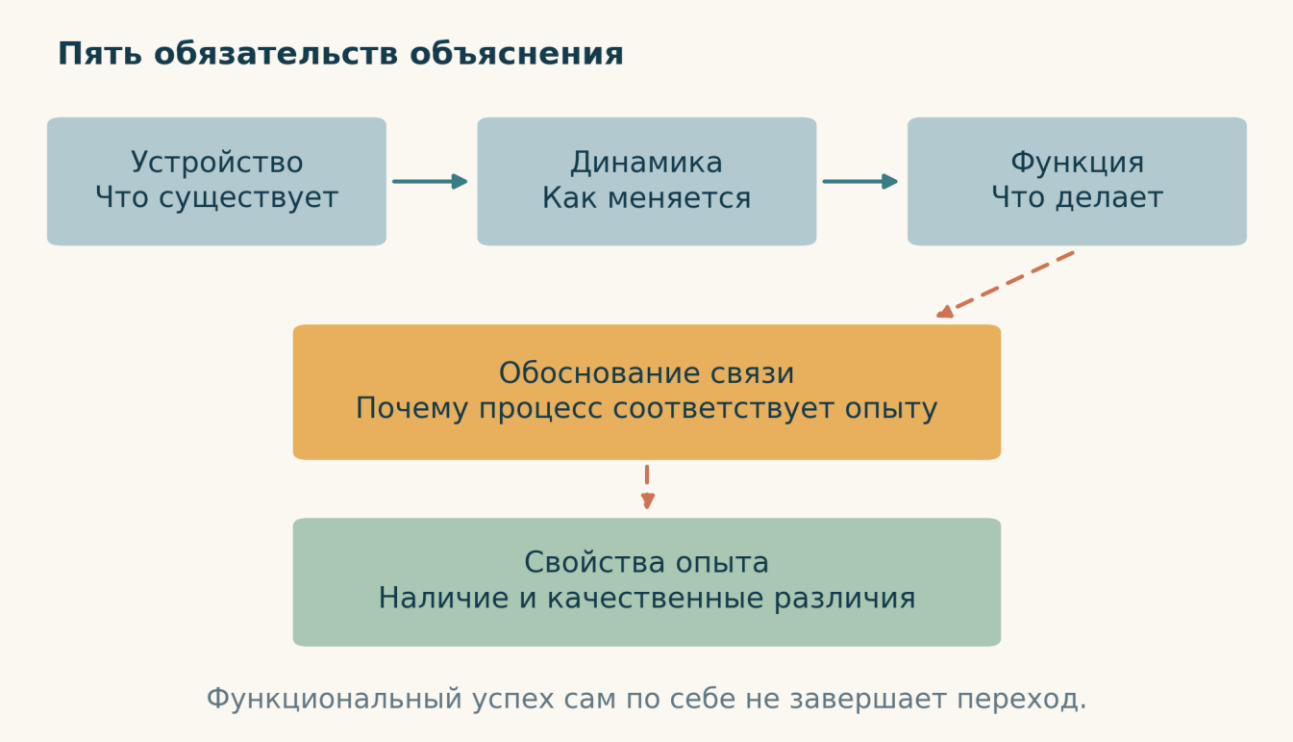

Рисунок 45. От устройства к объяснению опыта. Пять вопросов из аудита объяснения: что существует, как меняется, какую функцию выполняет, какие свойства опыта объясняются и почему процесс связан с ними. Пунктир отмечает переход, требующий собственного обоснования. Объяснение функции не устанавливает автоматически наличие опыта или его качественную структуру.

Пять предложений выполняют разную объяснительную работу

Описание связанных элементов отвечает, что существует в устройстве. Правило их обновления показывает динамику. Успешное различение предметов описывает функцию. Отношения между сообщаемыми качествами задают феноменальное требование. Принцип соответствия объясняет, почему выбранная организация считается достаточной или существенной для опыта. Если оставить только первые три пункта и назвать выход переживанием, последний переход не исчезнет. Если же предложить явное тождество процесса и опыта, оно получает содержательные обязательства: структуру соответствий, границы применения и различающие последствия.

Что теперь установлено

Мы определили, какого объяснения ещё недостаёт после хорошей функциональной модели. Теперь попробуем сформулировать этот остаток точно, не отказываясь от уже полученного результата.

Практика 20 Аудит объяснения

Зачем выполнять: Разобрать одно объяснение по его отдельным обязательствам и найти сильного конкурента.

Что понадобится: Необходимость, достаточность, корреляция и предполагаемое тождество из главы 5. Различайте наличие опыта, его качество, единство и принадлежность.

Ориентир по времени: 30–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Возьмите конкретное объяснение из глав 13–15. Запишите по одному предложению об устройстве, динамике, функции, объясняемом свойстве опыта и связи между ними.
2. Рядом укажите основание каждого предложения: определение, данные, вычисление или теоретическое предположение.
3. Выберите конкурента, сохраняющего часть тех же функциональных результатов. Укажите, какое дополнительное требование различает объяснения.
4. Сформулируйте возможное различающее наблюдение. Если оно пока не определено, точно опишите, какой метод или вывод ещё требуется.

Что сохранить: Пять связанных утверждений, конкурирующее объяснение и точный оставшийся вопрос.

Подсказка: Объяснение отношений между качествами является содержательным результатом, даже если ещё не установлено наличие переживания у новой реализации.

Исследовательское продолжение: Предложите собственный ограниченный принцип соответствия. Проверьте его на простом регуляторе и на системе с одинаковыми ответами, получаемыми другим устройством.

Разбор и ответы — приложение Е, глава 20, с. 303

Вопросы для самопроверки

1. Почему объяснение различения цветов может не решать вопрос о наличии опыта?
2. Обязательно ли вводить дополнительную сущность для объяснения переживания?
3. Что считается прогрессом, если окончательного объяснения ещё нет?

Прежде чем выбирать устройство будущего кандидата, рассмотрим данные о разных носителях опыта. Это покажет, почему критерии нужно применять к конкретным механизмам и условиям.

Часть 6

Возможные носители опыта

Глава 21

Сознание и самоконтроль у животных

Теоретический принцип должен сохранять смысл при переходе от привычного человека к другому организму. При этом буквальное повторение человеческого задания может требовать неподходящей речи, движения или сенсорного канала. Рассмотрим исследования животных как отдельные аргументы о памяти, содержании и состояниях после повреждения. Их сила состоит в сочетании подходящих измерений и контроля альтернатив, а не во внешнем сходстве ответа с человеческой фразой.

Какие вопросы мы задаём животному

Вопрос о возможном опыте животного требует различать несколько целей. Феноменальное сознание означает наличие переживания. Самоотнесение касается связи информации с собственным телом, действием или состоянием. Метакогнитивный контроль использует сведения о надёжности познавательных процессов. Явная рефлексия делает собственные основания предметом целенаправленного рассмотрения. Эти возможности нельзя считать взаимозаменяемыми, даже если в некоторых задачах они взаимодействуют.

Животное, избегающее неопределённого выбора, может оценивать качество памяти, использовать усвоенный внешний признак или следовать более простому правилу подкрепления. Животное, защищающее повреждённую часть тела, демонстрирует другой набор функций. Ни одно наблюдение не следует автоматически переводить в утверждение о человеческом способе мыслить о себе. Нужен профиль конкретных оснований и открытых альтернатив.



Рисунок 46. Разные организмы и разные исследовательские вопросы. Предметная иллюстрация: птица, осьминог, рыба и пчела. Масштаб условный. Сопоставлять возможных носителей опыта следует по конкретным признакам и исследованиям; положение и размер животного на этой странице не задают шкалу сознания.

Общая оценка и конкретный эксперимент

Нью-Йоркская декларация 2024 года сообщает о сильных научных основаниях приписывать опыт другим млекопитающим и птицам и о реалистичной возможности опыта у всех позвоночных и многих беспозвоночных. Это заявление подписантов, а не отдельный эксперимент и не одинаковая доказательная база для каждого вида. Его полезно использовать как ориентир исследовательской неопределённости, сохраняя различия между конкретными данными. [33]

Общая оценка способна опираться на множество линий исследований, но отдельный учебный разбор должен оставаться проверяемым. Поэтому ниже приведены несколько первичных работ с разными предметами: память, зрительное содержание и состояния, связанные с повреждением. Они не образуют соревнование животных за единственный балл сознательности. Каждая работа ограничивает свой круг объяснений.

Память и зрительное содержание

В исследовании Хэмптона две макаки могли принять или отклонить тест памяти. Выбор зависел от условий запоминания и был связан с последующей

точностью; результаты животных различались по полноте выполнения критериев. Работа дала основания обсуждать контроль доступности собственной памяти. Она не показала словесной рефлексии и не исключила разом все возможные ассоциативные объяснения любых подобных задач. [34]

В исследовании Нидера, Вагенер и Риннерта активность нейронов переднего мозга ворон была связана с поведенческим сообщением о наличии или отсутствии слабого зрительного стимула. Авторы интерпретировали результат как нейронный коррелят сенсорного сознания. Исследование расширяет сравнительную базу, но не определяет все свойства птичьего опыта и не устанавливает наличие человеческого биографического самосознания. [35]

Две работы отвечают на разные вопросы. Первая помогает исследовать управление решением по доступному знанию. Вторая связывает нейронную активность с различением сообщаемого зрительного содержания. Совместное рассмотрение полезно именно благодаря различию. Если назвать оба результата просто «самосознанием», исчезнет возможность понять, какой механизм объясняет каждый эффект.

Повреждение и отрицательное переживание

В работе Крук с осьминогами изучались избегание места, связанного с повреждающим воздействием, и предпочтение места, связанного с облегчением после него. Поведенческие и нейрофизиологические результаты автор интерпретировала как основания предполагать аффективный компонент боли. Это аргумент о переживаемом отрицательном состоянии в изученных условиях, не прямое измерение чужих квалиа и не доказательство одинакового опыта всех головоногих. [36]

Барр и Элвуд в 2024 году исследовали, как береговые крабы соотносят избегание неприятного воздействия с избеганием яркого света. Изменение выбора в зависимости от конкурирующих условий соответствовало предсказаниям гибкого оценивания. Авторы рассматривали результат как совместимый с болью. Более узкий непосредственно показанный вывод касается компромисса между разными основаниями выбора. [37]

В опубликованном исследовании Манзи и коллег 2026 года домовые сверчки дольше ухаживали за усиком после повреждающего теплового воздействия, чем после контрольных воздействий. Изменение условий, задуманных как менее и более стрессовые, значимого эффекта не дало. Работа усиливает основания обсуждать направленную самозащиту и возможные состояния, подобные боли; она не демонстрирует метакогнитивного анализа или явной рефлексии. [38]

Здесь важно различать регистрацию повреждения, организованное избегание и отрицательное переживание. Простой защитный ответ даёт меньше информации, чем устойчивое изменение предпочтений с контролем альтернатив. Но даже сложность поведения не является самостоятельной единицей измерения боли. Объяснение должно показать, почему совокупность наблюдений лучше согласуется с предполагаемым состоянием, чем с ближайшими конкурирующими механизмами.

Сравнительная карта оснований

Таблица объединяет рассмотренные типы данных. Она указывает направление проверки и границу переноса, а не присваивает видам численные вероятности сознания.

Группа	Основание в данном разборе	Существенный нерешённый вопрос
Млекопитающие	Общая оценка сильных оснований и опыт с памятью макак	Какие формы контроля переносятся между задачами
Птицы	Общая оценка и нейронное исследование ворон	Насколько результат обобщается на другие содержания и виды
Другие позвоночные	В декларации признаётся реалистичная возможность опыта	Какие конкретные механизмы и показатели сопоставимы
Головоногие	Сочетание предпочтений и нейрофизиологии у осьминогов	Как устроены качество и организация предполагаемого опыта
Десятиногие ракообразные	Компромиссы выбора у береговых крабов	Какие альтернативы оцениванию переживаемого состояния исключены
Насекомые	Направленная самозащита у домовых сверчков	Как связаны защитный контроль и возможная феноменальность

В строке других позвоночных отдельный эксперимент здесь не разбирается. Это граница выбранного материала, которую нельзя переводить в вывод об отсутствии опыта. Аналогично данные одного вида не следует распространять на всю группу без дополнительного сравнительного обоснования. Объём исследований и свойства исследуемых организмов — разные вещи.



Рисунок 47. Матрица вопросов и пробелов свидетельств. Матрица относится только к исследованиям, разобранным в главе: памяти макак, зрительного сообщения ворон, избегания и облегчения у осьминогов, компромиссов выбора крабов и самозащиты сверчков. Точка обозначает предмет работы, а не доказанность переживания. Прочерк означает отсутствие соответствующего разбора, а не отсутствие способности у животного. Данные одного вида не представляют всю группу.

Как усиливать вывод

Сильный профиль строится из сходящихся данных с различными ограничениями. Если наблюдаются перенос обучения, чувствительность к изменению ценности результата и соответствующие причинные изменения, простому фиксированному правилу становится труднее объяснить всё сразу. Однако альтернативу нужно формулировать конкретно. Фраза «это просто инстинкт» не является механизмом, пока не указаны входы, внутренние состояния и правила изменения поведения.

Точно так же выражение «животное понимает» требует уточнения. Что именно представлено? Как знание проявляется в новой ситуации? Что изменится, если убрать внешнюю подсказку? Может ли предполагаемая оценка собственной ошибки влиять на действие до получения результата? Ответы делают спор о богатстве объяснения продуктивным.

Для явной рефлексии планка утверждения выше, чем для контроля неопределённости. Нужно показать рассмотрение собственных оснований и их пересмотр, исключая доступные более узкие правила. Отсутствие такого доказательства не опровергает возможного феноменального опыта. Оно ограничивает конкретное приписывание рефлексивной способности.

При сравнении следует проверить, одинаково ли понятна сама задача. Если одному виду предлагается привычный способ выбора, а другому — трудно выполнимое движение, разница точности будет смешивать интересующую способность с устройством теста. Нужен контроль доступности сигнала, мотивации и возможности выполнить ответ. Иногда более справедливое сравнение использует разные внешние действия, но одинаковое отношение между информацией и выбором.

Представим две учебные процедуры: нажатие кнопки и выбор прохода. В обеих можно оценить, насколько решение зависит от качества доступных сведений. Сопоставлять следует эту зависимость, а не внешнее сходство движения. Такой перенос тоже требует обоснования, однако он показывает, почему буквально одинаковый тест не всегда является одинаковым испытанием механизма.

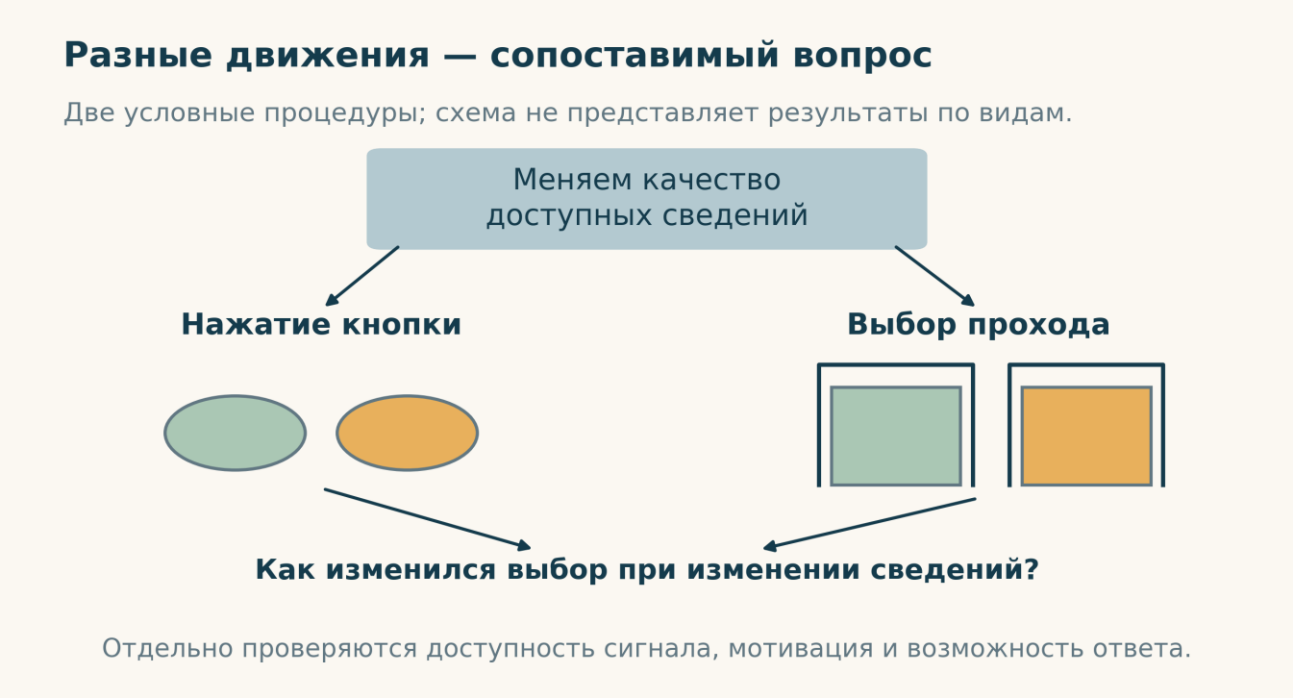

Рисунок 48. Сопоставимые задачи с разными внешними ответами. Нажатие кнопки и выбор прохода могут проверять сходную зависимость решения от доступной информации. При сравнении нужно обосновать сопоставимость этой зависимости и отдельно проверить сенсорные, мотивационные и двигательные требования. Совпадение внешнего результата не делает внутренние механизмы одинаковыми и не даёт общей шкалы сознательности.

Что добавляет отказ от трудного теста

Животное чаще отклоняет тест после условий плохого запоминания, а принятые тесты выполняет точнее. Это совместимо с контролем доступности памяти. Конкурент может предполагать выбор по внешнему признаку, например по задержке. Чтобы различить версии, нужно менять качество запоминания при сопоставимой задержке или вводить другое подходящее условие. Поэтому важны конкретные критерии и данные каждого животного. Общее слово «метакогниция» не заменяет этот разбор и не приписывает животному словесную взрослую рефлексию.

Что теперь установлено

Отдельные работы отвечают на разные вопросы. Профиль оснований позволяет объединять их, сохраняя различие между исследованным свойством и неизвестным.

Практика 21 Профиль одного изученного вида

Зачем выполнять: Составить профиль одного исследования без переноса его вывода на всю группу организмов.

Что понадобится: Паспорт исследования, различие функционального и феноменального вывода. Декларация научной позиции и конкретный эксперимент имеют разный статус.

Ориентир по времени: 30–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Выберите работу главы о памяти, зрительном содержании или состоянии после повреждения. На первом проходе заполните поля по тексту книги.
2. Укажите вид, задачу, измерение, контроль и основной результат. Если число животных не дано в доступном описании, пометьте пропуск.
3. Предложите конкретный альтернативный механизм и условие, в котором он ожидает другой результат.
4. Сформулируйте один сильный вывод и один более широкий вывод, который эта работа пока не разрешает.

Что сохранить: Паспорт одного исследования и проверяемая альтернатива без общего балла сознательности.

Подсказка: Отсутствие исследований выбранного свойства в данном обзоре не является наблюдением его отсутствия у организма.

Исследовательское продолжение: Откройте первичную статью и уточните выборку, правила исключения и индивидуальные результаты. Проверьте, действительно ли предусмотрен предложенный вами контроль.

Вопросы для самопроверки

1. Почему признание возможной боли не означает признания явной рефлексии?
2. Можно ли распространять результат одного вида на всех его родственников?
3. Что делать с объяснением через простое правило?

Теперь проверим, какие функции доступны системам без нейронов. Это поможет отличить память и адаптацию в широком смысле от дополнительной гипотезы о переживаемости.

Глава 22

Границы между адаптацией и переживанием

Сложная координация и влияние прошлого встречаются за пределами нервных систем. Чтобы понять значение таких данных, начнём с измеряемой реакции и нескольких возможных механизмов. Ослабление ответа может возникнуть из специфического привыкания или из общего истощения. Сравнение с новым сигналом и паузой позволит развести эти версии. Вопрос об опыте потребует следующего, отдельно обоснованного шага.

Почему системы без нейронов важны

Если мы хотим объяснить условия опыта, нельзя заранее ограничивать поиск существами, похожими на человека. Одновременно нельзя считать любую полезную реакцию признаком квалиа. Исследование растений, одноклеточных и других систем без нейронов полезно именно потому, что заставляет разделять обработку сигналов, сохранение истории, обучение и предполагаемую переживаемость.

Начальный вопрос должен касаться измеряемого явления. Изменилась ли реакция после повторения воздействия? Сохранилось ли изменение? Зависит ли оно от конкретного сигнала? Может ли организм использовать связь между событиями? Лишь после ответов возникает следующий вопрос: какие основания связывают найденный механизм с опытом? Такой порядок оставляет исследование открытым и одновременно защищает от чрезмерного вывода.

Сигнализация и координация

Тойота и коллеги показали у растения резуховидки связь повреждения с глутаматным сигналом, распространением изменений концентрации кальция и активацией защитных реакций в удалённых частях растения. Это первичное исследование механизма дальней связи внутри организма. Сходство отдельных молекул и способов передачи сигнала с известными нейронными процессами не устанавливает тождество общей организации или наличие боли как переживания. [39]

На собственном примере представим систему нескольких резервуаров. Изменение в одном вызывает сигнал, который закрывает клапаны в других. Здесь есть локальное обнаружение, передача информации и согласованное

защитное действие. Если сигнал зависит от прежних воздействий, появится ещё и влияние истории. Все эти свойства можно исследовать без приписывания системе представлений о себе или отрицательного опыта.

Аналогия не объясняет растение целиком. Она показывает логическую границу: наличие передачи сигнала ещё не определяет, какие дополнительные свойства реализованы. Для более сильного утверждения потребуются данные о взаимодействии процессов, гибкости выбора, внутреннем состоянии и теоретическом основании связи с переживанием. Слово «защита» описывает функцию и не обязано означать осознанное намерение защищаться.

Память как влияние прошлого

В минимальном функциональном смысле память существует там, где история меняет текущую реакцию при одинаковом нынешнем входе. Такая память может быть распределена по состоянию системы и не содержать отдельной записи о событии. Нужно различать этот широкий смысл и воспоминание как переживаемое возвращение к прошлому. Совпадение слова не делает функции одинаковыми.

Построим учебную модель. В ней есть переменная привыкания, первоначально равная нулю. После каждого предъявления сигнала она приближается к единице на пятую часть оставшегося расстояния. Реакция равна разности между единицей и текущим привыканием. Первые реакции составят единицу, восемь десятых, шестьдесят четыре сотых и далее. Во время паузы переменная каждый шаг уменьшается на пятую часть, поэтому реакция восстанавливается.

Все числа заданы нами ради простоты. Модель демонстрирует, как сохранение состояния создаёт уменьшение ответа и восстановление после перерыва. Здесь нет нейронов, словесных убеждений и наблюдателя внутри модели. Но по одной кривой реакции нельзя узнать, совпадает ли такое правило с реальным биологическим механизмом. Для выбора нужны дополнительные вмешательства и измерения.

Добавим два сигнала и две отдельные переменные привыкания. Тогда повторение первого сигнала уменьшает реакцию только на него, а второй остаётся эффективным. Если вместо этого использовать одну общую переменную ресурса, снизятся оба ответа. Таким образом, проверка нового сигнала помогает различать специфическое изменение и общее истощение. Она всё ещё не является тестом наличия феноменального опыта.

Что показано у слизевика

Буассо, Фогель и Дюссутур исследовали слизевик Physarum polycephalum, организм без нейронов. При повторных воздействиях химических веществ реакция ослабевала, а после перерыва восстанавливалась; анализировалась и специфичность изменения. Авторы интерпретировали результаты как привыкание, простую форму обучения. Исследование показывает возможность такой пластичности без нервной системы. Оно не измеряет квалиа, самосознание или явную рефлексию. [40]

Здесь полезно отделить признание результата от расширения его смысла. Если критерии привыкания выполнены, нет необходимости отрицать обучение только потому, что система непривычна. Но признание обучения не обязывает приписывать переживание. Следующий научный шаг — выяснить механизм хранения истории, правила переноса и ограничения пластичности, а затем сравнить их с требованиями конкретной теории сознания.

Ещё один вопрос касается единицы анализа. У распределённой системы можно измерять отдельный участок, весь организм или взаимодействие с изменённой им средой. Если следы прошлой активности сохраняются снаружи и влияют на дальнейшее движение, часть функциональной памяти может находиться во взаимодействии со средой. Такое описание требует явно указать границу системы, иначе внешний след легко ошибочно представить как внутреннее воспоминание.

Почему спор об обучении требует воспроизведения

Маркел в 2020 году не воспроизвёл ранее заявленный эффект ассоциативного обучения у гороха. Работа поставила под вопрос устойчивость конкретного результата. Отрицательная попытка воспроизведения не доказывает невозможность всех форм обучения у растений; положительное первоначальное сообщение также не решает вопрос о квалиа. Для оценки нужны сопоставимость условий, контроль альтернатив и новые независимые проверки. [41]

В задаче на ассоциацию важно показать, что результат зависит именно от связи двух событий. Если группа получила больше света, тепла или времени для роста, изменение может иметь другое объяснение. Контроль должен по возможности сохранять количество воздействий, меняя их отношение. Иначе исследователь сравнивает разные физические условия и приписывает весь результат обучению связи.

Даже успешная проверка ассоциации отвечает на ограниченный вопрос. В нашей собственной программе можно задать вес между двумя сигналами,

увеличивающийся при совместном появлении. Программа начнёт прогнозировать второй по первому. Это ясный функциональный результат. Для перехода от прогнозирования к переживаемому ожиданию потребуется дополнительное обоснование, применимое не только к понравившемуся примеру.

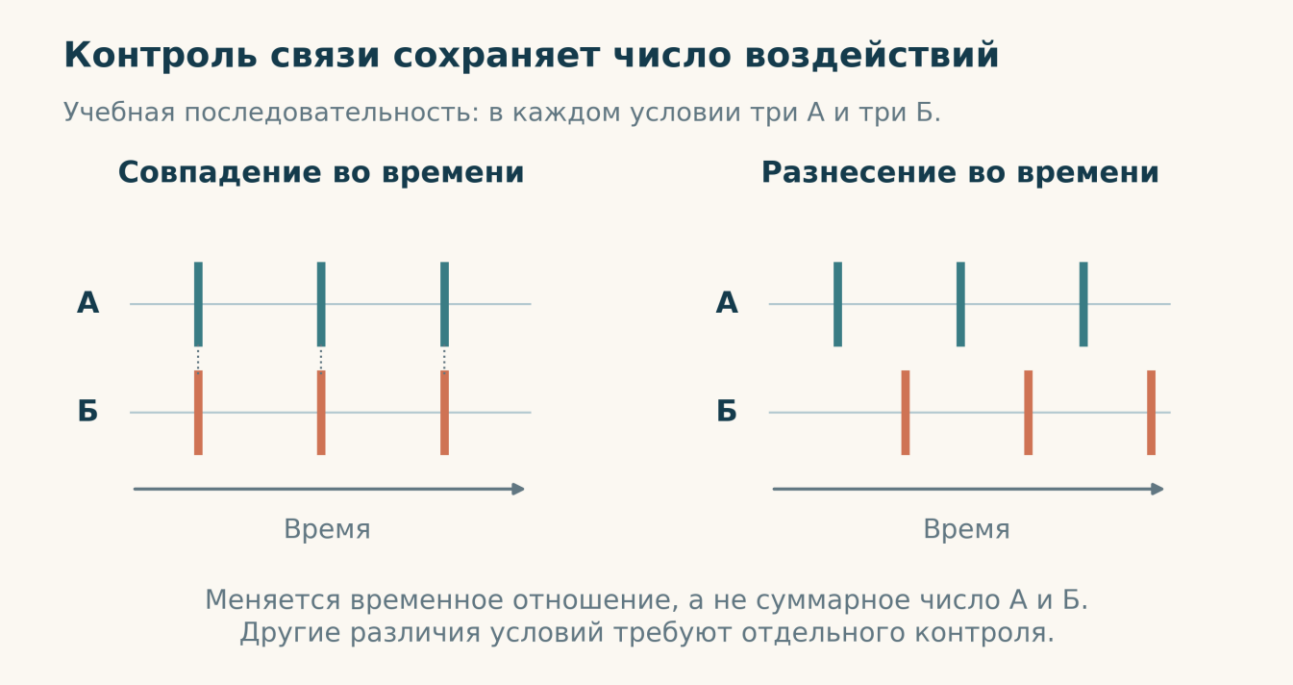

Рисунок 49. Разделение числа событий и отношения между ними. В двух вымышленных последовательностях одинаковое число сигналов, но различно их временное отношение. Такой контроль проверяет роль совпадения, сохраняя количество воздействий; сдвинутые события всё ещё могут образовывать другую предсказуемую связь. Для вывода об обучении нужно проверить именно заявленное правило, а не приписывать весь эффект различия условий ассоциации.

Что должна добавить гипотеза об опыте

Предположим, исследователь утверждает, что достаточно любой памяти и любой обратной связи. Нужно применить такое правило к контрольным устройствам, где оба свойства реализованы простым способом. Если правило приписывает опыт и им, автор должен признать это следствие или уточнить исходное условие. Нельзя сохранить широкую формулировку для привлекательного биологического случая и незаметно сузить её для неудобного инженерного примера.

Другая гипотеза может требовать особого объединения сигналов и внутренней перспективы. Тогда следует показать, как эти свойства реализуются без нейронов и чем отличаются от распределённой регуляции. Выражение «организм действует в собственных интересах» само по себе не задаёт такой

механизм. Нужно указать переменные, зависимости и условия нарушения функции.

Наконец, полезно различать положительную гипотезу и отсутствие запрета. Из того, что теория не исключает опыт растения, не следует, что она его предсказывает. Для положительного вывода должны быть установлены соответствующие условия. Эта разница помогает сохранять исследовательскую открытость без необоснованного расширения круга подтверждённых носителей опыта.

Две модели с одинаковым началом

В модели привыкания для каждого сигнала есть своё h, первоначально 0. Ответ равен 1 минус h, затем h увеличивается на 0,2 оставшегося расстояния до 1. В модели истощения есть общий ресурс r, первоначально 1. Ответ равен r, затем ресурс умножается на 0,8. На первых пяти предъявлениях А ответы обеих моделей равны 1; 0,8; 0,64; 0,512; 0,4096. На новом Б специфическая модель отвечает 1, а общий ресурс уже равен 0,32768.

Во время каждого шага паузы все h умножаются на 0,8, а r увеличивается на 0,2 от недостающего до 1. Сначала записываем ответ, потом обновляем состояние. Это полностью заданные учебные правила, а не измеренные параметры растения или слизевика.

Что теперь установлено

Одной нисходящей кривой недостаточно, чтобы выбрать объяснение. Новый сигнал проверяет специфичность, пауза проверяет восстановление. Оба результата нужно вычислять по одному и тому же заранее заданному порядку.



Рисунок 50. Привыкание и истощение различаются переносом. Расчёт по явным правилам главы: для предъявленного стимула h ← h + 0,2(1 − h); общий ресурс после каждого события R ← 0,8R; при паузе h ← 0,8h и R ← R + 0,2(1 − R). Ответ регистрируется до изменения состояния. После пяти А первый Б даёт 1 у раздельного механизма и 0,32768 у ресурсного. Это условные механизмы, а не опыт на организме.

Практика 22 Две модели одного уменьшения реакции

Зачем выполнять: Развести специфическое привыкание и общий ресурс на одинаковой последовательности воздействий.

Что понадобится: Правило изменения состояния из главы 9 и логика контроля из главы 7. До практики ниже заданы оба алгоритма в одинаковом порядке.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Начните с h для А и Б, равных нулю, и общего ресурса 1. Подайте А пять раз, после каждого ответа обновляя состояние.
2. Предскажите ответ на Б до его расчёта. Затем предъявите Б один раз.
3. Выполните пять шагов паузы по правилам выше, затем снова предъявите А. Сравните две модели.
4. Сверьте вычисления через python teaching_lab.py --chapter 22. Отдельно запишите, что результат говорит о механизмах и чего он не устанавливает об опыте.

Что сохранить: Пошаговая таблица двух моделей, прогноз нового сигнала и вывод о специфичности.

Подсказка: При новом Б состояние h для А не становится состоянием Б. Во время паузы сигналы не предъявляются, поэтому ответ для этих шагов не выдумывается.

Исследовательское продолжение: Сопоставьте контроль с одной первичной работой главы. Укажите, какие наблюдения учебная модель учитывает, а какие вообще не представляет.

Разбор и ответы — приложение Е, глава 22, с. 304

Вопросы для самопроверки

1. Может ли функциональная память существовать без нервной системы?
2. Почему уменьшение реакции ещё не доказывает обучение?
3. Делает ли отсутствие нейронов вопрос о квалиа бессмысленным?

Функциональная память оказалась шире нервной ткани. Следующая глава рассмотрит искусственные и биологические реализации, для которых важно установить устройство, границы и источники самоописания.

Глава 23

Искусственные и биологические кандидаты

Мы рассмотрели разные организмы и увидели, почему материал сам по себе не определяет ни механизм, ни опыт. Теперь перейдём к искусственным и гибридным кандидатам. Прежде чем оценивать их самоотчёты, нужно описать доступные входы, сохраняемое состояние и использование информации. Базовая практика этой главы использует готовую минимальную схему. Полный агент будет введён позже, в главе 28; обращаться к нему заранее не требуется.

Вопрос о носителе

Может ли опыт возникнуть в программе, роботе или выращенной нейронной системе? Ответ нельзя получить из названия материала. Программа описывает способ организации вычислений, робот — соединение вычисления с физическим действием, а нейронная культура — биологический состав определённой системы. Ни одна из этих характеристик сама по себе не задаёт полный механизм сознания.

Для исследования нужно различить носитель, организацию, историю взаимодействия и доступные показатели. Две системы из одинаковых клеток могут иметь разные связи. Одна программа может работать на разном оборудовании. Робот может управляться локальным контроллером или получать команды извне. Если не уточнить границу рассматриваемого устройства, свойства всей установки легко ошибочно приписать одной её части.

Что можно вывести из теории

Батлин и коллеги предложили оценивать искусственные системы через признаки, выведенные из научных теорий сознания, с последующей проверкой их реализации. Статья опубликована онлайн в ноябре 2025 года и включена в журнальный выпуск 2026 года. Метод не представляет собой готовый универсальный тест: вывод зависит от обоснованности выбранных теорий, корректного перевода их требований и фактического устройства системы. [42]

Проверка признака должна отвечать на три вопроса. Является ли он необходимым, достаточным или только сопутствующим согласно данной версии теории? Какая часть устройства его реализует? Каким вмешательством

можно показать, что он действительно выполняет предполагаемую роль? Если признак существует лишь в описании разработчика, это основание изучить архитектуру, а не засчитать условие выполненным.

В главах 13–15 мы рассмотрели разные направления требований. Для рабочего пространства важна доступность выбранного содержания нескольким процессам. Для рекуррентного подхода существенна определённая организация обратных взаимодействий. Для теорий высшего порядка — представление собственного состояния. Для подходов, чувствительных к внутренней физической причинности, функциональное сходство может оказаться недостаточным. Эти различия нужно сохранить при оценке кандидата, а не объединять в безразличную сумму «признаков разумности». [16] [17] [19] [20]

Программа и её самоописание

Вообразим программу, которая после любого сбоя выводит фразу «я растеряна». Другой вариант той же программы выводит «код ошибки 7». Если замена текста не меняет вычисление, сообщение само по себе не доказывает наличия переживаемой растерянности. Оно является частью наблюдаемого поведения, происхождение которого нужно исследовать. И наоборот, отсутствие эмоционального словаря не устанавливает отсутствие всех потенциально значимых механизмов.



Рисунок 51. Раздельный аудит информации и самоописания. Дизайн независимо меняет доступ программы к внутренним сведениям и формулировку сообщения. Если контроль зависит от первого фактора, а слова — от второго, их причинные роли различаются; это предварительное условное предсказание, а не результат уже выполненного опыта. Необходимо исключить доступ к скрытому правильному ответу из среды.

Более содержательный кандидат способен прогнозировать собственную ошибку до получения результата, учитывать пределы памяти и изменять стратегию при обнаружении систематического сбоя. Это уже можно проверять функционально. Однако рассказ о причинах решения может формироваться отдельно от самого решения. Тогда правдоподобное объяснение не гарантирует доступа к фактически действовавшим основаниям.

В собственном эксперименте полезно раздельно изменять доступ к внутренней информации и шаблон ответа. Если качество контроля зависит от внутреннего состояния, а текстовые объяснения — только от шаблона, функции различаются. Если программа использует скрытый правильный ответ из среды, успешная уверенность вообще может быть утечкой данных. Поэтому аудит входов столь же важен, как анализ полученных фраз.

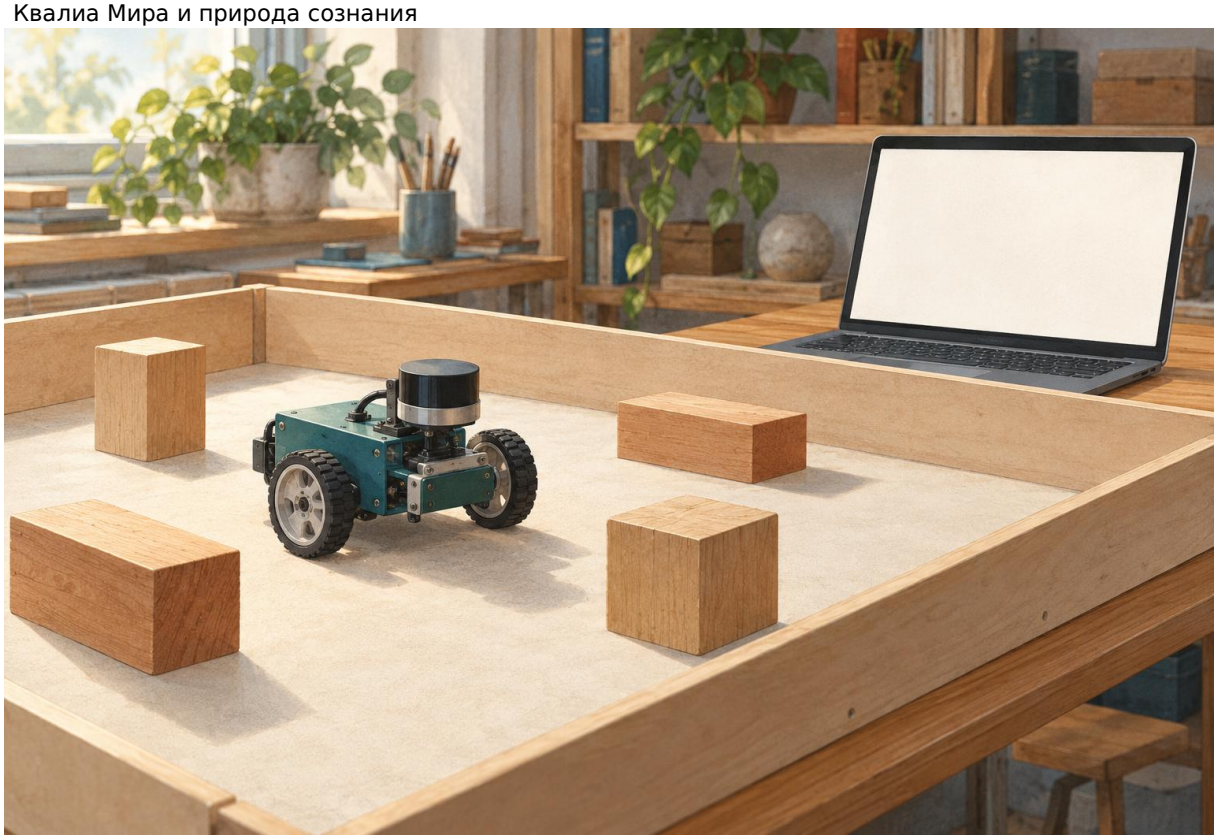

Рисунок 52. Искусственная система в физической среде. Предметная иллюстрация мобильного устройства с датчиками и вычислительным оборудованием. При оценке такого кандидата необходимо установить границы системы: где формируется команда, какая информация доступна контроллеру и какие последствия действия он получает.

Что добавляет тело

Робот позволяет замкнуть взаимодействие: действие изменяет мир, новое наблюдение зависит от действия, а модель может уточнять связь между командой и эффектом. Такое устройство удобно для исследования самоотнесения и адаптации. Оно не делает телесность достаточным условием сознания. Нужно установить, какие именно функции физического тела нельзя без потери заменить иной организацией взаимодействия.

Рассмотрим учебный робот с камерой и подвижной платформой. Поворот меняет изображение. Если устройство предсказывает изменение, оно может различать движение объекта и следствие собственного поворота. При незаметной смене управления потребуется обновить модель действия. В этом примере тело задаёт конкретные причинные зависимости, а не символическое право употреблять слово «я».

Если теория утверждает необходимость определённой телесной регуляции, её проверка требует указать, какая переменная регулируется и почему она существенна. Запас батареи, отображённый числом, не становится автоматически переживаемой потребностью. Нужно исследовать его

интеграцию с выбором, обучением, памятью и предложенным принципом связи с опытом. Без этого мы лишь переименуем обычный управляющий сигнал.

Нейронные культуры, органоиды и гибриды

Нейронная культура представляет биологические нервные клетки в искусственно организованных условиях. Органоид — выращенная трёхмерная система, моделирующая определённые стороны развития и организации ткани; это название не означает готовый мозг. Гибрид соединяет биологический компонент с вычислительным управлением, датчиками или исполнительными устройствами. Подробные экспериментальные примеры и ограничения будут разобраны в главе 31.

Для нашей задачи важен одинаковый принцип анализа: обучение, электрическая активность и успешное управление следует измерять отдельно от предполагаемого опыта. Биологическое происхождение делает некоторые механистические гипотезы более предметными, но не сообщает автоматически об их реализации. Необходимо выяснить устройство связей, границы обратной связи, роль внешнего алгоритма и временную устойчивость взаимодействий.

Особенно показателен гибрид, в котором внешний контроллер выбирает вознаграждение и преобразует ответы ткани в команды. Изменение результата всей установки может происходить из-за настройки контроллера. Чтобы приписать обучение биологической части, нужны соответствующие контрольные сравнения. Чтобы приписать опыт, потребуется ещё и отдельная теоретическая аргументация. Это два последовательных вопроса, которые не следует объединять одним названием проекта.

Класс системы	Что нужно описать прежде всего	Частая ошибка при выводе
Программа	Реальные состояния и зависимости вычислений	Считать самоописание прямым измерением опыта
Робот	Связь восприятия действия и состояния тела	Принимать наличие корпуса за достаточное условие
Нейронная культура	Организацию клеточной сети и внешнего управления	Приписывать сети свойства всей установки
Органоид	Степень организации и доступные взаимодействия	Отождествлять модель ткани с целым мозгом

Класс системы	Что нужно описать прежде всего	Частая ошибка при выводе
Гибрид	Границы частей и распределение причинных функций	Не различать обучение контроллера и биологического компонента

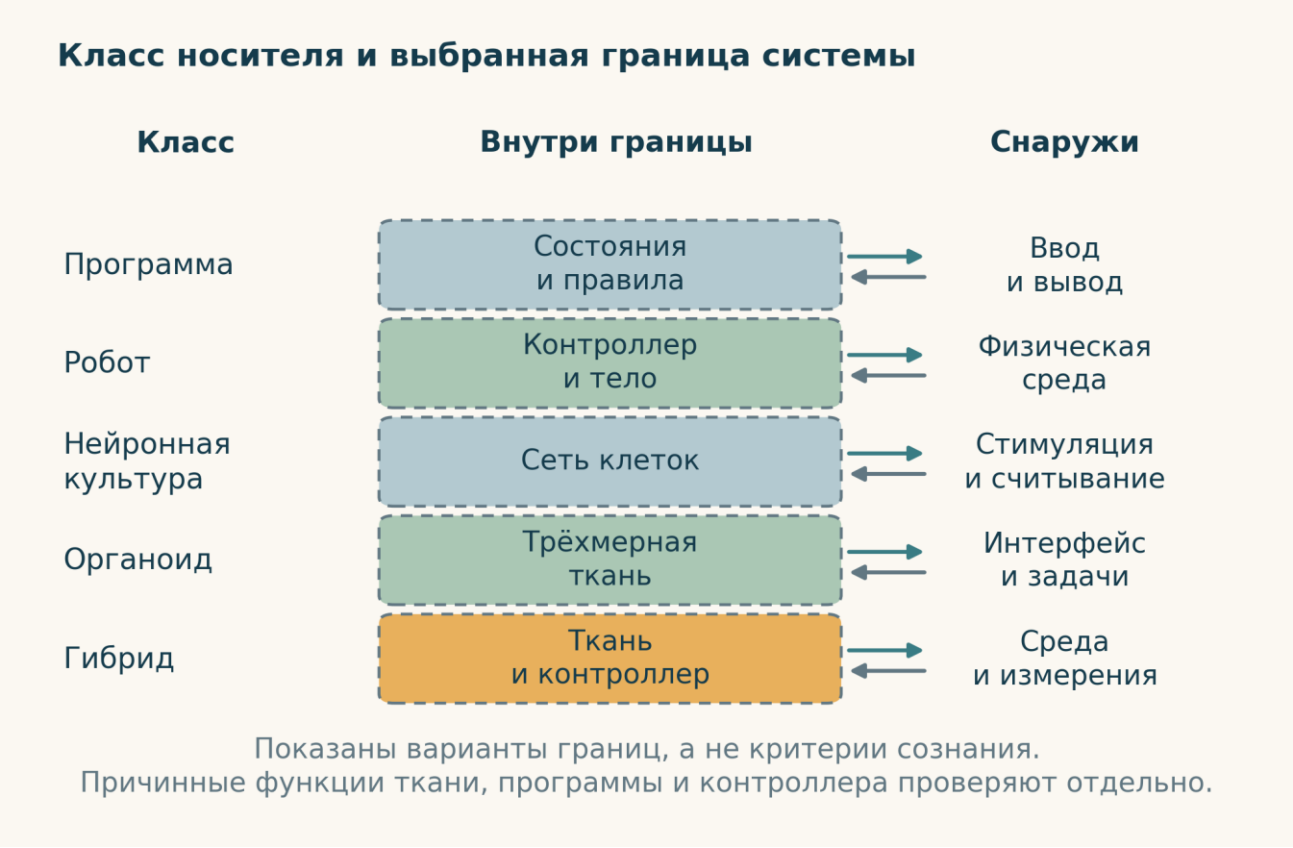

Рисунок 53. Носитель не определяет границу исследования. Пять примеров выбора границы системы. Пунктир отделяет рассматриваемую часть от внешнего взаимодействия. Для программы, робота, культуры, органоида и гибрида необходимо установить, где выполняются обучение, хранение состояния и управление. Это варианты описания установки, а не обязательное устройство всех систем соответствующего класса и не рейтинг вероятности сознания.

Как не превратить признаки в рейтинг

Представим два кандидата. У одного есть хорошо проверенный контроль ошибки, но неизвестна организация распространения содержания. У второго есть общий канал обмена, но оценка собственной надёжности формируется по готовому правилу. Простое сложение галочек скроет различие. Для одной теории отсутствие определённого признака критично, для другой тот же признак не обязателен.

Поэтому результат аудита лучше представить как профиль. Для каждого требования указываются теория, интерпретация, измерение, результат и

остающаяся альтернатива. Отдельно отмечаются неизвестные свойства. Если оборудование недоступно, описание программы не заменяет сведения о физической причинности, когда именно она существенна для проверяемого подхода.

Проверяйте также, действительно ли предполагаемый признак используется. Модуль может быть подключён к общей памяти и никогда не читать интересующее содержание. Другой модуль способен хранить оценку уверенности, которая не влияет ни на один выбор. Наличие полей в коде тогда преувеличивает функциональную организацию. Избирательное удаление, искажение и восстановление состояния помогают установить причинную роль.

Однако отсутствие изменения после удаления не всегда даёт окончательный отрицательный вывод. Возможно, задача не требовала этой функции или её выполнил дублирующий механизм. Следует заранее предусмотреть условие, где функция должна быть востребована, и проверить такое объяснение компенсации. Тогда архитектурный аудит становится исследованием работающего устройства.

Минимальная схема доступная до полного практикума

Пусть устройство хранит вероятность цели справа p и вероятность прямого управления q. Модуль выбора читает оба числа и выбирает команду. Монитор читает те же числа и вычисляет ожидаемый успех. Текстовый отчёт можно либо формировать из монитора, либо подставлять по фиксированному шаблону. Изменение шаблона при сохранённых состояниях не должно менять выбор команды. Искажение q, напротив, способно изменить и выбор, и прогноз. Это схема причинных проверок; полные правила обновления будут разобраны в главе 28. Уже сейчас можно отличить архитектурное требование от впечатления, создаваемого текстом.

Что теперь установлено

Мы можем составить профиль даже при неполном знании реализации, если неизвестные свойства так и обозначены. Наличие памяти, телесного корпуса или нейронных клеток не заменяет проверки конкретных связей.

Практика 23 Паспорт кандидата

Зачем выполнять: Раздельно оценить доступ к состояниям и словесное самоописание на готовой схеме.

Что понадобится: Модель мира и модель действия из главы 15; внутренний доступ и внешний журнал из главы 14. Название компонента не является проверкой его реализации.

Ориентир по времени: 25–35 минут. Можно делать с перерывами; это не норматив скорости.

1. Нарисуйте четыре компонента: оценки p и q, выбор, монитор, отчёт. Проведите только указанные в разборе связи.
2. Для каждой связи укажите отправителя, получателя и последствия её отключения. Не добавляйте скрытую истинную цель в монитор.
3. Сравните два вмешательства: смену шаблона фразы и искажение q при прежнем внешнем входе. Запишите ожидаемые изменения выбора, прогноза и текста отдельно.
4. Составьте профиль по двум теориям: что схема моделирует, что не проверено и какое основание для вывода об опыте ещё потребуется.

Что сохранить: Карта связей и профиль проверенных, непроверенных и отсутствующих свойств.

Подсказка: По рисунку не следует, что все компоненты реально реализованы. Когда перейдёте к коду, статус каждой связи потребуется подтвердить.

Исследовательское продолжение: После главы 28 повторите профиль по действующему агенту и его журналу. Не заменяйте проверку реализации совпадением с названием блока.

Разбор и ответы — приложение E, глава 23, с. 305

Вопросы для самопроверки

1. Почему успешный разговор о переживаниях недостаточен для вывода о них?
2. Гарантирует ли биологический носитель сознание?
3. Чем профиль кандидата лучше количества выполненных признаков?

Оценка чужих кандидатов подготовила язык требований. Теперь выберем ограниченное объясняемое различие и построим собственную модель, которую можно сравнить с содержательными конкурентами.

Часть 7

Разработка собственных объяснений

Глава 24

Требования к будущей модели

До сих пор мы разбирали имеющиеся объяснения и реализации. Теперь начнём собственный проект. Возьмём конкретный вопрос о влиянии истории источника на оценки принадлежности. Он достаточно узок для построения механизма и одновременно связан с более широкой темой субъекта. Сначала определим, что модель обязана объяснить, какие факты уже опубликованы и какие результаты пока являются только нашим ожиданием.

От интересного вопроса к обязательствам объяснения

Представьте, что вы управляете изображением руки на экране. Движение обычно соответствует вашей команде, но иногда запаздывает. После нескольких минут устойчивой связи короткая задержка кажется случайной помехой. Если связь постоянно менялась, та же задержка может восприниматься иначе. Возникает содержательный вопрос: зависит ли переживаемая принадлежность только от текущего совпадения сигналов или также от истории предполагаемого источника? Это мысленный учебный пример, а не сообщение о проведённом эксперименте. Его назначение — превратить общую идею о непрерывности себя в определённые требования к модели.

Модель начинается с обязательства объяснить различие. Фраза «мозг строит себя» допускает почти любой результат. Утверждение «при одинаковом текущем согласовании история устойчивого источника изменяет последующую динамику принадлежности» уже ограничивает ожидания. Теперь можно спросить, сколько длится влияние истории, переносится ли оно между источниками и сохраняется ли при изменении способа ответа. Чем точнее поставлены эти вопросы, тем яснее, какой результат будет успехом объяснения.

Определите предмет с нужной точностью

Первоначальный предмет проекта — переживаемая телесная принадлежность. В эксперименте Калькерта и Эрссона принадлежность искусственной руки и ощущение авторства движения могли расходиться. Поэтому принадлежность нельзя определять исключительно успешным произвольным управлением. Модель должна объяснять это различие, сохраняя отдельные показатели принадлежности, управления и оценки правильности действия. [4]

Предмет можно записать одним предложением: «Мы объясняем изменение переживаемой принадлежности при изменении истории согласованных телесных сигналов». Следующая строка уточняет область применения: взрослые участники, краткие эпизоды, определённый способ предъявления и отчёта. Такая запись сохраняет возможность расширения. Она не требует уже сейчас объяснять цвет, сновидения, биографическое самоощущение и первоначальное возникновение сознания одной переменной.

При расширении будет меняться само объяснительное обязательство. Сначала модель предсказывает оценки принадлежности. Затем потребуется обосновать их связь с переживаемой организацией. Наконец, более сильная гипотеза должна сказать, почему предлагаемая организация составляет или обеспечивает наличие опыта. Эти переходы станут отдельными задачами, а не незаметным увеличением громкости одного вывода.

Различайте данные и желаемые результаты

Полезно завести два перечня. Первый содержит опубликованные результаты, которые модель должна учитывать. Второй — ещё не проверенные ожидания нашего проекта. Если смешать их, собственное предположение начнёт выглядеть установленным фактом. Например, утверждение о переносе истории между двумя похожими изображениями пока относится ко второму перечню, если именно такого сравнения мы не нашли в литературе.

Работа Samad и коллег показывает, как принадлежность можно исследовать через модель вывода об общей причине зрительных, проприоцептивных и тактильных сигналов. Это ближайший опубликованный ориентир, а не наша новая идея. Собственный вклад придётся искать в определённой динамике, условиях переноса или различении альтернативных механизмов. [15]

Каждая запись первого перечня должна включать задачу, показатель и границу вывода. «Сигналы объединяются» слишком широко. «В таких условиях изменилась оценка принадлежности, одновременно измеряли положение руки» уже позволяет выяснять, какое именно наблюдение

воспроизводится. Также отмечайте, совпадают ли результаты нескольких показателей. Их расхождение может оказаться главным требованием к будущей модели.

Четыре требования к нашему проекту

Первое требование касается настоящего. Модель должна реагировать на текущую совместимость сигналов: их согласование не может быть совершенно несущественным. При этом необходимо определить, какие признаки совместимости доступны системе. Геометрическое совпадение, временная близость и зависимость от действия — разные сведения; нельзя заменять их нерасшифрованной переменной «согласованность».

Второе требование касается прошлого. Влияние истории должно иметь вычислимую динамику. Недостаточно добавить память и затем объявлять любое запаздывание её подтверждением. Нужно заранее указать, что сохраняется: отдельные события, оценка устойчивости среды, ожидание общей причины или история определённого источника. Каждое решение предлагает собственный механизм забывания и переноса.

Третье требование касается идентичности. Если история принадлежит определённому источнику, модель должна объяснять, почему она возвращается при его повторном появлении. Одновременно нужен конкурент, который отслеживает внешние объекты тем же общим способом. Иначе преимущество памяти источника будет ошибочно названо доказательством особого механизма самосознания.

Четвёртое требование касается измерения. Одинаковое внутреннее состояние может давать разные ответы при изменении формулировки вопроса или критерия согласия. Поэтому модель формирования оценки должна быть отделена от модели источника. Такое разделение позволит проверить, где именно возник эффект: в обработке сигналов, в их отнесении или в способе сообщения результата.



Рисунок 54. Карта обязательств модели. Методическая схема. Четыре требования относятся к устройству и проверке модели. Пунктир обозначает отдельную аргументативную задачу: объяснить связь механизма с переживанием; он не означает установленный переход от точного прогноза к наличию опыта.

Сформулируйте несколько конкурентов

Назовём M0 моделью текущего согласования. Она использует только выбранное текущее окно. M1 сохраняет общую историю подтверждений и противоречий без отдельных записей об источниках. M2 ведёт различаемые истории предполагаемых источников. Эти названия обозначают семейства будущих вычислений; пока для каждого семейства нужно выбрать конкретные правила.

Добавим M1И — общую модель отслеживания идентичности объектов. Она нужна, потому что M2 может выигрывать благодаря обычному разделению потоков. Сравнение M2 с M1И должно показать, требуется ли дополнительная телесная организация или достаточно механизма, применимого к любому устойчивому объекту. Возможный ответ «достаточно общего механизма» будет результатом исследования природы самоотнесения.

Не делайте конкурента намеренно слабым. Если M0 получает меньше текущей информации, а M2 дополнительно видит скрытые метки, их сравнение проверяет преимущество входа. Если M2 имеет значительно больше

настраиваемых параметров, нужна проверка на новых данных. Требование справедливого сравнения входит в спецификацию модели с самого начала.

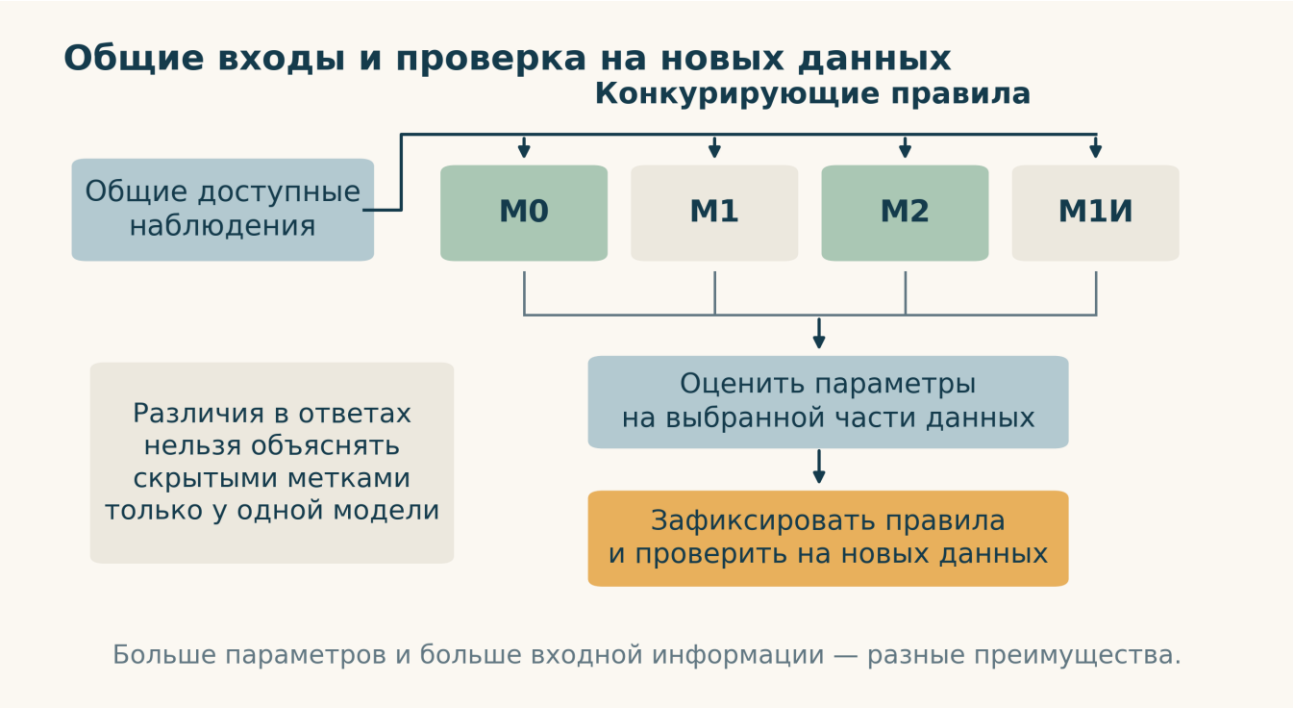

Рисунок 55. Справедливое сравнение вычислительных конкурентов. M0, M1, M2 и M1И сравнивают при общем наборе доступных наблюдений, сохраняя различия их правил памяти. Параметры оценивают отдельно от проверки на новых данных, а скрытую метку источника нельзя предоставлять только предпочитаемому варианту. Преимущество в таком сравнении относится к заданной задаче и ещё не подтверждает специальную природу самоотнесения.

Задайте положительный критерий успеха

Предположим, модель должна предсказывать восстановление принадлежности после возвращения знакомого источника. Успехом будет не просто совпадение средней оценки. Нужны предсказанные направление, скорость изменения и условия, при которых эффект исчезает. Особенно ценно, если параметры, оценённые в одной задаче, помогают предсказать другую без новой произвольной настройки.

Для учебного проекта можно задать требование: модель должна различать две последовательности с одинаковым числом согласованных событий, но разной организацией источников. Это ещё не экспериментальная закономерность. Это выбранная проверка, которая заставляет теорию объяснить роль порядка. В следующей главе мы построим простую последовательность, где вычислительные варианты действительно расходятся.

Критерий успеха полезно дополнить условием остановки. Если доступные данные не различают M1 и M2, проект пока не отвечает на вопрос об истории

источника. Тогда следует изменить задачу, а не увеличивать уверенность в любимой модели. Хорошая спецификация помогает вовремя увидеть, какое исследование нужно сделать следующим.

Как общий замысел превращается в обязательство

Замысел «история помогает ощущать себя» допускает слишком много результатов. Уточнение связывает исходное различие с процедурой: при одинаковой тестовой последовательности должна по-разному восстанавливаться оценка принадлежности после возвращения знакомого или нового источника. Дальше нужно указать, что означает знакомство, какой признак распознаёт источник и как измеряется динамика ответа. Отдельно записывается конкурент, который хранит историю любых внешних объектов. Если он даёт тот же прогноз, специальная телесная функция пока не выделена. Такое ограничение помогает выбрать следующий опыт, а не обесценивает проект.

Что теперь установлено

Хорошее требование ограничивает и ожидаемый успех, и условия пересмотра. Теперь можно составить спецификацию, в которой прошлое, настоящее, идентичность и измерение получают отдельные роли.

Практика 24 Составьте требования к одной модели

Зачем выполнять: Подготовить требования к одной модели прежде чем выбирать удобные правила расчёта.

Что понадобится: Различение принадлежности и авторства; модель общей причины; честное сравнение конкурентов. Успешность прогноза ответа и наличие опыта требуют разных оснований.

Ориентир по времени: 35–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Одним предложением определите объясняемое различие и область условий. Для базовой работы используйте возвращение источника после вмешавшегося другого источника.
2. Разделите опубликованные ограничения и собственные ожидания. Для опубликованного результата укажите, что именно измерялось.
3. Выпишите четыре требования: текущие признаки, сохранение истории, распознавание идентичности, формирование ответа. Назовите M0, M1, M2 и общий трекер как будущих конкурентов.

4. Определите результат, который поддержит выбранное различие, и результат, после которого специальная версия окажется избыточной. Пока не назначайте произвольные человеческие средние.

Что сохранить: Одна спецификация с предметом, требованиями, конкурентами и критерием следующего решения.

Подсказка: Увеличение числа параметров и добавление доступных сведений являются разными преимуществами. Проверять нужно оба.

Исследовательское продолжение: Предложите вторую задачу, в которую модель должна перенести параметры без свободной настройки. Объясните, почему она проверяет то же отношение.

Разбор и ответы — приложение Е, глава 24, с. 305

Самопроверка

1. Почему сохранение информации о прошлом само по себе не подтверждает специальный механизм самоотнесения?
2. Что показывает преимущество M2 перед M0, если первая получает скрытую метку источника, а вторая её не получает?
3. Зачем разделять установленное ограничение и собственное ожидание?

Требования ещё не определяют следующий шаг модели. В главе 25 зададим вход, состояние и обновление и проследим один расчёт до конца.

Глава 25

От идеи к механизму

В главе 24 мы потребовали объяснить влияние истории определённого источника. Теперь нужно задать процедуру, которая по старому состоянию и новому наблюдению однозначно вычисляет следующее состояние. Это станет первым исполнимым звеном собственного проекта. Мы подробно разберём обновление одной вероятности, а затем изменим только организацию памяти: отсутствие переноса, общую запись и отдельные записи источников.

Механизм должен позволять следующий шаг

В предыдущей главе мы потребовали, чтобы модель объясняла влияние истории определённого источника. Теперь нужно сделать идею исполнимой. Если сообщить модели текущее состояние и новое наблюдение, она должна определить следующее состояние. Фраза «система объединяет важную информацию» этого не обеспечивает: неизвестно, что считается важным, как происходит объединение и что сохраняется между событиями.

Начнём с небольшой конструкции. Она не изображает устройство мозга во всех подробностях. Её задача — показать, как разные правила обращения с историей дают различные предсказания. Все числа и последовательности этой главы являются учебными. Они выбраны для прозрачного расчёта и не описывают измеренные человеческие переживания.

Разведите вход, состояние и выход

На вход поступает событие двух типов: поддерживающее общую причину сигналов или противоречащее ей. В реальном проекте такую информацию придётся получать из времени, положения и связи с действием. Здесь мы временно заменим многомерное событие двоичным признаком. Это явное упрощение, область применимости которого нужно будет проверить.

Скрытое состояние модели — вероятность того, что текущие сигналы относятся к общей причине. Назовём её убеждением. Это техническое слово: модель хранит число и обновляет его по указанному правилу. В неё не добавляется внутренний наблюдатель, который понимает смысл числа. Выход пока представляет само обновлённое убеждение. Связь с человеческой оценкой принадлежности будет отдельным преобразованием.

Наконец, различайте истину среды и оценку модели. Экспериментатор может знать, какие события действительно связаны. Модель получает лишь

несовершенные признаки. Если передать ей истинную причинную метку, исчезнет задача вывода. Поэтому каждый вход следует проверить вопросом: каким наблюдаемым способом система могла получить это сведение?

Постройте одно правило обновления

Пусть поддерживающее событие встречается с вероятностью 0,75 при общей причине и с вероятностью 0,25 при разных причинах. Это выбранная нами модель надёжности признака. При исходном убеждении 0,5 одно поддерживающее событие даёт новое убеждение 0,75. Противоречащее событие при том же исходном состоянии даёт 0,25.

Между событиями будем ослаблять старое убеждение в направлении неопределённости. В учебной версии сохраняются восемь десятых отклонения от 0,5. Поэтому состояние 0,75 перед следующим наблюдением превращается в 0,70. Ослабление может интерпретироваться как допущение изменения скрытой ситуации. Это выбранная динамика модели, а не установленный закон памяти.

```
def update(previous, supports_common_cause, retention=0.8):
    prior = 0.5 + retention * (previous - 0.5)
    likelihood_common = 0.75 if supports_common_cause else 0.25
    likelihood_separate = 1.0 - likelihood_common
    common_weight = prior * likelihood_common
    separate_weight = (1.0 - prior) * likelihood_separate
    return common_weight / (common_weight + separate_weight)
```

Этот фрагмент определяет переход между состояниями. Он не вычисляет квалиа. Его можно выполнять вручную, на калькуляторе или в Python. Каждый параметр имеет отдельную роль: надёжность признака задаёт силу нового свидетельства, сохранение — влияние прошлого. Если результат меняется, теперь можно установить, какое правило вызвало изменение.

Один переход в общем виде

Обозначим сохранённую вероятность общей причины буквой p, коэффициент удержания r, надёжность положительного признака a. Для базового примера r = 0,8 и a = 0,75. Перед наблюдением получаем p_до = 0,5 + r × (p − 0,5). Значение r = 0 полностью убирает предыдущее отклонение, r = 1 сохраняет его.

Для положительного признака вес общей причины равен p_до × a, вес разных причин равен (1 − p_до) × (1 − a). Новое p получается делением первого веса на их сумму. Для отрицательного признака правдоподобия меняются местами. Такой выбор предполагает симметричную надёжность двух исходов и обязательно входит в описание модели.

В расчёте различаются три числа. Сохранённое p относится к предыдущей обработке; p_до учитывает возможное изменение причины; новое p учитывает уже полученный признак. Для предсказания самого признака используем ещё одну величину: P(+) = a × p_до + (1 − a) × (1 − p_до). Именно она записывается до наблюдения, когда позднее оценивается качество прогноза.

Числа будущей последовательности не должны поступать в правило. Даже если среда известна разработчику, модель получает только разрешённые признаки и метку наблюдаемого источника. Когда мы позже усложним распознавание идентичности, безошибочная метка будет заменена отдельной оценкой, а не добавлена как скрытая подсказка.

У M0, M1 и M2 меняется только выбор записи для этого обновления. Поэтому различия можно проследить до конкретного правила памяти. Это намного сильнее общего утверждения о том, что история важна: теперь известны случаи совпадения, расхождения и условия, при которых полученный эффект должен исчезнуть.

Три способа пользоваться одним обновлением

M0 перед каждым событием начинает с 0,5. Таким образом, поддерживающее событие всегда приводит к 0,75. M1 сохраняет одно общее убеждение для всей последовательности. M2 хранит отдельное убеждение для каждого различаемого источника. При появлении источника модель обновляет связанную с ним запись. В этой упрощённой версии ослабление происходит при следующем обращении к записи, а не на каждом календарном шаге.

Рассмотрим три события: источник А поддерживает общую причину; источник Б противоречит ей; источник А снова поддерживает её. Для M1 состояния последовательно составят 0,75, затем 0,4375 и приблизительно 0,7105. Для M2 получим 0,75 в записи А, 0,25 в записи Б и 0,875 после возвращения А. M0 выдаст соответственно 0,75, 0,25 и 0,75.

Различие возникло из обращения с историей. При возвращении А модель M2 использует его прошлое подтверждение; M1 несёт также недавнее противоречие от Б. Мы получили конкретное предсказание, которое можно проверить вычислением. Но пока это предсказание внутренней динамики нашей конструкции. Чтобы связать его с данными человека, потребуется модель измерения и эмпирическая проверка.

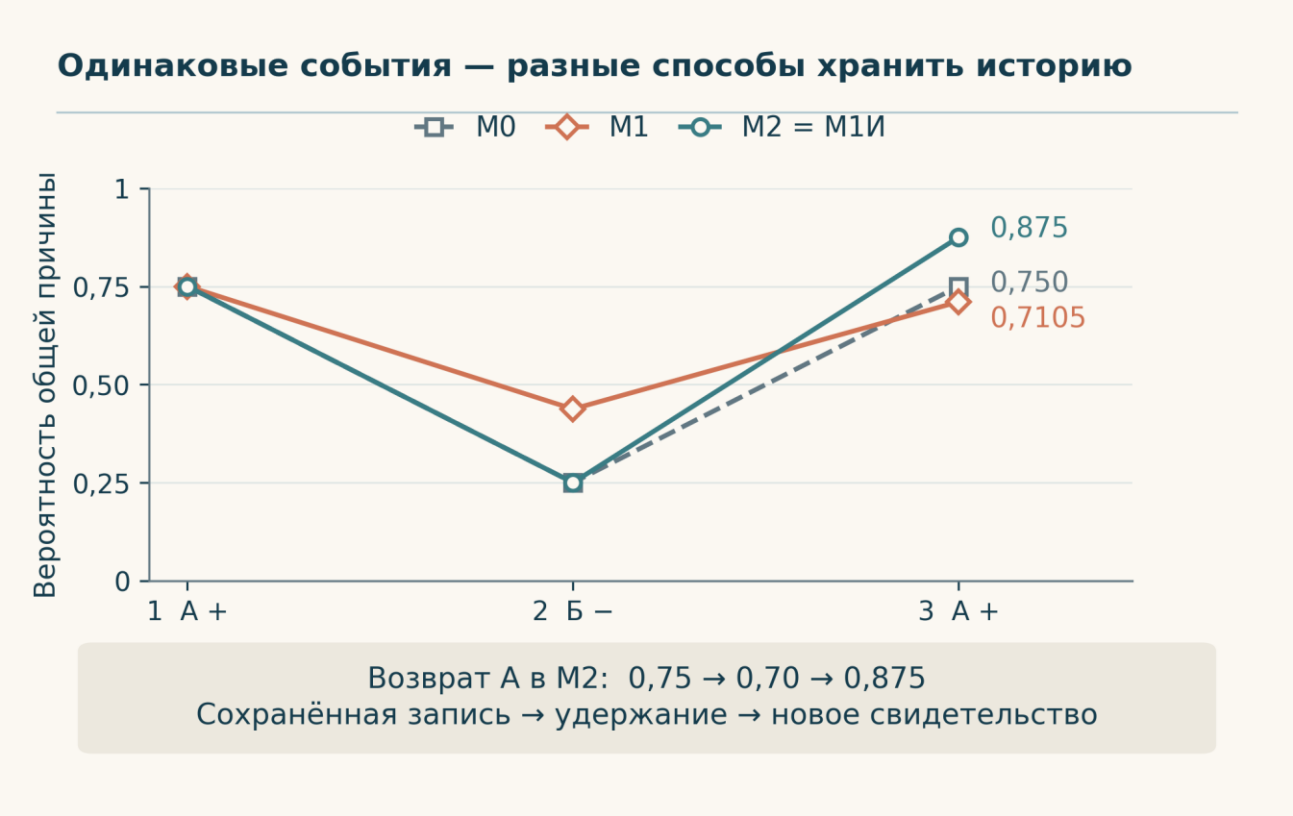

Рисунок 57. Возвращение к истории источника. Точный учебный расчёт для последовательности А+, Б−, А+. Надёжность признака 0,75; удержание 0,8 применяется перед обновлением соответствующей записи. M0 начинает с 0,5, M1 переносит общую историю, M2 и M1И ведут одинаковые отдельные записи. Это состояния вычислительных моделей, а не измеренные переживания или оценки участников.

Откуда берётся идентичность источника

Самый опасный обходной путь — передать M2 безошибочное имя источника и затем приписать ей способность самостоятельно обнаруживать идентичность. В учебном расчёте буквы А и Б обозначают идеализированное различение. В полном проекте необходимо объяснить его происхождение: например, источник распознаётся по устойчивому сочетанию положения, внешнего вида и ожидаемых последствий действия.

Такое распознавание тоже может быть неопределённым. Если событие с вероятностью 0,7 относится к А и с вероятностью 0,3 к Б, можно распределять обновление между записями или хранить несколько возможных историй. Эти решения не эквивалентны. Первое сглаживает различия, второе сохраняет конкурирующие версии. Выбор должен основываться на предсказаниях и вычислительных ограничениях, а не на удобстве названия.

Затем понадобится общий конкурент M1И из предыдущей главы. Он использует те же признаки идентичности для внешних объектов. Если M1И объясняет данные столь же хорошо, дополнительного специального механизма

телесного самоотнесения пока не требуется. Механизм может быть общим по устройству и при этом участвовать в содержании самосознания.

Добавьте измерение, не пряча в нём объяснение

Предположим, участник отмечает принадлежность на шкале от нуля до десяти. Самая простая учебная связь — умножить убеждение модели на десять. Тогда последнее событие последовательности даст M0 значение 7,5, M1 примерно 7,1, M2 — 8,75. Это условные прогнозы ответа, а не результаты эксперимента и не количественная мера самого переживания.

Более реалистичная связь может включать индивидуальный критерий ответа, нелинейность шкалы и случайную изменчивость. Но каждый добавленный элемент увеличивает гибкость объяснения. Если любое внутреннее состояние разрешено преобразовать в любой ответ, сравнивать механизмы становится невозможно. Поэтому способ формирования ответа нужно определить заранее и проверить, переносится ли он между условиями.

Полезно проектировать дополнительную задачу, которая помогает оценить надёжность сенсорного различения независимо от принадлежности. Тогда модель не будет одновременно угадывать и качество входа, и силу истории, и критерий ответа по одному числу. Несколько источников данных могут ограничивать разные параметры, хотя окончательная идентифицируемость требует отдельной проверки.

Проверьте граничные случаи

При сохранении, равном нулю, каждое обновление начинается с 0,5. Различия между использованием прошлых убеждений исчезают. При сохранении, равном единице, старое состояние переносится без ослабления. Если события перестают различать общую и разные причины, новое наблюдение не должно менять убеждение. Эти случаи следуют из выбранных правил и служат проверкой реализации.

Затем измените порядок событий, сохраняя их количество. Переставьте имена А и Б везде одновременно: результат должен измениться только по названиям записей. Если переименование само по себе меняет прогноз принадлежности, в программе, вероятно, действует лишнее правило. Такой тест проверяет инвариантность к обозначениям, а не существование субъективной перспективы.

Не ограничивайтесь красивой последовательностью. Найдите условия, где модели совпадают. Если источник всегда один, M1 и упрощённая M2 ведут себя одинаково. Следовательно, эксперимент только с одним источником не

сможет проверить ценность отдельных историй, сколько бы наблюдений он ни содержал.

Почему одинаковое свидетельство не даёт одинаковую оценку

Перед повторным появлением А его сохранённая вероятность равна 0,75. Удержание 0,8 сохраняет восемь десятых отклонения от 0,5: предварительная вероятность становится 0,70. Положительный признак имеет правдоподобие 0,75 при общей причине и 0,25 при разных. Получаются веса 0,525 и 0,075. Их сумма 0,600 нужна для нормировки, поэтому новая вероятность общей причины равна 0,875.

Эта вероятность не равна прогнозу следующего положительного признака. Для прогноза нужно ещё учесть ненадёжность наблюдения: 0,875 умножить на 0,75 плюс 0,125 умножить на 0,25, то есть 0,6875, если скрытая причина до следующего наблюдения не менялась. Если допускается смена причины, сначала применяется соответствующий переход. Эти две величины нельзя незаметно оценивать по одному и тому же исходу.

Что теперь установлено

Новые числа появились из явно выбранных правил. Теперь можно проверить, какая именно сохранённая запись участвует в обновлении и что произойдёт при её замене.



Рисунок 56. Нормировка одного обновления. Полный шаг главы 25 при надёжности 0,75. Предварительное состояние после удержания, ненормированные веса и итоговая вероятность подписаны отдельно. Вероятность скрытой причины не подменяет прогноз наблюдаемого признака.

Практика 25 Проследите причинную цепочку

Зачем выполнять: Восстановить четыре события для трёх вариантов памяти и общего трекера.

Что понадобится: Нормировка вероятностей и различие наблюдения и скрытой причины из главы 11. Удержание истории здесь применяется при обращении к записи; календарное время рассмотрено отдельно в главе 33.

Ориентир по времени: 35–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Используйте последовательность А положительный, Б отрицательный, А положительный, Б положительный. Везде начальное убеждение 0,5; надёжность 0,75; удержание 0,8.
2. Для M0 каждый раз начинайте с 0,5. Для M1 переносите одну запись. Для M2 обновляйте только запись распознанного источника. Общий трекер использует те же отдельные записи.
3. На каждом шаге выпишите выбранную запись, состояние после удержания, два веса и итог нормировки. Сначала рассчитайте сами, затем откройте разбор.
4. Повторите с удержанием ноль и проверьте переименование А и Б во всей последовательности.

Что сохранить: Четыре пошаговых расчёта и два контроля с заранее известным результатом.

Подсказка: У M2 на четвёртом шаге возвращается отрицательная история Б. Раздельная память не обязана давать наибольшую уверенность.

Исследовательское продолжение: Замените безошибочные метки источника наблюдаемыми признаками. Опишите, как неопределённость распознавания изменит выбор записи, и сравните сильного общего трекера на тех же входах.

Разбор и ответы — приложение Е, глава 25, с. 306

Самопроверка

1. Почему M2 сильнее поддерживает общую причину при третьем событии учебной последовательности?
2. Что потеряется при использовании только одного источника во всём исследовании?
3. Почему преобразование убеждения в оценку принадлежности является отдельной частью модели?

Механизм памяти стал исполнимым. Теперь необходимо явно сформулировать более сильное утверждение: какая организация, по авторской гипотезе, могла бы быть связана с качеством и наличием опыта.

Глава 26

От механизма к переживанию

Мы умеем вычислять историю причин и предсказывать определённые ответы модели. Центральная гипотеза этой главы добавляет самостоятельный шаг: предполагает связь некоторой физически реализованной взаимной организации с минимальным опытом. Разберём этот шаг последовательно. Сначала определим исследуемый механизм без феноменальных слов, затем сильное утверждение достаточности, затем структурные следствия и способы возражения. Такой порядок сохраняет исследовательскую смелость и делает её обязательства обсуждаемыми.

Какой вопрос ещё не получил ответа

Модель из предыдущей главы умеет сохранять историю, различать источники и менять вероятность общей причины. Допустим, она хорошо предсказывает человеческие оценки принадлежности. Мы узнаем нечто существенное о механизме этих оценок. Однако остаётся вопрос: почему соответствующие процессы вообще сопровождаются переживанием? В философской постановке Чалмерса объяснение способности различать и сообщать не является автоматически объяснением наличия субъективного опыта. Здесь мы используем это различение как требование к аргументу, а не как заранее доказанный запрет на объяснение. [2]

Чтобы двигаться дальше, нужен принцип соответствия: явное утверждение о том, как организация системы связана с качеством и наличием опыта. Оно может оказаться ошибочным. Но если такого утверждения вообще нет, технический проект никогда не выйдет за пределы того, что в нём непосредственно запрограммировано. Сформулируем сильную рабочую гипотезу и выясним, какие обязательства она создаёт.

Определите организацию без слова переживание

Предлагаем исследовать процессы, в которых несколько различений взаимно определяют друг друга. Один компонент различает свойства наблюдаемого объекта. Второй оценивает источник изменения. Третий задаёт, какие дополнительные наблюдения или внутренние преобразования сейчас возможны. Их состояния обновляются совместно: предполагаемый источник меняет интерпретацию признака, признак меняет ожидание источника, а выбранная проверка перестраивает оба представления.

Взаимность здесь имеет причинный смысл. Для каждой предполагаемой связи заранее указывают направленные вмешательства и ожидаемые изменения в обоих направлениях: например, как изменение оценки источника изменит интерпретацию признака и как изменение представленного признака изменит оценку источника. Общей активности или корреляции недостаточно. Кроме того, система должна различать альтернативы, которые не происходят сейчас: что изменится при другом наблюдении, другой команде или отказе от прежней интерпретации. Эти альтернативы должны влиять на обработку, а не существовать только в объяснении разработчика.

Наконец, отдельные различения образуют общую динамику. Нельзя заменить каждый компонент независимым воспроизведением его обычных выходов и сохранить ответы на все заранее определённые вмешательства. Такая проверка отличает взаимное определение от совпадения нескольких записанных потоков. Она не определяет сознание по названию; она задаёт исследуемое устройство системы.

Рабочий принцип соответствия

Сформулируем гипотезу: **определённый класс физически реализованных процессов взаимного определения различий достаточен для минимального субъективного опыта; качественная организация этого опыта соответствует устойчивым отношениям между доступными процессу различениями и альтернативами**. Это авторская исследовательская гипотеза для данной книги. Её самостоятельность относительно существующих подходов и её истинность ещё предстоит установить.

В рабочей версии класс ограничен тремя требованиями: представлены несколько раздельно изменяемых признаков или степеней свободы содержания; их изменения причинно связаны в замкнутой динамике; отношения между ними используются для выбора и пересмотра альтернатив. Раздельная изменяемость не означает статистическую независимость: как раз взаимодействие признаков представляет предмет проверки. Одного счётчика, записи истории или самоназвания недостаточно для выполнения этих требований. Однако наличие всех трёх тоже пока не является установленным признаком опыта: именно достаточность организации составляет сильную часть гипотезы.

Достаточность не означает необходимости: принцип не утверждает, что без этой организации невозможен любой опыт. Предсказания качественных изменений должны следовать из его структурной части и условий конкретного вмешательства.

В принципе не требуется словесное «я» или постоянное внешнее движение. Выбор может касаться внутреннего уточнения представления. Самопринадлежность появляется как дополнительное содержание, когда система связывает изменения с устойчивым источником собственного воздействия. Поэтому M2 может стать частью предполагаемого механизма самосознания, но сама таблица отдельных историй ещё не реализует всю предложенную организацию.

Почему такая гипотеза заслуживает проверки

Первое основание состоит в единстве объяснения. Взаимные ограничения могут одновременно задавать, какие содержания совместимы, как они сохраняются и почему изменение одного меняет остальные. Вместо отдельных необъяснённых ярлыков для качества, единства и времени появляется один исследуемый процесс. Это достоинство замысла, а не доказательство его истинности.

Второе основание связано с перспективой. Состояние имеет значение для последующих изменений самой системы: оно определяет, какие различия та сможет проверить и какие ошибки исправить. Внешний наблюдатель может описать этот процесс, но его причинная организация не создаётся описанием. Гипотеза предлагает считать такую внутреннюю определённость возможной основой субъективной перспективы.

Третье основание — возможность вывести последствия за пределами исходной задачи. Если предложенная организация связана с опытом, разрыв определённых взаимных связей должен менять его организацию предсказуемым способом, даже когда простое различение сохраняется. Согласование таких предсказаний с независимыми данными людей даст основание рассматривать реализацию как кандидата. Переход к искусственной системе всё равно потребует аргумента о переносимости принципа.

Качество через отношения и инварианты

Пусть учебная система различает оттенок, яркость и источник зрительного изменения. Значение одной координаты в её векторе не является готовой «краснотой». Нужно исследовать отношения: какие состояния она смешивает, какие противопоставляет, как изменение яркости влияет на распознавание оттенка и какие вмешательства сохраняют различие. Полученная структура намного содержательнее отдельного числового кода.

Если мы только переименовали состояния или сменили координаты, сохранив все существенные переходы, рабочая гипотеза предсказывает сохранение

качественной организации. Если изменили отношения различимости и их взаимное влияние, она предсказывает изменение организации. Так возникает инвариант: содержание зависит от определённых отношений, а не от выбранных разработчиком обозначений.

Это ещё не устанавливает, какое именно человеческое цветовое переживание соответствует состоянию машины. Для такого сопоставления нужны дополнительные основания. Даже полное совпадение структуры сходств не разрешает автоматически вопрос об абсолютном качественном соответствии. Вместе с тем объяснение отношений между качествами представляет самостоятельный содержательный результат и ограничивает возможный будущий ответ.

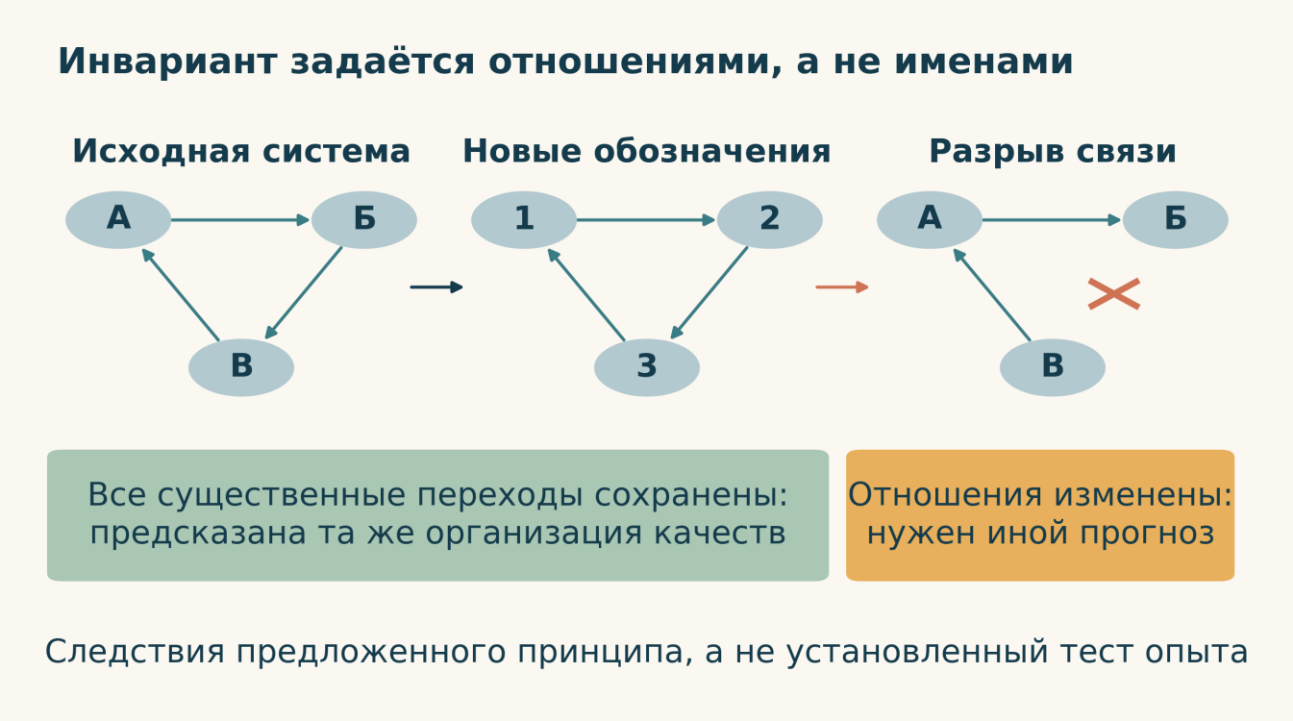

Рисунок 58. Переименование и изменение причинной структуры. Схематический пример инварианта рабочей гипотезы главы 26: переименование сохраняет существенные отношения, а удаление связи меняет предмет прогноза. Принцип предлагает достаточную организацию для опыта; из нарушения отдельной связи не следует необходимое исчезновение опыта.

Дойти до наличия опыта

Сильная часть принципа утверждает достаточность организации, а не только соответствие её структуры самоотчётам. Мы принимаем это как проверяемую в широком исследовательском смысле гипотезу отождествления: некоторый процесс взаимного определения и минимальный опыт представляют один конкретный процесс, описанный с разных сторон. Такое отождествление нельзя получить простым определением слова «опыт».

Его обоснование должно складываться из независимых связей: объяснения уже доступных переживаний, предсказания их изменений, устранения конкурирующих механизмов и последовательного применения к новым реализациям. Если теория объясняет только тексты машины, основание остаётся слабым. Если один механизм связывает разнообразные феноменальные различия с определёнными причинными изменениями, аргумент становится сильнее, хотя логическая возможность альтернатив сохраняется.

Критик вправе спросить, почему взаимная определённость должна вообще ощущаться. Наш принцип не выводит ответ из нейтральной арифметики: он предлагает конкретное тождество и берёт на себя его следствия. Дальнейшая работа должна показать, что это тождество лучше организует объяснение и предсказания, чем конкуренты. Так мы действительно ставим вопрос о существовании опыта, не подменяя его успехом одной функции.

Уточните реализацию и границы

Слово «физически» не решает вопрос автоматически: любой запущенный код имеет носитель. Требуется установить, какая причинная организация должна существовать в самом исполнении. Если существенна только определённая вычислительная структура, разные подходящие устройства могут её реализовать. Если существенны дополнительные физические взаимодействия, совпадения программы недостаточно. Эти версии принципа нужно разделить до сравнения результатов.

IIT 4.0 представляет важный пример теории, явно связывающей предложенные свойства опыта с физическими постулатами. Наша заготовка не является вычислением IIT и не получает её основания по сходству слов «причина» или «единство». Сравнение должно касаться конкретных требований и расходящихся следствий. [19]

Особенно опасен пример простого регулятора. Он может иметь обратную связь и несколько состояний. Поэтому необходимо проверить, действительно ли наши три требования исключают слишком простые конструкции, и по какому основанию. Если исключение приходится вводить только потому, что результат нам не нравится, принцип недоопределён. Такой контрпример помогает улучшить определение исследуемого класса.

Три этапа проверки сильной гипотезы

Первый этап устанавливает организацию. Нужны несколько различаемых состояний, определённые правила их взаимного изменения и участие альтернатив в дальнейшей обработке. Для каждой предполагаемой связи

задаётся вмешательство. Наличие корреляции, общего источника или красивой петли на схеме недостаточно. Результат этого этапа может быть положительным независимо от принятия гипотезы об опыте.

Второй этап связывает организацию с качественными отношениями. Пусть теория предполагает совместную принадлежность содержания источнику. Тогда она должна предсказать, как вмешательство в оценку источника изменит различение неоднозначных содержаний и какие отношения останутся прежними. Нужно сопоставить это с альтернативой, меняющей только критерий ответа при сохранении переживания. Обычная правильность классификации такой вопрос не решает.

Третий этап касается наличия минимального опыта и переноса между реализациями. Здесь сильный принцип предлагает тождество или достаточность определённого класса процессов. Аргумент поддерживается системой независимых соответствий и предсказаний, анализом альтернатив и последовательным применением границы. Организационный тест из первого этапа не становится независимым доказательством достаточности только потому, что повторяется много раз.

Такой порядок позволяет продвигаться без обещания окончательного решения одной программой. Можно строго показать нужную причинную взаимность, обнаружить неверное структурное следствие или установить, что предложенное понятие слишком широко. Каждый результат меняет исследовательскую программу определённым способом.

Вычислимый пример двух направлений

Для проверки логики построим ещё один авторский учебный контур. Пусть p и q находятся между нулём и единицей. Они обозначают оценку неоднозначного признака и оценку источника. Введём функцию s(z) = 1 / (1 + exp(−z)), переводящую любое вещественное число в этот интервал. Обе новые оценки вычисляются из прежних одновременно: p_нов = s(x + b × (2q − 1)), q_нов = s(h + d × (2p − 1)). Здесь x и h являются доступными внешними свидетельствами, b и d силами двух направлений.

При x = h = 0, b = d = 1 и исходных p = q = 0,5 оценки остаются 0,5. Теперь при неизменном входе принудительно поставим q = 0,8 перед обновлением. Следующий p станет примерно 0,645656. Отдельное вмешательство p = 0,8 меняет следующий q на ту же величину. Если b = 0, первое направление исчезает; если d = 0, исчезает второе. Так условие взаимности получает конкретную матрицу проверок.

Вариант выбора проверки можно задать отдельно: запрос включается, когда 1 − |2p − 1| превышает выбранную стоимость. Это условная политика, используемая ради прозрачности. Она показывает, как изменение представления может менять следующий шаг. Словесное название запроса не добавляет к нему переживания.

Команда python teaching_lab.py --chapter 26 воспроизводит базовое состояние и направленные вмешательства. Пример проверяет причинную организацию определённых чисел. Он ещё не реализует всё сильное требование о качественной структуре и минимальном опыте. Для этого нужно определить содержание различений, вывести соответствующие феноменальные следствия и обосновать физическую реализацию.

Сильный альтернативный контроллер должен получать те же наблюдения и сопоставимые ресурсы. Если он сохраняет всю матрицу ответов, выбранный тест пока не выделяет нашу организацию. Это содержательная возможность результата. Вывод об уникальности нельзя получать победой только над заранее записанной неподвижной последовательностью.

Как проверить взаимность не определяя опыт по результату теста

Пусть p оценивает неоднозначный признак, q его источник. В учебном контуре оценка источника влияет на следующий p, а оценка признака на следующий q. При неизменном внешнем входе принудительно изменим q. Если p меняется по заданному правилу, проверено одно направление. Затем независимо изменим p и проверим q. Одного из этих опытов недостаточно для утверждения о двух направлениях.

Даже успех обоих вмешательств пока устанавливает причинную организацию. Принцип соответствия дополнительно предполагает её отношение к опыту. Для качественного вывода нужно заранее назвать, какие отношения между содержаниями изменятся и какие сохранятся. Для переноса на иной носитель нужно уточнить, какое устройство действительно должно существовать в физическом исполнении. Поэтому следующая практика начинает с аудита готового принципа, а самостоятельное предложение становится вторым уровнем работы.

Что теперь установлено

Мы разделили три результата: наличие взаимной связи, предсказание структуры различий и сильное утверждение о наличии опыта. Они могут входить в один проект, но требуют разных оснований.

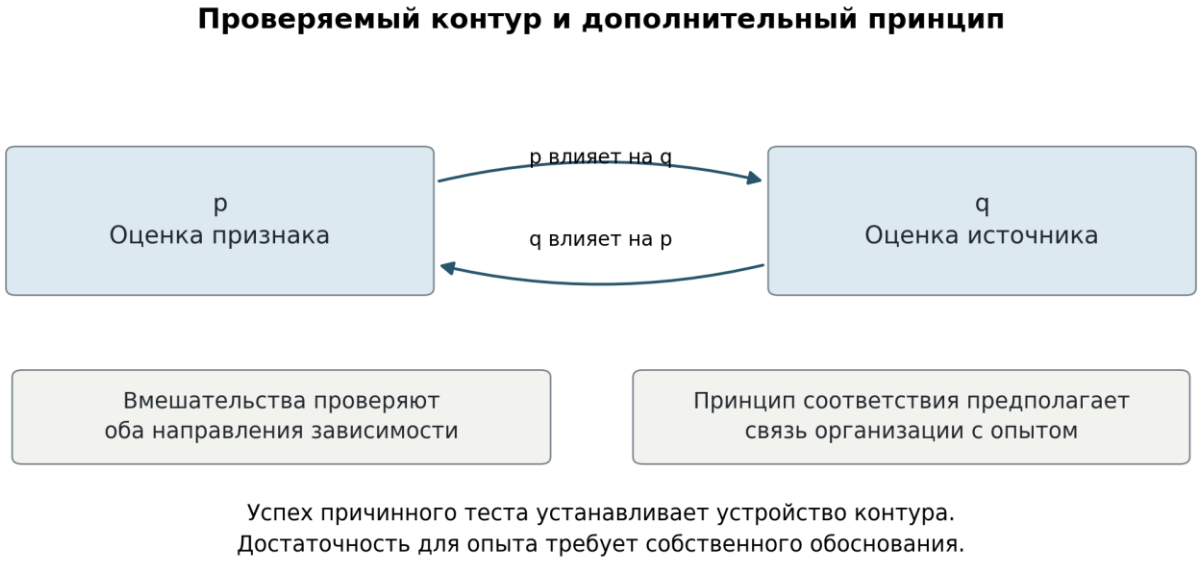

Рисунок 59. От причинной организации к гипотезе об опыте. Учебный контур p и q реализован в teaching_lab.py --chapter 26 и проверяется двумя направленными вмешательствами с односторонними контролями. Связь такой организации с опытом является отдельной авторской гипотезой. Нарушение предложенного достаточного условия не доказывает отсутствие любого опыта.

Практика 26 Защитите один принцип

Зачем выполнять: Проверить обязательства принципа соответствия и отличить достаточность от необходимости.

Что понадобится: Причинное вмешательство, достаточность и необходимость; раздельная история из главы 25. Сначала прочитайте полный разбор контура и принципа, затем переходите к собственному варианту.

Ориентир по времени: 40–60 минут. Можно делать с перерывами; это не норматив скорости.

1. Перепишите принцип главы в трёх предложениях: организация, предполагаемая связь с опытом, структурное следствие для качеств.
2. Для каждого требования укажите наблюдение или вмешательство, позволяющее проверить выполнение организационного условия без слов «сознание» и «переживание».
3. Рассмотрите три контроля: независимые записи выходов, односторонний контроллер и простой регулятор. Укажите, какие условия принципа они выполняют, а какие неизвестны.
4. Для одного вмешательства отдельно запишите функциональный прогноз и ещё требующее обоснования феноменальное следствие. Не выводите отсутствие любого опыта из нарушения достаточного условия.

Что сохранить: Карточка принципа, три контрпримера и два раздельных уровня предсказаний.

Подсказка: Фраза «этот процесс достаточен» разрешает возможность других достаточных процессов. Отсутствие выбранного процесса поэтому не является логическим отрицанием всякого опыта.

Исследовательское продолжение: Предложите собственную уточнённую версию. Выведите инвариант, изменение существенного отношения и условие пересмотра. Сопоставьте её с ближайшими научными подходами до заявления о самостоятельности идеи.

Разбор и ответы — приложение E, глава 26, с. 306

Самопроверка

1. Чем сильный принцип этой главы отличается от хорошего предсказания отчётов?
2. Почему переименование координат не должно само по себе менять предсказанное качество?
3. Можно ли считать предлагаемую гипотезу доказанной после сборки системы с тремя указанными свойствами?

Принцип получил отдельные обязательства. Следующая глава превратит различия между механизмами в проект, который может поддержать одну версию и ограничить другую.

Глава 27

Конкуренты и рискованные предсказания

Рабочая гипотеза полезна, когда возможный результат способен изменить её содержание. Теперь подготовим различающий проект. Первый этап касается истории источников и прогнозов оценок; следующий взаимных причинных отношений и качественной организации. Не будем требовать от одного сравнения решения всех вопросов. Для каждого этапа определим собственный показатель, сильного соперника и условие пересмотра.

Различие должно быть опасным для идеи

Предсказание полезно, когда возможный результат способен изменить наше отношение к модели. Утверждение «память иногда влияет на решения» легко согласовать почти с любым наблюдением. Утверждение «при возвращении источника А определённая история изменит траекторию оценки сильнее, чем история другого источника, при одинаковом текущем сигнале» создаёт риск. Можно получить другую траекторию, отсутствие переноса или эффект, полностью объясняемый общим обучением.

В этой главе мы подготовим экспериментальный проект. Он пока не проведён. Его числовые примеры относятся к учебной модели, а будущие ответы людей остаются неизвестными. Такая постановка позволяет сделать прогноз до того, как появится соблазн подстроить объяснение под результат.

Начните с совпадения моделей

М2 из главы 25 хранит отдельные истории А и Б. Общий идентификационный трекер М1И способен делать то же самое. Если он получает те же признаки и использует те же правила, две модели вычислительно совпадают. Никакая величина выборки не различит одинаковые предсказания. Нужно либо признать, что предложен общий механизм, либо определить дополнительное содержательное различие.

Это не поражение проекта. Общий механизм отслеживания причин может участвовать в телесном самоотнесении. Ошибкой было бы объявить его специфически телесным только потому, что в задаче изображена рука. Первый этап исследования вправе отвечать на более узкий вопрос: требуется ли раздельная история источников для объяснения измеренных оценок?

Для сильной гипотезы из главы 26 понадобится следующий этап. Там значение имеет участие истории в общей взаимной динамике восприятия и выбора проверки. Простое хранение источников и их включение в такой контур — разные организации. Их нужно сравнивать по заранее выведенным последствиям, сохраняя различие между функциональным результатом и феноменальным выводом.

Спроектируйте четыре условия

Предлагаемый человеческий эксперимент использует обычные зрительные и двигательные задания с виртуальным указателем или изображением руки. Исследование потребует профессионально составленного протокола и добровольного участия; здесь разрабатывается его логика. Участник сначала взаимодействует с источником А при устойчивой либо изменчивой связи сигналов. Затем предъявляется одинаковая тестовая последовательность от прежнего либо нового источника Б.

Получаются четыре сочетания: устойчивая история и прежний источник; устойчивая история и новый источник; изменчивая история и прежний источник; изменчивая история и новый источник. Общая длительность, число событий и характеристики тестовой последовательности сопоставляются. Источник различается по заранее выбранному наблюдаемому признаку, а не по ответу участника о принадлежности.

Важно не определять устойчивость только числом совпадений. Две последовательности могут иметь одинаковое количество подтверждений и противоречий, но разный порядок. Модель должна получить реальные последовательности целиком и выдать временной прогноз. Если исходные режимы отличаются по нескольким характеристикам, каждая из них включается в перечень возможных объяснений.

Измерьте несколько разных результатов

Основным результатом первого этапа будет заранее выбранная динамика оценок принадлежности. Способ получения оценок должен минимально различаться между условиями. Дополнительно можно измерять различение текущей задержки, распознавание источника, оценку авторства действия и уверенность. Эти показатели помогут понять, какой процесс изменился, но не должны бесконтрольно превращаться в новые основные исходы.

Не существует внешнего правильного числа для переживаемой принадлежности данного человека. Поэтому качество сенсорного различения нельзя использовать как автоматическую проверку истинности отчёта. Оно отвечает на другой вопрос: насколько доступно участнику физическое

различие, которое модель считает входом. Распознавание источника также помогает проверить, получали ли участники информацию, необходимую механизму M2.

Формулировки вопросов и порядок измерений нужно сохранить. Если сначала спросить об управлении, а затем о принадлежности, первая оценка может изменить способ второй. Такой эффект является возможностью дизайна, которую следует проверить пилотной работой, а не предполагать отсутствующей. Пилот помогает улучшить процедуру; подтверждающая проверка затем использует отдельно зафиксированный план.

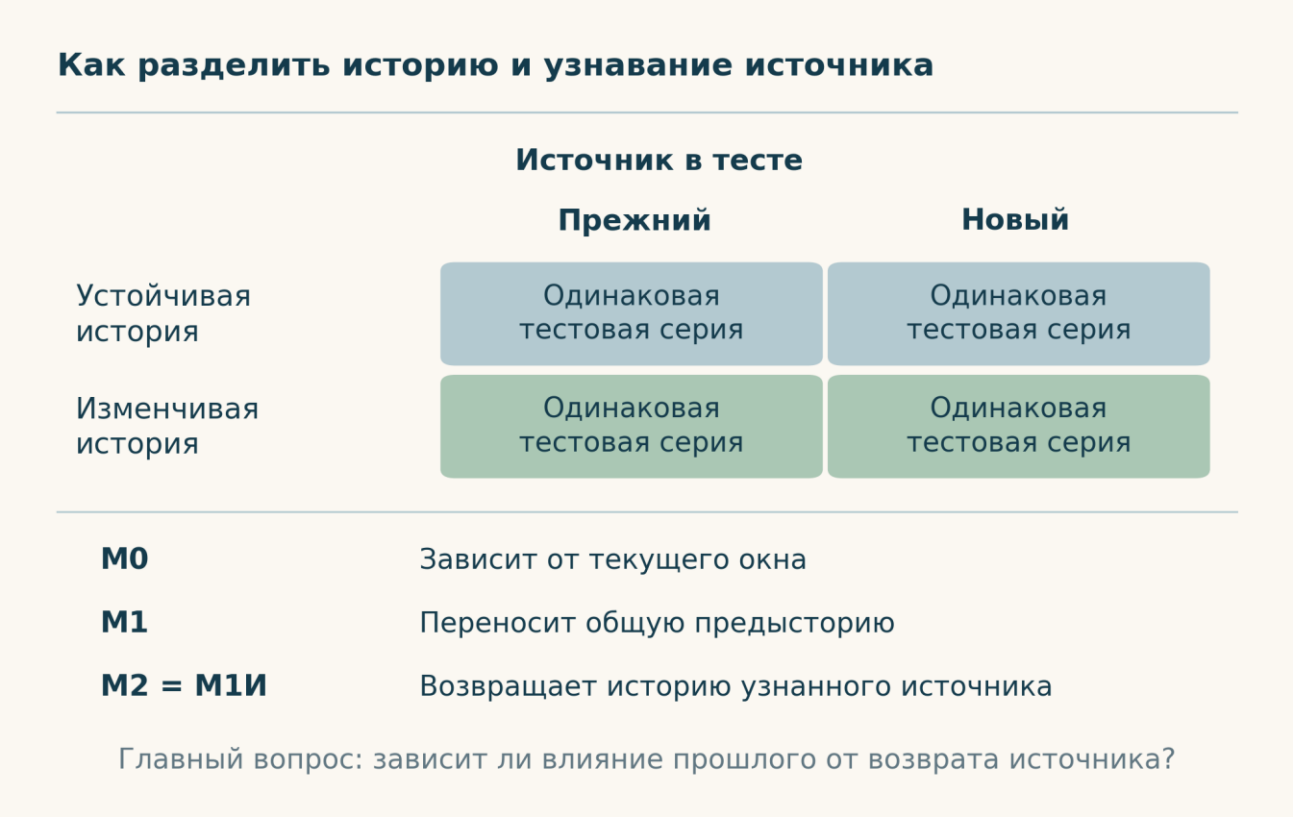

Рисунок 60. Четыре условия различающего проекта. Проект ещё не проведён. Два фактора — характер предыстории и возвращение прежнего либо появление нового источника — сочетаются в четырёх условиях при сопоставимой тестовой серии. Внизу указаны качественные различия правил, а не ожидаемые человеческие средние. В выбранной реализации M2 и общий трекер M1И совпадают; этот дизайн сам по себе их не различает.

Получите конкретные прогнозы

M0 предсказывает одинаковое внутреннее состояние при одинаковом текущем окне. M1 допускает влияние предыстории, но переносит её без разделения источников. M2 предсказывает возвращение соответствующей отдельной истории. Поэтому главный различающий эффект — зависимость влияния

прошлого от того, вернулся ли прежний источник. Направление и величина должны вычисляться из полностью определённых правил.

Учебная последовательность А положительное, Б отрицательное, А положительное даёт M0 последнее состояние 0,75, M1 приблизительно 0,7105, M2 — 0,875. Эти значения не являются ожидаемыми человеческими средними без дополнительных допущений. Они показывают, что различие между механизмами можно получить до наблюдения результатов. Модель ответа и индивидуальные параметры потребуют отдельной оценки.

Общий трекер M1И на этой последовательности совпадает с M2. Поэтому выигрыш над M0 и M1 поддержит пользу учёта истории источника в рассматриваемом объяснении, но ещё не специальную телесную природу механизма. Для такого вывода нужен перенос между телесной задачей и сопоставимой задачей отслеживания внешнего объекта.

Проверьте взаимное определение

Для следующего этапа создадим две программные реализации. Первая использует оценку источника для выбора дополнительного наблюдения; результат проверки меняет восприятие, историю и следующий выбор. Вторая воспроизводит привычные выходы компонентов, но разрывает одну из этих обратных связей. При обычной последовательности их поведение может выглядеть сходным.

Вмешательство должно сделать различие заметным. Например, изменим надёжность одного признака после формирования истории источника. Взаимная система должна перестроить выбор проверки в соответствии со своей изменённой неопределённостью. Воспроизведение записанного потока этого не обеспечит. Необходимо также сравнение с полноценным альтернативным контроллером, чтобы победа над записью не стала единственным достижением.

Такой опыт проверяет наличие определённой причинной организации. Структурная часть гипотезы главы 26 дополнительно требует предсказать, какие отношения между переживаемыми содержаниями изменятся при выбранном вмешательстве. Проверка потребует независимых данных о переживаниях и аргумента переноса. Из одной достаточности организации не следует, что её разрыв исключает любой опыт. Одинаковые человеческие тексты на выходе программы не дают отдельного феноменального основания.

Сравнивайте на новых данных

Если параметры подобраны по всей последовательности, успешное её описание ещё не является успешным прогнозом. Разделите данные: по одной части оцените параметры, по другой сравните предсказания. При наличии нескольких участников или биологических образцов разделение должно учитывать их структуру. Случайное перемешивание тесно связанных событий способно сделать проверку слишком лёгкой.

Вычислительное исследование также требует проверки, может ли выбранная процедура различить модели, когда породившая данные модель известна. Wilson и Collins рассматривают восстановление параметров и моделей как важные проверки вычислительного анализа. В нашем проекте такие проверки будут проводиться на собственных искусственных последовательностях. [43]

Сильное различие желательно искать там, где соперники расходятся, а не там, где всем легко. Однако выбирать условия после просмотра основных человеческих результатов нельзя выдавать за заранее предсказанный успех. Такой выбор может стать полезным исследовательским открытием и основанием следующего самостоятельного эксперимента.

Запишите условия пересмотра

Если M1 предсказывает новые оценки столь же хорошо, отдельная история источника пока не оправдана. Если M2 выигрывает, но M1И полностью объясняет преимущество, специальная телесная версия избыточна. Если преимущество возникает только после свободной перенастройки преобразования ответа в каждом условии, неясно, объясняет ли его механизм или гибкость измерительной части.

Для сильной гипотезы отдельной проблемой станет отсутствие тех феноменальных изменений, которые её структурная часть предсказывала при конкретном вмешательстве. Нужно выяснять, удалось ли вмешательство и не работал ли другой путь. Но нельзя бесконечно объявлять любой неудобный результат недостаточно настоящим нарушением. Критерий реализации вмешательства фиксируется до проверки.

Результат должен сузить множество объяснений

Наблюдение	Что поддержано	Следующий тест
M1 и M2 одинаково точны	Отдельная история пока не обоснована	Создать конфликт историй источников
M2 лучше M1; M1И равна M2	Общая память идентичности	Сопоставить тело и внешний объект
Выигрыш лишь при свободной шкале	Неясен вклад механизма	Закрепить шкалу; новые данные

Возможные исходы будущего проекта

Рисунок 61. От возможного исхода к различающему тесту. Таблица задаёт три возможных исхода будущего сравнения и разные решения после них. Совпадение M2 с общим трекером M1И оставляет вопрос о специальном самоотнесении открытым, даже если раздельная память полезна.

Почему больше данных не различит одинаковые правила

M2 и общий трекер хранят отдельные истории и в минимальной реализации обновляют их одинаково. Если дать им одинаковые наблюдения и начальные состояния, они вычислят те же прогнозы на любой длине такого опыта. Увеличение выборки уменьшит случайную неточность оценок данных, но не создаст отсутствующего различия между моделями. Поэтому первый проект может проверить пользу отдельной истории относительно общей памяти, а специальное телесное объяснение потребует дополнительного механизма или другого сравнения.

Что теперь установлено

Различающий результат определяется предсказаниями, а не названием модели. Теперь выпишем условия и выводы до появления основных данных.

Практика 27 Составьте различающий проект

Зачем выполнять: Подготовить четыре условия первого этапа и отдельно обозначить следующую причинную проверку.

Что понадобится: M0, M1, M2 и общий трекер из глав 24–25; принцип соответствия из главы 26. Человеческий эксперимент здесь проектируется, а не считается уже выполненным.

Ориентир по времени: 35–50 минут. Можно делать с перерывами; это не норматив скорости.

1. Составьте четыре сочетания: устойчивая или изменчивая история и прежний или новый источник в тесте. Укажите доступные признаки и одинаковую тестовую последовательность.
2. Для каждого условия выпишите, какую историю используют M0, M1 и M2. Общий трекер сохраняйте самостоятельным контролем, даже если он совпадает с M2.
3. Назовите основной результат и дополнительные показатели: принадлежность, распознавание источника, авторство и сенсорное различение не объединяйте одним числом.
4. Задайте одно направленное вмешательство для следующего этапа и контроль его побочного действия. Сохраните результат, при котором пересмотрите предпочитаемую версию.

Что сохранить: Проект с четырьмя условиями, различимыми прогнозами и явной границей первого вывода.

Подсказка: Числа 0,75, 0,710526 и 0,875 из расчёта не становятся человеческими средними без модели измерения и оценки параметров.

Исследовательское продолжение: Подготовьте две реализации с одинаковыми обычными выходами, но разными направленными влияниями. Сравните также полноценного альтернативного контроллера и проверьте перенос на новые сценарии.

Разбор и ответы — приложение E, глава 27, с. 306

Самопроверка

1. Можно ли увеличить выборку, чтобы различить вычислительно одинаковые M2 и M1И?
2. Почему победа над воспроизведением записанных выходов является ограниченным результатом?
3. Что делать с интересным эффектом, замеченным после анализа основных данных?

Проект задаёт, что нужно проверить. Теперь возьмём действующего агента, у которого видны входы, память, прогноз и действие, и воспроизведём функциональные проверки по коду.

Часть 8

От слов к практике

Глава 28

Агент в среде с проверяемой задачей

Мы переходим от проектирования к действующей программе. Она будет оценивать цель, учиться собственному управлению и выбирать дополнительное наблюдение с учётом его цены. Все эти функции уже подготовлены предыдущими главами. Первый проход можно выполнить без изменения кода: понять задачу, записать прогноз и воспроизвести готовый запуск. Собственные новые условия появляются после проверки исходного результата.

Первый агент на обычном компьютере

Откройте каталог практикума и найдите файл agent_lab.py. Перед нами небольшой агент, который выбирает действие в условиях неполной информации. Вся его работа доступна для просмотра: какие сведения поступили, что сохранилось в памяти, как изменилось предположение и какой прогноз был сделан перед действием. Это удобная отправная точка для проектирования более сложных систем. Мы можем менять один механизм и наблюдать последствия, не полагаясь на правдоподобие словесного ответа программы.

Для начала не требуется понимать каждую строку Python. Сначала разберите задачу и предскажите результат. Затем запустите готовую программу, сопоставьте ожидание с числами и только после этого переходите к реализации. Код здесь служит точной записью модели. Его можно читать постепенно: входы, состояние, обновление, действие, обратная связь. Эти пять вопросов позволяют разобраться даже в незнакомой программе.

В установочном приложении приведены команды для Windows, macOS и Linux. Нужен Python версии 3.10 или новее. Программа использует стандартную библиотеку, выполняется локально и не требует API. В каталоге с файлом выполните команду ниже. В Windows замените python3 на py -3, если ваша установка использует этот запускатель.

```
python3 agent_lab.py --steps 1000 --runs 30 --seed 1729 --json my_results.json
```

Параметр steps задаёт число шагов одного сценария, runs — число повторений, seed — начальное значение генератора случайности. Последний параметр задаёт файл результата. При одинаковых параметрах программа повторяет тот же вычислительный эксперимент. Это помогает проверить изменения кода, но для оценки устойчивости нужны и другие сценарии.

Как устроена среда

На каждом шаге существует цель слева или справа. Агент её непосредственно не видит. Между соседними шагами сторона сохраняется с вероятностью 0,95. Поэтому недавняя информация обычно полезна, хотя иногда уже устаревает. Обычное наблюдение правильно указывает сторону в 75 процентах случаев и отсутствует на четверти шагов. За небольшую цену можно получить дополнительное наблюдение с точностью 90 процентов.

Агент отправляет команду влево или вправо. В первой половине сценария она обычно вызывает эффект на соответствующей стороне, а во второй половине связь инвертируется. Отдельная моторная ошибка с вероятностью 0,05 также меняет эффект. Получается два различных источника неопределённости: где находится цель и что произойдёт после команды. Надёжное восприятие цели не спасёт, если модель управления устарела.

Смена правила происходит с шага 501 при тысяче шагов. Это знает экспериментатор. Агент не получает ни флага переключения, ни расписания, ни истинной цели, ни seed. После команды ему доступен видимый эффект. Сведения об успешности подсчитывает среда и возвращает исследователю после завершения сценария.

Почему не передать агенту успех сразу? В данной двоичной задаче сочетание известного эффекта и успеха раскрыло бы истинную цель. Тогда память и шумное наблюдение проверялись бы уже в другой информационной ситуации. Запрет относится к конкретной модели исследования. В иной задаче обратная связь о правильности может быть законной частью интерфейса, как в работе о пересмотре стратегии в главе 30.

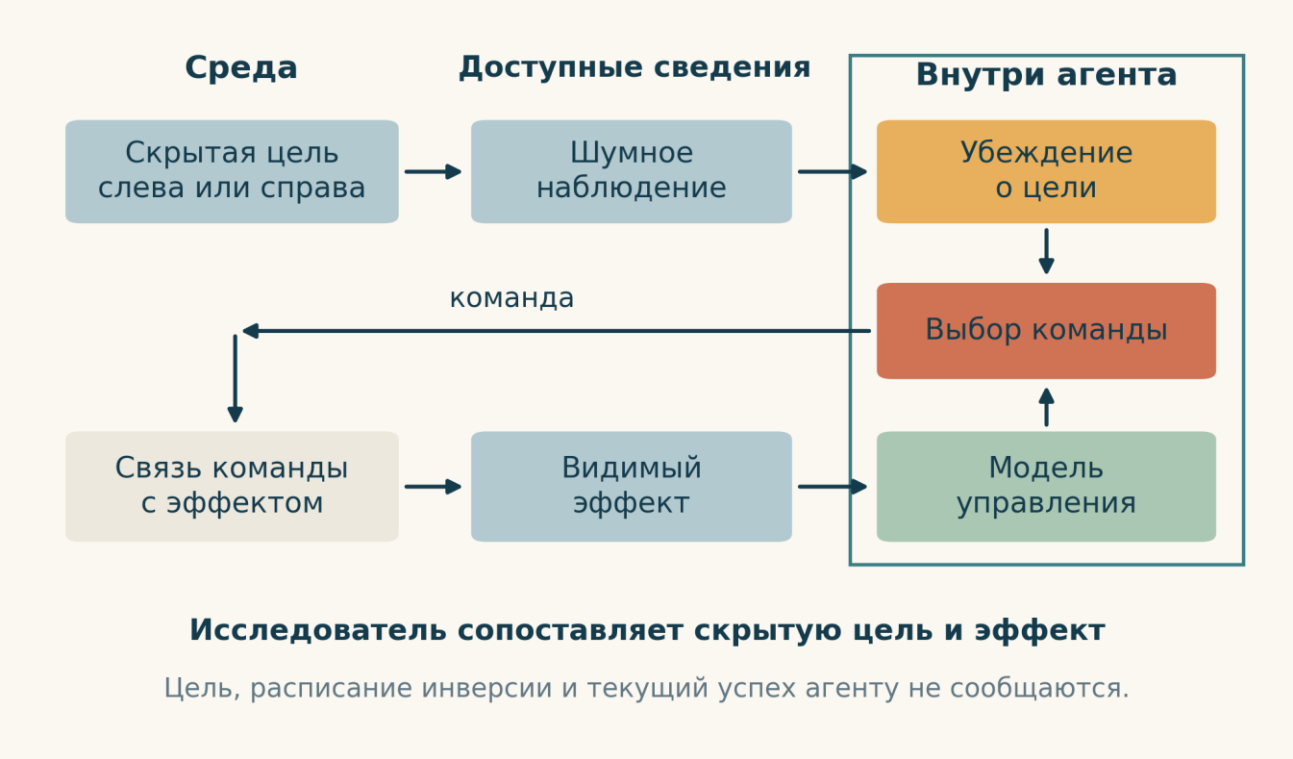

Рисунок 62. Два контура одного агента. Учебная схема интерфейса первой работы. Наблюдение обновляет убеждение о цели, а видимый эффект команды — модель управления. Истинная цель, момент смены правила и текущий успех не входят в доступные агенту сведения. Дополнительный платный запрос даёт ещё одно наблюдение цели и для ясности отдельно не показан.

Память в форме вероятностного убеждения

Состояние target_right_probability хранит вероятность правой цели по текущей модели. Число 0,8 означает, что при данных предпосылках и наблюдениях правая сторона считается более вероятной. Это утверждение модели, которое может оказаться ошибочным. Истинное положение не входит в вычисление.

Перед новым наблюдением агент учитывает возможность движения цели. Если вероятность правой стороны была 0,8, а цель сохраняется с вероятностью 0,95, следующий прогноз получается из двух возможностей: цель была справа и осталась там либо была слева и переместилась. Их вероятности складываются. В данном примере новый прогноз равен 0,77. Затем наблюдение уточняет его по правилу Байеса.

```
old_belief = 0.8
stay = 0.95
prior = stay * old_belief + (1 - stay) * (1 - old_belief)
print(prior)  # 0.77
```

Память здесь не представляет подробную автобиографию. Прошлое сжато до одного числа, достаточность которого определяется выбранной вероятностной

моделью. Такой способ хранения позволяет ясно показать причинный вклад истории. Если наблюдение пропущено, агент всё ещё имеет прогноз; если история отключена, начальное убеждение о цели возвращается к 0,5.

Вторая переменная, direct_mapping_probability, оценивает вероятность прямой связи команды с эффектом. После действия агент сравнивает команду с видимым результатом. Несовпадение может быть следствием моторной ошибки или изменения управления. Байесовское обновление меняет их относительную правдоподобность. Модель допускает небольшую вероятность смены связи на каждом шаге и поэтому способна пересматривать ранее надёжное предположение.

Зачем оценивать успех до действия

Из двух убеждений можно вычислить вероятность успеха для каждой команды. Агент объединяет варианты положения цели и возможного эффекта, выбирает более перспективную команду и записывает прогноз до действия. Если записывать уверенность после получения правильного ответа, показатель перестанет отражать предшествующую неопределённость.

Дополнительное наблюдение покупается, когда ожидаемое улучшение текущего решения превышает цену. Программа рассматривает оба возможных ответа датчика, пересчитывает убеждение для каждого и усредняет лучший ожидаемый результат. Это расчёт ближайшей пользы. Будущая ценность информации на следующих шагах отдельно не планируется, поэтому мы не называем такую политику оптимальной для всех возможных горизонтов.

Стоимость нужна, чтобы различать точность и полезность. Получать больше сведений иногда выгодно, но это зависит от их цены и того, помогут ли они изменить действие. Если главная неопределённость связана с управлением, очередное уточнение цели может мало улучшить решение. Состояние мира и состояние собственной модели действия следует учитывать совместно.

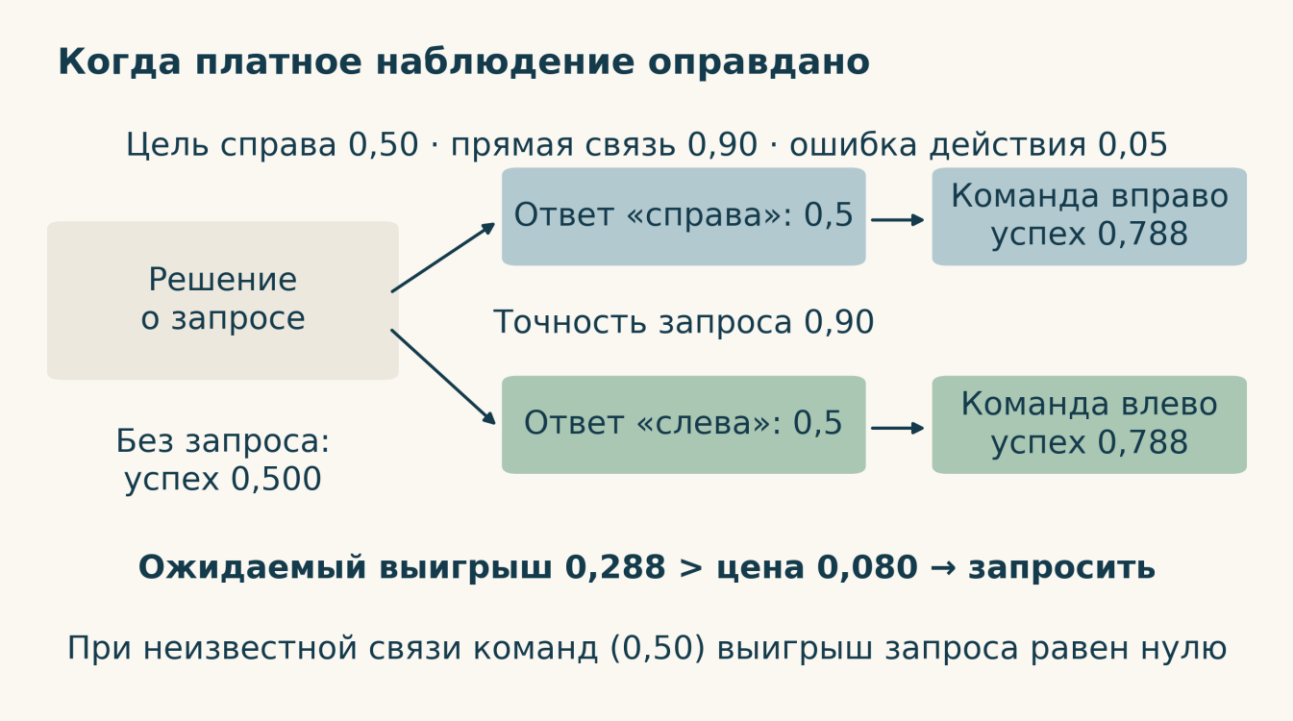

Рисунок 63. Условная ценность дополнительного наблюдения. Расчёт по правилам agent_lab.py для выбранного учебного состояния при первом пропущенном наблюдении; это не средние результаты запуска. Оба возможных ответа дают ожидаемый успех 0,788, поэтому прирост 0,288 превышает цену 0,08. Если вероятность прямой связи команд равна 0,5, знание цели само по себе не улучшает выбор действия.

Четыре режима и честное сравнение

Полный режим использует все описанные механизмы. В no_world_memory отключается только перенос убеждения о цели между шагами. Обучение модели действия сохраняется. В fixed_action_model вероятность прямой связи остаётся равной исходным 0,90, без обновления. В no_extra_query запрещена покупка дополнительного наблюдения. Такое точное описание важно: название режима само по себе не объясняет, какие вычисления были изменены.

Все режимы получают одинаковые потенциальные наблюдения и моторные ошибки для соответствующего сценария. Случайность разделена на независимые потоки, а значения заранее подготовлены. Покупка наблюдения открывает дополнительное значение, не сдвигая последующие случайные события. Различные команды могут вызвать различные эффекты, но моторное искажение на данном шаге общее.

В приложенном запуске выполнено тридцать сценариев по тысяче шагов для каждого режима. Это реальные результаты программы, а не данные о людях или животных. Цена запроса равна 0,08; успех приносит единицу, неуспех — ноль. Полезность получается после вычитания расходов на запросы.

Режим	Успех в процентах	Полезность на шаг	Brier
Полный	87,94	0,8619	0,1054
Без памяти мира	85,36	0,7767	0,1202
Фиксированная модель действия	49,97	0,4841	0,3444
Без дополнительного запроса	77,91	0,7791	0,1589

Brier — среднее квадрата разности прогноза успеха и фактического исхода, записанного как ноль или единица. Меньшее значение означает более точный вероятностный прогноз. Оно зависит и от калибровки, и от информативности прогнозов. Называть его чистой мерой калибровки, самосознания или качества переживания было бы ошибкой.

Общая успешность фиксированной модели около половины скрывает резкое различие двух периодов: 87,80 процента до инверсии и 12,13 процента после неё. У полного режима соответствующие значения 88,06 и 87,83 процента. Для измерения скорости переобучения нужны короткие временные окна: средняя всей второй половины сглаживает переходный процесс.

Есть и менее очевидное сравнение. Версия без памяти действует точнее версии без запросов, но покупает значительно больше информации. В результате её полезность немного ниже. Поэтому фраза «этот агент лучше» нуждается в указании критерия. Точность ответа, стоимость и надёжность прогноза могут изменяться по-разному.

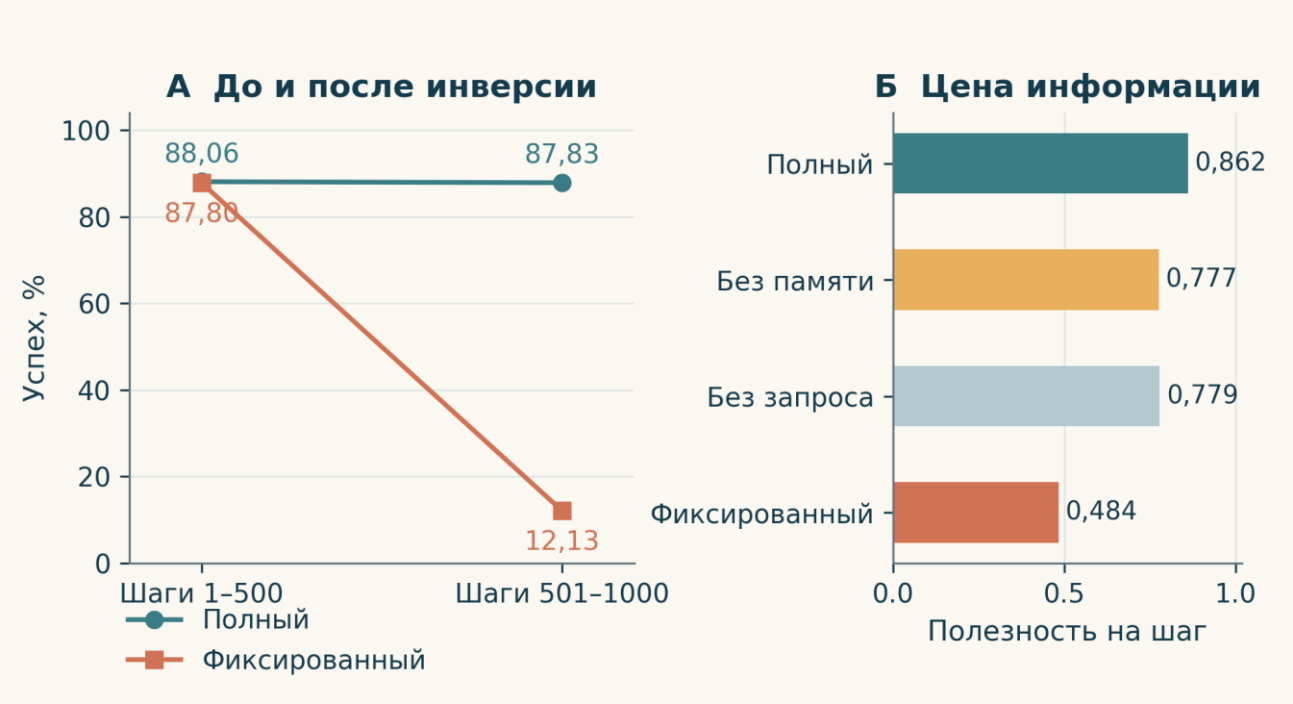

Рисунок 64. Успех и полезность отвечают на разные вопросы. Результаты программы agent_lab.py: 30 общих случайных сценариев по 1000 шагов, seed 1729. Слева показаны средние двух периодов, а не скорость адаптации внутри периода. Справа из успеха вычтена стоимость дополнительных запросов. Данные: lab01_results.json.

Как прочитать один шаг агента

Сначала агент учитывает возможное перемещение цели в сохранённом убеждении. Затем получает доступный шумный сигнал и обновляет оценку. Из неопределённости цели и управления он вычисляет полезность команд и возможного запроса. Прогноз успеха записывается до действия. После команды видимый эффект обновляет модель управления. Истинную цель, расписание инверсии и текущий успех оценщик не открывает агенту. В этой двоичной задаче успех вместе с эффектом раскрыл бы цель и изменил исследуемый интерфейс.

Что теперь установлено

Мы знаем порядок информации и можем отличить прогноз от позднего знания. Числа режима теперь нужно читать вместе с ценой запросов и с периодом до или после инверсии.

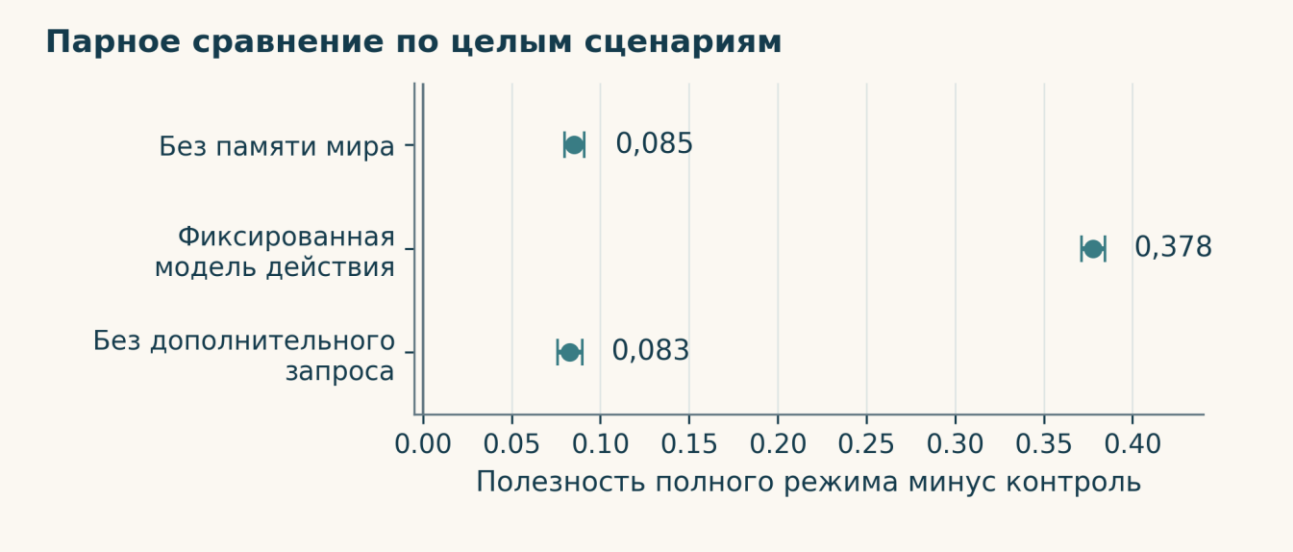

Рисунок 65. Неопределённость различий между режимами. Результат анализа приложенных данных. Точки — средние парные разности полезности; отрезки — приближённые 95-процентные процентильные интервалы из 4000 повторных выборок 30 целых сценариев. Единицей повторения служит сценарий. Интервалы относятся к отдельным сравнениям и не дают совместной гарантии 95 процентов. Данные: paired_analysis.json.

Практика 28 Предскажите результат изменения среды

Зачем выполнять: Воспроизвести базовую работу и объяснить один результат через устройство механизма.

Что понадобится: Вероятностные модели мира и действия, Brier score и единица независимого повторения. Практикум использует Python 3.10 или новее и стандартную библиотеку.

Ориентир по времени: 30–45 минут. Можно делать с перерывами; это не норматив скорости.

1. Распакуйте практикум и выполните python run_practice.py --self-test. До основного запуска запишите ожидаемое поведение фиксированной и обучающейся модели после смены управления.
2. Выполните python run_practice.py --lab 1. Сохраните созданный каталог, параметры и контрольные суммы.
3. В таблице результатов отдельно сравните успех до и после переключения, полезность на шаг и качество прогноза. Не выбирайте лучший режим без указания показателя.
4. Проследите один эпизод по журналу: доступное наблюдение, прежнее состояние, прогноз, команда, видимый эффект, новое состояние.

Что сохранить: Сохранённый запуск и короткий вывод с механизмом, показателем и областью условий.

Подсказка: Режим без памяти мира всё ещё учится связи команды с эффектом. Выключена конкретная функция, а не вся способность хранить прошлое.

Исследовательское продолжение: Измените ровно одно условие, заранее сформулировав прогноз: цену запроса, надёжность датчика или изменчивость цели. Затем повторите сравнение на новых сценариях.

Разбор и ответы — приложение Е, глава 28, с. 307

Самопроверка

1. Почему режим без памяти сохраняет обучение связи команды с эффектом?
2. Почему общая успешность фиксированного агента вводит в заблуждение без разделения периодов?
3. Что изменилось бы, если дать агенту правильный ответ до записи уверенности?

Первый агент показал совместную работу двух убеждений. Теперь отдельно проверим устройство представления, повторную обработку и передачу содержания, не смешивая их с успешностью всей системы.

Глава 29

Восприятие и общий доступ

Успешность агента сама по себе не объяснила, какие отношения сохраняет представление и почему определённая обработка помогает. Три работы этой главы разложат задачу: геометрия различий, смешивание шумных каналов и доставка содержания. Каждая имеет собственный контроль и собственный итог. После каждого опыта остановимся на том, что установлено, прежде чем переносить вывод к научной теории сознания.

Что именно должна представлять система

В предыдущей главе агент различал две стороны. Теперь нас интересует устройство представления: какие состояния система считает похожими, как изменяет их при повторной обработке и каким модулям они становятся доступны. Эти вопросы относятся к разным свойствам. Удобная геометрия не гарантирует доступности для планирования, а широкая передача информации не объясняет, почему один оттенок воспринимается похожим на другой.

Начнём с пяти точек на плоскости. Каждая точка будет представлять некоторый объект, а расстояние между точками — выбранную модель различия. Числа здесь заданы конструктором; они не являются измерениями человеческих ощущений. Их преимущество в прозрачности: мы можем точно указать, какие отношения сохраняются после преобразования и какие меняются. После этого станет проще оценивать более сложные модели качественной организации.

Выполните вторую лабораторную работу. Все дополнительные работы находятся в файле mechanisms_lab.py; рядом должен лежать agent_lab.py, который используется составным прототипом. Результат сохраняется в JSON, пригодный для просмотра обычным текстовым редактором.

```
python3 mechanisms_lab.py --lab 2 --json geometry.json
```

Программа поворачивает исходные точки на угол 0,73 радиана и переносит их на одинаковый вектор. Затем она отдельно растягивает одну координату вдвое. Сравниваются расстояния каждой пары до и после преобразования. Небольшое отличие в последних знаках при повороте связано с представлением вещественных чисел в компьютере. В базовом запуске максимальная разность составляет приблизительно 0,00000000000000044.

Почему инварианты важнее названий

Инвариант — свойство, сохраняющееся при оговорённом преобразовании. Если модель использует только попарные евклидовы расстояния, общий поворот и перенос не должны менять её предсказания о сходстве. Названия координат при этом могут стать совершенно другими. Утверждение «эта координата и есть краснота» не выдерживает такой проверки без дополнительного объяснения, почему выбран именно этот способ записи.

Растяжение одной оси обычно меняет отношения расстояний. Значит, оно существенно для модели, основанной на этих расстояниях. Однако вывод зависит от выбранного показателя. Модель, использующая лишь порядок объектов по первой координате, может сохранить некоторые ответы после положительного растяжения. Перед экспериментом следует назвать, что считается существенным: абсолютное расстояние, порядок соседства, направление, отношение расстояний или возможность различить категории.

Учебная геометрия не даёт готовой карты квалиа. Для такой карты потребуются данные о различении или сходстве, решение о масштабе измерения и проверка на новых сочетаниях. Важно также отличать расстояние между стимулами от отношения между сообщаемыми переживаниями. Один и тот же физический шаг по параметру не обязан означать одинаковый субъективный шаг. Формулировка исследовательского вопроса предшествует красивой картинке.

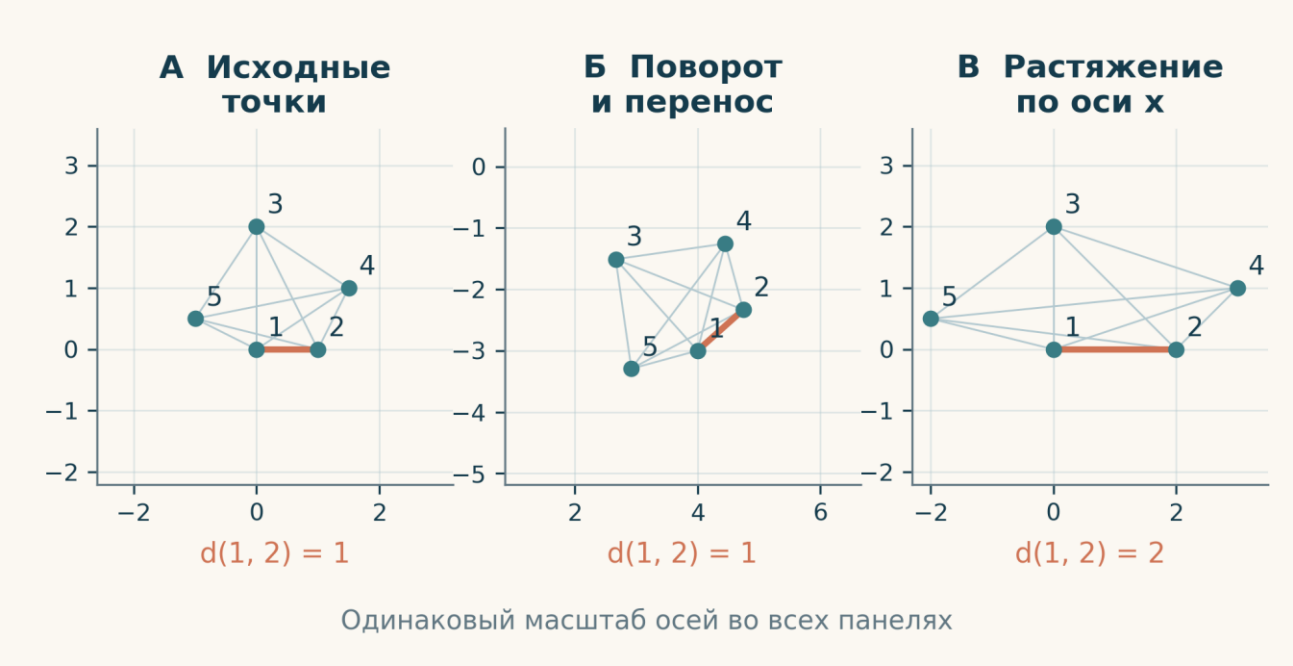

Рисунок 66. Что сохраняется при преобразовании представления. Учебная геометрия второй работы. Поворот на 0,73 радиана и перенос на (4; −3) сохраняют все десять попарных расстояний с машинной погрешностью; растяжение x меняет часть расстояний. Выделена одна пара, а расчёт проверяет все пары. Номера обозначают точки, не виды переживания. Исходные координаты заданы в mechanisms_lab.py.

Повторная обработка на простом механизме

В третьей работе три канала получают шумные значения одного скрытого сигнала. Истинное значение равно минус единице или единице, к каждому наблюдению добавляется шум. Одна версия сразу использует первый канал. Другая четыре раза смешивает значения каналов: половина нового состояния сохраняет прежнее значение данного канала, а по четверти поступает от двух соседей. Для результата читается первый канал.

Повторная обработка постепенно делает каналы похожими. Это можно понять без формулы. Если один канал случайно сильно завышен, а остальные ближе к цели, обмен уменьшает одиночное отклонение. Но общей ошибки всех каналов такой обмен не исправит. Он перераспределяет уже имеющуюся информацию; новый независимый датчик после каждой итерации не появляется.

```
python3 mechanisms_lab.py --lab 3 --json recurrence.json
```

Важный контроль — четыре последовательно развёрнутых слоя. Каждый выполняет ту же операцию, что одна итерация рекуррентной схемы. При одинаковом входе результат совпадает точно. В первом случае программа повторяет один шаг в цикле, во втором явно записывает четыре вложенных вызова. Смена формы записи не создаёт нового вычислительного свойства.

Вариант в 2000 учебных испытаниях	Средняя квадратичная ошибка
Один канал без смешивания	1,892736
Четыре рекуррентных шага	0,632419
Четыре развёрнутых слоя	0,632419
Прямое среднее трёх каналов	0,632465

Числа получены запускаемой программой с seed 1729. Они относятся к этой задаче. Прямое среднее — ещё одна полезная альтернатива, которая использует все входы без четырёх шагов смешивания. Его вычислительные затраты отличаются, поэтому это не контроль с равным числом операций. Практически равный результат показывает, что для выбранной задачи может существовать значительно более простой способ обработки.

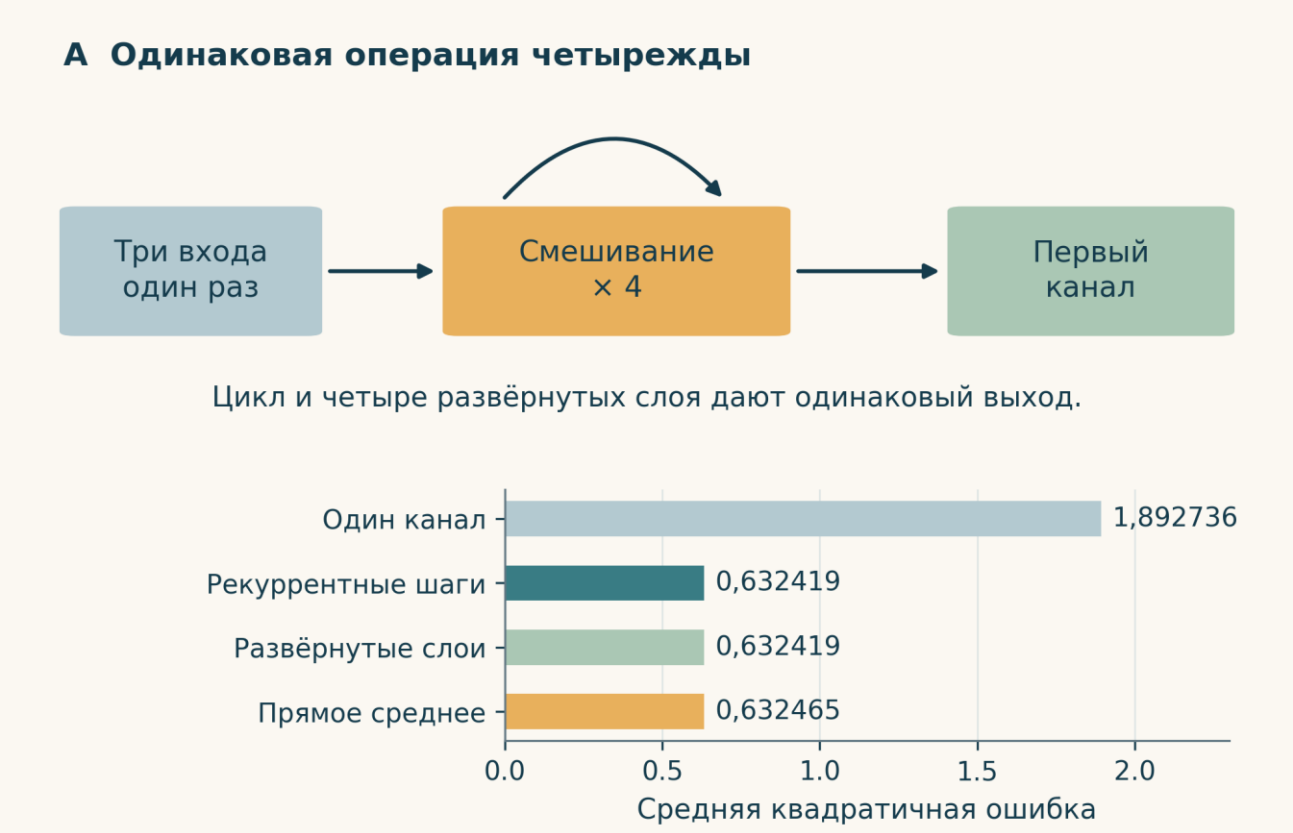

Рисунок 67. Рекуррентность и сопоставимые альтернативы. Учебная схема и результаты третьей работы: 2000 испытаний, seed 1729. При смешивании канал сохраняет половину своего значения и получает по четверти от двух других. Рекуррентная и развёрнутая реализации выполняют одинаковые операции и совпадают точно. Прямое среднее служит дополнительной альтернативой с другими вычислительными затратами. Данные: mechanisms_results.json, работа 3.

Как поставить более сильную проверку рекуррентности

Конечную последовательность вычислений часто можно представить развёрнутой схемой. Поэтому вопрос «есть ли в коде цикл» слишком груб для исследования. Нужно уточнить, какое свойство цикла предполагается существенным: повторное использование параметров, продолжительность обработки, поддержание состояния, взаимодействие с меняющимся входом или физическая взаимная причинность элементов.

Если гипотеза относится к вычислительной функции, равенство конечных реализаций становится важным ограничением. Если она относится к физической организации, текст программы ещё не описывает полностью, как процессы реализованы в устройстве. Эти различия были введены в теоретической части; здесь мы получаем возможность проверить хотя бы вычислительное утверждение до более сильного вывода.

Для расширения работы можно заранее выбрать задачу с различным временем предъявления или с поступлением новых наблюдений между итерациями. Тогда необходимо дать конкуренту те же сведения и сопоставимые ресурсы. Иначе выигрыш рекуррентной системы может объясняться более длинным наблюдением. Содержательный эксперимент возникает, когда мы точно понимаем, какое различие архитектур осталось после выравнивания условий.

Рабочее пространство как передача содержания

В четвёртой работе используются три сообщения: цвет, форма и отвлекающее значение. Цвет и форма принимают значения от нуля до двух. Правильный ответ определяется суммой цвета и формы по модулю три. Правило известно программе заранее. Нам не нужно обучать классификатор, потому что изучается передача необходимых сведений.

Рабочее пространство в этой модели вмещает два сообщения. Полная версия выбирает цвет и форму и передаёт их модулю ответа. При отсутствии передачи тот получает пустое представление и использует нулевые значения. При случайном отборе часть места занимает отвлекающее сообщение. Дополнительный конкурент имеет прямой доступ ко всем входам.

```
python3 mechanisms_lab.py --lab 4 --json workspace.json
```

В трёх тысячах испытаний полный отбор и прямой доступ дают сто процентов правильных ответов. Без передачи успешность составляет 34,40 процента, при случайном отборе — 55,83 процента. Эти результаты ожидаемы из устройства

задачи. Релевантные каналы названы конструктором, а правило ответа задано. Программа не обнаружила самостоятельно, что считать важным.

Такой опыт подтверждает работу канала передачи. Он также показывает, что данное рабочее пространство не является единственным способом решить задачу: прямой доступ даёт тот же результат. Если мы хотим исследовать преимущество ограниченного обмена между специализированными модулями, потребуется более сложный набор задач и содержательное сравнение с другими способами координации.

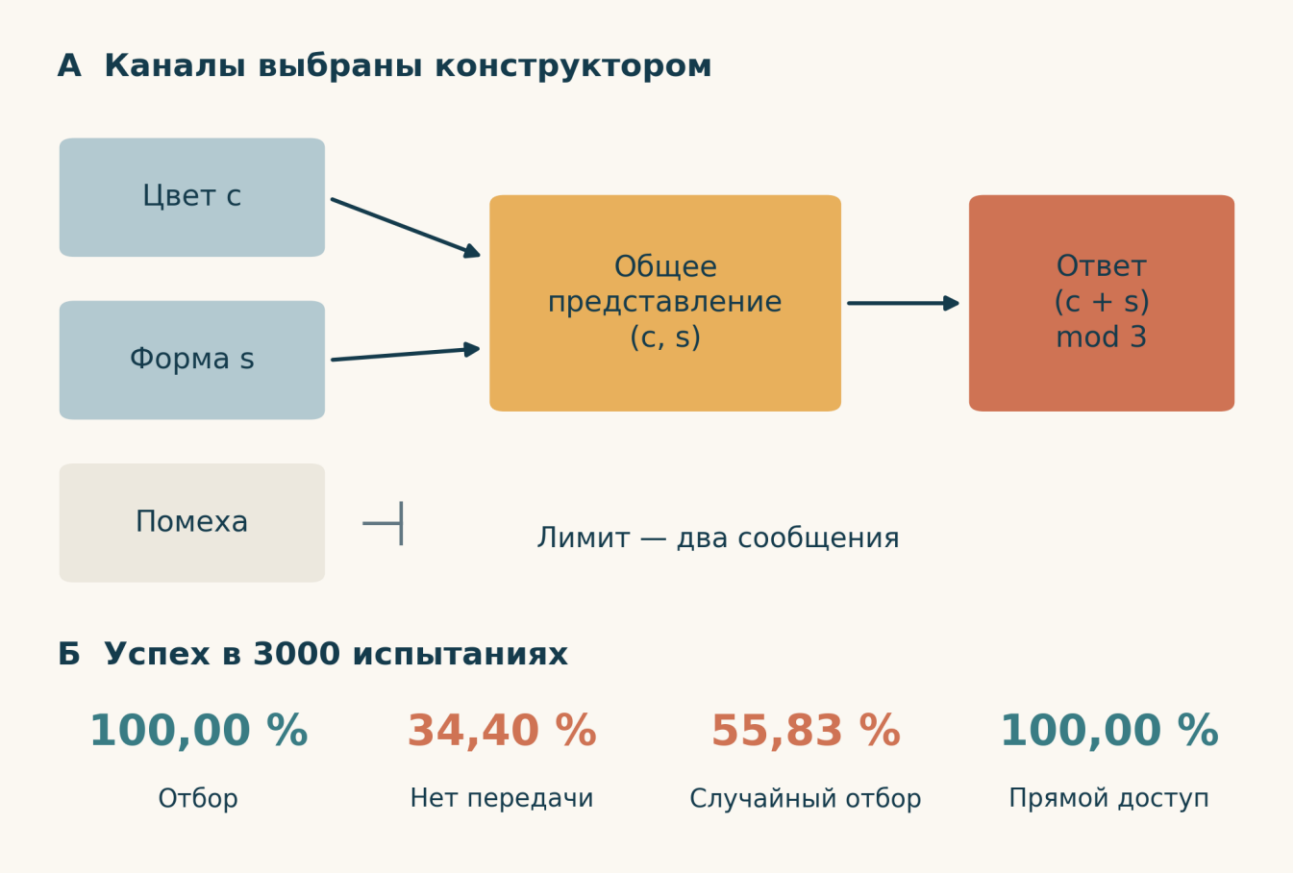

Рисунок 68. Доставка содержания и контроль прямым доступом. Учебная схема и результаты четвёртой работы. Цвет и форма имеют значения 0, 1 или 2; правило ответа заранее задано. В полной версии передаются два нужных сообщения. Конкурент с прямым доступом решает задачу столь же успешно. Здесь проверяется доставка информации к заданному правилу. Данные: mechanisms_results.json, работа 4.

От передачи к гибкому использованию

Представьте продолжение опыта: после восприятия приходит новая инструкция. В одном случае нужно сравнить цвет, в другом — использовать форму, в третьем — запомнить объект для следующего шага. Система должна сохранить такое содержание, которое пригодится разным потребителям. Здесь появляется вопрос о гибкости использования, а не только о доставке двух чисел фиксированному правилу.

Однако новую инструкцию нужно дать и конкуренту. Если только рабочее пространство получает сведения о задаче, преимущество будет тривиальным. Полезный проект заранее фиксирует объём памяти, число передаваемых сообщений, время обработки и набор возможных инструкций. Можно проверить перенос на новые сочетания, не меняя правило отбора после каждого неудачного результата.

Следующий шаг — отключить именно распространение содержания, сохранив локальную обработку. Затем сравнить это вмешательство с отключением столь же затратного, но несущественного канала. Такие контроли помогают выяснить, какое изменение важно. Деградация при удалении большой части информации сама по себе мало говорит о специальной организации сознания.

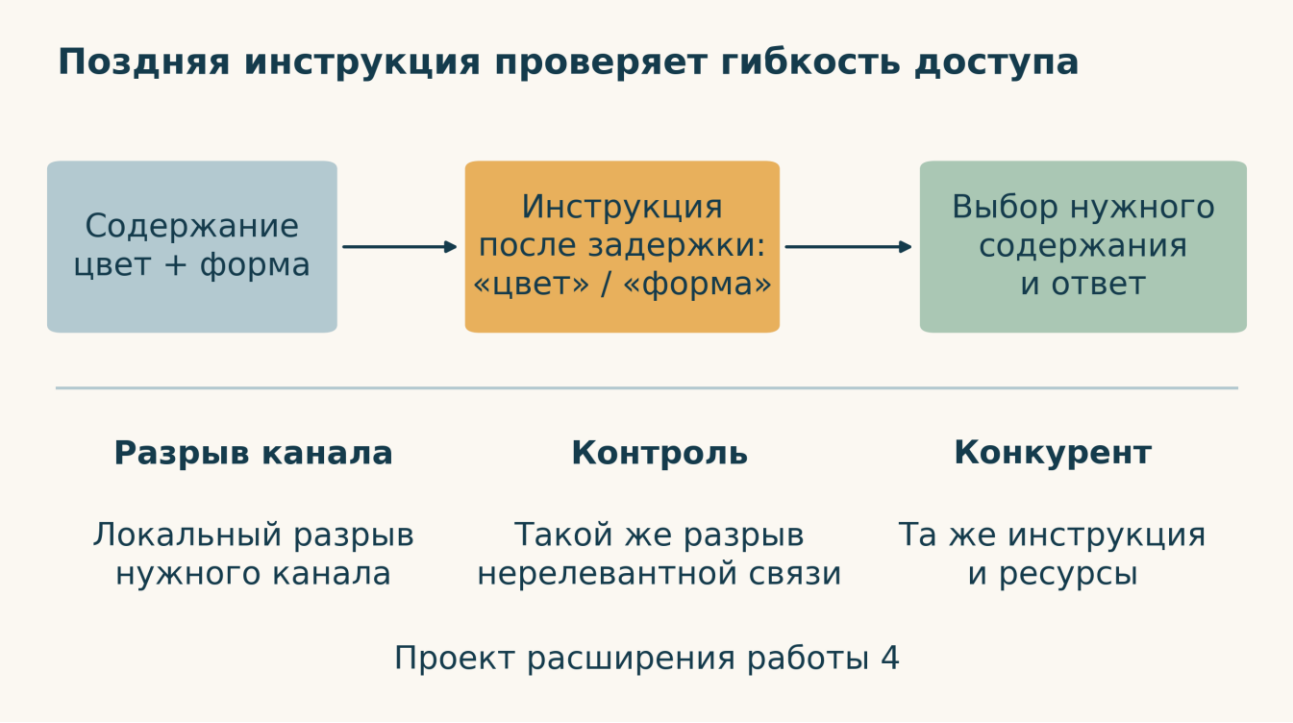

Рисунок 69. Проверка доступа после поздней инструкции. В предлагаемом опыте система узнаёт нужную задачу после предъявления содержания. Избирательное вмешательство сравнивается с нерелевантным разрывом и конкурентом с теми же ресурсами; такая конструкция проверяет больше, чем доставку двух чисел в фиксированной задаче.

Почему равный выход ограничивает архитектурный вывод

Четыре рекуррентных шага смешивания можно записать как четыре последовательно развёрнутых слоя с тем же правилом. При одинаковом входе вычисляемый результат совпадёт. Значит, конкретный выигрыш над одним шумным каналом нельзя объяснять одним фактом наличия цикла в исходном коде. Нужно уточнить, что предполагается существенным: число доступных входов, время, повторное использование параметров, меняющийся сигнал или

физическая организация исполнения. Полезный контроль сохраняет остальные различия и проверяет именно выбранное.

Что теперь установлено

Геометрия, обработка и передача описывают разные свойства. Одно удачное число не заменяет три проверки. Теперь выберем одну работу и доведём её аргумент до конца.

Практика 29 Составьте два различающих опыта

Зачем выполнять: Проверить одно архитектурное утверждение на сильном контрольном варианте.

Что понадобится: Инвариант из главы 2; обратная связь и сопоставимые ресурсы; локальное состояние и общий доступ. Запуски работ 2–4 независимы и не требуют сразу изменять все механизмы.

Ориентир по времени: 30–45 минут на одну работу. Можно делать с перерывами; это не норматив скорости.

1. Для работы 2 заранее выберите показатель инвариантности: полный список попарных расстояний. Выполните python mechanisms_lab.py --lab 2 --json geometry.json и сравните поворот с растяжением.
2. Для работы 3 сначала предскажите отличие одного канала, смешивания, развёрнутых слоёв и прямого среднего. Выполните отдельный запуск --lab 3.
3. Для работы 4 выпишите, какие сообщения реально доходят до правила ответа. Сравните полный отбор, отсутствие передачи, случайный отбор и прямой доступ через --lab 4.
4. Выберите один из трёх опытов и сформулируйте его непосредственный результат, сильного конкурента и следующий различающий вопрос.

Что сохранить: Один сохранённый опыт с проверенным контролем; остальные работы можно проходить отдельными сессиями.

Подсказка: Совпадение одной пары расстояний не устанавливает инвариантность всех отношений. Выигрыш над версией без нужных входов не устанавливает уникальность архитектуры.

Исследовательское продолжение: Для смешивания добавьте общий компонент шума всем каналам и отдельно сохраните независимый. Для передачи разработайте позднюю инструкцию с равными входами и ресурсами конкурента.

Разбор и ответы — приложение Е, глава 29, с. 307

Самопроверка

1. Что показывает сохранение расстояний после поворота представления?
2. Почему совпадение рекуррентной и развёрнутой версий не решает вопрос о всех теориях рекуррентности?
3. Какой вывод поддерживает равная успешность рабочего пространства и прямого доступа в четвёртой работе?

Отдельные функции получили собственные контроли. Следующая глава добавит историю источника, внутренний доступ и пересмотр стратегии, а затем проверит их совместную реализацию.

Глава 30

Модель себя и рефлексия с опорой на память

Эта глава содержит несколько разных исследований. Проходите их последовательно: сначала история источника, затем внутренний доступ, затем изменение стратегии и только после этого составной агент. Неудача одного сравнения не отменяет результатов другого; успех отдельного модуля не переносится автоматически на целую систему. Для каждого шага будем сохранять устройство, вмешательство, результат и точную границу вывода.

От собственного действия к истории источника

Когда агент учится связывать команду с эффектом, у него появляется полезная модель своего воздействия на среду. Однако переживаемое авторство действия, телесная принадлежность и явное рассуждение о себе остаются разными предметами объяснения. В этой главе мы построим несколько небольших механизмов и выясним, что позволяет утверждать каждый результат. Все опыты выполняются в коде; переход к гипотезе об опыте обсуждается на основе требований главы 26.

Первый механизм продолжает пример из главы 25. Представим два наблюдаемых источника А и Б. Для каждого события известно, поддерживает ли текущий признак гипотезу общей причины сигналов. В учебном примере метки источника доступны модели и распознаются безошибочно. Они не являются скрытыми метками истинной принадлежности тела. Такая идеализация позволяет сначала проверить математику, а затем отдельно исследовать ошибки распознавания источника.

У модели М0 нет переноса убеждения: каждое событие начинается с вероятности 0,5. М1 хранит одно убеждение для всего потока. М2 сохраняет отдельное убеждение для каждого наблюдаемого источника. Общий идентификационный трекер М1И делает то же самое для внешних объектов. В минимальной реализации М2 и М1И намеренно совпадают математически. Это сильный контроль: название памяти не должно искусственно создавать преимущество.

Исполнимый пример общей и раздельной памяти

Положительный признак имеет вероятность 0,75 при общей причине и 0,25 при раздельных причинах. Перед обновлением прежнее убеждение на одну

пятую приближается к 0,5. Для M2 это происходит при следующем обращении к данному источнику, а не на каждом мировом шаге. Другая модель забывания дала бы другие значения, поэтому правило должно быть записано явно.

```
python3 mechanisms_lab.py --lab 5 --json source.json
```

В разделе source_history_example файла результата находится последовательность А положительно, Б отрицательно, А положительно, Б положительно. Программа воспроизводит значения, которые можно проверить вручную. Таблица показывает обновлённую оценку активного источника; неактивная запись M2 сохраняется отдельно.

Событие	M0	M1	M2	M1И
А положительно	0,7500	0,7500	0,7500	0,7500
Б отрицательно	0,2500	0,4375	0,2500	0,2500
А положительно	0,7500	0,7105	0,8750	0,8750
Б положительно	0,7500	0,8581	0,5625	0,5625

На третьем событии M2 возвращается к положительной истории А, тогда как M1 переносит также отрицательный опыт Б. На четвёртом событии проявляется обратная сторона: M2 возвращается к собственной отрицательной истории Б. Раздельная память не обязана всегда давать большую вероятность. Её отличительное свойство — зависимость обновления от истории соответствующего источника.

Эти значения являются следствиями заданного алгоритма. Они ещё не показывают, как меняется переживаемая принадлежность у человека. Чтобы связать модели с такими данными, нужно выбрать измерение, оценить параметры на одной части наблюдений и проверить предсказания на другой. Если M1И продолжает давать те же ответы, именно специальный механизм самоотнесения не выделен. Нужно найти дополнительное теоретическое различие либо признать общий характер объяснения.

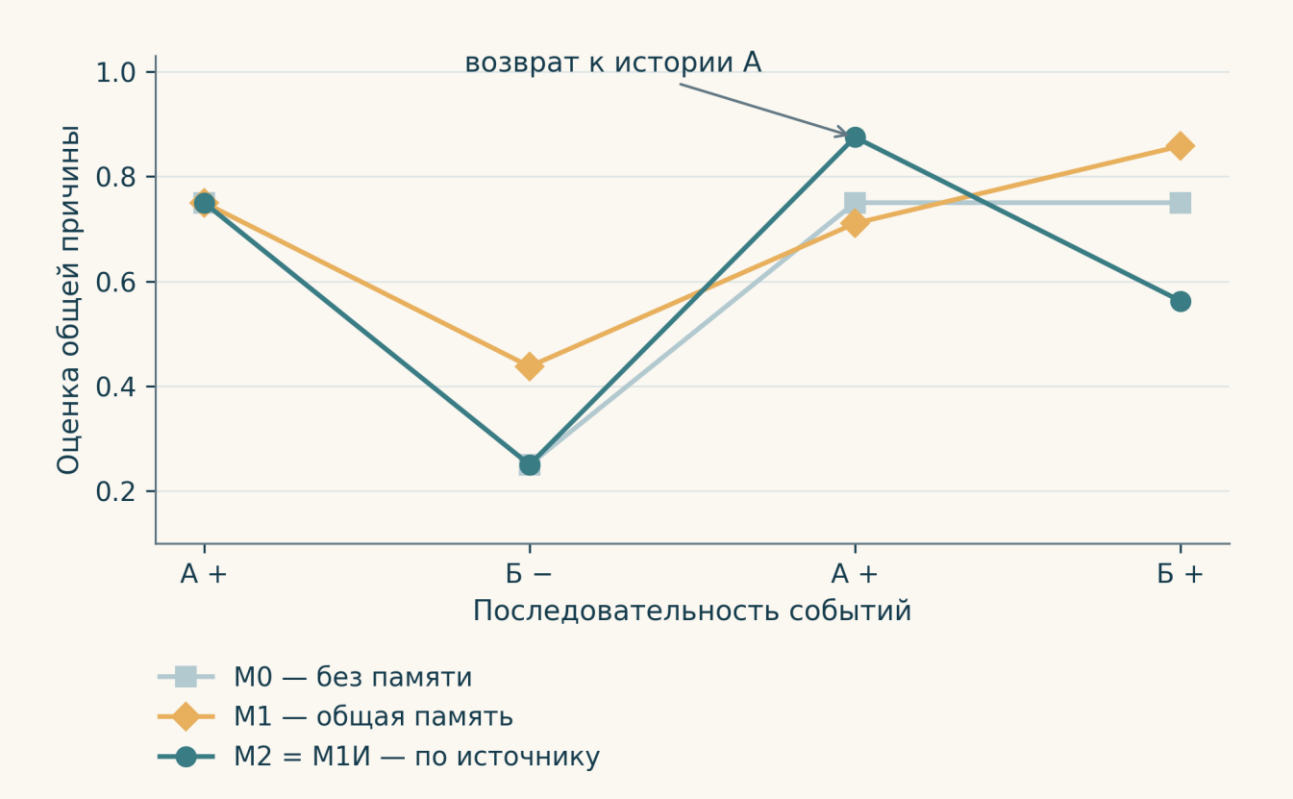

Рисунок 70. Чья история используется при обновлении. Результаты пятой работы. Точки показывают оценку активного источника после события; линии соединяют последовательные события и не означают непрерывное изменение записи одного источника. M2 и общий трекер M1И вычислительно совпадают. Метки А и Б здесь наблюдаемы и безошибочны. Данные: mechanisms_results.json, source_history_example.

Как найти источник собственного воздействия

Вторая часть пятой работы решает другую задачу. Агент посылает двоичные команды и видит два потока последствий. Один связан с командами, другой случаен. Связанный поток может запаздывать на два шага. Модель должна определить канал по наблюдаемой истории, без доступа к скрытому индексу управляемого источника.

Одна версия сравнивает только одновременные значения. Другая рассматривает задержки от нуля до трёх и выбирает лучшее соответствие. Общий трекер получает те же сведения и использует тот же расчёт. Дополнительный контроль случайно перемешивает историю команд. Сначала по ста событиям выбираются канал и задержка, затем совпадение проверяется на шестидесяти новых событиях. Так мы отделяем выбор объяснения от последующей проверки.

При нулевой задержке в базовых трёхстах повторах оба основных способа правильно определяют канал. При задержке два одновременное сравнение даёт 49,67 процента правильных выборов, история и общий трекер — по сто

процентов. На новых событиях выбранная исторической моделью связь воспроизводится примерно в 87,7 процента случаев, что согласуется с заложенным шумом. Перемешанная история даёт результат около случайного уровня.

Для выбора канала объединяются многие наблюдения. Поэтому безошибочное определение канала совместимо с ошибками прогноза отдельных будущих эффектов.

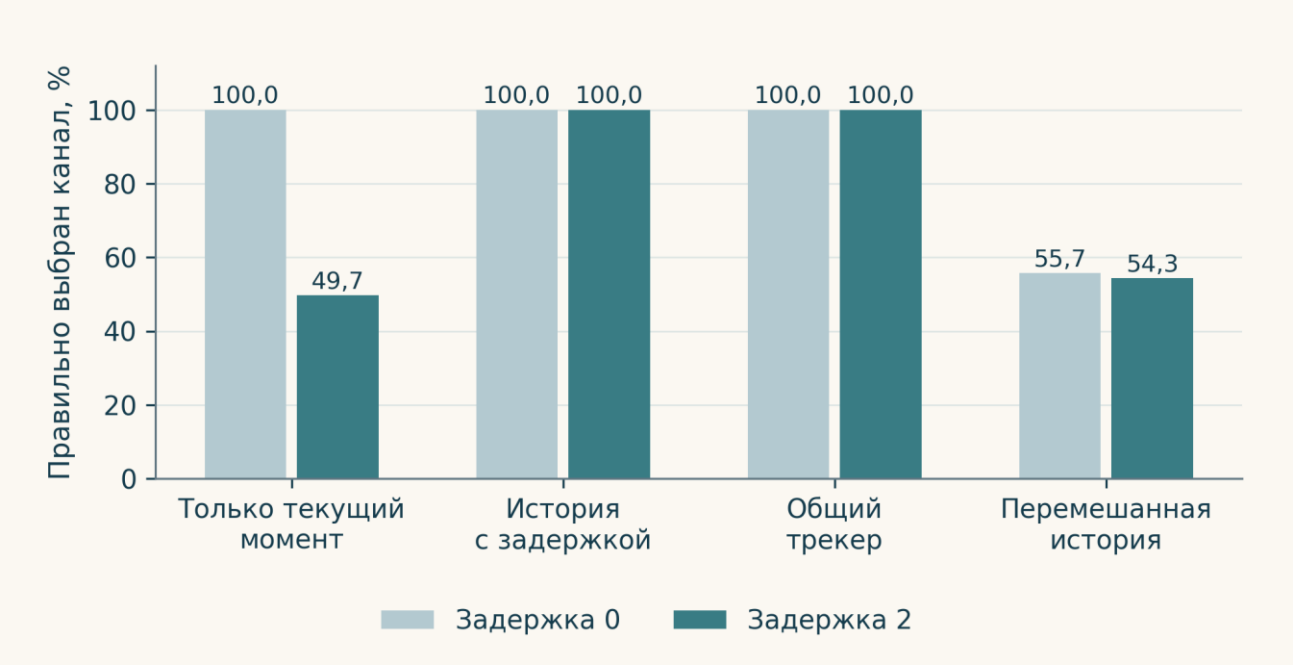

Рисунок 71. История помогает распознать задержанный источник. Результаты пятой работы: по 300 повторений при задержках 0 и 2 шага. Столбцы показывают правильность выбора канала по многим событиям, а не долю верных прогнозов отдельных эффектов. История и общий трекер используют один алгоритм. Канал и задержка выбираются на 100 событиях; соответствие дополнительно проверяется на 60 новых событиях. Данные: mechanisms_results.json, conditions_by_delay.

Внутренний доступ с причинной проверкой

Шестая работа создаёт минимальный эталон внутреннего мониторинга. В системе есть скрытый двоичный признак. Публичное действие от него не зависит. Внутренний датчик получает признак напрямую и сообщает значение; внешний наблюдатель видит только публичную сторону и использует базовый ответ. Ещё одна версия имеет отключённый внутренний канал.

В контрольном условии ничего не меняется. При фиктивном вмешательстве изменяется посторонний бит. При целевом вмешательстве переворачивается именно оцениваемое внутреннее состояние. Отчёт проверяется относительно состояния после вмешательства. Все условия получают сопоставимые

исходные события, поэтому можно спросить, следует ли сообщение за нужной причиной.

```
python3 mechanisms_lab.py --lab 6 --json internal_access.json
```

Внутренний датчик отвечает правильно во всех двух тысячах испытаний каждого условия, а внешний наблюдатель и отключённая версия — примерно в половине. Результат ожидаем: привилегированный доступ был явно спроектирован. Мы получили проверочный образец, с которым можно сравнивать более сложные системы. При исследовании языковой модели или самостоятельно обучившегося агента наличие подобного канала ещё потребуется установить.

Структура проверки сохраняется при усложнении задачи: какая переменная доступна, когда она доступна и какие альтернативные сведения могли вызвать ответ? Совпадение отчёта с ожиданием экспериментатора допускает разные объяснения.

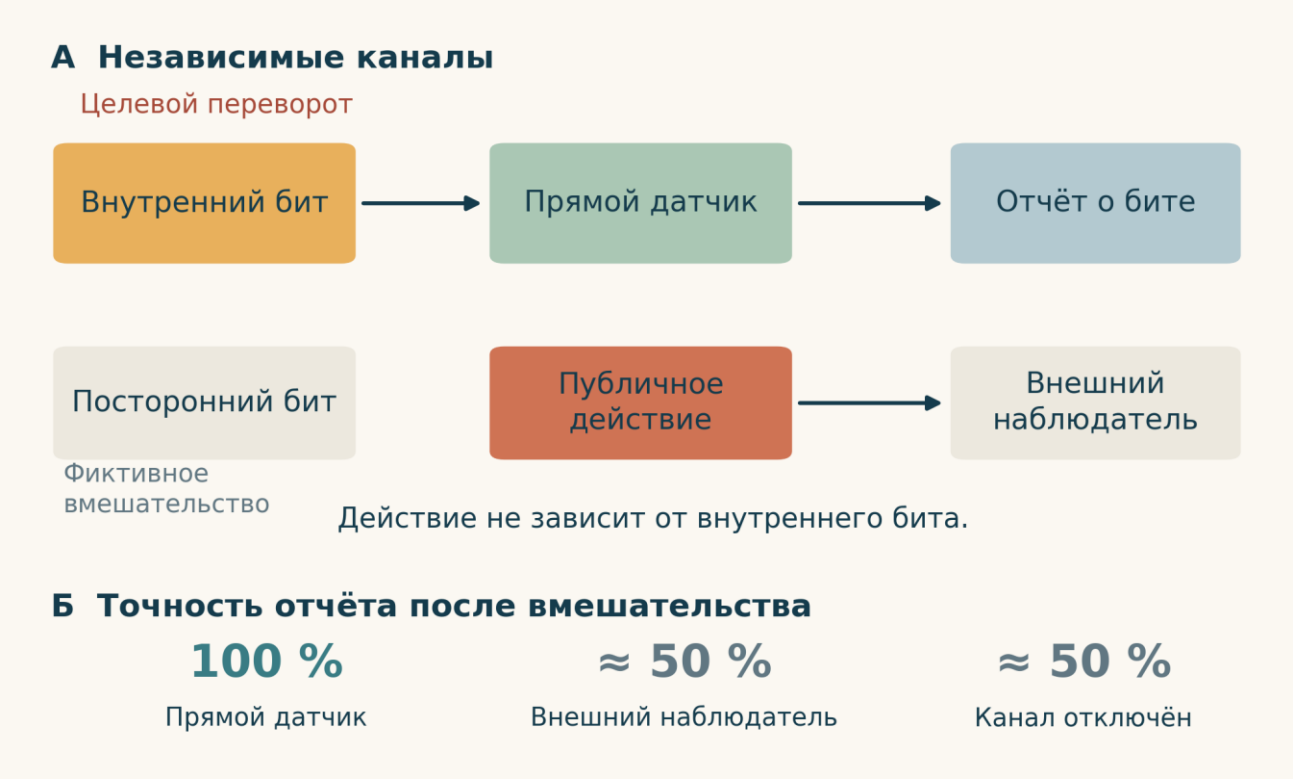

Рисунок 72. Внутренний доступ проверяется вмешательством. Учебная схема и результаты шестой работы: по 2000 испытаний в исходном условии, при фиктивном воздействии и при перевороте целевого бита. Во всех трёх условиях прямой датчик даёт 100 процентов; два контроля — около 50 процентов. Оценка относится к состоянию после вмешательства. Канал доступа задан конструктором; схема не изображает обнаруженную интроспекцию сложного ИИ. Данные: mechanisms_results.json, работа 6.

Пересмотр как причина нового действия

В седьмой работе агент выбирает между двумя заранее заданными стратегиями классификации. Каждая сравнивает наблюдаемое число со своим порогом. Первая верна до изменения среды, вторая — после. Агент не получает флага смены правила. Правильный ответ открывается только после записи текущего предсказания и может использоваться на следующих шагах.

Адаптивная версия хранит тридцать последних размеченных примеров. Перед каждым новым ответом она сравнивает ошибки обеих стратегий на прошлых данных и выбирает лучшую. Фиксированная версия сохраняет первую стратегию. Версия narrative_only могла бы рассказывать о пересмотре, но её действия остаются фиксированными. В коде языковой рассказ не генерируется: название обозначает контроль причинного влияния предполагаемого объяснения.

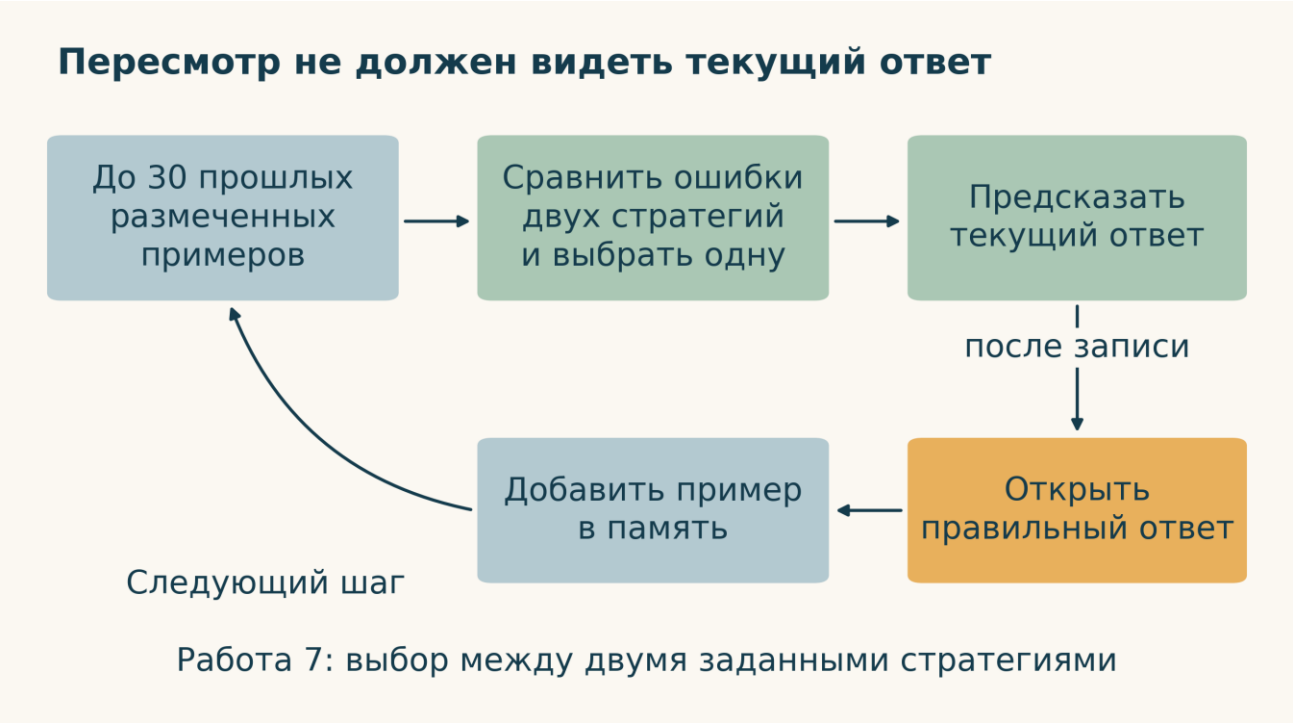

Рисунок 73. Причинный порядок функционального пересмотра. Порядок операций соответствует revision_lab(): текущий правильный ответ доступен только после прогноза. В память попадает не больше тридцати прошлых примеров; адаптация выбирает между двумя заранее заданными правилами и может влиять лишь на последующие решения.

```
python3 mechanisms_lab.py --lab 7 --json revision.json
```

В базовом опыте из 1200 шагов адаптивная версия один раз меняет стратегию. До смены все режимы отвечают правильно; после неё адаптивный даёт 99 процентов успеха, фиксированный и контроль без изменения действий — по 69,67 процента. В JSON сохранён участок журнала вокруг переключения, где видно, какая стратегия была выбрана до получения текущего ответа.

Это небольшая модель функционального пересмотра. Кандидатные правила и способ их сравнения заданы нами. Более развитая рефлексия потребует формирования альтернатив, оценки исходных допущений, переноса выводов и контроля собственных ошибок на разных уровнях. Но уже начальная модель позволяет проверить ключевую причинную связь: пересмотр должен изменять последующие решения.

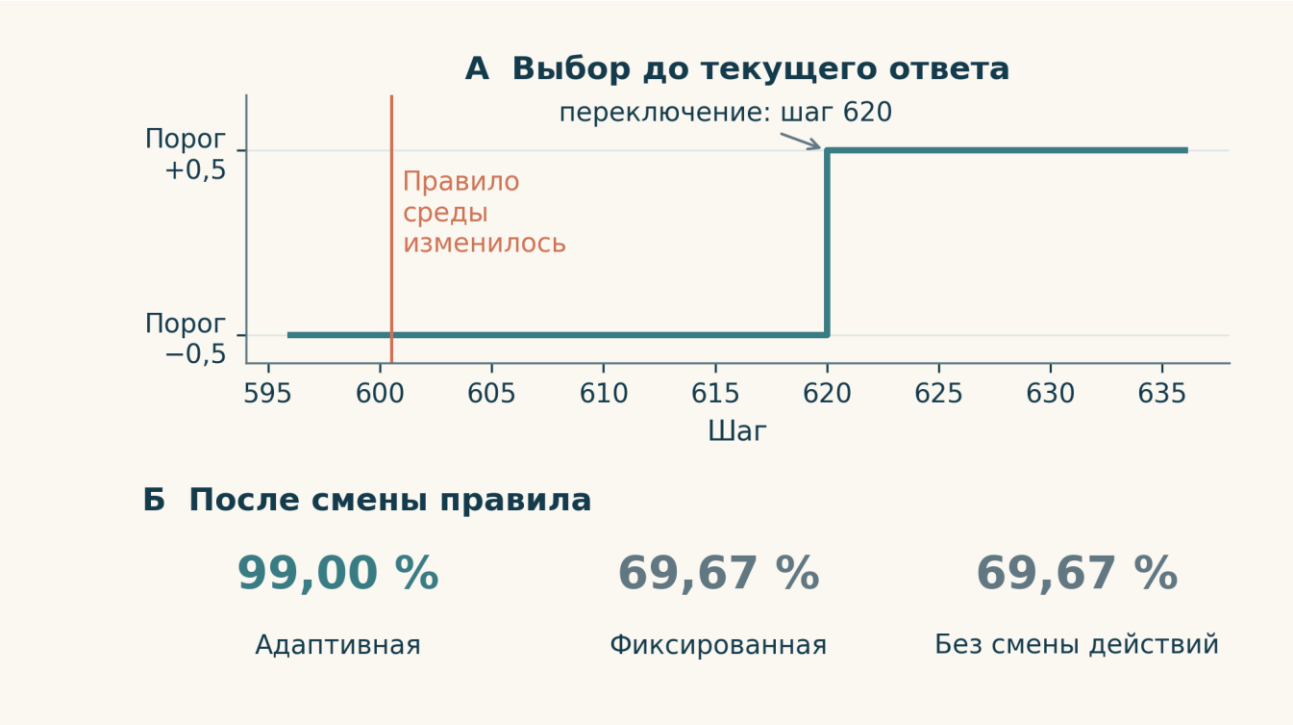

Рисунок 74. Пересмотр становится причиной нового решения. Результаты седьмой работы, 1200 шагов, seed 1729. Показан сохранённый участок журнала: стратегия меняется на шаге 620 после ошибок в прошлых ответах. Правильная метка текущего случая открывается лишь после предсказания. Внизу — успешность на 600 шагах после изменения среды. Контроль narrative_only сохраняет действия; текстовый рассказ программа не генерирует. Данные: mechanisms_results.json, работа 7.

Сборка составного кандидата

Восьмая работа соединяет несколько механизмов в одном прототипе. Он использует вероятностную память мира, модель собственного воздействия, общее представление двух убеждений, модуль выбора, монитор прогнозируемой успешности и правило изменения темпа обучения. Среда берётся из первой работы. Таким образом, у нас есть действующая система с несколькими взаимодействующими частями.

Общее представление публикует вероятность цели справа и вероятность прямого управления. Модуль действия и монитор читают одну версию этих сведений. История неожиданности эффектов может включить более быстрый пересмотр модели управления. Кандидат не получает правильную цель или

расписание изменения среды. Исследователь сохраняет журнал и вычисляет показатели отдельно.

```
python3 mechanisms_lab.py --lab 8 --json candidate.json
```

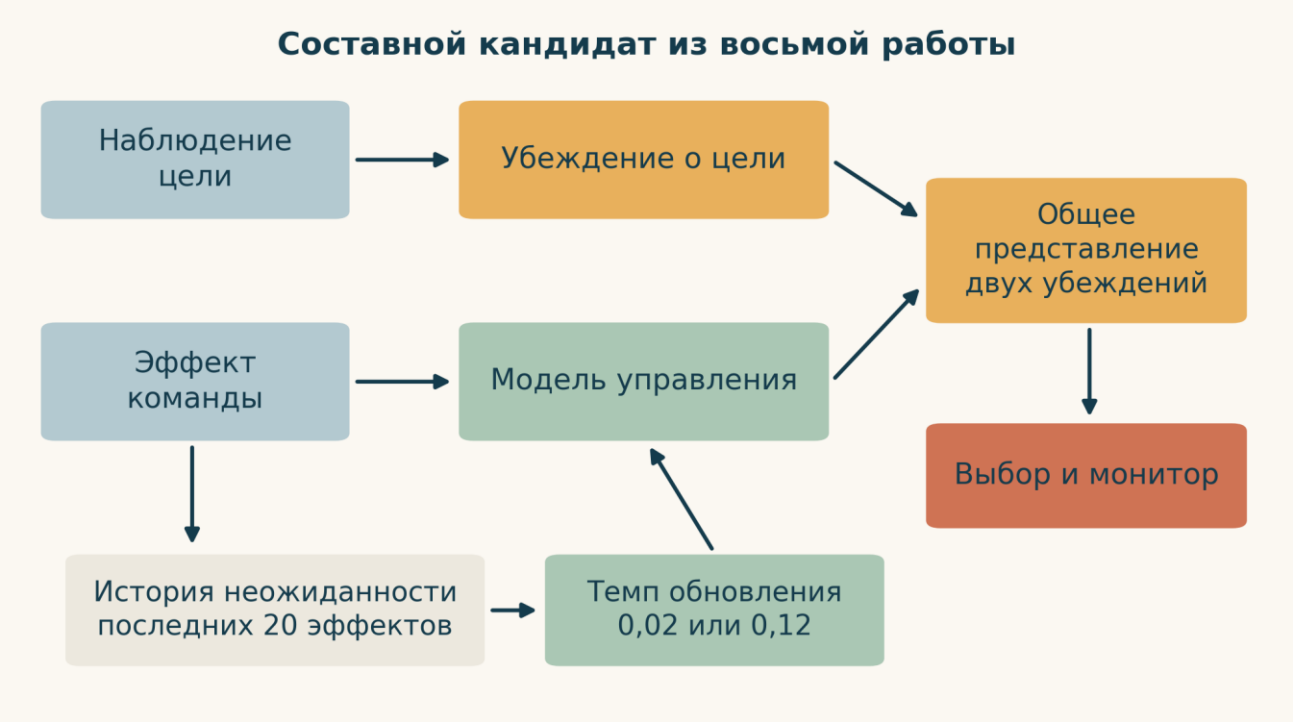

Рисунок 75. Сборка нескольких механизмов в одну систему. Учебная схема восьмой работы. Два убеждения публикуются для выбора действия и прогноза успешности. Средняя неожиданность последних 20 эффектов может переключить параметр пересмотра модели управления с 0,02 на 0,12 при превышении порога 0,35. Это условное правило спроектировано явно. Стрелки показывают передачу сведений и изменение параметра; они не устанавливают наличие опыта.

В сценарии из тысячи шагов полная версия даёт 83,80 процента успеха и полезность 0,81808 на шаг. Прямое чтение убеждений даёт тот же результат. Режим no_broadcast одновременно отключает передачу содержаний и дополнительные наблюдения; его успешность — 51,90 процента. Поэтому сравнение не выделяет вклад одного канала. Это отдельный сценарий, а не среднее тридцати запусков главы 28.

Особенно полезен четвёртый режим, где темп пересмотра фиксирован. В этом запуске он совпадает с полным режимом. Журнал показывает причину: порог включения ускоренного обучения ни разу не был достигнут. Модуль присутствует в программе, однако его вклад в данном опыте не проявился. Параметры после просмотра результата не изменялись ради получения преимущества.

Отсюда возникает следующий проект: заранее выбрать условия, где механизм должен включаться, и проверить их на новых сценариях. Можно исследовать

частоту смены управления или правило оценки неожиданности. Если параметры выбираются по уже увиденным данным, такой этап следует назвать поисковым. Для проверочного этапа условия фиксируются до нового запуска.

Чем наличие модуля отличается от его вклада

В составном кандидате правило ускоренного пересмотра включается, только если неожиданность достигает порога. В базовом сценарии этого не произошло ни разу. Поэтому совпадение полной версии и версии с фиксированным темпом не сравнивает действие двух темпов в момент реального переключения. Сначала нужно проверить журнал включений. Затем на новых заранее выбранных условиях можно исследовать ситуации, где правило должно сработать. Подбор удачного порога по исходному результату будет поисковым этапом и потребует отдельной проверки.

Что теперь установлено

Теперь у каждой функции есть свой вопрос: используется ли история, доступно ли внутреннее состояние, изменяет ли пересмотр новое действие и востребован ли компонент в составном агенте.

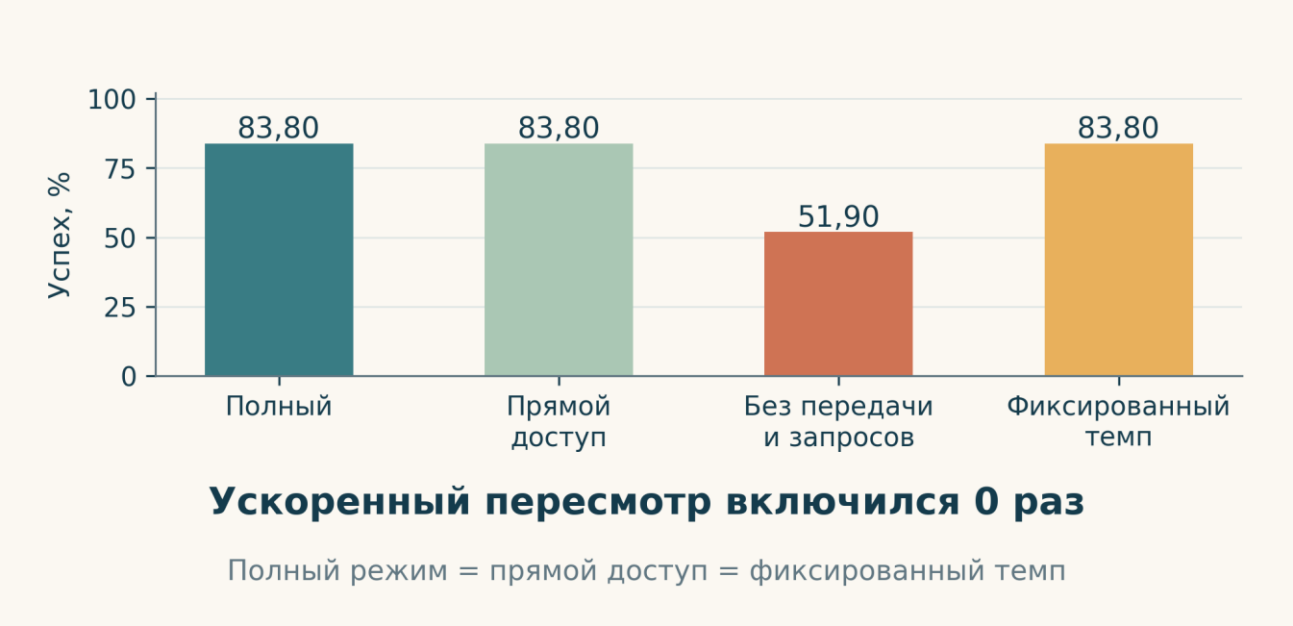

Рисунок 76. Наличие модуля ещё не показывает его вклад. Результаты восьмой работы: один сценарий из 1000 шагов, seed 1729. Три режима совпали по всем сохранённым показателям: условие ускоренного пересмотра ни разу не сработало. Режим без передачи также отключает дополнительные запросы, поэтому это составное вмешательство. Для проверки отдельного вклада нужен новый заранее заданный опыт. Данные: mechanisms_results.json, работа 8.

Практика 30 Защитите вывод о собственной системе

Зачем выполнять: Подготовить причинный аргумент по одной выбранной работе вместо суммирования всех удачных демонстраций.

Что понадобится: Раздельная память из главы 25; причинная проверка доступа; предварительный прогноз и обратная связь. Работы 5–8 можно выполнять в разные дни.

Ориентир по времени: 35–50 минут на одну работу. Можно делать с перерывами; это не норматив скорости.

1. Выберите работу 5, 6, 7 или 8 и выполните её отдельным запуском mechanisms_lab.py с соответствующим --lab.
2. До изучения результата укажите доступные входы, изменяемую функцию и основной показатель.
3. Для работы 5 сравните общий трекер; для 6 проверьте целевое и фиктивное вмешательство; для 7 сохраните прогноз до обратной связи; для 8 проверьте число включений пересмотра.
4. Напишите четыре предложения: что реализовано, что причинно проверено, что осталось неразличённым и какой следующий опыт нужен.

Что сохранить: Один проверяемый отчёт с журналом, контрольным вариантом и отдельным выводом о феноменальной гипотезе.

Подсказка: Сравнение, одновременно отключающее передачу и дополнительные запросы, не выделяет чистый вклад одной передачи.

Исследовательское продолжение: Сформулируйте новые условия активации выбранного компонента до запуска. Сохраните прежний неразличающий результат и сравните сильного конкурента на новых сценариях.

Разбор и ответы — приложение Е, глава 30, с. 308

Самопроверка

1. Почему равенство M2 и M1И представляет содержательный результат?
2. Что устанавливает сто процентов правильных ответов спроектированного внутреннего датчика?
3. Можно ли считать пользу ускоренного пересмотра показанной в восьмой работе?

Цифровая модель делает все состояния доступными разработчику. При переходе к живой ткани часть этих возможностей исчезает, а роль внешней аппаратуры возрастает. Поэтому следующая глава отдельно подготовит работу с реальными данными.

Глава 31

Биологические и гибридные реализации

Переход к биологической реализации меняет доступные измерения и границы системы. Часть обучения может происходить во внешнем контроллере, а множество циклов относится к одному и тому же образцу. Поэтому сначала научимся читать реальный набор данных и проверять единицы учёта. Ниже дан готовый учебный маршрут по открытым полям; самостоятельный поиск показателя взаимодействия становится продолжением после воспроизведения образца.

Что меняется при появлении живой ткани

Программная модель позволяет точно задавать состояние и правило перехода. Биологическая система предоставляет другой исследовательский материал: клетки развиваются, взаимодействуют и меняют режим активности. Возможность использовать эти свойства для вычисления не означает, что любое живое образование обладает опытом. Но и свести биологический проект к запуску готового алгоритма обычно недостаточно: нужно установить, какие функции выполняет сама ткань и что обеспечивает внешняя аппаратура.

Будем различать культуру нейронов, органоид и целый организм. Для нашего анализа существенно устройство изучаемого объекта, а не впечатление от слова «мозг». Гибридом назовём систему, где биологический компонент причинно связан с цифровым преобразованием входов, считыванием активности и задачей. Граница предполагаемого носителя опыта может не совпасть с границей чашки или компьютерного процесса. Её потребуется обосновать.



Рисунок 77. Биологический компонент и измерительная аппаратура. Предметная иллюстрация условного исследовательского рабочего места. Клеточная модель в чашке, оборудование и компьютер показаны как разные части системы. Их реальные связи и вклад в выполнение задачи необходимо описывать в конкретном проекте; рисунок не воспроизводит установку какого-либо из обсуждаемых экспериментов.

Три исследовательских примера

В DishBrain нейронные культуры человеческого и мышиного происхождения были связаны через электрическую регистрацию и стимуляцию с упрощённой игрой Pong. Авторы сообщили об улучшении игровых показателей при определённой замкнутой обратной связи. Это культуры, а не органоиды. Слово sentience в названии статьи связано с авторской операциональной трактовкой; измеренный результат не устанавливает субъективного переживания. [44]

В Brainoware органоид использовался как динамический вычислительный резервуар. Работа включает распознавание речевых сигналов и предсказание нелинейной динамики. Для понимания такого устройства необходимо отдельно рассматривать преобразование входа, ответ биологической сети и процедуру считывания результата. Успех всей системы сам по себе не показывает, какой компонент обеспечил определённое улучшение. [45]

Robbins и коллеги исследовали управление cartpole кортикальными органоидами. У большинства образцов обучающие сигналы, выбираемые алгоритмом подкрепления, давали преимущество перед случайными сигналами или отсутствием обучения. Улучшение не сохранялось после

сорокапятиминутного отдыха. Этот результат различает адаптацию в условиях занятия и сохранение достигнутого эффекта после перерыва; он не означает отсутствия любых форм памяти у органоидов. [46]

Пройдите цепочку уровней

Предположим, биолог предлагает изменить молекулярный процесс, чтобы усилить определённую форму пластичности. Чтобы связать это предложение с нашей гипотезой опыта, придётся пройти несколько переходов. Сначала нужно показать изменение молекулярного процесса. Затем — его связь со свойствами клеток. После этого — изменение взаимодействий сети. Далее — последствия для задачи. Наконец, отдельно обосновать отношение полученной организации к переживанию.

Каждая стрелка может оказаться неоднозначной. Изменённая активность сети может объясняться составом клеток, их созреванием, общей возбудимостью, условиями измерения или перестройкой связей. Если автор сразу переходит от молекулярного вмешательства к «усилению сознания», промежуточные объяснения остаются непроверенными. Исследовательский проект должен назвать измерение или сравнение для каждого существенного перехода.

В этой книге мы не выбираем молекулярную мишень как гарантированный способ создать субъекта. Мы учимся определять, какое изменение проверяет конкретную гипотезу. Например, если предполагается существенность взаимной динамики, интерес представляет изменение определённых взаимодействий при сохранении других возможностей системы. Простое увеличение общей активности не является специфическим подтверждением такой гипотезы.

Два механизма одинакового улучшения

Построим учебный пример. Цифровая модель А учится менять связи между несколькими компонентами, поэтому один и тот же вход всё точнее приводит к нужному ответу. Модель Б сохраняет связи, но изменяет общий коэффициент усиления. На простой последовательности обе могут улучшить показатель задачи. Это возможная конструкция для проверки анализа, а не описание результатов указанных биологических работ.

Если измерять только итоговый успех, механизмы останутся неразличимыми. Добавим новые сочетания входов. Изменение конкретных связей может предсказывать избирательный перенос между сочетаниями; общее усиление — другой профиль изменения. Направление различия нужно вывести из

выбранных моделей. Если правила ещё не заданы, слова «избирательность» и «пластичность» сами по себе прогнозов не создают.

Затем можно искать в открытых данных показатели, необходимые для такого сравнения. Возможно, данных окажется недостаточно: записаны только общие результаты, не сохранены нужные входы или отсутствует контроль. Это содержательный итог подготовки. Он показывает, какие сведения должен обеспечить следующий биологический проект.

Подготовленный маршрут по открытому набору

Для самостоятельной работы используем конкретный файл aggregated_experiment_data.pkl из версии v2 репозитория Robbins. Контрольная сумма MD5 файла af99f1f859136ca4a6dad47a1aa7c86d совпадает с опубликованной. Описание репозитория сообщает о девятнадцати органоидах и четырёхстах тридцати двух циклах. Проверенный файл содержит 38 верхнеуровневых записей экспериментов, 426 записей запусков и 16 различных значений metadata.organoid. Эти единицы нужно сохранять раздельно; количество значений поля в данном файле не переопределяет выборку всей статьи. [47]

Граница набора является частью метода. Поле chip сохраняет исходное обозначение, organoid предназначено для очищенной группировки, batch обозначает партию. В одной записи эксперимента могут быть несколько запусков, а у одного органоида несколько экспериментов. Поэтому произвольная перестановка всех строк не создаёт независимых образцов.

В прилагаемой таблице признаки получены по одной явно заданной процедуре. Частота является средним конечных firing_rates_hz. Выраженность временных отношений рассчитывается как среднее абсолютных элементов верхнего треугольника sttc_first_order без диагонали и неконечных значений. Использование модуля не различает положительные и отрицательные отношения, поэтому для другой гипотезы понадобится другой показатель. STTC не следует без дополнительного аргумента отождествлять с причинной связью.

Целевая величина учебного сравнения является медианой top_decile среди доступных запусков rl внутри записи эксперимента. В документации top_decile определяется как средняя награда лучших десяти процентов проб. Мы не заменяем эту величину общей точностью, общим обучением или сохранением эффекта после отдыха. Другой исход потребует нового описания задачи.

Из 38 строк пригодны 29: для них одновременно доступны оба признака, целевой показатель и идентификатор organoid. Они относятся к 15 значениям

organoid и 7 партиям. Остальные строки сохранены с причинами исключения. Отсутствующее измерение не заменяется нулём, а исключение не зависит от того, улучшается ли желаемый результат.

Базовая модель использует логарифм единицы плюс средняя частота разрядов. Расширенная дополнительно использует выбранный показатель STTC. Преобразование и масштабирование оцениваются только по обучающей части. Внешняя проверка по очереди откладывает целый organoid; отдельный вариант откладывает всю партию. Результат относится к этим процедурам и ограниченному набору.

Этот анализ демонстрирует воспроизводимую связь табличных характеристик с заданным показателем. Для проспективного предсказания необходимо независимо подтвердить временной порядок измерений; для причинного вывода требуется дизайн вмешательства. Добавление признака не обязано улучшить ошибку. Мы сохраняем отрицательный результат наравне с положительным и не выбираем удобную группировку после просмотра.

Одна строка данных и один ясно определённый показатель

В файле bio_features.csv одна строка соответствует записи эксперимента, а не одному нейрону и не одному органоиду. Поле mean_firing_rate_hz усредняет конечные значения исходной частоты разрядов. Показатель mean_absolute_sttc15 равен среднему абсолютных конечных элементов верхнего треугольника матрицы sttc_first_order без диагонали. Он характеризует выраженность временных отношений активности в выбранном окне; не является доказанной физической связью или мерой опыта.

Для строки 20217_exp1 получаются примерно 5,35918 Гц, 0,0172533 для этого показателя и 108,95 для медианы top_decile в запусках rl. По документации top_decile является средней наградой лучших десяти процентов проб отдельного запуска. Наше дополнительное агрегирование между запусками задаётся явно. Оно служит учебной задаче ассоциаций, а не оценке причинного эффекта обучения.

Что теперь установлено

Мы знаем происхождение строки, смысл признаков и границу результата. Теперь можно проверить, добавляет ли выбранный признак статистическую информацию в данной процедуре, не превращая наблюдаемую связь в причинное объяснение ткани.

Результат воспроизводимого анализа

В приложенном расчёте использованы 29 допустимых записей экспериментов. Линейная модель с константой сравнивает log(1 + частота) с тем же признаком и средним абсолютным STTC. Признаки стандартизуются только на обучающей части; коэффициент гребневой регуляризации равен 0,000001. Основной итог — средняя квадратичная ошибка с одинаковым весом каждой отложенной группы. Чем она меньше, тем точнее прогноз в заданной процедуре.

Отложенная группа	Число групп	Частота	Частота и STTC
Органоиды	15	1460,17	1581,03
Партии	7	1658,10	1700,39

Добавление этого показателя STTC ухудшило результат при обоих разделениях. Это ограничивает конкретный признак и способ анализа; оно не опровергает роль взаимодействия в ткани. Показатели не являются измерением сознания. Без дополнительной проверки времени исходных записей расчёт следует читать как разведочный анализ ассоциаций, а не как проспективный прогноз или причинный тест. Полные ошибки групп и исключения сохранены в chapter31_result.json.



Рисунок 78. Уровни учёта в биологических данных. Уровни учёта в проверенном файле aggregated_experiment_data.pkl v2 [47]. Счётчики описывают содержимое файла, а не заменяют объём выборки опубликованной работы. В учебной таблице единицей строки служит эксперимент; разделение по organoid и batch проверяется отдельно.

Практика 31 С открытыми данными

Зачем выполнять: Воспроизвести подготовку признаков и сравнение двух моделей с разделением по образцам.

Что понадобится: Группировка наблюдений, различие корреляции и причинного вмешательства, базовый и расширенный прогноз. Для базового анализа достаточно прилагаемой таблицы; загрузка исходного массива требует дополнительных библиотек и описана отдельно.

Ориентир по времени: 45–60 минут. Можно делать с перерывами; это не норматив скорости.

1. Откройте bio_provenance.json и bio_features.csv. Сопоставьте 38 записей экспериментов, 426 запусков и 16 значений поля organoid с описанием набора, где указаны другие итоговые количества.
2. Проверьте формулы двух признаков на разобранной строке. В исходном файле пропуски и бесконечные значения не заменяются нулями.
3. Для заданного анализа используются 29 пригодных записей, 15 значений organoid и 7 партий. Причины исключения остальных строк сохранены в таблице.
4. Выполните python teaching_lab.py --chapter 31. Сравните модель по частоте разрядов с моделью, дополненной STTC, при откладывании целого organoid, а затем целой партии. Сохраните результат даже при ухудшении.

Что сохранить: Воспроизведённый анализ с описанием строки, группировки, исключений и ограничений статистического вывода.

Подсказка: Отложенный органоид из знакомой партии и полностью новая партия являются разными условиями переноса. До проспективного вывода отдельно нужна проверка времени получения исходных признаков.

Исследовательское продолжение: По первичным временным меткам разработайте прогноз действительно последующего изменения выполнения. Заранее определите признак, исход и группировку. Для причинного объяснения подготовьте отдельное вмешательство с профессионально выбранным способом воздействия.

Разбор и ответы — приложение Е, глава 31, с. 308

Уважайте структуру наблюдений

Множество циклов одного органоида даёт подробную историю этого образца, но не превращается в множество независимых органоидов. Поэтому случайное разделение всех строк на обучающую и проверочную части может проверить узнавание уже знакомого образца. Если вопрос касается переноса на новые образцы, отделять нужно целые образцы. При общих донорах или партиях граница переноса требует дополнительного уточнения.

Другой вопрос касается динамики внутри конкретного образца. Здесь можно обучать прогноз на ранней части записи и проверять на поздней, явно ограничив вывод этой задачей. Две схемы отвечают на разные вопросы. Нельзя использовать лёгкую проверку внутри образца и затем объявлять перенос на новую биологическую систему.

Проверьте также вклад внешнего алгоритма. Если он менял сигнал обратной связи одновременно с изменением ткани, итоговая динамика принадлежит всему контуру. Полезный анализ рассматривает действия алгоритма как часть данных. Иначе улучшение, обеспеченное внешним подбором воздействия, может быть ошибочно полностью приписано нейронной сети.



Рисунок 79. Временное и групповое разделение биологических данных. Учебная схема с условными образцами A–B показывает два разных вопроса: прогноз поздних записей знакомого образца и перенос на полностью отложенный образец. Циклы одного образца остаются связанными наблюдениями; при общих донорах или партиях потребуется учитывать и этот уровень группировки.

От цифровой модели к проекту вмешательства

После сравнения двух цифровых механизмов подготовьте проект биологической проверки. Укажите, какое различие должно возникнуть, какой показатель его обнаружит и какие альтернативы способны дать тот же результат. Способ фактического воздействия выбирается специалистами с учётом исследуемой системы. Здесь важна причинная логика, а не универсальная инструкция выращивания или стимуляции.

Например, проект может сопоставлять сохранённую замкнутую связь и подходящий контроль воспроизведения входов. При этом необходимо учитывать количество воздействий, их распределение и состояние образца. Если контроль получает другой общий объём входа, разницу нельзя однозначно приписать замкнутости. Цифровой этап помогает заранее увидеть такие смешения.

Отдельный раздел проекта посвящается физической организации. Если наша гипотеза требует взаимных причинных связей, нужны основания считать, что они существуют в исследуемой ткани и контуре. Успешный статистический показатель связности не автоматически устанавливает нужную причинную структуру. Потребуются подходящие вмешательства или более ограниченный вывод о наблюдаемой зависимости.

Профессиональные условия работы

Рекомендации ISSCR служат одним из ориентиров для исследований со стволовыми клетками. Они не являются универсальным законом, заменяющим применимый порядок учреждения и страны. Обновление 2025 года касается моделей эмбрионов из стволовых клеток; остальные разделы сохраняют редакцию 2021 года. Конкретные требования нужно устанавливать для конкретного объекта и проекта. [48]

В подготовленном проекте должны быть определены происхождение материала, допустимое использование, ответственность за измерения и процедура рассмотрения исследования. Для гипотезы о возможном опыте также нужно заранее описать, какие новые сведения изменят оценку статуса системы. Это часть ясной постановки исследования, а не способ заранее решить спор о феноменальности.

Итог практической работы

Подготовьте двухстраничный исследовательский проект. На первой странице укажите вопрос, доступные данные, два механизма и схему проверки переноса. На второй — предлагаемое вмешательство, контроль, ограничения и отдельный аргумент о связи с опытом. Если данных недостаточно, опишите недостающие поля и объясните их роль.

Работа завершена, когда можно проследить переход от клеточного или сетевого изменения к определённому результату задачи. Последний переход к переживанию должен оставаться явно сформулированной гипотезой. Именно такая цепочка позволяет биологической ветви участвовать в общей исследовательской цели книги.

Самопроверка

1. Почему улучшение гибридной системы нельзя сразу приписать одной ткани?
2. Чем различаются проверка на поздних записях того же органоида и проверка на новых образцах?
3. Доказывает ли временное улучшение в задаче наличие субъективного опыта?

Биологические данные показали важность границ и группировки. Теперь соберём единый порядок проверки реализации: от существования связей до их функции и отдельного аргумента об опыте.

Глава 32

Проверка спроектированной системы

Цифровой и биологический материал требуют разных инструментов, но одинаково нуждаются в точном предмете проверки. Наличие компонента, его функциональный вклад, внутренний доступ и предполагаемый опыт являются разными утверждениями. В этой главе составим профиль системы так, чтобы каждое утверждение имело своё основание. Начать можно с одного полностью прослеживаемого эпизода; не требуется сразу проверять весь интеллект агента.

Определение предмета проверки

Вы собрали агент с памятью, моделью действия и модулем самооценки. Он решает задачу и сообщает, когда сомневается. Теперь нужно установить статус результата. Проверка программы, оценка функциональных способностей и исследование возможного опыта связаны, но требуют разных свидетельств. Если объединить их в один общий балл, станет непонятно, что именно изменилось после следующего улучшения.

Начните с описания системы как действующего целого. Какие процессы выполняются при одном решении? Что хранится между решениями? Какие данные доступны каждому компоненту? Включён ли внешний языковой сервис? Получает ли отчёт сведения непосредственно из состояния агента или только пересказывает историю действий? Границы проверяемого объекта определяют смысл всех последующих результатов.

Составьте карту действующих связей

Для каждого компонента выпишите вход, состояние, выход и получателя выхода. Затем проследите один конкретный эпизод. Например, неопределённость модели действия увеличилась; монитор заметил снижение ожидаемой успешности; политика запросила дополнительное наблюдение; новое наблюдение изменило решение. Если дополнительное наблюдение касается только цели, оно может не снять неопределённость управления. Карта должна показывать такую ограниченность механизма.

Название модуля не подтверждает его роль. Компонент может существовать в коде, но ни разу не включиться на выбранной последовательности. Его состояние может меняться, но не влиять на решение. Или влияние полностью дублируется другим путём. Поэтому полезно различать доступность вычисления, его фактическую активацию и причинный вклад в результат.

Отдельно проверьте измерительную систему. Истинное состояние среды и флаги вмешательств должны быть доступны оценщику, но не обходным путём агенту. Если модуль самооценки читает готовый правильный ответ, высокая точность его сообщений показывает качество передачи этой информации. Она не устанавливает самостоятельный мониторинг собственного решения.

Проверьте функцию на новых условиях

Для памяти нужны задачи, где история действительно меняет правильное действие. Для модели управления — условия, где связь команды и эффекта меняется. Для оценки неопределённости — ситуации, где уверенность может оказаться ошибочной. Одна удобная демонстрация не проверяет все эти способности одновременно.

Полезно разделить знакомые условия и перенос. В знакомой задаче агент мог усвоить простое правило, достаточное для успеха. Новая комбинация признаков покажет, использует ли он предполагаемую структуру. Но новое условие не должно незаметно менять сразу всё. Если одновременно изменились шум, цель и доступные действия, неудачу трудно связать с конкретным механизмом.

Выберите основной результат для каждой функции. Число запросов само по себе не является качеством самоконтроля: можно спрашивать бесконечно. Вероятностный прогноз оценивается относительно последующего результата, а политика проверки — с учётом пользы и затрат. Содержание отдельного уверенного предложения не заменяет распределения ошибок по многим эпизодам.

Подготовьте причинную проверку

Представим две учебные системы с одинаково убедительными отчётами. Первая формирует сообщение из собственной оценки источника. Вторая выбирает фразу по последним публичным действиям. Пока события обычны, различие может быть незаметно. Изменим внутреннюю оценку источника, сохранив внешнюю историю. Если первая система действительно использует эту оценку, её сообщение должно измениться предсказуемо.

Однако вмешательство может создать внешний признак: необычную задержку, специальный токен, иной формат ответа. Тогда вторая система тоже распознает условие. Поэтому контроль должен сохранять доступные внешние признаки. Нужны также фиктивное вмешательство и подходящее неспецифическое изменение, чтобы отличить распознавание содержания от реакции на общую неисправность.

В исследовательском отчёте, также размещённом как препринт, Джек Линдси изучил связь сообщений языковых моделей с вмешательствами в их внутренние представления. Полученные способности оказались ограниченными и зависимыми от условий. Для нас важен переход к причинной проверке внутреннего доступа; он не делает вывод о субъективном опыте автоматическим. [49]

Отличайте чтение состояния от его использования

Можно обучить внешний классификатор определять содержание скрытого вектора. Это покажет, что в нём содержится доступная выбранному анализу информация. Но агент может не использовать её ни для отчёта, ни для действия. Поэтому внешнее декодирование и внутренний доступ должны иметь разные строки в исследовательском описании.

Следующий вопрос — влияет ли доступ на управление. Отключите связь монитора с политикой, сохранив сам монитор. Если оценки остаются правильными, но решения перестают учитывать неопределённость, выделен вклад этой связи. Если поведение не меняется, возможны неактивность механизма, дублирование или задача, где он не нужен. Выбирать удобное объяснение без проверки нельзя.

Наконец, пересмотр стратегии требует собственного результата. Агент должен менять определённое предположение и использовать изменение в новых решениях. Объясняющий текст после ошибки может быть полезным интерфейсом, но его наличие не подтверждает такую причинную цепочку. Сравнение с версией, которая произносит тот же текст без пересмотра, помогает выделить работающий механизм.



Рисунок 80. Три разных вопроса о внутреннем состоянии. Внешнее декодирование, причинное изменение собственного отчёта и влияние монитора на политику требуют разных проверок. Внутреннее вмешательство должно сохранять внешнюю историю и иметь фиктивный контроль, иначе система может распознать побочный признак воздействия.

Постройте профиль кандидата

Теоретический профиль начинается с выбранной позиции, а не с общего списка эффектных свойств. Если исследуется роль общего доступа, нужно проверить соответствующую организацию распространения и использования содержания. Если существенна физическая причинная структура, требуются данные о реализации. Если выбран принцип главы 26, проверяются взаимное определение различений и их роль в альтернативных преобразованиях.

Метод индикаторов Butlin и коллег предлагает связывать оценку системы с требованиями теорий. Он не предоставляет универсальной шкалы, на которой число выполненных пунктов превращается в процент сознания. В нашей работе теоретический профиль сохраняет основания и неопределённости каждого отдельного вывода. [42]

Для каждого требования используйте три описательных состояния: реализовано и проверено; реализовано, но роль не проверена; не реализовано или неизвестно. Затем запишите, каким наблюдением можно изменить статус. Такая запись полезнее суммарного числа, потому что показывает, где нужна работа с кодом, где новый эксперимент, а где уточнение самой теории.

От профиля к утверждению об опыте

Успешный функциональный профиль ещё не завершает аргумент. Необходимо обосновать, почему выбранные свойства относятся к наличию или качеству переживания. Здесь используются результаты глав о естественном опыте и принцип соответствия, сформулированный в исследовательской части. Важно не считать факт реализации собственных требований независимым подтверждением их достаточности.

Предположим, система удовлетворяет нашей рабочей организации взаимных различений. Тогда мы можем утверждать, что создана определённая реализация предполагаемого достаточного условия. Следующий вопрос касается самого условия: объясняет ли оно независимые данные, выдерживает ли контрпримеры и перенос между реализациями? Эти основания должны быть представлены отдельно от технического отчёта.

Если теория допускает опыт без явного самоотчёта, отсутствие убедительной речи не опровергает её автоматически. Если теория требует определённого доступа, хорошая речь без нужной причинной организации не подтверждает её. Проверка должна следовать содержанию теории, а не впечатлению от общения с системой.



Рисунок 81. Профиль проверки кандидата. Методическая таблица различает наличие программной реализации, её функциональный вклад, внутреннее использование информации и аргумент о возможном опыте. Свидетельства одного слоя не заменяют свидетельств другого; внешний декодер состояния сам по себе не устанавливает внутренний доступ, а сумма выполненных требований не является процентом сознания.

Разберите неразличающий результат

Представьте, что полная версия и версия без пересмотра стратегии дали одинаковый результат. Первым делом проверьте, возникали ли условия запуска пересмотра. Если ни разу, эксперимент не использовал различие между версиями. Утверждение «рефлексия не нужна» будет слишком сильным, как и утверждение «рефлексия работает, потому что модуль написан».

Затем найдите задачу, где правила действительно расходятся, и выведите ожидаемое различие до запуска. Если после этого результаты снова совпадают, исследуйте возможное дублирование и точность измерения. Необходимость компонента устанавливается относительно определённой реализации и задачи. Другая система может обеспечивать тот же результат иным способом.

Такой анализ имеет прямую ценность для проектирования кандидатов на опыт. Он препятствует накоплению декоративных компонентов и помогает выбрать механизмы с доказанным функциональным вкладом. Феноменальная

гипотеза получает более точный предмет, даже если окончательный ответ о наличии опыта ещё не найден.

Четыре вопроса к одной записи состояния

В модели M2 запись A действительно существует: это проверяется состоянием программы. Её влияние на повторное A проверяется сравнением с удалённой историей при прежнем входе. Доступ к записи у отдельного монитора потребует проверить канал чтения; наличие записи само по себе его не создаёт. Наконец, если модель даёт нужные числа, ещё необходимо объяснить, почему эта организация относится к переживаемому содержанию. Таким образом, один журнал может поддерживать несколько разных проверок, но каждый переход требует собственного наблюдения.

Что теперь установлено

Профиль помогает не суммировать несопоставимые признаки. Статус неизвестного свойства указывает конкретную следующую работу вместо общего впечатления о сознательности.

Практика 32 Проведите аудит одного эпизода

Зачем выполнять: Разобрать один эпизод реализации на четыре самостоятельных утверждения.

Что понадобится: Причинная карта, абляция как отключение конкретного компонента, различие декодирования и использования состояния. Сохраняйте одинаковые доступные сведения у сравниваемых версий.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Выберите повторное A в модели источников или один шаг агента. Выпишите вход, старое состояние, изменение, выход и получателя выхода.
2. Назовите связь для отключения и ожидаемый результат до проверки. Убедитесь, что отключается именно она.
3. Составьте четыре строки профиля: реализация, функция, внутренний доступ, возможный опыт. Для каждой укажите выполненную проверку или неизвестность.
4. Предложите одну задачу, в которой функция действительно востребована, и контроль, сохраняющий сопоставимые сведения и ресурсы.

Что сохранить: Карта одного эпизода, причинный тест и профиль с раздельными статусами.

Подсказка: Внешний классификатор может извлекать содержание, которого агент не использует. Для внутреннего доступа нужен получатель информации внутри исследуемого контура.

Исследовательское продолжение: Проверьте текстовый отчёт при независимом изменении внутреннего состояния и внешней истории. Сравните с объяснением, формирующимся только по публичным действиям.

Разбор и ответы — приложение Е, глава 32, с. 309

Самопроверка

1. Доказывает ли наличие класса Monitor в программе работающий самоконтроль?
2. Чем внешнее декодирование состояния отличается от внутреннего доступа?
3. Почему выполнение всех самостоятельно выбранных требований не доказывает достаточность этих требований для опыта?

Мы научились проверять конкретную реализацию. Теперь уточним правила формализации и выясним, способен ли выбранный анализ вообще различить механизмы, которые мы хотим сравнить.

Часть 9

Исследовательская работа

Глава 33

Формализация и симуляция

Проверка программы не заканчивается правильным вычислением одного примера. Два набора параметров могут давать похожий итог, а разные модели могут оказаться неразличимыми в выбранной среде. В этой главе формально разделим генератор, доступные наблюдения, прогноз и измерение. Затем выполним восстановление параметров и моделей на искусственных данных, где породивший механизм известен. Это подготовка анализа, который позднее будет работать с неизвестной причиной.

Формализация делает обязательства видимыми

Словесная модель допускает полезную неопределённость на раннем этапе. Но для проверки потребуется решить, что именно происходит между двумя событиями. Сохраняется ли история, пока источник отсутствует? Изменяется ли вероятность скрытой причины каждый момент или только при новом наблюдении? Обновляет ли противоречие одну запись или всю систему? Эти решения способны изменить предсказание, хотя краткое описание модели остаётся прежним.

Формализация означает не обязательное усложнение математики, а устранение таких неоднозначностей. Вход, состояние, переход и измерение должны образовать процедуру, которую другой исследователь сможет воспроизвести. Программа помогает проверить последствия правил. При этом совпадение кода со спецификацией и пригодность спецификации для объяснения опыта остаются разными вопросами.

Задайте учебную среду отдельно от модели

Расширим пример двух источников А и Б. Каждый имеет скрытое двоичное состояние: общая причина либо разные причины. В учебной среде состояние источника сохраняется с вероятностью 0,9 при следующем обращении к нему. Наблюдаемый положительный признак появляется с вероятностью 0,75 при общей причине и 0,25 при разных причинах. Это сконструированная среда, не описание физиологического механизма.

Экспериментатор задаёт скрытые состояния и наблюдения. Модель получает только источник и признак. Поэтому можно отдельно измерять прогноз следующего признака и оценку скрытой причины. Если модель будет видеть скрытую причину напрямую, задача фильтрации исчезнет. Разделение среды и наблюдателя необходимо сохранить даже тогда, когда они написаны в одном файле.

Вероятность сохранения 0,9 соответствует коэффициенту удержания 0,8 из главы 25. После возможной смены причины прежнее убеждение сдвигается к 0,5 по правилу, которое мы уже использовали. Теперь у этого правила появился определённый смысл в генераторе данных. Другая среда могла бы требовать иной динамики.

Определение предсказываемого результата

Убеждение 0,8 означает вероятность общей причины. Оно не равно вероятности следующего положительного признака. При надёжности 0,75 следующий признак может оказаться отрицательным даже при общей причине. Прогноз положительного признака получается смешением двух возможностей: вероятность общей причины умножается на 0,75, вероятность разных причин — на 0,25; результаты складываются.

```
def predict_positive(prior_common, reliability=0.75):
    return (
        prior_common * reliability
        + (1.0 - prior_common) * (1.0 - reliability)
    )
```

При убеждении 0,8 прогноз положительного признака равен 0,65. Это проверяемое числовое различие. Если сравнить убеждение о причине с частотой признака напрямую, модель может ошибочно показаться плохо откалиброванной. Поэтому название каждого выходного числа должно указывать событие, вероятность которого оно представляет.

Человеческая оценка принадлежности добавит ещё один уровень. Она не является ни истинной причинной меткой, ни следующим сенсорным признаком. Связь внутреннего состояния с такой оценкой требует собственной измерительной модели. Формализация помогает заметить эти переходы до анализа результатов.

Запишите время модели

В нашем простом M2 удержание применяется при следующем обращении к определённому источнику. Если A долго отсутствует, его запись не ослабляется на каждом промежуточном событии Б. Это удобное правило для

учебного расчёта. Но гипотеза о человеческой памяти может требовать зависимости от прошедшего времени.

Вторая версия будет ослаблять запись за каждый временной шаг. При коэффициенте 0,8 и двух пропущенных шагах влияние прежнего отклонения последовательно умножится на 0,8 два раза. Теперь длительность отсутствия станет самостоятельным предиктором. Две версии можно сравнить, предъявляя одинаковое число событий с разными промежутками.

Не исправляйте первую версию незаметно, увидев неудобный результат. Изменение времени модели означает новую гипотезу. Сохраните прежние правила, опишите причину пересмотра и получите новые предсказания. Это особенно важно для связи с опытом: субъективная непрерывность не должна объясняться произвольно меняющимся понятием времени.

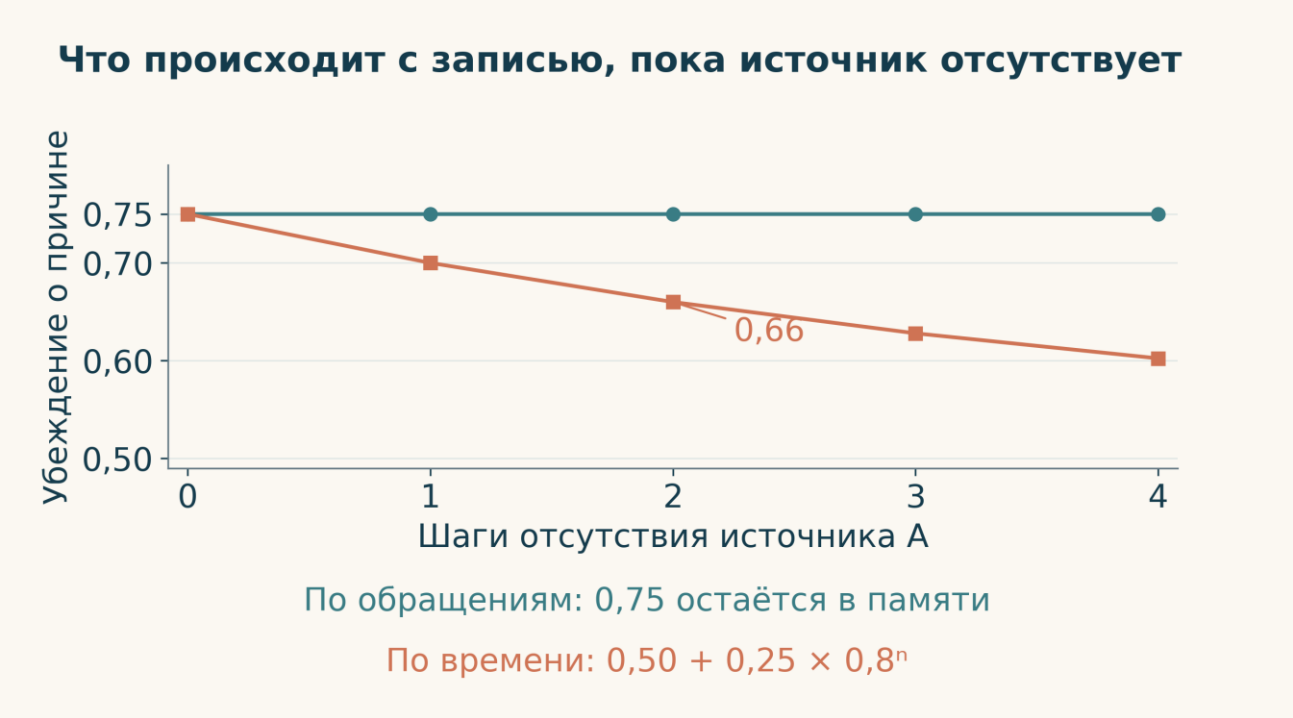

Рисунок 82. Две явно разные модели времени памяти. Точный учебный расчёт при начальной записи 0,75 и удержании 0,8. Показано сохранённое состояние во время отсутствия A, до нового обращения и до нового наблюдения: за два пропущенных шага временная версия приходит к 0,66, а версия по обращениям хранит 0,75.

Проверьте значения параметров

Два механизма могут дать одно и то же конечное число. Убеждение 0,75 при надёжности положительного признака 0,75 обновляется до 0,9. Исходное убеждение 0,5 при надёжности 0,9 тоже обновляется до 0,9. По одному итоговому ответу нельзя определить, связано ли высокое значение с сильной историей или сильным новым свидетельством.

Чтобы различить параметры, нужны другие условия. Можно независимо оценивать различение текущего признака, менять длину предыстории и добавлять противоречащие события. Затем проверить, остаются ли разные наборы параметров практически неразличимыми. Неопределённость параметров является результатом анализа, а не препятствием, которое нужно скрыть выбором одного красивого набора.

Попробуйте восстановление на искусственных данных. Задайте параметры, сгенерируйте последовательность, затем поручите анализу найти их без доступа к исходным значениям. Если процедура регулярно путает удержание с надёжностью, исследование человека столкнётся с тем же затруднением. Увеличение количества однотипных событий может не устранить структурную неоднозначность.

Проверьте восстановление модели

Теперь создайте данные отдельно с помощью M0, M1 и M2, добавив заранее заданную изменчивость наблюдения. Анализ должен сравнить все модели для каждого набора. В идеале он предпочитает породивший данные механизм, но это нельзя предполагать. Возможно, условия почти не различают конкурентов или правило выбора систематически предпочитает более гибкую модель.

Восстановление параметров и моделей рассматривается в методологической работе Wilson и Collins. Наши конкретные генераторы и последовательности служат самостоятельным учебным упражнением по этой логике. Успешная проверка здесь показывает способность анализа работать в известных условиях; она не доказывает, что реальные данные порождены одной из рассмотренных моделей. [43]

Полезно добавить четвёртый генератор, который не совпадает ни с одним кандидатом. Тогда можно увидеть, как процедура выбирает лучшую из неверных моделей. Даже убедительная победа в ограниченном наборе конкурентов не означает, что найден действительный механизм. Проверка абсолютного качества прогноза должна дополнять относительное сравнение.

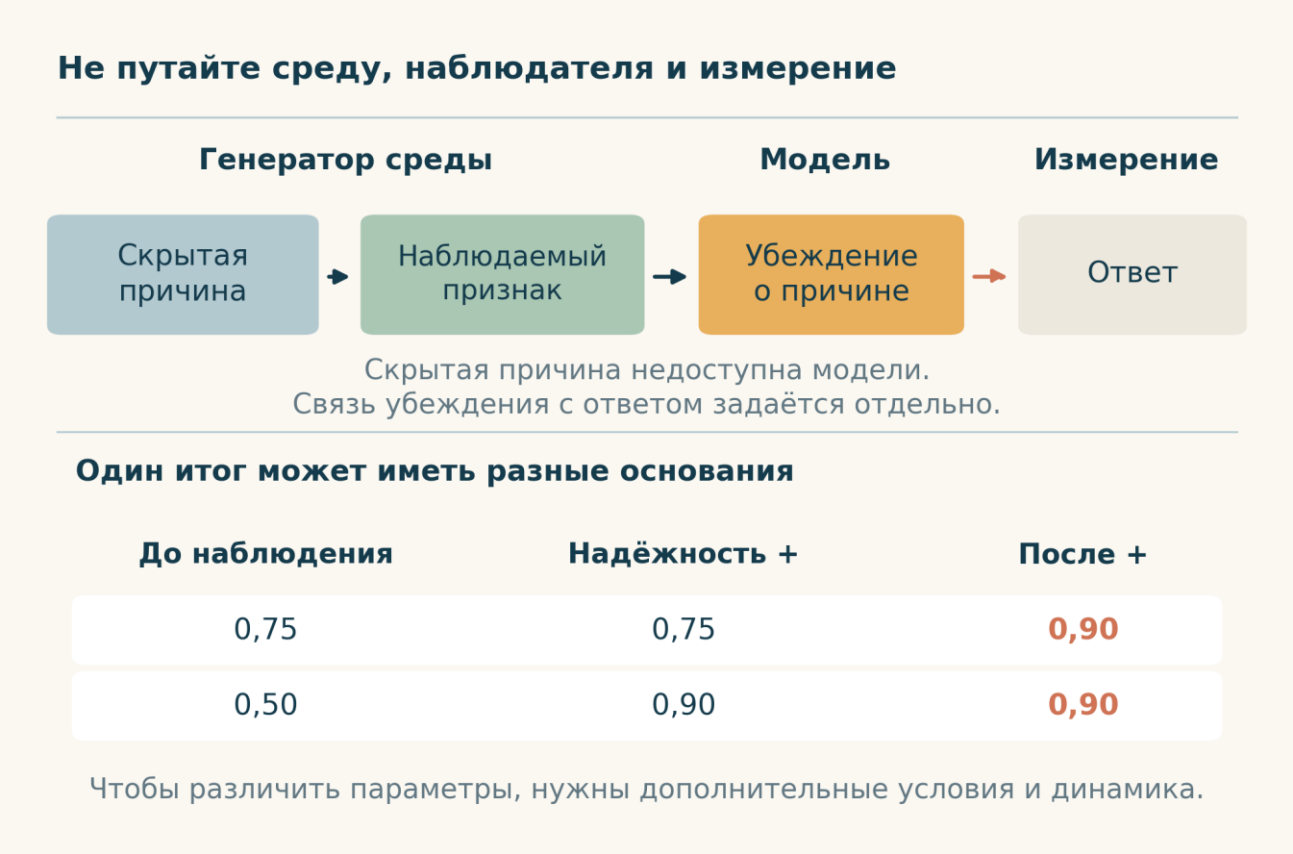

Рисунок 83. Граница доступной информации и неоднозначность параметров. Учебная схема и точный расчёт. Верхняя часть отделяет скрытую причину в генераторе среды от доступного модели наблюдения; пунктир обозначает отдельно задаваемую модель формирования ответа. В нижнем примере положительное свидетельство с надёжностью 0,75 при предварительном убеждении 0,75 и с надёжностью 0,90 при убеждении 0,50 даёт один итог 0,90. Здесь предварительное убеждение задано непосредственно перед наблюдением: дополнительное ослабление памяти не применяется. Один итог не определяет оба параметра.

Сохраните воспроизводимость

Для каждого запуска записывайте версию правил, параметры, исходное состояние и порядок событий. При случайной генерации фиксируйте начальные значения генераторов. Но одного seed недостаточно, если изменилась последовательность обращений к случайности. Дополнительное наблюдение не должно непреднамеренно сдвигать весь будущий сценарий сравниваемой версии.

Разделяйте генерацию условий и решения агента. Можно заранее подготовить потенциальные наблюдения на каждом шаге, а действие агента будет определять, какое из них откроется. Если же действие физически меняет будущую среду, это изменение должно быть частью причинной модели. Сравнимость не означает искусственного запрета на последствия действия.

Сохраняйте также необработанные результаты и программу их сводки. Иначе исправление одного фильтра может незаметно изменить числа в тексте. В итоговом наборе читатель должен видеть, как из последовательности получились прогнозы, из прогнозов — ошибки, из ошибок — вывод. Эта цепочка важнее декоративной сложности интерфейса.

От симуляции к исследовательскому выводу

Симуляция может установить, что механизм реализуем, какие следствия он имеет и при каких условиях модели различаются. Она также может показать невозможность желаемого результата при выбранных правилах. Но искусственные данные, сгенерированные самой моделью, не являются независимым подтверждением её применимости к человеку или биологической сети.

Для гипотезы главы 26 необходимо дополнительное описание исполнения. Если существенна вычислительная организация, покажите, что программа действительно её реализует, включая ответы на вмешательства. Если существенны определённые физические взаимодействия, потребуется обоснование на этом уровне. График симуляции не подменяет вопрос о том, какой процесс фактически происходит в носителе.

Исполнимая процедура восстановления

В дополненном практикуме генератор M0 создаёт независимые двоичные признаки с вероятностью 0,5. Генератор M1 хранит одну скрытую причину для всей последовательности, генератор M2 отдельную причину каждого чередующегося источника. При обращении причина сохраняется с вероятностью 0,9, а положительный признак появляется с вероятностью 0,75 для причины 1 и 0,25 для причины 0. Первоначальные причины равновероятны. Так генерирующая среда задана независимо от анализирующей модели.

Восстановление моделей использует 30 независимых последовательностей по 120 событий для каждого из трёх генераторов. Анализирующие M0, M1 и M2 получают одинаковые метки источников и признаки, но не скрытые причины. В каждой пробе сначала сохраняется вероятность положительного признака, затем открывается его значение и обновляется убеждение. Основной критерий здесь средняя квадратичная ошибка. При точном равенстве результатов сохраняется ничья; случайное назначение победителя не добавляет различимости.

Для восстановления параметров M2 используется заранее заданная сетка удержания от 0 до 1 с шагом 0,1 и надёжности от 0,55 до 0,95 с шагом 0,05.

Для каждой пары вычисляется ошибка полной траектории на известных искусственных данных. Этот небольшой поиск делает видимыми соседние почти равные решения. Он является учебной процедурой, а не универсально лучшим методом оценки параметров.

Контроль с одним источником обязан дать равенство M1 и M2. Дополнительно сохраняются все серии, где выбранный анализ не предпочёл породивший механизм. Матрица таких выборов показывает способность именно этого объёма данных и критерия различать именно эти генераторы. Для другой среды, другого шума или другой сетки проверка выполняется заново.

После проверки инструмента можно задать основной эксперимент главы 34. Использование известных искусственных причин здесь законно для оценщика, который проверяет восстановление. Передача этих причин самому прогнозирующему кандидату нарушила бы задачу и сделала успех неинформативным.

Один итог не определяет происхождение уверенности

Убеждение 0,75 при правдоподобии положительного признака 0,75 обновляется до 0,9. Исходное 0,5 при надёжности 0,9 тоже даёт 0,9. Одного конечного числа недостаточно, чтобы восстановить силу прошлого и нового свидетельства. Поэтому анализу нужна траектория, противоречащие признаки и разные условия. Проверка на данных известного генератора показывает, где процедура путает параметры, а где путает целые модели. Если конкуренты вычислительно одинаковы, такая проверка должна обнаружить равенство, а не искусственно назначить победителя.

Что теперь установлено

Симуляция позволяет исследовать следствия и различимость правил. Она не является независимым подтверждением их применимости к человеческому опыту. Теперь проверим сам инструмент сравнения.

Практика 33 Исследуйте неопределённость механизма

Зачем выполнять: Воспроизвести восстановление на известных генераторах и найти область неоднозначности.

Что понадобится: Обновление M0-M2, различие вероятности скрытой причины и вероятности признака, единица независимого повторения. В практикуме добавлен исполнимый пример сравнения и восстановления.

Ориентир по времени: 40–60 минут. Можно делать с перерывами; это не норматив скорости.

1. Выполните python teaching_lab.py --chapter 33. Генераторы и сетка параметров описаны перед практикой; все прогнозы записываются до открытия текущего признака.
2. Рассмотрите ошибки M0, M1 и M2 на каждой серии, затем таблицу выбора при известном генераторе. Сохраните все ошибки классификации моделей.
3. Для M2 сравните несколько значений удержания и надёжности. Найдите соседние сочетания, близкие по ошибке, и проверьте промежуточные состояния.
4. В одноканальном условии проверьте равенство M1 и M2. Сравните время по обращениям и календарным шагам на последовательности с отсутствующим A.

Что сохранить: Результат восстановления, известный контроль равенства и одно условие, усиливающее различение параметров.

Подсказка: Не отбрасывайте серии, в которых победил не породивший данные кандидат. Они характеризуют выбранную процедуру и объём информации.

Исследовательское продолжение: Добавьте генератор, не совпадающий ни с одним кандидатом, и оцените абсолютное качество лучшей модели. Проверьте восстановление при другой изменчивости и количестве наблюдений, зафиксировав новый план до запуска.

Разбор и ответы — приложение E, глава 33, с. 309

Самопроверка

1. Почему вероятность общей причины отличается от вероятности положительного признака?
2. Доказывает ли восстановление M2 на данных самой M2 её истинность для человека?
3. Почему изменение правила забывания требует отдельной версии модели?

Теперь известны ограничения самого анализа. В следующей главе зафиксируем, как он будет применяться к новым данным, чтобы результат не выбирал правила собственной проверки.

Глава 34

Протокол до сбора данных

Предварительный план связывает вопрос, будущие данные и решение по результату. Он нужен не только для защиты от произвольного выбора удачного графика: с ним становится видно, проверяет ли процедура именно нужный механизм. Здесь подготовим исполнимый протокол сравнения общей и раздельной истории. Его логика затем переносится на человеческие и биологические проекты с их дополнительными уровнями группировки и требованиями измерения.

Зачем описывать решения заранее

Представьте, что вы получили несколько графиков. На одном M2 лучше M1, на другом различие почти исчезло. Можно выбрать первый график, изменить временное окно и объяснить, почему именно этот показатель всегда казался главным. Такая история может звучать убедительно даже для самого автора. Предварительный протокол сохраняет сведения о том, что вы действительно ожидали и собирались проверять.

Предварительная регистрация не запрещает исследовать неожиданные результаты. Она разделяет проверку заранее сформулированного ожидания и создание новой гипотезы по увиденным данным. Это различение лежит в центре методологического предложения Nosek и коллег. Для нашей работы существенна прозрачность решений, а не церемония размещения документа. [10]

Ниже приведён учебный образец протокола. Он не является сообщением о зарегистрированном или выполненном исследовании. Его можно заполнить для собственной будущей проверки, сохранив дату и версию. Если данные уже изучены, это нужно указать и выбрать новый независимый материал для подтверждающего этапа.

Начните с одного главного вопроса

Наш вопрос: помогает ли раздельная история источников предсказывать их следующие признаки лучше общей истории? Это вычислительный этап проекта, подготовленный в главе 33. Он не проверяет непосредственно человеческую принадлежность и не устанавливает наличие опыта у программы. Зато позволяет точно определить механизм, который затем можно сопоставлять с более сложными данными.

Главная гипотеза утверждает преимущество M2 перед M1 в среде с двумя независимыми устойчивыми источниками. Контрольная гипотеза утверждает совпадение этих моделей, когда источник один и правила удержания одинаковы. M0 служит ориентиром без переноса истории. M1И, реализованный теми же отдельными записями, совпадает с M2 и не считается отдельной эмпирически различимой альтернативой.

Запишите также основания гипотезы: M1 смешивает свидетельства от разных источников, а M2 сохраняет их раздельно. Это механизм предполагаемого преимущества. Слова «более умная модель» ничего к нему не добавляют. Если условия не создают конфликтующих историй, ожидаемое различие может оказаться очень маленьким.

Полностью определите будущие данные

В учебном протоколе планируется двести независимых последовательностей по сто двадцать событий. Эти числа задают воспроизводимый объём упражнения и не являются нормой достаточности научной выборки. Половина последовательностей содержит один источник, половина — два. В варианте двух источников обращения чередуются. Первоначальные скрытые состояния выбираются равновероятно.

Каждый источник при следующем обращении сохраняет скрытое состояние с вероятностью 0,9. Вероятность положительного наблюдения составляет 0,75 при общей причине и 0,25 при разных причинах. Моделям известны эти параметры, но не скрытые состояния. Такой вариант проверяет механизм в благоприятно определённых условиях; устойчивость при неверных предположениях станет отдельным исследованием.

До генерации закрепите алгоритм случайности и список начальных значений. Удобно использовать последовательные значения от одного до двухсот и сопоставить условие с номером заранее. Не подбирайте удачный seed после просмотра результатов. Если требуется распределение условий вперемешку, правило этого распределения тоже фиксируется до запуска.

Определите прогноз и момент его записи

Перед каждым событием модель вычисляет вероятность положительного признака. Сначала старое убеждение обновляется с учётом возможной смены причины, затем из него получается прогноз наблюдения. Только после сохранения прогноза открывается текущий признак и изменяется убеждение. Этот порядок препятствует использованию ответа при предсказании самого ответа.

Основной показатель — средняя квадратичная ошибка вероятностного прогноза, обычно называемая Brier score. Для положительного события фактическое значение равно единице, для отрицательного — нулю. Из прогноза вычитается фактическое значение, разность возводится в квадрат; затем ошибки усредняются. Меньшее значение означает лучший вероятностный прогноз в выбранном смысле.

Например, прогноз 0,7 перед положительным событием даёт ошибку 0,09, а прогноз 0,4 — 0,36. Эти числа иллюстрируют вычисление, а не являются результатами будущей проверки. Показатель относится к наблюдаемому признаку. Его нельзя назвать мерой квалиа или напрямую использовать как шкалу достоверности человеческого самоощущения.

Зафиксируйте главное сравнение

Для каждой последовательности вычислите разность ошибок M1 и M2. Положительная разность означает преимущество M2. Основной итог — среднее парное различие в условии двух источников с оценкой неопределённости по независимым последовательностям. События внутри одной последовательности связаны историей, поэтому их нельзя считать таким же количеством независимых повторений всего исследования.

В условии одного источника ожидается точное совпадение выбранных реализаций M1 и M2. Это контроль логики программы. Если результаты различаются, сначала нужно искать различие входов, обновлений или начальных состояний. Такой случай не должен немедленно превращаться в неожиданное открытие о памяти.

Сравнение M2 с M0 и графики по отдельным шагам можно заранее обозначить дополнительными. Если после просмотра данных появляется новое временное окно, его анализ остаётся полезным, но получает исследовательский статус. Главное сравнение не меняется задним числом только потому, что дополнительное оказалось красивее.

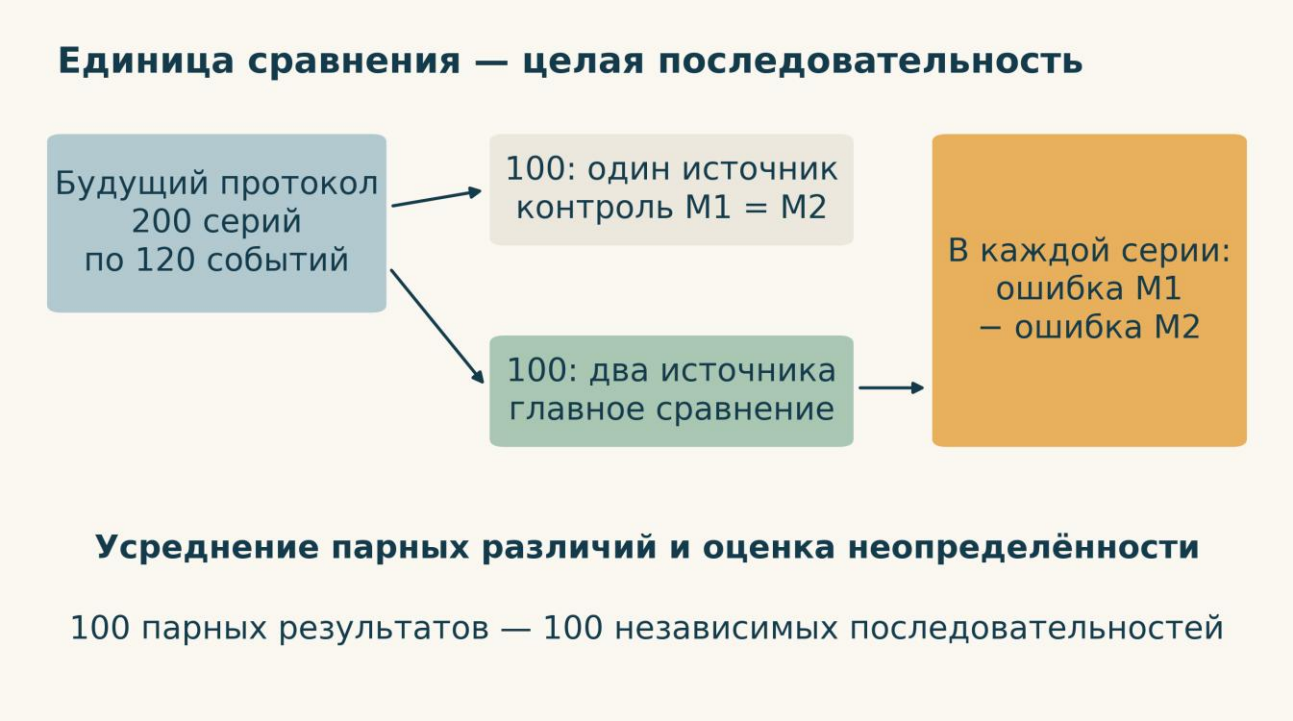

Рисунок 84. Структура данных предварительного протокола. Числа взяты из будущего учебного протокола главы 34 и задают объём упражнения, а не норму научной выборки. Для основного сравнения неопределённость оценивается по ста независимым последовательностям с двумя источниками, сохраняя парность результатов M1 и M2.

Укажите критерий содержательного результата

Положительное среднее различие может быть настолько маленьким, что почти не влияет на использование модели. Поэтому кроме направления задайте интересующую величину эффекта или требуемую точность оценки. В учебном проекте можно заранее выбрать разницу Brier 0,01 как ориентир для обсуждения. Это условный порог упражнения, а не установленная универсальная граница значимости.

Для серьёзного проекта порог выводится из задачи: какой выигрыш оправдывает усложнение, какое различие способны разделить конкурирующие объяснения? Затем симуляция помогает выбрать объём данных для нужной точности. Если заранее определённый бюджет недостаточен, вывод должен сохранить эту неопределённость, а не подменять её словами «эффекта нет».

Нельзя устанавливать правило остановки «добавлять данные, пока модель не победит». Можно заранее выбрать фиксированный объём или обоснованную последовательную процедуру. В обоих случаях правило входит в протокол. Причины вынужденной остановки и фактический объём должны остаться видимыми в итоговом отчёте.

Подготовьте правила обработки ошибок

В цифровой задаче следует заранее определить, какие результаты являются технически невозможными: вероятность вне допустимого диапазона, отсутствие события, несовпадение сценариев сравниваемых моделей. Такой случай требует исправления реализации и повторного запуска с сохранением версии. Исключение неудобного, но допустимого результата не является исправлением ошибки.

В исследовании человека или биологической сети будут другие основания: незавершённая процедура, утрата записи, нарушение заранее заданного качества измерения. Правила необходимо связать с пригодностью данных, а не с направлением эффекта. Для биологических данных дополнительно фиксируется единица независимого сравнения и доступная информация об образцах.

Укажите, кто и когда принимает решение об исключении. Если решение зависит от знания группы или результата, это следует раскрыть. Протокол не делает исследование безошибочным; он позволяет оценить, как решения могли повлиять на вывод.

Сохраните место для открытия

Допустим, M2 особенно выигрывает после возвращения источника с противоречащей историей. Этот рисунок не был главным прогнозом, но может подсказать новый механизм. Запишите его, проверьте технические объяснения и сформулируйте следующий различающий вопрос. Затем подготовьте новую последовательность, на которой обновлённая идея рискует оказаться неверной.

При таком подходе исследовательский анализ не становится второсортным. Он выполняет другую работу: создаёт содержание будущей проверки. Проблема возникает только при смешении двух этапов, когда обнаруженное различие представляют заранее ожидаемым и тем самым завышают силу свидетельства.

Как один вопрос определяет анализ

Вопрос о пользе отдельных историй требует хотя бы двух различаемых источников. Для каждого независимого сценария можно вычислить среднюю ошибку общей памяти и раздельной памяти, а затем их разность. Положительная разность означает преимущество M2. Если усреднять все события как независимые повторы, исчезнет общая история внутри сценария и точность может быть завышена. Поэтому в учебном протоколе повторной

единицей служит целая последовательность, а одноканальная среда является контролем равенства.

Что теперь установлено

План должен позволять выполнить главное сравнение без выбора показателя после просмотра графика. При этом он сохраняет отдельное место для последующих открытий и новых гипотез.

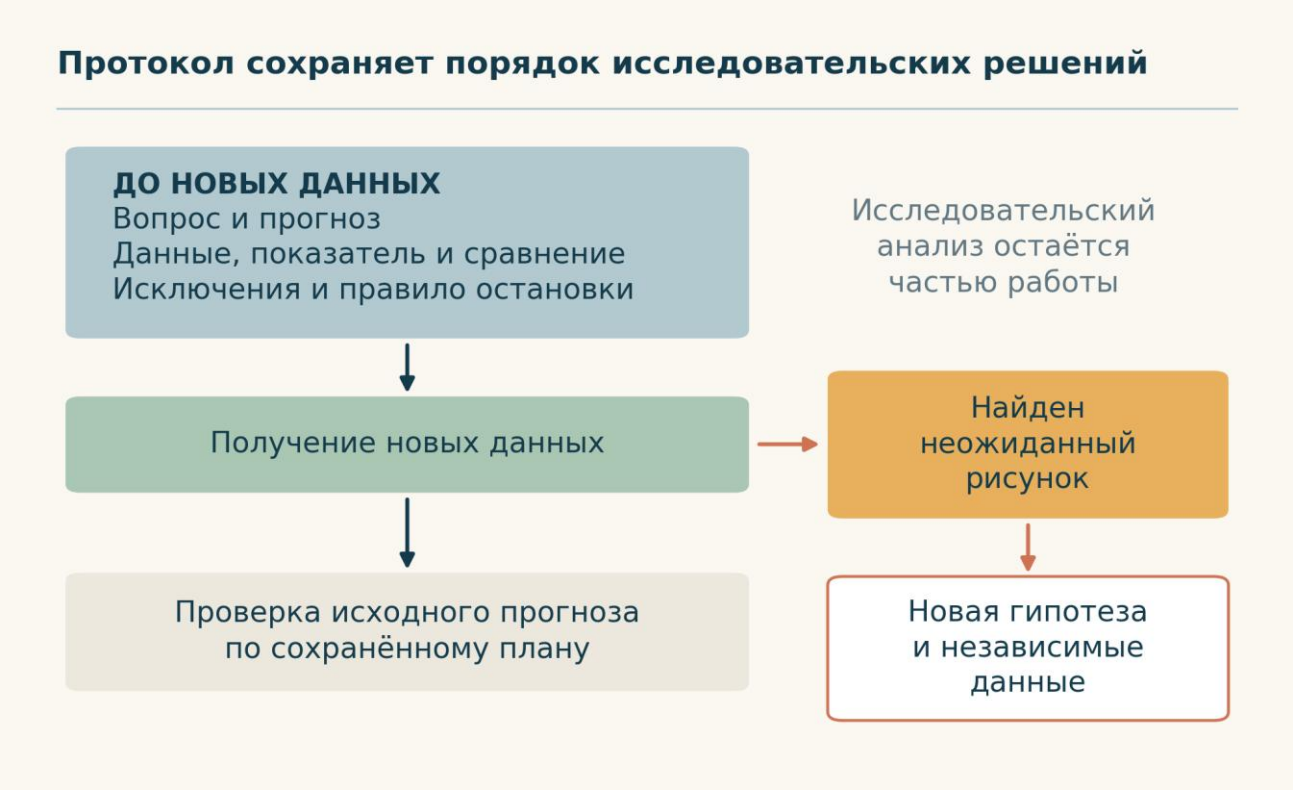

Рисунок 85. Предварительный план и последующее открытие. Методическая схема: заранее сохранённый протокол определяет подтверждающую проверку. Неожиданный рисунок в полученных данных может породить новую гипотезу и следующий проект с независимыми данными. Эта ветвь полезна, но не превращает найденный после анализа эффект в заранее предсказанный. Схема не сообщает о выполненной регистрации или эксперименте.

Практика 34 Заполните собственный протокол

Зачем выполнять: Составить протокол, который другой человек способен исполнить без дополнительных решений за автора.

Что понадобится: Генератор и анализ главы 33; главный показатель, парное сравнение и единица независимого повторения. Различайте новый подтверждающий материал и уже просмотренные результаты.

Ориентир по времени: 35–50 минут. Можно делать с перерывами; это не норматив скорости.

1. Используйте образец protocol_source_history.txt в практикуме. Задайте вопрос о парной разности Brier для M1 и M2 в двухисточниковой среде.
2. Зафиксируйте 200 независимых серий по 120 событий: 100 с одним источником и 100 с двумя. Назовите правила генерации, параметры, seed и момент прогноза.
3. Сохраните главный показатель, условный содержательный ориентир 0,01, единицу анализа, обработку технических ошибок и правило остановки до нового запуска.
4. Укажите, какие данные уже видели, и отдельно перечислите поисковые анализы. При собственном новом исследовании используйте новые сценарии.

Что сохранить: Датированная версия полного протокола и явная граница подтверждающего и поискового этапов.

Подсказка: Протокол, записанный после знакомства с результатом, не становится предварительным от смены названия. Сохраните происхождение решений.

Исследовательское продолжение: Подготовьте соответствующий протокол для реальных измерений, добавив критерии пригодности и уровни участников, образцов, доноров или партий. Обоснуйте требуемую точность относительно вопроса.

Разбор и ответы — приложение E, глава 34, с. 310

Самопроверка

1. Запрещает ли предварительная регистрация неожиданный анализ?
2. Почему прогноз нужно сохранять до открытия текущего признака?
3. Являются ли двести последовательностей универсально достаточным объёмом исследования?

После сохранения плана появятся результаты. Глава 35 покажет, как они меняют вывод, когда преимущество мало, компонент не включался или лучшие кандидаты всё равно ошибаются.

Глава 35

Проверка и пересмотр

Исследование должно закончиться решением, связанным с полученным результатом. Иногда данные поддержат механизм; иногда ограничат величину эффекта; иногда окажутся неразличающими. Начнём с восстановления информационного порядка и известных контролей, затем оценим размер и неопределённость. Только после этого будем менять объяснение. Такая последовательность помогает не принять техническое улучшение числа за научное продвижение.

Результат должен изменить наши ожидания

Исследование заканчивается не получением числа, а уточнением того, что теперь допустимо утверждать. Положительный результат может поддержать механизм, выявить полезную инженерную возможность или показать лишь преимущество одной реализации. Отрицательный результат может опровергнуть конкретный прогноз, обнаружить нечувствительность задачи или остаться неопределённым. Для каждого случая нужен свой следующий шаг.

Возьмём проект раздельной памяти источников. Мы ожидали, что M2 будет лучше общей памяти M1 прогнозировать признаки двух устойчивых источников. После запуска следует спросить, что именно сравнивалось: одинаковые наблюдения, одинаковый момент прогноза, одинаковые возможности обновления? Только затем можно обсуждать размер преимущества и его отношение к исходной идее.

Сначала восстановите причинную последовательность

Самая убедительная ошибка — та, которая улучшает результат. Представьте, что программа сначала обновила убеждение по текущему признаку, а затем сохранила его как прогноз этого же признака. Ошибка прогноза уменьшится, но модель воспользуется ответом, который должна была предсказать. Большой выигрыш здесь свидетельствует о нарушении процедуры.

Выберите несколько событий и восстановите их полностью: старое состояние, доступная информация, предварительный прогноз, открывшийся признак, новое состояние. Проверяйте не только итоговый график, но и порядок этих переходов. Для активного агента добавьте действие, его стоимость и фактически полученное дополнительное наблюдение. Именно такая запись позволяет установить, какой информацией располагала система.

Затем выполните контроль, для которого ответ известен из определения. В одноканальной среде наши М1 и М2 должны совпадать. При нулевом удержании все версии забывают прежнее состояние. Нарушение этих равенств требует объяснения в правилах реализации. Исправление ошибки сопровождается новой версией и повторным получением результатов; прежние числа сохраняются как отменённые.

Оцените величину и неопределённость

Пусть в четырёх искусственных последовательностях разность ошибок М1 и М2 составила 0,018, 0,006, −0,004 и 0,012. Это специально придуманные учебные значения, а не результат выполненного опыта. Средняя разность равна 0,008: в среднем М2 лучше, но одна последовательность дала противоположный результат. Одного знака среднего недостаточно для содержательного заключения.

В протоколе предыдущей главы условным ориентиром служила разность 0,01. Наше иллюстративное среднее меньше этого ориентира. При этом четыре последовательности не позволяют объявить отсутствие интересующего преимущества без оценки неопределённости. Необходимо выяснить, какие величины совместимы с данными и насколько точна оценка относительно заранее выбранного вопроса.

Для последовательностей с общей внутренней историей единицей повторения остаётся целая последовательность. Нельзя превратить четыре запуска по сто событий в четыреста независимых повторов механизма. Если проводится повторная выборка для оценки неопределённости, переносить нужно целые независимые последовательности вместе с их парными результатами. Для людей, доноров и клеточных образцов соответствующие уровни группировки определяются устройством исследования.

Отчёт должен содержать величину различия, её неопределённость и смысл выбранного масштаба. Формулировка «данные недостаточно точны, чтобы различить нулевое и практически интересующее преимущество» полезнее заявления «модели одинаковы». Равенство требует собственного обоснования. Иногда оно, как для М2 и М1И в нашем примере, следует непосредственно из правил, и тогда статистическое сравнение излишне.

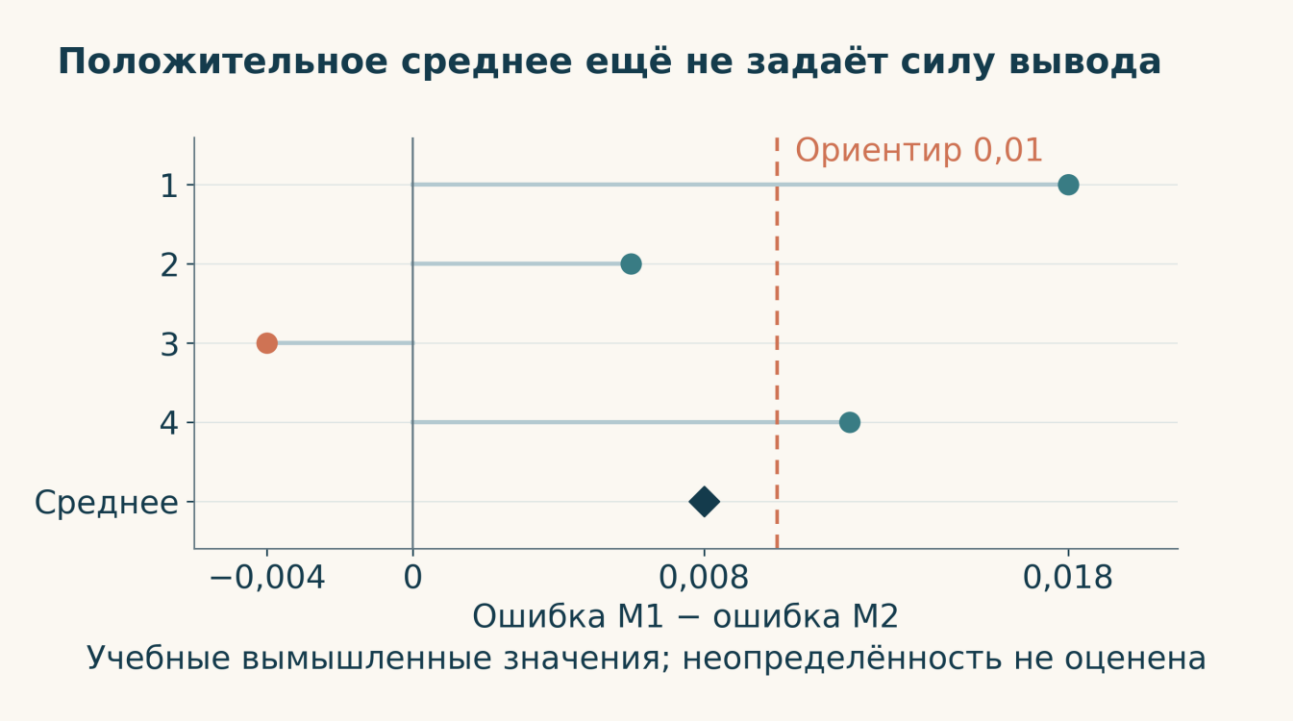

Рисунок 86. Иллюстративные парные различия и условный ориентир. Четыре специально придуманных значения главы 35 дают среднее 0,008 при условном ориентире 0,01. График не содержит рассчитанного доверительного интервала: по одному знаку среднего или его положению относительно ориентира нельзя заключить ни равенство моделей, ни отсутствие интересующего преимущества.

Поиск источника преимущества

Среднее по всем событиям объединяет разные ситуации. Модель может особенно выигрывать при возвращении источника, проигрывать сразу после его смены и совпадать с конкурентом на длинных однородных участках. Такой рисунок помогает проверить объяснение преимущества. Если раздельная история важна именно при конфликте историй, соответствующее условие должно иметь особое значение.

Разделите заранее запланированный анализ и найденную после просмотра закономерность. Оба могут быть содержательными. Во втором случае сформулируйте новое предсказание и проверьте его на независимых последовательностях. Не следует удалять неудобные участки и затем описывать оставшийся выигрыш как общий результат первоначального теста.

Проверьте также абсолютное качество. Лучший из трёх прогнозов может оставаться систематически неверным. Например, все кандидаты могут приписывать высокую устойчивость источнику, который в новой среде часто меняется. Тогда преимущество M2 не устраняет общий дефект. Следующий шаг — исследовать модель изменения самой устойчивости, а не просто увеличить число записей памяти.

Отличите вклад компонента от его наличия

Абляция отвечает на вопрос о конкретной реализации в конкретной задаче. Если удаление памяти ухудшило результат, это показывает её вклад при сохранённых условиях сравнения. Отсюда ещё не следует, что без именно такой памяти задача вообще неразрешима. Другой механизм может хранить необходимые различия в иной форме.

Обратная ситуация возникает, когда удаление компонента ничего не меняет. Проверьте, использовался ли он. Если условие пересмотра стратегии ни разу не выполнилось, совпадение полной версии с фиксированной не оценивает полезность пересмотра в ситуациях, где он действительно происходит. Такой случай разобран в инженерном практикуме главы 30. Он выявляет границу выполненной проверки.

Чтобы получить различающий тест, задайте среду, для которой заранее ожидается включение компонента, и отдельно измерьте частоту включений. Затем сравните результат с контролем, имеющим сопоставимый доступ к информации и вычислительным ресурсам. Новые условия нужно назвать новым исследованием. Предыдущий запуск не становится положительным задним числом.

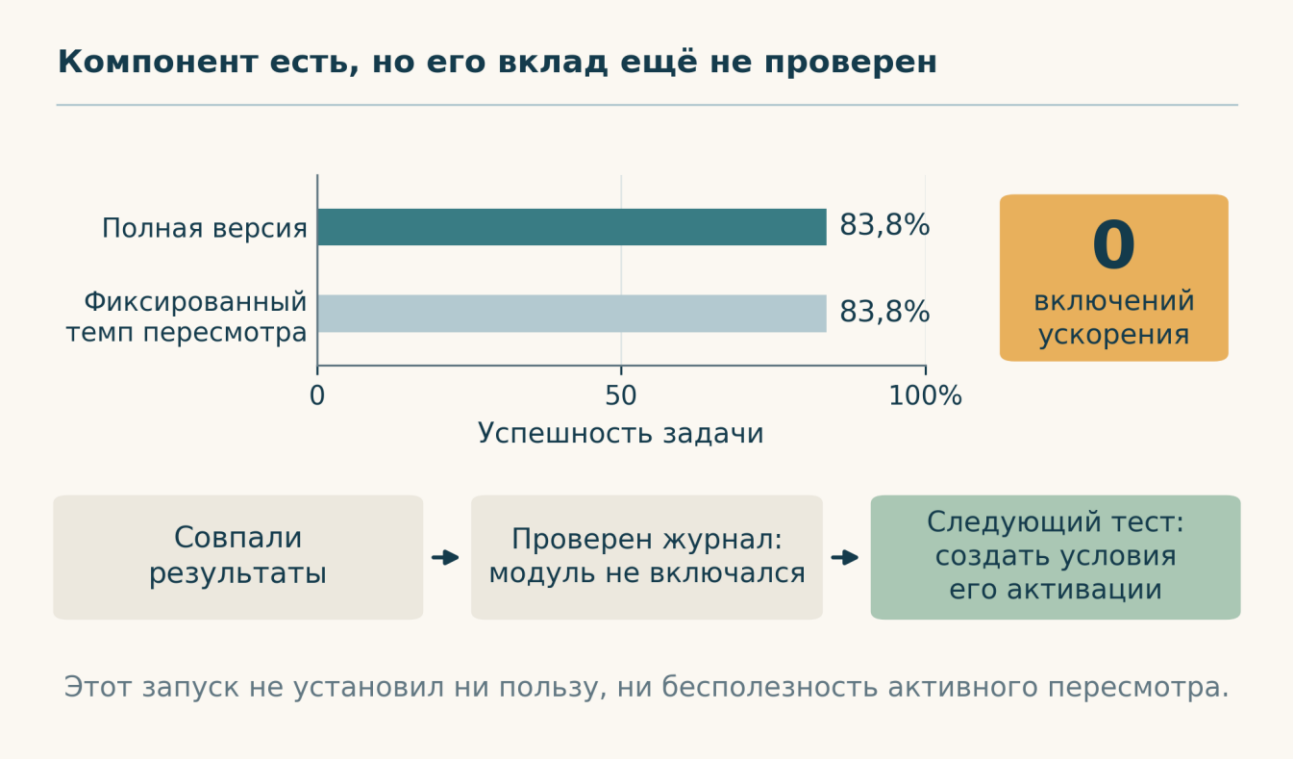

Рисунок 87. Честный неразличающий результат работы 8. Результат сохранённого программного запуска, mechanisms_results.json, работа 8: 1000 шагов одного сценария. Полная версия и версия fixed_revision имеют по 838 успешных шагов (83,8%); в полной версии ускоренное обучение не включилось ни разу. Это точное совпадение данного запуска, а не статистическая эквивалентность механизмов во всех средах. Следующий тест с активацией компонента будет новым исследованием. Показатели не оценивают квалиа.

Пересматривайте механизм открыто

Предположим, при увеличении времени отсутствия источника его влияние ослабевает сильнее, чем предсказывает M2. Один путь — каждый раз подбирать коэффициент удержания заново. Более содержательный путь — сформулировать новую модель времени, в которой запись изменяется и без обращения к источнику. Глава 33 показала, почему эти правила различаются.

У новой версии появляется рискованное следствие: при одинаковом количестве наблюдений длительность паузы должна самостоятельно влиять на результат. Конкурентная версия может объяснять ослабление числом вмешавшихся событий. Поэтому следующий эксперимент разводит продолжительность паузы и количество таких событий. Пересмотр становится научным, когда создаёт новое различие, а не только приближает старую кривую.

Сохраняйте таблицу версий: какое наблюдение мотивировало изменение, какое правило заменено, что теперь ожидается в новом условии. Эта история защищает содержательную преемственность проекта. Неудачная версия

может остаться полезной как контроль или частный случай, даже если перестала быть главным объяснением.

Вернитесь к исходному феномену

Теперь представим явно гипотетический результат человеческого исследования: M2 предсказывает динамику оценок принадлежности лучше M1, но общая модель идентичности M1И предсказывает её столь же хорошо. Обоснованный вывод состоит в поддержке раздельной истории причин или объектов. Утверждение об особом механизме самости пока не отделено от общего узнавания источника.

Следующий вопрос должен касаться этого различия. Можно сопоставить перенос истории телесного источника и обычного внешнего объекта, уравняв доступные признаки и трудность различения. Если вводятся дополнительные параметры только для тела, их основания задают заранее. Иначе «особая самость» превращается в название дополнительной свободы подгонки.

Для сильного принципа главы 26 требуется ещё один шаг: проверка предсказанных изменений качественной организации при направленных вмешательствах во взаимную динамику. Успех модели отчётов предоставляет исходный материал для такого проекта. Он не завершает аргумент о наличии опыта. Исследовательская программа продвигается, когда следующий опыт различает именно те объяснения, которые предыдущий оставил совместимыми.

Что разрешает средняя разность 0,008

Четыре специально придуманные разности M1 минус M2 равны 0,018; 0,006; минус 0,004; 0,012. Сумма 0,032 и среднее 0,008. Знак среднего поддерживает направление преимущества в этой иллюстрации, но одна серия имеет обратный знак, а среднее ниже учебного ориентира 0,01. Без оценки неопределённости нельзя объявить ни установленное важное преимущество, ни отсутствие такого преимущества. Следующее действие определяется тем, какой диапазон эффектов ещё совместим с данными.

Что теперь установлено

Вывод связан с размером, точностью и областью условий. Новый механизм должен порождать новый различающий прогноз, а не только удобнее описывать уже увиденную кривую.

Практика 35 Напишите решение по результату

Зачем выполнять: Принять обоснованное решение по одному сохранённому результату.

Что понадобится: Предварительный протокол, Brier, парная разность и активация компонента. Отсутствие убедительного эффекта не равно доказанному равенству.

Ориентир по времени: 30–40 минут. Можно делать с перерывами; это не норматив скорости.

1. Возьмите один выполненный запуск и восстановите несколько шагов: доступные данные, предварительный прогноз, открытый исход и новое состояние.
2. Проверьте контроль с известным результатом. Для истории это одноканальное равенство; для компонента с порогом это журнал активации.
3. Запишите величину различия, единицу повторения и доступную оценку неопределённости. Не заменяйте отсутствующий интервал одним знаком среднего.
4. Выберите решение: проверить перенос, уточнить правило, изменить неразличающие условия или отказаться от конкретной версии. Назовите два будущих исхода, ведущих к разным решениям.

Что сохранить: Короткое решение по результату с основанием и следующим тестом.

Подсказка: Если параметры пришлось изменить после просмотра, сохраните старую версию и назовите новую проверку отдельным исследованием.

Исследовательское продолжение: Разведите влияние длительности паузы и числа вмешавшихся событий в новой модели памяти. Получите прогноз до сбора новых данных и сравните с прежней версией.

Разбор и ответы — приложение Е, глава 35, с. 310

Самопроверка

1. Что означает совпадение версий с модулем и без него, если модуль ни разу не включился?
2. Почему преимущество M2 перед M1 не доказывает особый механизм самости?
3. Чем содержательный пересмотр отличается от произвольной подгонки?

Теперь можно собрать собственную небольшую работу, в которой вопрос, механизм, метод и вывод согласованы. Последняя глава предложит маршрут от воспроизведения к самостоятельному изменению.

Глава 36

Собственное исследование

Маршрут книги завершается исследованием, которое другой человек может повторить и содержательно оспорить. Для начала достаточно небольшого механизма: его ценность определяется ясностью вопроса и силой проверки. Раздельная история источников даст завершённый вычислительный пример. Следующий проект добавит взаимное определение признака и источника и покажет, какие новые основания потребуются для более сильного вывода об опыте.

От учебного результата к собственной задаче

Цель книги состоит в том, чтобы научиться строить и проверять объяснения опыта, а затем проектировать системы, которые на этих основаниях могут стать кандидатами на субъективный опыт. Итоговая работа должна соединить замысел, механизм и свидетельство. Масштаб реализации может быть небольшим, если вопрос определён точно и результат действительно меняет наше понимание выбранного механизма.

Ниже приведено завершённое учебное вычислительное мини-исследование. Оно показывает форму аргумента, которую можно воспроизвести и расширить. Его числовые результаты следуют из заданных правил; они не получены от людей или клеточных культур. После него мы сформулируем следующий проект, направленный к более сильной гипотезе о качественной организации и наличии опыта.

Учебное исследование раздельной истории источников

Вопрос. Достаточно ли раздельного хранения истории, чтобы одно и то же новое свидетельство по-разному влияло на оценку причины после возвращения источника? Дополнительный вопрос: требует ли полученное различие специального механизма самости или воспроизводится общей памятью идентичности объектов?

Метод. Задана последовательность четырёх наблюдений: А с положительным признаком, Б с отрицательным, снова А с положительным и снова Б с положительным. Метки источников доступны модели и в этом упражнении безошибочны. Положительный признак имеет вероятность 0,75 при общей причине и 0,25 при разных причинах. Начальная вероятность общей причины составляет 0,5.

Перед обработкой признака прежнее убеждение сдвигается к 0,5 с коэффициентом удержания 0,8. Затем применяется правило обновления главы 25. M0 каждый раз начинает с 0,5. M1 переносит одно убеждение через все события. M2 хранит отдельное убеждение каждого источника; удержание применяется при следующем обращении к этой записи. M1И реализует те же отдельные записи как обычный трекер объектов.

Это детерминированный разбор заданной последовательности. Здесь нет оценки среднего эффекта в популяции последовательностей и нет статистической выборки участников. Каждое число можно проверить вручную либо функцией source_history_example() инженерного практикума. Выход модели означает вероятность общей причины после текущего наблюдения, а не прогноз самого наблюдения.

Результаты и их объяснение

Шаг и событие	M0	M1	M2	M1И
1 А положительный	0,750000	0,750000	0,750000	0,750000
2 Б отрицательный	0,250000	0,437500	0,250000	0,250000
3 А положительный	0,750000	0,710526	0,875000	0,875000
4 Б положительный	0,750000	0,858108	0,562500	0,562500

На третьем шаге M2 возвращается к истории А. Его прежняя оценка 0,75 после удержания превращается в предварительную вероятность 0,7. Положительное свидетельство повышает её до 0,875. M1 приходит к этому шагу после отрицательного события Б, поэтому даёт приблизительно 0,710526. Различие возникло из правил переноса истории, хотя текущий признак одинаков.

Четвёртый шаг особенно полезен для понимания механизма. Общая история M1 теперь преимущественно поддерживающая, и новое положительное свидетельство даёт 127/148, приблизительно 0,858108. M2 возвращается к прежней отрицательной истории Б и получает лишь 0,5625. Раздельная память не просто увеличивает уверенность: она сохраняет разные основания и поэтому может понижать оценку относительно общей памяти.

Во всех четырёх событиях M1И совпадает с M2. Это равенство следует из одинаковых правил, а не из случайного недостатка данных. Переименование словаря объектов в «модель себя» не создаёт дополнительного механизма.

Вычислительный результат устанавливает роль разделения историй и одновременно ограничивает его интерпретацию.

Контроли и вывод мини-исследования

Переименование А и Б не должно менять числовую последовательность, если вместе с метками перенесены их истории. Если заменить все метки одной, М1 и М2 совпадут. Если установить удержание равным нулю, все версии будут давать 0,75 после положительного признака и 0,25 после отрицательного. Эти следствия позволяют проверить, действительно ли реализация соответствует заявленной модели.

Вывод. Раздельное хранение истории достаточно для воспроизведения описанного вычислительного эффекта возвращения источника. Пример не устанавливает преимущество на любой среде и не отличает специальное самоотнесение от общего отслеживания идентичности. Он даёт точный механизм, который можно проверить на независимых данных, и ясную альтернативу, с которой придётся работать дальше.

Ограничения здесь содержательны. Безошибочная метка облегчает задачу; реальная система должна распознавать источник по наблюдениям. Правило времени выбрано по обращениям, а не по физической длительности. Убеждение о причине пока не связано с качественной структурой восприятия. Каждое ограничение указывает отдельную следующую гипотезу, а не отменяет уже установленный результат.

Следующий проект взаимного определения содержания

Расширим задачу так, чтобы источник влиял на интерпретацию неоднозначного признака, интерпретация меняла оценку источника, а система могла выбирать дополнительную проверку. Получится конкретный кандидат на часть организации, предложенной в главе 26. Для начала зададим два раздельно изменяемых свойства содержания: оценку признака и оценку источника. Третья переменная будет определять ожидаемую пользу проверки.

До запуска опишем направленные вмешательства. Принудительное изменение оценки источника при сохранённом сенсорном входе должно сдвигать интерпретацию признака и выбор проверки. Изменение представления признака должно предсказуемо менять последующую оценку источника. Разрыв одной обратной связи должен нарушать соответствующее направление влияния при сохранении доступности исходного наблюдения. Размеры и временные окна эффектов задаются конкретными правилами реализации.

Теперь появляется различающий проект. Полная система сравнивается с системой, имеющей те же наблюдения и записи истории, но только одно направление влияния. Отдельный контроль воспроизводит обычные выходы без причинной связи между компонентами. Проверяется не только успешность задачи, но и матрица ответов на вмешательства. При этом нужен достаточно сильный альтернативный контроллер: победа над заведомо беспомощной записью не объяснит уникальность механизма.

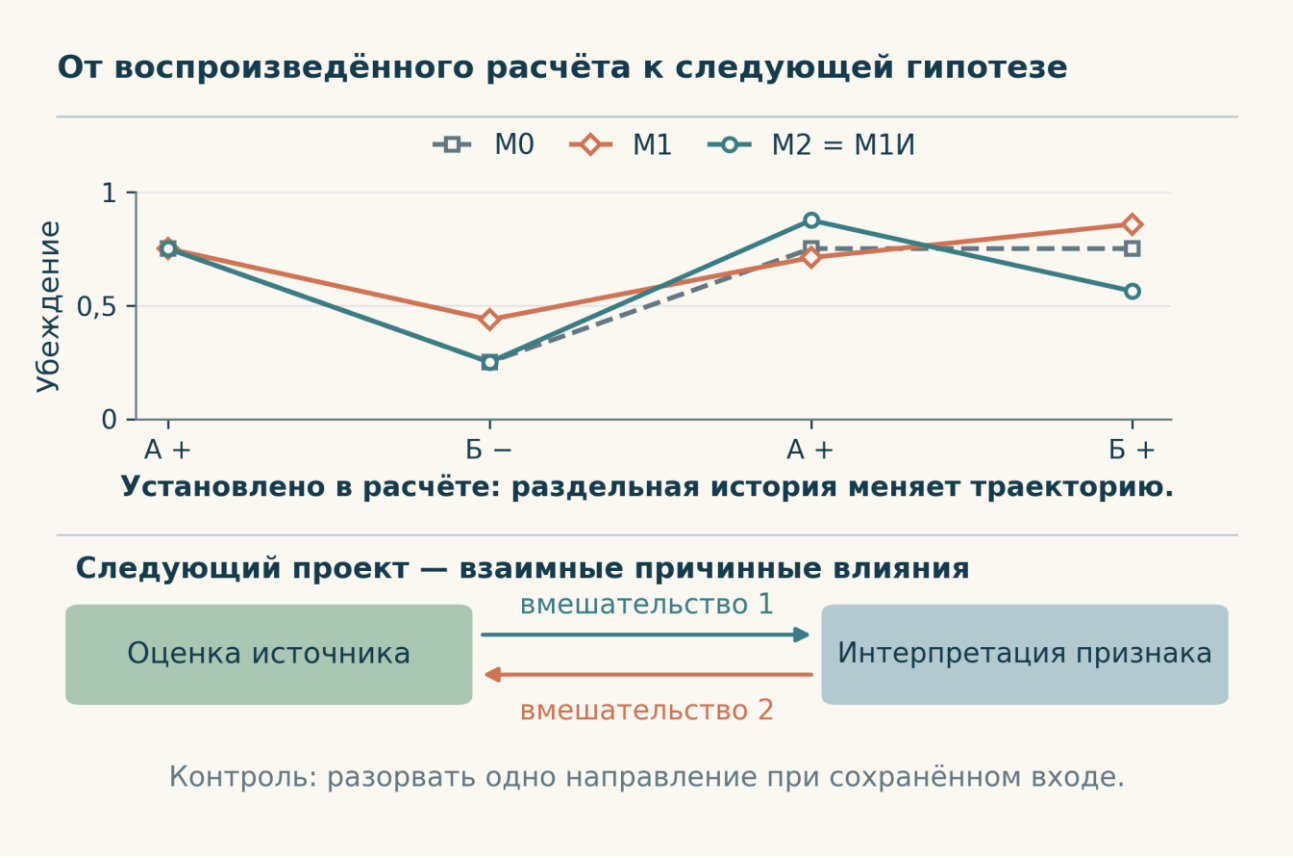

Рисунок 88. Завершённый расчёт и следующий причинный проект. Вверху точный учебный расчёт четырёх событий А+, Б−, А+, Б+. На четвёртом шаге M0 даёт 0,75, M1 — 127/148 (приблизительно 0,858108), а M2 и совпадающий с ним M1И — 0,5625. Внизу показан ещё не выполненный проект: вмешательства в оценку источника и интерпретацию признака должны иметь заранее определённые последствия в противоположном компоненте. Односторонний контроль проверяет конкретную причинную связь. Расчёт сам по себе не реализует всю сильную гипотезу главы 26 и не устанавливает наличие опыта.

Как проект приблизится к вопросу об опыте

Принцип главы 26 связывает качественную организацию с устойчивыми отношениями различений и альтернатив. Поэтому будущая модель должна предсказать структуру сходств, смешений и изменений содержания, а не только правильный ответ. Например, вмешательство в источник должно менять некоторые отношения между неоднозначными содержаниями,

сохраняя заранее указанные другие отношения. Это уже более сильное обязательство, чем повышение общей точности.

Для сопоставления с человеком понадобятся независимые данные об этих различиях: оценки сходства переживаний, их динамика, различение источника и контроль трудности. Предполагаемый механизм сравнивается с объяснением через общую память объектов и с объяснением через изменение решения при неизменном переживании. Эти конкуренты должны получить собственные предсказания до сбора подтверждающих данных.

Сильное утверждение о наличии опыта сохраняется как гипотеза: определённая реально осуществлённая организация достаточна для минимального субъективного процесса. Реализация и функциональные проверки проверят выполнение предложенных условий. Обоснование достаточности потребует связи с феноменальными данными, уточнения физической реализации и анализа альтернатив. Так задача создания субъекта остаётся прямой исследовательской целью, а каждый этап имеет собственный проверяемый результат.

Каждое утверждение требует своей проверки

Утверждение	Действие проверки	Содержательный контроль
История источника влияет на ответ	Удалить или заменить именно эту запись	Общий трекер той же истории
Имеются два направления влия	Вмешаться в p и отдельно в q	Два односторонних контура
Показатель улучшает прогноз	Проверить полностью отложенные группы	Исходная активность без нового признака
Организация связана с опытом	Уточнить принцип и расходящееся следствие	Иная организация с теми же выходами

Контрольное сравнение ограничивает определённый вывод, а не все вопросы сразу.

Рисунок 89. От утверждения к проверке и контролю. Учебная матрица завершения проекта. Каждая строка связывает утверждение с конкретным действием и сильным контролем. Последняя строка сохраняет самостоятельность теоретического аргумента об опыте; функциональные проверки выше его автоматически не заменяют.

Что должно войти в итоговую работу

Подготовьте связный текст с вопросом, основаниями, механизмом, методом, результатами и обсуждением. При цифровой реализации приложите исполнимый код, описание среды, параметры, сохранённые данные и способ повторного получения таблиц. При биологическом направлении допустимы

анализ открытых данных, цифровая модель и профессиональный проект эксперимента с ясно указанной границей уже выполненной работы.

Включите хотя бы одного содержательного конкурента, один причинный или различающий тест и один контроль с заранее известным результатом. Отдельно предъявите принцип соответствия, если делаете утверждение о переживании. Укажите, какие его положения проверены, какие предполагаются и какие пока недоопределены. Авторское предложение требует проверки близких опубликованных идей; самостоятельная формулировка сама по себе не устанавливает научного приоритета.

Неразличающий результат тоже может стать полноценной итоговой работой. Важно показать, почему он возник и какой следующий тест устранит неопределённость. Наличие модуля, его красивое описание и убедительный самоотчёт не заменяют демонстрацию его роли. Такой вывод особенно полезен, когда позволяет отказаться от лишней сложности и точнее спроектировать следующего кандидата.

Как превратить расчёт в маленькую статью

Вопрос примера: изменяет ли отдельная история обновление при возвращении источника? Метод задаёт четыре события, надёжность, удержание и правила памяти. Результат предъявляет все траектории и совпадение M2 с общим трекером. Обсуждение связывает эффект с разделением записей и сохраняет ограничения меток, времени и отношения к опыту. Следующий шаг меняет одно предположение и получает новый прогноз. Такая структура позволяет воспроизвести вычисление и отдельно критиковать объяснительный переход.

Что теперь установлено

У нас есть завершённый функциональный результат и открытая исследовательская цель. Следующий проект не должен стирать эту границу: он должен добавить новое определённое обязательство и подходящую проверку.

Практика 36 Завершите собственную статью

Зачем выполнять: Подготовить воспроизводимую собственную работу с одним проверяемым изменением механизма.

Что понадобится: Спецификация механизма, сильный конкурент, предварительный план и решение по результату. Новизна формулировки не устанавливает научного приоритета.

Ориентир по времени: Несколько отдельных сессий после воспроизведения примера. Можно делать с перерывами; это не норматив скорости.

1. Воспроизведите четыре события M0-M2 и совпадение с общим трекером. Сохраните код, параметры и исходные числа.
2. Выберите одно изменение: время забывания, ошибки распознавания источника или взаимное влияние признака и источника. До запуска запишите отличие от исходной модели.
3. Подготовьте сильного конкурента, различающий тест и контроль с известным ответом. Выполните проверку и сохраните все исходы.
4. Напишите работу с вопросом, основаниями, механизмом, методом, результатом и обсуждением. В обсуждении отделите установленный факт, оставшуюся неопределённость и следующий вопрос об опыте.

Что сохранить: Текст исследования, исполнимый пример, данные и ясная история версий модели.

Подсказка: Неразличающий результат допустим, если понятно, почему он возник и какое новое условие разведёт объяснения. Не заменяйте этот анализ красивым самоотчётом системы.

Исследовательское продолжение: Для сильной гипотезы задайте вмешательства в обоих направлениях и заранее выведите изменения отношений между содержаниями. Сопоставление с феноменальными данными и перенос на новую физическую реализацию остаются отдельными этапами программы.

Разбор и ответы — приложение Е, глава 36, с. 311

Самопроверка

1. Что доказал приведённый расчёт четырёх событий?
2. Почему для сильного проекта нужны вмешательства в обоих направлениях?
3. Может ли итоговая работа содержать существенную неопределённость относительно опыта?

Дальнейшая работа может идти по цифровой, сравнительной или биологической ветви. Сохраняйте конкретный вопрос, полную историю изменений и возможность получить результат, который заставит пересмотреть объяснение.

Приложение A Рабочий словарь

Этот словарь закрепляет употребление терминов в книге. При чтении другого автора сравнивайте определения: одинаковое слово может обозначать разные исследовательские задачи.

Квалиа. Качественные стороны переживания, то, как нечто воспринимается или чувствуется. В рабочее определение не включены заранее безошибочность описания, невыразимость или независимость от физического мира.

Субъективный опыт. Переживаемая сторона состояния. Вопрос о наличии опыта отличается от вопроса о том, какие сведения система различает или сообщает.

Сознание. Термин с несколькими употреблениями. В книге каждый раз уточняется, обсуждаются ли наличие опыта, его содержание, состояние бодрствования или доступность информации для дальнейшего использования.

Субъект. Предполагаемый носитель субъективного опыта. Определение не требует человеческой биографии, речевого интерфейса или принадлежности к заранее выбранному виду.

Самосознание. Семейство форм отношения системы к себе. Телесная принадлежность, авторство действия, представление о собственной истории и явное суждение о себе исследуются отдельно.

Рефлексия. Рассмотрение собственных суждений, оснований или способов действия с возможностью их оценки и пересмотра. Простой численный прогноз успешности ещё не реализует все эти возможности.

Метакогниция. Мониторинг и регулирование познавательных процессов. В конкретной задаче нужно назвать, что оценивается и как оценка влияет на поведение.

Интроспекция. Обращение к собственным состояниям для их замечания или описания. Природа доступа и надёжность отчёта требуют исследования.

Функциональная интроспекция. Исследуемая способность системы использовать сведения о своих внутренних состояниях. Сам термин не решает вопрос о переживании этих состояний.

Принадлежность тела. Переживание или оценка того, что определённая часть тела относится к себе. В модели необходимо указать измерение, связанное с этой характеристикой.

Авторство действия. Отношение между собой, командой и её последствием. Успех управления, причинное влияние и переживаемое авторство нельзя считать одним показателем.

Модель себя. Представление о собственных состояниях, возможностях, границах или истории. Нужно уточнять содержание и причинную роль конкретной модели.

Модель мира. Представление внешней ситуации, позволяющее оценивать скрытые состояния или прогнозировать события.

Скрытое состояние. Переменная, известная среде или исследователю, но непосредственно недоступная агенту в данной задаче.

Наблюдение. Сведение, которое система действительно получает через предусмотренный интерфейс. Наличие переменной в коде среды не делает её наблюдением агента.

Убеждение. В вероятностной модели — оценка возможных состояний с учётом предпосылок и доступных сведений. Это техническое употребление не приписывает программе переживаемой убеждённости.

Априорная вероятность. Оценка перед учётом рассматриваемого свидетельства. Она может включать результаты предыдущего обучения.

Апостериорная вероятность. Оценка после обновления на основании нового свидетельства в пределах выбранной модели.

Правдоподобие. Вероятность наблюдаемых данных при предполагаемом состоянии или модели. Его следует отличать от вероятности самой гипотезы после наблюдения.

Неопределённость. Ограниченность знания о состоянии, параметре или объяснении. Разные источники неопределённости могут требовать разных действий.

Калибровка. Соответствие вероятностных прогнозов частоте исходов в определённых группах случаев. Её оценка зависит от данных и способа группировки.

Brier score. Средний квадрат разности вероятностного прогноза двоичного исхода и результата, записанного как ноль или единица. Характеризует качество прогноза, а не только калибровку.

Полезность. Заданный способ объединить результаты и затраты. В первой работе это успехи за вычетом цены дополнительных наблюдений.

Ценность информации. Ожидаемое улучшение решения благодаря дополнительным сведениям с учётом цели и принятой модели. Уменьшение неопределённости не всегда одинаково полезно.

Обратная связь. Возвращение сведений о результате в дальнейшую работу системы. Нужно указать, какие именно сведения и с какой задержкой возвращаются.

Рекуррентная обработка. Обработка с возвратом влияния: результат изменяет работу элементов, уже участвовавших в его формировании. Конкретная вычислительная и физическая организация должна быть описана отдельно.

Рабочее пространство. Организация отбора и общего использования содержания несколькими процессами. Простая таблица переменных может моделировать отдельные функции, но не автоматически всю научную теорию.

Абляция. Удаление или отключение компонента для проверки его вклада в определённой реализации и задаче.

Причинное вмешательство. Направленное изменение переменной или связи, предназначенное для проверки её влияния. Побочные изменения и доступность вмешательства другим процессам входят в анализ.

Фиктивное вмешательство. Контрольная процедура, сохраняющая существенную исследуемую переменную при воспроизведении других сторон воздействия.

Коррелят. Признак, статистически связанный с изучаемым явлением. Связь сама по себе не устанавливает необходимость или достаточность.

Необходимое условие. Условие, без которого предполагаемое явление не возникает в оговорённой области. Отсутствие эффекта после вмешательства требует проверки специфичности самого вмешательства.

Достаточное условие. Условие, наличие которого должно обеспечивать явление при явно заданных предпосылках. Оно может не быть единственным возможным способом получения явления.

Инвариант. Свойство, сохраняющееся при определённом классе преобразований. Класс преобразований и способ измерения нужно назвать заранее.

Идентифицируемость. Возможность различить модели или параметры по доступным данным. Разные названия математически одинаковых алгоритмов не создают идентифицируемости.

Переобучение. Подстройка к особенностям использованных данных, которая не обеспечивает ожидаемого качества на новых случаях.

Контрольная выборка. Данные, используемые для оговорённой проверки после разработки или оценки модели. Повторная настройка по ним меняет их исследовательскую роль.

Предварительная регистрация. Сохранённое до соответствующего этапа описание вопроса, сравнений, показателей и анализа. Она делает изменения плана заметными, но не гарантирует качества исходной идеи.

Единица наблюдения. Объект или случай, к которому относится измерение. Её следует отличать от независимой единицы повторения эксперимента.

Псевдоповторение. Ошибочное обращение с зависимыми измерениями как с независимыми повторениями. Много эпизодов одного образца не заменяют много независимых образцов.

Бутстреп. Получение повторных выборок из имеющихся данных для оценки изменчивости статистики. Способ повторного выбора должен соответствовать структуре зависимости наблюдений.

Препринт. Доступная версия исследовательского текста до или вне завершённой журнальной публикации. Статус и версию нужно проверять отдельно.

Нейронная культура. Экспериментальная система с выращиваемыми нервными клетками. Её организация зависит от конкретной модели и условий.

Органоид. Трёхмерная клеточная модель, воспроизводящая некоторые особенности организации или развития ткани. Органоид не следует автоматически считать целым органом или организмом.

Гибридная система. Система, в которой биологические и технические компоненты причинно взаимодействуют. Анализ должен показывать вклад каждого компонента и выбранную границу целого.

Феноменальный вывод. Утверждение о наличии или характере субъективного опыта, сделанное на основании данных и теоретических предпосылок.

Принцип соответствия. Предлагаемая связь между описанием организации системы и описанием опыта. В авторской гипотезе он требует самостоятельного аргумента и определения области применения.

Правдоподобие: Вероятность определённых данных при выбранной версии. Оно не равно вероятности самой версии после данных; для последней нужны предварительная вероятность и нормировка.

Условная независимость: Допущение, при котором вероятность нового события при заданной причине не зависит от уже известных событий. Копирование сигнала это допущение обычно нарушает.

Ошибка Брайера: Средний квадрат разности между вероятностным прогнозом двоичного наблюдения и его исходом ноль или один. Прогноз записывается до появления оцениваемого наблюдения.

Восстановление модели: Проверка выбора механизма на искусственных данных, где механизм генерации известен. Ошибки восстановления показывают ограничения процедуры различения.

Отложенная группа: Все записи выбранного органоида или партии, которые целиком исключены из обучения для конкретной проверки переноса.

STTC: Spike time tiling coefficient — показатель временного совпадения разрядов. В учебном анализе используется среднее абсолютное значение конечных внедиагональных элементов матрицы при окне 15 мс. Это не доказанная причинная связь.

Достаточное условие: Если условие выполнено, предполагаемый вывод следует в рамках принятого утверждения. Невыполнение такого условия само по себе не исключает других оснований того же вывода.

Приложение Б Запуск и устройство практикума

Подготовка

Распакуйте qualia_mira_practice.zip в отдельный каталог. Внутри находятся программы, инструкции и сохранённые результаты. Начните с файла README.md. Не запускайте Python-файлы прямо из окна просмотра архива: им нужны соседние файлы. Для работ 2–8 agent_lab.py и mechanisms_lab.py должны находиться рядом. Удобно сохранять новые результаты под другими именами, чтобы исходные примеры оставались доступными для сравнения.

Откройте терминал в этом каталоге. В Windows это можно сделать через меню проводника или перейти командой cd. В macOS и Linux используйте обычный терминал. Проверьте версию Python: требуется 3.10 или новее. Если Python ещё не установлен, используйте официальный дистрибутив python.org для своей системы. Название команды зависит от установки.

```
python3 --version
python3 agent_lab.py --self-test
python3 mechanisms_lab.py --self-test
```

Для Windows типичная форма первой команды — py -3 --version. Аналогично замените python3 в остальных командах. Если система сообщает, что команда не найдена, это ошибка запуска интерпретатора, а не результат эксперимента. Если не найден agent_lab.py, проверьте текущий каталог. Если mechanisms_lab.py не может импортировать agent_lab, проверьте, что оба файла распакованы вместе.

Для первого знакомства можно выполнить все восемь работ и парный анализ общей командой. Она создаёт отдельный каталог results_... с результатами и файлом run_manifest.json, в котором записаны параметры и контрольные суммы программ.

```
python3 run_practice.py
```

Одну работу можно выбрать параметром --lab, например python3 run_practice.py --lab 6. В Windows используйте py -3 вместо python3. Общий запуск проверок без сохранения нового опыта: python3 run_practice.py --self-test.

Как читать файлы результата

JSON хранит данные в виде именованных полей и списков. Не нужно осваивать весь формат сразу. В lab01_results.json раздел summary содержит общие

результаты режимов, а per_run — результаты отдельных сценариев. В mechanisms_results.json работы находятся внутри labs под номерами от 2 до 8. Подробные объяснения полей включены в инструкции архива и главы 28–30.

Поля success_rate и utility_per_step означают долю успеха и полезность на шаг. Значение 0,8 для доли успеха соответствует восьмидесяти процентам. Поле queries содержит число дополнительных запросов, а brier_score — показатель качества вероятностного прогноза. При сравнении таблиц всегда проверьте, одинаковы ли число шагов, параметры и способ усреднения.

Сначала воспроизведите исходный результат. Затем создайте новый файл с одним изменённым параметром. Сохраните рядом текст команды и предварительное ожидание. После проверки на исходном seed используйте другой. Случайное зерно обеспечивает повторяемость конкретного сценария; оно не превращает единичный сценарий в достаточное исследование.

Восемь работ

Работа	Программа	Основной материал
1	agent_lab.py	Память мира, модель действия, ценность наблюдения
2	mechanisms_lab.py --lab 2	Геометрия представлений и инварианты
3	mechanisms_lab.py --lab 3	Повторная обработка и развёрнутый контроль
4	mechanisms_lab.py --lab 4	Ограниченный отбор и общий доступ
5	mechanisms_lab.py --lab 5	M0, M1, M2, общий трекер и задержка источника
6	mechanisms_lab.py --lab 6	Спроектированный внутренний доступ и вмешательства
7	mechanisms_lab.py --lab 7	Выбор стратегии по прошлым ошибкам
8	mechanisms_lab.py --lab 8	Составной кандидат и контроль вклада компонентов

Все работы 2–8 можно выполнить одним запуском без параметра --lab. Команда ниже создаст новый файл результата. Она ничего не отправляет внешнему сервису и не изменяет код.

```
python3 mechanisms_lab.py --json my_mechanisms.json
```

Программы предназначены для изучения небольших механизмов. Их подробное устройство описано в книге. Объединение названий функций в одну

систему не заменяет отдельной проверки того, как компоненты взаимодействуют. Восьмая работа специально включает сравнения, в которых добавленный механизм не даёт преимущества в базовом условии.

Дополнительные пошаговые модели

Файл teaching_lab.py дополняет исходные восемь работ. Он использует стандартную библиотеку Python и при анализе главы 31 читает лежащий рядом bio_features.csv. Выполняйте команды из папки practice распакованного комплекта. Команда python обозначает доступный интерпретатор Python 3; в некоторых системах он запускается как python3 или py -3. Базовые проверки и учебные модели не требуют установки внешних пакетов.

```
python teaching_lab.py --self-test
python teaching_lab.py --chapter 13
python teaching_lab.py --chapter 25
python teaching_lab.py --chapter 31
python teaching_lab.py --chapter 33
python teaching_lab.py --chapter 34
```

Каждый запуск сохраняет JSON с правилами или числовыми результатами. Для главы 31 готовая таблица позволяет повторить основной анализ без загрузки большого массива. Скрипт extract_bio_features.py принимает официальный файл [47], проверяет его контрольную сумму и повторяет извлечение признаков с явными исключениями. Для этого отдельного шага нужна NumPy; подробности и атрибуция находятся в README практикума.

Приложение B Парное сравнение результатов

В первой работе все четыре режима используют общие случайные сценарии. Поэтому естественная единица сравнения — разность показателей двух режимов внутри одного сценария. Если сначала объединить все шаги и объявить их независимыми повторениями, получится неоправданно сильное впечатление точности: соседние шаги связаны состоянием цели и обучением агента.

Программа analyze_lab01.py вычисляет разности полезности полного режима и каждого контрольного режима по тридцати сценариям. Затем она четыре тысячи раз повторно выбирает целые разности с возвращением и строит процентильный интервал. Метод описывает изменчивость при принятой структуре повторений; точное покрытие девяносто пять процентов для любой выборки не гарантируется.

```
python3 analyze_lab01.py lab01_results.json --json paired_analysis.json
```

В сохранённом анализе среднее преимущество полного режима по полезности на шаг составляет 0,08516 относительно версии без памяти, 0,37772 относительно фиксированной модели действия и 0,08280 относительно версии без дополнительного запроса. Соответствующие интервалы приблизительно равны от 0,07975 до 0,09080, от 0,37133 до 0,38434 и от 0,07583 до 0,08997. Эти значения получены из приложенных данных и алгоритма анализа.

Интервалы относятся к конкретным сравнениям, метрике и искусственной среде. Несколько отдельных интервалов не дают автоматической совместной гарантии девяносто пять процентов. Если после просмотра многих показателей выбрать только благоприятный, нужно учитывать поисковый характер такого решения. В полноценном исследовании основной критерий и семейство сравнений описываются заранее.

Для самостоятельной работы выберите главным показателем полезность, а точность и стоимость используйте для объяснения механизма различий. Сначала сформулируйте ожидаемое направление эффекта. Затем посмотрите разности отдельных сценариев: устойчив ли результат, есть ли необычные случаи и связаны ли они с конкретной последовательностью событий? Не удаляйте неудобный сценарий без заранее установленной причины.

Если среднее различие мало, возможны разные объяснения: механизмы действительно похожи в выбранной задаче, данных недостаточно или показатель плохо отражает ожидаемое свойство. Более широкий интервал не выбирает между ними автоматически. Следующий опыт должен быть связан с конкретной причиной неопределённости: новой задачей, дополнительными независимыми повторами или более подходящим измерением.

Приложение Г Бланки самостоятельной работы

Запись переживаемого эпизода

Укажите ситуацию и момент события. Отметьте, когда сделана запись: во время эпизода, сразу после или позднее. Опишите то, что удалось заметить, и отдельно сохраните неясные места. Затем запишите своё объяснение и хотя бы одну альтернативу. Закончите вопросом о том, какое наблюдение помогло бы уточнить вывод. Поля можно заполнять свободной прозой; отсутствие замеченной детали следует оставлять отсутствием, а не достраивать по образцу.

Карточка научной статьи

Запишите авторов, год, название и ссылку. Сформулируйте исследовательский вопрос одним предложением. Укажите участников или образцы, условия, измерения и сравнение. Отдельно изложите полученный результат и вывод авторов. Затем выпишите допущение, которое связывает их, возможную альтернативу и следующий различающий тест. Если статья теоретическая, вместо выборки и измерения опишите исходные положения и выводимые следствия.

Карточка модели

Назовите объясняемое явление и область применения. Перечислите доступные входы, скрытые состояния, память и правило обновления. Объясните, как модель производит наблюдаемый ответ. Укажите параметры и способ их определения. Завершите описанием ближайшего конкурента и условия, где прогнозы расходятся. Для гипотезы о переживании добавьте принцип соответствия и аргумент в его пользу.

Запись программного опыта

Сохраните вопрос, версию кода, команду, параметры, seed, основной показатель и ожидаемое различие. После запуска добавьте результат, путь к данным и объяснение отклонений. Отметьте, какие решения были приняты до просмотра данных и какие после. Следующий проверочный запуск должен быть отделён от этапа настройки.

Паспорт итогового проекта

Укажите предмет объяснения, предложенный механизм, способ реализации, доступные данные, план вмешательств и сопоставимые альтернативы. Опишите положительный критерий успеха и результат, который потребует изменения гипотезы. Затем перечислите фактически выполненные проверки. В заключении отдельно сформулируйте функциональный результат, теоретическую интерпретацию и основания для возможного вывода об опыте. Эти абзацы должны согласовываться, сохраняя различие своих задач.

Приложение Д Указатель иллюстраций

Рисунки расположены рядом с объяснениями. Подписи указывают статус изображения и происхождение чисел. Схемы раскрывают выбранные отношения; предметные иллюстрации помогают представить эпизод или исследуемый объект. Номер в указателе ведёт к рисунку.

1. Три области описания одного эпизода 13
2. Время записи как часть наблюдения 14
3. Сравнение поверхностей в одном эпизоде 19
4. Названия материалов и сходство прикосновений 21
5. От попарных оценок к карте сходства 22
6. Принадлежность без движения и управление внешним объектом 28
7. Четыре отношения к себе в одном эпизоде 29
8. Уверенность при одинаковой успешности 36
9. Проверка логического вывода контрпримером 294
10. Причинное влияние и общая причина 41
11. Три исследовательские роли в устройстве лампы 42
12. Участник и наблюдатель в исследовании 46
13. Согласование разных данных по времени 48
14. Время вопроса как условие процедуры 50
15. Распределение материала и порядка условий 54
16. Общий бюджет и распределение времени подготовки 55
17. Величина эффекта и точность его оценки 60
18. Ответы и участники как разные уровни данных 62
19. Динамика накопления и сброса 67
20. Распределение входов и отключение сохранения состояния 69
21. Два способа использовать контекст 75
22. Избирательное нарушение канала использования 76
23. Рука инструмент и наблюдаемый результат 80
24. Последовательное обновление причинных версий 81
25. Различение причин изменением команды 84
26. Информация до и после обратной связи 88
27. Прогнозы и ошибка двух моделей 90
28. Рекуррентность и общий доступ 96
29. Независимое отключение двух механизмов 100
30. Два разных предмета объяснения 105
31. Внешний журнал и внутреннее использование 107
32. От предварительной оценки к обновлённой вероятности 111
33. Раздельные модели мира и действия 112
34. Отчёт и границы измерения 118
35. Почему отсутствие значимости не равно точному нулю 120
36. Три траектории одного учебного механизма 124
37. Усреднение и динамика отдельного эпизода 127
38. Развитие в повседневном взаимодействии 130
39. Разложение задачи визуального самоузнавания 132
40. Профиль развития без общей лестницы 134
41. Среда задаёт разные задачи поведения 137

42. Компромисс памяти вместо лестницы сложности 138
43. Две истории одного распределения признака 139
44. Объяснение присутствия и объяснение качества 144
45. От устройства к объяснению опыта 147
46. Разные организмы и разные исследовательские вопросы 151
47. Матрица вопросов и пробелов свидетельств 154
48. Сопоставимые задачи с разными внешними ответами 155
49. Разделение числа событий и отношения между ними 161
50. Привыкание и истощение различаются переносом 163
51. Раздельный аудит информации и самоописания 167
52. Искусственная система в физической среде 168
53. Носитель не определяет границу исследования 170
54. Карта обязательств модели 176
55. Справедливое сравнение вычислительных конкурентов 177
56. Нормировка одного обновления 185
57. Возвращение к истории источника 183
58. Переименование и изменение причинной структуры 191
59. От причинной организации к гипотезе об опыте 195
60. Четыре условия различающего проекта 199
61. От возможного исхода к различающему тесту 202
62. Два контура одного агента 206
63. Условная ценность дополнительного наблюдения 208
64. Успех и полезность отвечают на разные вопросы 210
65. Неопределённость различий между режимами 211
66. Что сохраняется при преобразовании представления 215
67. Рекуррентность и сопоставимые альтернативы 216
68. Доставка содержания и контроль прямым доступом 218
69. Проверка доступа после поздней инструкции 219
70. Чья история используется при обновлении 224
71. История помогает распознать задержанный источник 225
72. Внутренний доступ проверяется вмешательством 226
73. Причинный порядок функционального пересмотра 227
74. Пересмотр становится причиной нового решения 228
75. Сборка нескольких механизмов в одну систему 229
76. Наличие модуля ещё не показывает его вклад 230
77. Биологический компонент и измерительная аппаратура 233
78. Уровни учёта в биологических данных 237
79. Временное и групповое разделение биологических данных 239
80. Три разных вопроса о внутреннем состоянии 245
81. Профиль проверки кандидата 247
82. Две явно разные модели времени памяти 252
83. Граница доступной информации и неоднозначность параметров 254
84. Структура данных предварительного протокола 261
85. Предварительный план и последующее открытие 263
86. Иллюстративные парные различия и условный ориентир 267
87. Честный неразличающий результат работы 8 269
88. Завершённый расчёт и следующий причинный проект 275
89. От утверждения к проверке и контролю 276

Приложение E Разборы практик и самопроверки

Возвращайтесь к разбору после собственной попытки. Для наблюдений оценивайте ясность записи и различение вывода и предположения. Для расчётов сверяйте исходные правила и промежуточные шаги. Исследовательские продолжения допускают несколько обоснованных решений. Самопроверка сохраняет вопросы из главы; здесь находятся ответы.

Глава 1 Разбор практики

Единого списка ощущений здесь нет. Работа выполнена, если внешний факт, описание и объяснение можно различить по записи. Большее число деталей во второй попытке ещё не устанавливает обогащение самого опыта: измениться могли внимание, память и словарь. Хороший оставшийся вопрос называет альтернативы и следующее сравнение.

Ответы на самопроверку

1. Фраза сообщает об изменении ощущения и объясняет его физическим изменением предмета. Основания нужны как для утверждения о нагревании, так и для причинной связи. Одно переживаемое изменение допускает несколько объяснений.
2. Да, если говорится о том, как переживался конкретный эпизод. Нужно отличать это сообщение от устойчивого предпочтения или общего суждения о качестве вещи. Решение определяется смыслом высказывания.
3. Установлено различие между записями. Оно может отражать изменение опыта, внимания, памяти или языка. Для вывода о причине потребуются дополнительные сравнения. Само число подробностей не является мерой количества сознания.
Вернуться к главе 1

Глава 2 Разбор практики

Проверяется соответствие карты исходным отношениям. Если числа задавали только порядок, не нужно требовать точного совпадения всех расстояний. Неустойчивый ответ нельзя молча исправлять ради красивой геометрии. В исследовательском продолжении полезность модели оценивается по новым, заранее предсказанным отношениям.

Ответы на самопроверку

1. Название может объединять случаи по материалу, назначению или привычной категории. Описание переживания и попарное сравнение задают другие отношения. Нужно уточнять, по какому основанию производится группировка.
2. Её смысл зависит от инструкции. Она может обозначать только место в порядке оценок. Утверждение о точном отношении величин, например «вдвое сильнее», требует отдельного основания.
3. Следует заранее определить, какие отношения или новые случаи модель предскажет лучше. Затем использовать данные, не участвовавшие в выборе координат. Улучшение рисунка на тех же четырёх точках ещё не показывает способности объяснять новое.
Вернуться к главе 2

Глава 3 Разбор практики

Неподвижность руки не отменяет принадлежность, а управление указателем не делает его частью тела. Прошлое авторство заметки не тождественно нынешнему действию. Не требуется получать одинаковые переживания у разных людей; требуется сохранять различия вопросов и не выводить внутреннее отношение исключительно из внешнего события.

Ответы на самопроверку

1. Управление относится к связи команды с результатом. Принадлежность относится к тому, дан ли объект как часть собственного тела. Эти отношения могут сочетаться, но одно не определяет другое.
2. Фактическое авторство выясняют по причинной цепочке событий. Чувство авторства изучают как переживаемое отношение к действию и через соответствующие отчёты. Они могут расходиться, и это расхождение само требует объяснения.
3. Имя — элемент самоописания. Оно не объясняет пространственную перспективу, телесную принадлежность, переживаемое действие или проверку собственных суждений. Полезная модель должна указать, какие из этих задач она решает и каким механизмом.
Вернуться к главе 3

Глава 4 Разбор практики

Ответы: ластиков; наличие большой зелёной карточки не следует; последним пришёл Пётр. Во втором случае все зелёные могут быть маленькими круглыми, а большие круглые иметь другой цвет. Это контрпример, удовлетворяющий условиям и опровергающий обязательность вывода. Задача показывает связь прогноза с последующей проверкой, а не измеряет общий уровень самосознания.

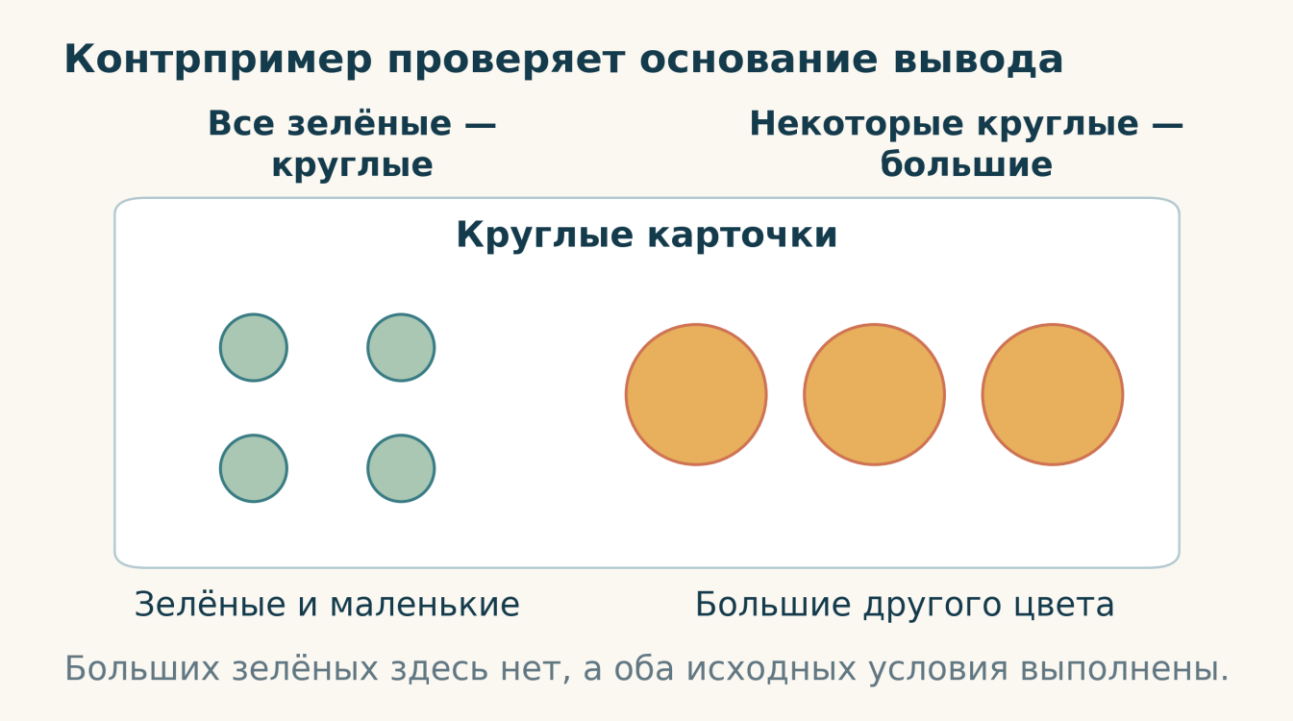

Рисунок 9. Проверка логического вывода контрпримером. Все зелёные карточки на схеме круглые, а некоторые круглые карточки большие. При этом больших зелёных нет: исходные условия не гарантируют такой вывод, и контрпример показывает конкретный пробел в рассуждении.

Ответы на самопроверку

1. Обратная связь добавляет информацию и позволяет построить новый способ объяснения. Нужно различать первоначально замеченные шаги, позднее воспоминание и реконструкцию с учётом правильного ответа.
2. Да. Можно почти всем ответам приписывать одинаковую уверенность. Общая успешность и способность направлять проверку на вероятные ошибки тогда будут разными характеристиками.
3. Указать тип задания, наблюдаемую ошибку и предполагаемый процесс, затем предложить конкретное изменение процедуры. После этого сравнить результаты на новых сопоставимых заданиях, учитывая возможные альтернативные причины улучшения.

Вернуться к главе 4

Глава 5 Разбор практики

Сравнение дневника с равной паузой начинает разделять предложенные версии. Оно всё ещё требует измерения, сопоставимых задач и контроля порядка. Если выбранный показатель отражает лишь увеличение длины описания, вывод об улучшении различения пока не обеспечен. Хорошая работа сохраняет это ограничение и предлагает следующее сравнение.

Ответы на самопроверку

1. Механизм показывает, как части взаимодействуют и производят результат, включая условия и последовательность изменений. Перечень деталей не объясняет их совместную работу.
2. Показатель может влиять на отчёт, зависеть от него или изменяться вместе с ним из-за другого процесса. Различение требует дополнительных данных и подходящих вмешательств.
3. Нужен обоснованный мост между реализованной функцией и переживанием: теория соответствия, область её действия и проверки альтернатив. Сам успех в распознавании устанавливает выполнение задачи, а дальнейший вывод требует дополнительного аргумента.

Вернуться к главе 5

Глава 6 Разбор практики

Проверка предмета устанавливает его внешние свойства. Первоначальный отчёт относится к своей процедуре и не переписывается по позднему знанию. Паспорт полезен, если другой человек может повторить инструкцию и определить, когда именно сведения стали доступны. Согласие нескольких показателей требует проверки общих причин, а не подсчёта голосов.

Ответы на самопроверку

1. Отсутствие воспоминания может возникнуть после эпизода или отражать трудность вопроса. Нужно различать переживание, сохранение сведений и возможность сообщить их в данный момент.
2. Оно показывает успешное использование информации в определённой задаче. Выводы о ясности, полноте опыта и других свойствах требуют дополнительных данных и обоснованной связи с ними.
3. Все три показателя могут зависеть от общего процесса, например подготовки ответа. Согласие становится более содержательным, когда проверены такие общие причины и показано, какие альтернативные объяснения исключаются сочетанием методов.

Вернуться к главе 6

Глава 7 Разбор практики

Длинная запись может объясняться словарём, тренировкой формата и временем. Для вывода о различении нужен соответствующий новый результат. Четыре распределения помогают разделить материал, порядок и способ работы. Если все исходы объявлены подтверждением, гипотеза пока не задаёт проверяемого различия.

Ответы на самопроверку

1. В одном условии человек может всё время читать, а в другом часть времени тратить на схему. Сравнение имеет определённый смысл при общем бюджете, но внутренние действия остаются различными.
2. Оно разрывает систематический выбор условия по известным особенностям участников и задаёт процедуру распределения. Однако в маленькой группе состав условий может случайно оказаться неодинаковым.
3. Новая гипотеза возникла с учётом уже увиденных закономерностей. Для независимой проверки нужны новые данные или заранее отделённая часть данных, которая не использовалась при её формулировании.

Вернуться к главе 7

Глава 8 Разбор практики

В четвёртом повторе разность равна минус двадцати пунктам, в шестом плюс двадцати. По условиям учебного генератора режимы одинаковы, а отдельные оценки колеблются. Неточность может потребовать новых независимых данных; точное исключение важного эффекта ограничивает гипотезу; зависимость от материала требует новой проверки условий. Число строк не заменяет число независимых исследуемых единиц.

Ответы на самопроверку

1. Процентные пункты — разность самих долей. Относительный рост вычисляют относительно исходного уровня: десять делят на семьдесят, получая примерно четырнадцать процентов.
2. Ответы общей модели зависят от одних параметров и истории обучения. Они помогают оценивать её поведение, но не заменяют сведения об изменчивости между независимо созданными системами.
3. Да, если измерение достаточно точно, чтобы исключить заранее значимое улучшение в указанных условиях. При широком диапазоне совместимых эффектов результат остаётся неопределённым, и отсутствие убедительного сигнала не доказывает отсутствия эффекта.

Вернуться к главе 8

Глава 9 Разбор практики

У А импульс появляется на третьем шаге. У Б порог не достигается; конечное состояние примерно 0,921754. При сохранении ноль каждый ненулевой вход даёт только 0,6, поэтому импульсов нет в обеих последовательностях. Объяснена динамика выбранного механизма, а не универсальное свойство нейронов или сознания.

Ответы на самопроверку

1. Правило использует не только текущий вход, но и сохранённое состояние. Разная история меняет расстояние до порога.
2. Состояние описывает текущую динамику при данных правилах. Обучение меняет сами параметры этих правил и тем самым поведение на будущих последовательностях.
3. Оно различает объяснение через общую сумму и объяснение через временное накопление с утечкой. Различие выходов поддерживает роль временной структуры именно в заданной модели.

Вернуться к главе 9

Глава 10 Разбор практики

Прямое правило и повторная обработка могут выдать одинаковую категорию. Само совпадение не выбирает механизм. Две одинаковые карточки с разными скрытыми метками показывают информационную неоднозначность. Для дополнительного вывода нужны определённые правила, равные доступные сведения и условие, на котором ответы расходятся.

Ответы на самопроверку

1. Отчёт зависит от формата вопроса, памяти, времени и последовательности сообщения. Для вывода об опыте нужна обоснованная связь с этими процессами.
2. Контекст может использоваться разными архитектурами. Нужно сравнивать модели с сопоставимыми входами и ресурсами на условиях, где их предсказания различаются.
3. Кому или чему именно доступна информация, для какой операции и в какой момент. Возможность классифицировать, сообщать, запоминать и проверять ответ нельзя автоматически считать одной и той же способностью.

Вернуться к главе 10

Глава 11 Разбор практики

История позволяет сопоставить эффект с прошлой командой. Для выбранного задержанного ряда связь восстанавливается с лагом два. Если два механизма порождают одинаковые наблюдения на всех проведённых шагах, текущих данных для выбора нет. Скрытая колонка нужна оценщику, но не должна участвовать в вычислениях исследуемой модели.

Ответы на самопроверку

1. Общую причину могут иметь сигналы от внешних объектов. Для телесной принадлежности нужны дополнительные отношения к представлению тела и данные о соответствующем переживании.

2. Новые вмешательства в команды и сравнение последствий. Если наблюдения при двух версиях одинаковы, одного пассивного совпадения недостаточно для их уверенного различения.
3. Улучшение может объясняться общей вместимостью памяти. Чтобы выделить организацию по источникам, нужно сравнивать её с подходящими альтернативами и проверять перенос на новые условия.
Вернуться к главе 11

Глава 12 Разбор практики

Первичные прогнозы равны примерно 0,500; 0,667; 0,750; 0,600; 0,667; 0,571. Их средняя квадратичная ошибка около 0,309, у постоянной модели 0,250. Это исход одной короткой серии. На продолжении нужно начинать с сохранённых счётчиков, а не повторять исходное сглаживание. Проверяется качество прогноза, а не глубина переживаемой уверенности.

Ответы на самопроверку

1. Он использует информацию, которая раньше отсутствовала. Проверяемый прогноз должен быть зафиксирован до обратной связи, иначе возникает подсказка из будущего.
2. Отдельные задания могут различаться по трудности и доступным признакам. Общая история оценивает среднюю надёжность выбранного класса, а конкретный случай требует дополнительных сведений.
3. Обновление следует заданному правилу. Явная рефлексия рассматривает основания, применимость или ограничения этого правила и может привести к предложению и проверке другого способа оценки.
Вернуться к главе 12

Глава 13 Разбор практики

Полная система даёт 10 из 10 по обоим показателям. Без возврата распознавание и память дают 7 из 10. Без передачи распознавание сохраняет 10 из 10, память даёт 4 из 10; при отключении обоих показателей соответственно 7 и 4. Прямой конкурент даёт 10 из 10. Эти числа следуют из устройства, а не получены в исследовании людей.

```
Воспроизведение: python teaching_lab.py --chapter 13
```

Ответы на самопроверку

1. Почему глобальное распространение и рекуррентность не образуют строгую противоположность?
Рекуррентность описывает возвращение влияния, а глобальность — доступность содержания разным процессам. Общая система обмена может

использовать обратные связи. Различать нужно предполагаемые достаточные условия и функции конкретных взаимодействий.

2. Доказывает ли исчезновение отчёта после вмешательства исчезновение переживания?
Само по себе не доказывает. Вмешательство могло нарушить память, выбор ответа или выполнение инструкции. Вывод усиливают независимые показатели и контроль сохранности этих возможностей, хотя проблема измерения полностью не исчезает.

3. Зачем сравнивать механизмы при равном вычислительном бюджете?
Иначе преимущество можно ошибочно приписать особой архитектуре, хотя одна система просто получила больше шагов обработки. Контроль помогает выяснить, необходима ли организация взаимодействий для наблюдаемого результата.

Вернуться к главе 13

Глава 14 Разбор практики

Успех здесь означает разделение обязательств, а не выбор победителя. Паспорт должен позволять согласиться с устройством и отдельно спорить о связи с опытом. При одинаковых предсказаниях данных недостаточно для выбора; при отсутствии нужного формализма нельзя выдавать приближённую схему за вычисление критерия теории.

Ответы на самопроверку

1. Почему высокая корреляция между частями недостаточна для вывода об их взаимном причинном влиянии?
Обе части могут зависеть от третьего источника. Для причинного вывода нужны дополнительные предпосылки, вмешательства или специально обоснованная модель зависимостей.

2. Равна ли рефлексия любому представлению высшего порядка?
Нет. Развёрнутая рефлексия включает исследование собственных оснований, сравнение объяснений и пересмотр стратегии. Представление надёжности текущего сигнала может выполнять гораздо более узкую функцию.

3. Может ли теория быть математически точной и эмпирически спорной?
Да. Точность позволяет строго вывести последствия из допущений. Подтверждение того, что эти допущения и последствия описывают сознание, требует самостоятельных данных и сравнения с альтернативными объяснениями.

Вернуться к главе 14

Глава 15 Разбор практики

Для сигнала справа новая вероятность около 0,857. Для сигнала слева она равна 0,12 из суммы 0,12 и 0,32, примерно 0,273. При прямой связи команда совпадает со стороной цели, при обратной противоположна. Само совпадение ожиданий не доказывает переживаемое авторство; здесь проверяется вычислимая причинная модель.

```
Воспроизведение: python teaching_lab.py --chapter 15
```

Ответы на самопроверку

1. Является ли всякое предсказание сознательным ожиданием?
Нет. Предсказанием можно назвать вычисляемое распределение возможных сигналов. Наличие такого вычисления само по себе не устанавливает его переживаемость.

2. Чем модель действия отличается от модели мира?
Первая описывает связь собственных команд с последствиями, вторая — состояние и изменения окружающей ситуации. Их ошибки требуют разных исправлений, даже если внешне обе приводят к неудачному действию.

3. Почему отсутствие сигнала нужно отличать от отрицательного сигнала?
Неработающий датчик не сообщает, что цель отсутствует. Такое смешение создаёт необоснованную уверенность. Необходимо явно моделировать пропуск наблюдения и возможное изменение мира за время пропуска.

Вернуться к главе 15

Глава 16 Разбор практики

Правильный паспорт может содержать неизвестные поля. По отмене отчёта нельзя автоматически сделать вывод об отсутствии всех сопутствующих процессов. Декодирование показывает информацию, доступную анализу; необходимость и достаточность требуют иных оснований. Если следующий прогноз общий для двух теорий, он полезен для механизма, но пока не выбирает между теориями.

Ответы на самопроверку

1. Почему крупное исследование не обязано окончательно решить спор теорий?
Его масштаб повышает информативность определённых измерений, но не расширяет автоматически область проверенных утверждений. Значение имеют выбранные предсказания, задачи, показатели и чувствительность анализа.

2. Является ли условие без отчёта прямым измерением опыта?

Нет. Оно уменьшает некоторые требования к ответу, но нуждается в обосновании того, как определяется восприятие. Косвенные показатели также имеют предпосылки и ограничения.

3. Что сильнее меняет доверие к модели — ожидаемый результат или различающее предсказание?
Особенно информативен результат, который вероятен при одной модели и маловероятен при серьёзном конкуренте. Совпадение с общим ожиданием полезно, но слабее помогает выбрать объяснение.

Вернуться к главе 16

Глава 17 Разбор практики

Полная траектория по шагам 0–4: 0; 0,4; 0,6; 0,6; 0,6. Без поддержания после 0,6 получаются 0,3 и 0,15. Без первого входа получаются 0; 0; 0,4; 0,2; 0,1. Метка чтения состояния не меняет динамику, если отдельно не включена в правило. Все числа относятся к учебному механизму.

Ответы на самопроверку

1. Почему динамическое возникновение не равно эволюционному происхождению?
Первое описывает переходы состояния работающей системы. Второе касается истории появления и изменения её механизмов между поколениями. Они требуют разных данных и объясняют разные временные масштабы.

2. Доказывает ли устойчивая активность наличие сознательного эпизода?
Нет. Устойчивость объясняет сохранение состояния. Необходимо дополнительно объяснить содержание, его причинную роль и основание связи с переживанием.

3. Что означает расхождение нейронного и поведенческого показателей?
Оно ограничивает простое отождествление показателей и требует проверить несколько объяснений: чувствительность измерения, роль процесса в задаче и возможность компенсации. Расхождение само является полезными данными.

Вернуться к главе 17

Глава 18 Разбор практики

Например, отсутствие касания метки может зависеть от незамеченной детали или от трудности соответствующего движения. Контроль заметности и действия разделяет эти версии. Запрос помощи требует также социальной доступности помощи. Вывод о каждом достижении не равен датировке появления всех форм сознания.

Ответы на самопроверку

1. Почему возраст успешного самоузнавания нельзя считать возрастом появления сознания?
Самоузнавание требует определённого набора зрительных, двигательных и когнитивных возможностей. Сознательное переживание и остальные формы самоотнесения не тождественны выполнению этой конкретной задачи.

2. Чем запрос помощи отличается от явной рефлексии?
Запрос может регулироваться оценкой неопределённости без словесного анализа. Явная рефлексия предполагает, что человек делает собственные основания предметом рассмотрения, сравнивает их и способен целенаправленно пересматривать.

3. Что означает отсутствие ожидаемого показателя у младшей группы?
Оно допускает несколько объяснений: отсутствие функции, другую реализацию, дополнительные трудности задачи или недостаточную чувствительность измерения. Выбрать объяснение помогают контрольные задачи и наблюдение изменений во времени.

Вернуться к главе 18

Глава 19 Разбор практики

Обе учебные истории совместимы с указанными листьями. Польза памяти условна относительно изменчивости и затрат. Даже уверенное восстановление происхождения M не устанавливает опыт без обоснованной связи M с переживанием. Хорошая карта сохраняет все эти зависимости и указывает, какие данные нужны следующими.

Ответы на самопроверку

1. Доказывает ли адаптивная польза механизма наличие переживания?
Нет. Она объясняет, при каких условиях механизм мог поддерживаться отбором. Связь этого механизма с переживанием требует отдельного обоснования.

2. Почему близкородственные виды не дают полностью независимых подтверждений?
Их общие признаки могут восходить к одному историческому событию. Сравнительный анализ должен учитывать общую историю, иначе количество независимой информации будет завышено.

3. Зачем сохранять несколько историй происхождения?
При неполных данных они могут одинаково объяснять наблюдаемое распределение признаков. Явные альтернативы показывают, какие

дополнительные наблюдения действительно различат реконструкции и какие выводы пока преждевременны.

Вернуться к главе 19

Глава 20 Разбор практики

Нельзя требовать, чтобы небольшой проект сразу решил все вопросы сознания. Но его вывод должен точно соответствовать выполненной работе. Сильный принцип берёт на себя больше обязательств, чем предсказание отчёта; его нужно обосновывать последовательно, без произвольного исключения неудобных реализаций.

Ответы на самопроверку

1. Почему объяснение различения цветов может не решать вопрос о наличии опыта?
Оно показывает, как система классифицирует и сопоставляет сигналы. Утверждение о переживаемости этих процессов требует отдельной связи между вычислительной организацией и феноменальным свойством.

2. Обязательно ли вводить дополнительную сущность для объяснения переживания?
Нет. Возможна стратегия физического тождества. Но она должна обосновать, какой процесс тождествен опыту и почему, а не просто объявить два описания совпадающими.

3. Что считается прогрессом, если окончательного объяснения ещё нет?
Новое различающее предсказание, исключение серьёзной альтернативы, более точное соответствие структуры механизма структуре опыта и ясное сужение оставшихся допущений. Эти результаты можно проверять независимо от притязания решить весь вопрос.

Вернуться к главе 20

Глава 21 Разбор практики

Для исследования памяти допустим вывод об определённой зависимости выбора от доступных сведений при данных условиях. Для исследования повреждения регистрация воздействия, избегание и предполагаемый отрицательный опыт остаются разными уровнями. Ближайшую альтернативу нужно задать правилом или механизмом, а не словом «инстинкт».

Ответы на самопроверку

1. Почему признание возможной боли не означает признания явной рефлексии?

Отрицательное переживание, контроль состояния тела и рассмотрение собственных оснований — разные свойства. Данные о первом не становятся автоматически доказательством остальных.

2. Можно ли распространять результат одного вида на всех его родственников?
Только с дополнительным обоснованием, учитывающим механизмы, родство и условия задачи. Близость видов помогает строить гипотезу, но не заменяет проверку её границ.

3. Что делать с объяснением через простое правило?
Нужно записать его точно и проверить на тех же данных и новых условиях. Если правило объясняет результат, более сложный вывод ослабляется; если систематически ошибается, появляются основания для более богатой модели.

Вернуться к главе 21

Глава 22 Разбор практики

До нового Б реакции совпадают. На Б они равны 1 и 0,32768, поэтому наблюдение разделяет правила. После пяти шагов паузы ответ на А в специфической модели около 0,779695, у ресурса около 0,758219. Эти различия являются следствиями заданного расчёта и не устанавливают квалиа у системы.

```
Воспроизведение: python teaching_lab.py --chapter 22
```

Ответы на самопроверку

1. Может ли функциональная память существовать без нервной системы?
Да, состояние системы может сохранять влияние прошлых воздействий и менять будущую реакцию. Это широкий функциональный смысл памяти, который следует отличать от переживаемого воспоминания.

2. Почему уменьшение реакции ещё не доказывает обучение?
Его могут вызвать истощение, повреждение, изменение доступности входа или другие процессы. Нужны контрольные условия, проверка восстановления, специфичности и более точная модель альтернатив.

3. Делает ли отсутствие нейронов вопрос о квалиа бессмысленным?
Нет. Оно меняет круг возможных механизмов и требования к доказательству. Теория, допускающая иной носитель, должна указать достаточную организацию и способ проверки, а не ограничиваться общей сложностью системы.

Вернуться к главе 22

Глава 23 Разбор практики

Смена независимого шаблона может изменить текст без изменения вычислений. Целевое изменение используемого состояния способно менять действия, но его роль нужно проверить по правилам конкретной реализации. Схема поддерживает вопросы об архитектуре и доступе; сама по себе она не обосновывает переживаемую растерянность или иное качество.

Ответы на самопроверку

1. Почему успешный разговор о переживаниях недостаточен для вывода о них?
Текст может порождаться механизмом, не связанным с теми внутренними состояниями, о которых сообщает. Требуется исследовать происхождение ответа, доступ к информации и причинную роль предполагаемого самоконтроля.

2. Гарантирует ли биологический носитель сознание?
Нет. Материал не задаёт автоматически организацию, необходимую по рассматриваемой теории. Нужно исследовать конкретную систему, её связи и взаимодействие с внешним управлением.

3. Чем профиль кандидата лучше количества выполненных признаков?
Он сохраняет зависимость вывода от теории и различает необходимое, достаточное и сопутствующее. Простая сумма может скрыть отсутствие критического условия и создать необоснованную точность оценки.

Вернуться к главе 23

Глава 24 Разбор практики

Раздельная история может оказаться полезной, но совпадение с общим трекером ограничит вывод механизмом идентичности. Отсутствие различающих условий делает даже большую выборку неинформативной относительно этого вопроса. Готовая спецификация должна показывать, какие данные и какие вычислимые правила нужны следующими.

Ответы на самопроверку

1. История может использоваться общим механизмом обучения и отслеживания объектов. Для специального вывода нужен соответствующий конкурент.
2. Такое сравнение смешивает устройство модели с преимуществом доступной информации; вклад самого механизма пока не выделен.
3. Иначе гипотеза начинает служить собственным доказательством. Разделение показывает, что модель уже обязана учитывать и что ещё предстоит проверить.
Вернуться к главе 24

Глава 25 Разбор практики

M0 даёт 0,75; 0,25; 0,75; 0,75. M1 даёт 0,75; 0,4375; примерно 0,710526; примерно 0,858108. M2 и общий трекер дают 0,75; 0,25; 0,875; 0,5625. При удержании ноль все обновления начинаются с 0,5. Одновременное переименование источников меняет обозначения записей, но не числовые отношения.

```
Воспроизведение: python teaching_lab.py --chapter 25
```

Ответы на самопроверку

1. Она возвращается к истории A, сохранив его подтверждение отдельно от противоречия Б.
2. В данной реализации M1 и M2 совпадут, поэтому ценность отдельных историй проверить не удастся.
3. Ответ зависит также от шкалы и способа её использования. Связь вычисления с сообщением об опыте нельзя считать заданной самим названием переменной.

Вернуться к главе 25

Глава 26 Разбор практики

Независимая запись может воспроизвести обычную траекторию и не сохранить направленные ответы на вмешательства. Односторонний контроллер требует отдельного теста отсутствующего направления. Простой регулятор нельзя исключать одним впечатлением простоты: нужно проверить содержательные условия определения. Ни одна из этих проверок сама не доказывает достаточность организации для опыта.

```
Воспроизведение: python teaching_lab.py --chapter 26
```

Ответы на самопроверку

1. Он утверждает достаточность определённой организации для наличия опыта; предсказание отчётов устанавливает более ограниченную связь.
2. В гипотезе существенны отношения и переходы, сохраняемые при таком переименовании.
3. Нет. Сборка проверяет реализацию условий. Достаточность, границы реализации и перенос к феноменальному выводу требуют собственного обоснования.

Вернуться к главе 26

Глава 27 Разбор практики

M0 использует текущее окно, M1 общую историю, M2 историю узнанного источника. Если общий трекер совпадает с M2, выигрыш над M1 поддерживает отдельные истории, но не выделяет специальную телесную

организацию. Феноменальные следствия принципа требуют независимой проверки качественных отношений и не следуют из одного функционального сравнения.

Ответы на самопроверку

1. Нет. Нужны различающиеся модели или другой предмет сравнения.
2. Она показывает роль текущей причинной связи относительно слабого контроля, но не исключает полноценные альтернативные механизмы.
3. Описать его как обнаруженный в анализе, вывести новое предсказание и проверить на независимых данных.
Вернуться к главе 27

Глава 28 Разбор практики

Базовый полный режим имеет успех около 0,8794 и полезность 0,8619 на шаг. Фиксированная модель имеет средний успех около 0,4997, скрывающий 0,8780 до и 0,1213 после инверсии. Разница поддерживает функциональную пользу обновления управления в данной среде. Полезность и точность различаются из-за затрат на наблюдения.

Ответы на самопроверку

1. Мы отключаем один определённый механизм. Удаление всей памяти одновременно изменило бы несколько причин результата.
2. Среднее объединяет почти верное управление и почти систематическую ошибку после переключения.
3. Прогноз перестал бы измерять неопределённость до исхода. Возникла бы утечка ответа в оцениваемый показатель.
Вернуться к главе 28

Глава 29 Разбор практики

Поворот и перенос сохраняют расстояния с точностью численного представления, растяжение обычно меняет их. Рекуррентные шаги и те же развёрнутые слои совпадают. В задаче доставки полный отбор и прямой доступ дают 100 процентов: здесь показана работа канала, а не его исключительная необходимость или феноменальность.

Ответы на самопроверку

1. Выбранное преобразование сохраняет геометрию данной модели. Наличие переживания из этого не следует.
2. Проверены конкретные конечные вычисления. Теория может принимать дополнительные обязательства о времени, состоянии или физической реализации, которые требуют отдельной проверки.

3. Для этой задачи необходимые сведения могут использоваться по меньшей мере двумя способами. Вклад именно ограниченного общего пространства пока не выделен.
Вернуться к главе 29

Глава 30 Разбор практики

M2 совпадает с общим трекером по заданным правилам. Прямой внутренний датчик имеет доступ по проекту. Адаптивная стратегия меняет действия по прошлым ошибкам. У ускоренного пересмотра базового составного кандидата ноль включений. Это разные результаты, требующие разных следующих шагов; их нельзя сложить в доказательство опыта одной системы.

Ответы на самопроверку

1. Оно показывает, что минимальные вычисления пока не различают специальное самоотнесение и общий учёт истории источников.
2. Работу заданного канала доступа и измерительной проверки. Более сильные выводы требуют дополнительных оснований.
3. Нет. Механизм не активировался в базовом сценарии; нужен заранее определённый опыт, который проверит его предсказанную роль.
Вернуться к главе 30

Глава 31 Разбор практики

В учебной таблице сохраняются исходные идентификаторы и явные исключения. Фактическая структура скачанного файла не подменяется кратким описанием репозитория. Добавленный STTC может улучшить или ухудшить ошибку; допустимый вывод относится к выбранной процедуре и этому набору. Проверка по партиям строже проверки по органоидам относительно партийного переноса, но также не устраняет все зависимости и не устанавливает переживание.

```
Воспроизведение: python teaching_lab.py --chapter 31
```

Ответы на самопроверку

1. Результат могут совместно обеспечивать ткань, преобразование входа, алгоритм обратной связи и считывание.
2. Первая оценивает временной перенос внутри образца, вторая — перенос между образцами.
3. Нет. Оно устанавливает определённый функциональный результат; для феноменального вывода нужен дополнительный аргумент о механизме и принципе соответствия.
Вернуться к главе 31

Глава 32 Разбор практики

Профиль должен различать реализовано и проверено, реализовано без проверенной роли и неизвестно. Отсутствие эффекта отключения требует проверить активацию, чувствительность задачи и компенсацию. Даже полный функциональный профиль не подтверждает достаточность собственных теоретических требований независимым образом.

Ответы на самопроверку

1. Нет. Нужно проверить активацию, информационные зависимости и влияние на решение.
2. Первое показывает возможности внешнего анализа, второе — использование информации самой системой.
3. Реализация проверяет устройство кандидата, а достаточность является отдельной теоретической гипотезой и нуждается в независимых основаниях.

Вернуться к главе 32

Глава 33 Разбор практики

Восстановление не обязано быть безошибочным. Итог включает правило генерации, критерий сравнения, все исходы выбора и неопределённость. Однотипные дополнительные события не устраняют структурную неидентифицируемость. Победа среди неверных моделей может сочетаться с плохим абсолютным прогнозом.

```
Воспроизведение: python teaching_lab.py --chapter 33
```

В приложенном запуске для каждого генератора построено по 30 последовательностей из 120 событий. Выбор выполняется по наименьшей ошибке Брайера прогноза наблюдения; параметры кандидатов здесь фиксированы. Строка матрицы показывает заданный генератор, столбец — выбранного кандидата. Это диагностический учебный опыт, а не сравнение на неизвестной природе человека.

Генератор	Выбрана M0	Выбрана M1	Выбрана M2
M0	25	1	4
M1	3	24	3
M2	3	5	22

На выбранной сетке наименьшую ошибку Брайера 0,231314 дала пара параметров: сохранение истории 0,80, надёжность наблюдения 0,75. В этом запуске она совпала с заданной генератору парой. Один такой результат не является оценкой точности восстановления во всех условиях.

Ответы на самопроверку

1. Признак шумный и может возникать при разных скрытых причинах; нужен прогноз по обеим возможностям.
2. Нет. Оно проверяет анализ в известной искусственной ситуации.
3. Правило задаёт причинную динамику и новые предсказания. Незаметная замена затрудняет оценку того, что действительно было проверено.

Вернуться к главе 33

Глава 34 Разбор практики

В протоколе должны быть конкретные данные, порядок вычислений, основной результат и условия решения. Условный ориентир 0,01 является выбранным учебным критерием, а не универсальной научной нормой. Остановка после первого удачного различия без заранее заданной последовательной процедуры меняет смысл проверки.

```
Воспроизведение: python teaching_lab.py --chapter 34
```

Для 100 независимых последовательностей с двумя источниками среднее парное различие ошибки M1 минус M2 равно 0.016965; процентильный интервал 95 процентов при 4000 повторных выборках целых пар: [0.013687; 0.020376]. Заранее выбранный учебный порог равен 0,01. Интервал относится к генератору и процедуре этого опыта. Протокол сохранён до вычисления результата; это локальная фиксация, а не публичная регистрация независимого исследования.

Ответы на самопроверку

1. Нет. Такой анализ нужно отличать от заранее запланированной проверки.
2. Иначе в прогноз может попасть сам проверяемый ответ.
3. Нет. Это выбранный объём учебного примера; для конкретного исследования объём обосновывается требуемой точностью и различием моделей.

Вернуться к главе 34

Глава 35 Разбор практики

Среднее 0,008 в учебной иллюстрации не отвечает само по себе на вопрос о практически интересующем преимуществе 0,01. Совпадение версий при нуле включений компонента не проверяет его действие в активном режиме. Выбор сильного следующего опыта является содержательным результатом анализа.

Ответы на самопроверку

1. Выполненная задача не проверила последствия работы этого модуля; для оценки его вклада нужны различающие условия.

2. Тот же результат может дать общая память идентичности источников М1И, совпадающая с М2 по правилам.
3. Пересмотр задаёт явное новое правило и последствия для независимых условий; подгонка может ограничиваться улучшением уже увиденного результата.
Вернуться к главе 35

Глава 36 Разбор практики

Полноценная работа позволяет повторить путь от входов к выводу и не требует согласия с феноменальной гипотезой для проверки расчётов.
Самостоятельный вклад начинается с нового различимого следствия. Если предложенный принцип достаточности ещё не обоснован, это указывается точно и связывается с определённым следующим исследованием.

Ответы на самопроверку

1. Он установил следствия выбранного механизма раздельной памяти и его совпадение с общим трекером объектов.
2. Рабочая гипотеза требует взаимного причинного влияния; общая активность или одна передача сигнала его не устанавливают.
3. Да. Она должна точно определить выполненную проверку, сохранить исследовательскую цель и предложить следующий различающий шаг.
Вернуться к главе 36

Источники

[1] Dennett, Daniel C. (1988). Quining Qualia. In A. J. Marcel and E. Bisiach (eds.), Consciousness in Contemporary Science. Oxford University Press
[2] Chalmers, David J. (1995). Facing Up to the Problem of Consciousness. Journal of Consciousness Studies, 2(3), 200–219
[3] Shepard, Roger N. (1962). The analysis of proximities: Multidimensional scaling with an unknown distance function. I. Psychometrika, 27, 125–140
[4] Kalckert, Andreas; Ehrsson, H. Henrik (2012). Moving a Rubber Hand that Feels Like Your Own: A Dissociation of Ownership and Agency. Frontiers in Human Neuroscience, 6, 40
[5] Schwitzgebel, Eric (2008). The Unreliability of Naive Introspection. The Philosophical Review, 117(2), 245–273
[6] Fleming, Stephen M.; Lau, Hakwan C. (2014). How to measure metacognition. Frontiers in Human Neuroscience, 8, 443
[7] Machamer, Peter; Darden, Lindley; Craver, Carl F. (2000). Thinking about Mechanisms. Philosophy of Science, 67(1), 1–25
[8] Petitmengin, Claire (2006). Describing one’s subjective experience in the second person: An interview method for the science of consciousness. Phenomenology and the Cognitive Sciences, 5, 229–269
[9] Logothetis, Nikos K.; Pauls, Jon; Augath, Mark; Trinath, Torsten; Oeltermann, Axel (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature, 412, 150–157
[10] Nosek, Brian A.; Ebersole, Charles R.; DeHaven, Alexander C.; Mellor, David T. (2018). The preregistration revolution. Proceedings of the National Academy of Sciences, 115(11), 2600–2606
[11] Simmons, Joseph P.; Nelson, Leif D.; Simonsohn, Uri (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. Psychological Science, 22(11), 1359–1366
[12] Hodgkin, Alan L.; Huxley, Andrew F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. The Journal of Physiology, 117(4), 500–544
[13] Sperling, George (1960). The information available in brief visual presentations. Psychological Monographs: General and Applied, 74(11), 1–29
[14] Simons, Daniel J.; Chabris, Christopher F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28(9), 1059–1074
[15] Samad, Majed; Chung, Albert Jin; Shams, Ladan (2015). Perception of Body Ownership Is Driven by Bayesian Sensory Inference. PLOS ONE, 10(2), e0117178
[16] Dehaene, S.; Kerszberg, M.; Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. Proceedings of the National Academy of Sciences 95(24):14529–14534
[17] Lamme, V. A. F. (2006). Towards a true neural stance on consciousness. Trends in Cognitive Sciences 10(11):494–501
[18] Fahrenfort, J. J.; Scholte, H. S.; Lamme, V. A. F. (2007). Masking disrupts reentrant processing in human visual cortex. Journal of Cognitive Neuroscience 19(9):1488–1497
[19] Albantakis, L.; Barbosa, L.; Findlay, G.; Grasso, M.; Haun, A. M.; Marshall, W.; Mayner, W. G. P.; Zaeemzadeh, A.; Boly, M.; Juel, B. E.; Sasai, S.; Fujii, K.; David, I.; Hendren, J.; Lang, J. P.; Tononi, G. (2023). Integrated information theory (IIT) 4.0: Formulating the properties of phenomenal existence in physical terms. PLOS Computational Biology 19(10):e1011465
[20] Lau, H. C. (2008). A higher order Bayesian decision theory of consciousness. Progress in Brain Research 168:35–48

[21] Rao, R. P. N.; Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature Neuroscience 2(1):79–87
[22] Seth, A. K. (2013). Interoceptive inference, emotion, and the embodied self. Trends in Cognitive Sciences 17(11):565–573
[23] Cogitate Consortium; Ferrante, O.; Górska-Klimowska, U.; Henin, S.; Hirschhorn, R.; et al. (2025). Adversarial testing of global neuronal workspace and integrated information theories of consciousness. Nature 642:133–142
[24] Cohen, M. A.; Dembski, C.; Ortego, K.; Steinhibler, C.; Pitts, M. (2024). Neural signatures of visual awareness independent of postperceptual processing. Cerebral Cortex 34(11):bhae415
[25] He, L.; Zhang, Z.; Wang, Y.; Meng, Q.; Jiang, Y.; Xu, W.; Wu, J.; Sheng, J.; Ye, Z.; Ding, X. (2026). Report-modulated prefrontal activity and consistent posterior representations during conscious visual perception. Consciousness and Cognition 142:104075
[26] Zhang, Y.; Li, X.; Jin, Z.; Zhang, J.; Li, L. (2026). Distinct frontal lobe subregions mediate the emergence and reporting of visual consciousness. NeuroImage 334:121964
[27] Noorman, S.; Stein, T.; Zantvoord, J.; Fahrenfort, J.; van Gaal, S. (2025). A causal role of the NMDA receptor in recurrent processing during perceptual integration. eLife 13:RP100530
[28] Kouider, S.; Stahlhut, C.; Gelskov, S. V.; Barbosa, L. S.; Dutat, M.; de Gardelle, V.; Christophe, A.; Dehaene, S.; Dehaene-Lambertz, G. (2013). A neural marker of perceptual consciousness in infants. Science 340(6130):376–380
[29] Amsterdam, B. (1972). Mirror self-image reactions before age two. Developmental Psychobiology 5(4):297–305
[30] Goupil, L.; Romand-Monnier, M.; Kouider, S. (2016). Infants ask for help when they know they don't know. Proceedings of the National Academy of Sciences 113(13):3492–3496
[31] Felsenstein, J. (1985). Phylogenies and the Comparative Method. The American Naturalist 125(1):1–15
[32] Barron, A. B.; Klein, C. (2016). What insects can tell us about the origins of consciousness. Proceedings of the National Academy of Sciences 113(18):4900–4908
[33] Andrews, K.; Birch, J.; Sebo, J.; et al.; individual signatories (2024). The New York Declaration on Animal Consciousness. Официальное заявление, 19 апреля 2024
[34] Hampton, R. R. (2001). Rhesus monkeys know when they remember. Proceedings of the National Academy of Sciences 98(9):5359–5362
[35] Nieder, A.; Wagener, L.; Rinnert, P. (2020). A neural correlate of sensory consciousness in a corvid bird. Science 369(6511):1626–1629
[36] Crook, R. J. (2021). Behavioral and neurophysiological evidence suggests affective pain experience in octopus. iScience 24(3):102229
[37] Barr, S.; Elwood, R. W. (2024). Trade-Offs between Avoidance of Noxious Electric Shock and Avoidance of Bright Light in Shore Crabs Are Consistent with Predictions of Pain. Animals 14(5):770
[38] Manzi, O.; Lynch, K. E.; Allman, D. M.; Latty, T.; White, T. E. (2026). Flexible self-protection as evidence of pain-like states in house crickets. Proceedings of the Royal Society B 293(2070):20260609
[39] Toyota, M.; Spencer, D.; Sawai-Toyota, S.; et al. (2018). Glutamate triggers long-distance, calcium-based plant defense signaling. Science 361(6407):1112–1115
[40] Boisseau, R. P.; Vogel, D.; Dussutour, A. (2016). Habituation in non-neural organisms: evidence from slime moulds. Proceedings of the Royal Society B 283(1829):20160446
[41] Markel, K. (2020). Lack of evidence for associative learning in pea plants. eLife 9:e57614
[42] Butlin, P.; Long, R.; Bayne, T.; Bengio, Y.; Birch, J.; Chalmers, D.; Constant, A.; Deane, G.; Elmoznino, E.; Fleming, S. M.; Ji, X.; Kanai, R.; Klein, C.; Lindsay, G.; Michel, M.; Mudrik, L.; Peters, M. A. K.; Schwitzgebel, E.; Simon, J.; VanRullen, R. (2026). Identifying indicators of consciousness in AI systems. Trends in Cognitive Sciences 30(6):488–501

[43] Wilson R. C., Collins A. G. E. (2019). Ten simple rules for the computational modeling of behavioral data. eLife 8, e49547
[44] Kagan B. J. et al. (2022). In vitro neurons learn and exhibit sentience when embodied in a simulated game-world. Neuron 110(23), 3952–3969.e8
[45] Cai H. et al. (2023). Brain organoid reservoir computing for artificial intelligence. Nature Electronics 6, 1032–1039
[46] Robbins A. et al. (2026). Goal-directed learning in cortical organoids. Cell Reports 45(2), 116984
[47] Robbins A. (2025). Goal-Directed Learning in Cortical Organoids — Experimental Data and Code. Zenodo, version v2, doi:10.5281/zenodo.17684862
[48] International Society for Stem Cell Research (2025). Guidelines for Stem Cell Research and Clinical Translation. Version 1.2, August 2025 update
[49] Lindsey J. (2025). Emergent Introspective Awareness in Large Language Models. Исследовательский отчёт Transformer Circuits, 29 October 2025; препринт arXiv:2601.01828, 5 January 2026

Автор и проекты

Евгений Бренев

Чтобы сообщить о найденной ошибке, обсудить материал или познакомиться с проектами автора, используйте эти адреса.

ЭЛЕКТРОННАЯ ПОЧТА

info@brenev.com

САЙТ АВТОРА

brenev.com

САЙТ МИРЫ

mira.technology

ДРУГИЕ ПРОЕКТЫ

haraizen.tech

ТЕЛЕГРАМ

https://t.me/Eudgenee

Название книги отсылает к авторскому проекту искусственного интеллекта «Мира».