

ИСКУССТВО ПЕРЕМЕН



БОРИС КРИГЕР

БОРИС КРИГЕР

**ИСКУССТВО
ПЕРЕМЕН**

ALASPERA

© 2021 Борис Кригер

Опубликовано издательством Altaspera Publishing .

Борис Кригер — междисциплинарный философ, занимающийся вопросом о том, как разрозненные области знаний могут быть объединены в целостное видение человеческого существования. В своих работах он стремится преодолеть разделение философии и науки, этики и политики, индивидуального опыта и коллективных структур. Объединяя идеи экзистенциализма, социальной теории, когнитивной науки и технологических исследований, он разрабатывает способ мышления, который не является ни редукционистским, ни утопическим, а открыт сложности современного мира.

Искусство перемен

Мы часто стремимся переосмыслить себя, свои организации или общество, веря, что новый старт возможен, если только мы достаточно постараемся. Однако опыт неоднократно показывает, что мы редко начинаем с нуля. Ограничения — некоторые видимые, многие скрытые — формируют каждое новое начало.

Опираясь на философию, когнитивную науку, искусственный интеллект и исторические примеры, книга *«Искусство перемен»* исследует эту простую, но неуловимую истину: каждая отправная точка уже имеет структуру. Не существует истинного чистого листа, нет «нулевого года», нет нейтрального начала.

Вдохновленная мудростью: «Даруй мне спокойствие, чтобы принять то, что я не могу изменить, мужество, чтобы изменить то, что я могу, и мудрость, чтобы отличить одно от другого», — эта книга предлагает простую схему, помогающую отличить то, что можно изменить, от того, что изменить нельзя.

В основе лежит простая идея: каждое начальное условие содержит **предопределенные элементы** (то, что мы должны принять или обойти) и **свободные элементы** (то, что мы можем изменить). Распознавание трех распространенных ошибок — отношение к неизменяемому как к изменяемому, принятие изменяемого как фиксированного или путаница между одной областью и другой — может помочь нам избежать многих распространенных ошибок.

На повседневных примерах — от организационных реформ и разработки ИИ до личного роста и принятия политических решений — книга предлагает читателям провести собственный «анализ ограничений» и сосредоточить усилия там, где они действительно могут принести результаты.

Ключевые слова

начальные условия, структурные ограничения, *tabula rasa*, методологические ошибки, индуктивная предвзятость, зависимость от пути, моделирование системы

Содержание

ПРЕДИСЛОВИЕ	6
ГЛАВА 1: Соблазн чистого листа	16
ГЛАВА 2: Цена иллюзии.....	34
ГЛАВА 3: Паралич ложной неподвижности	49
ГЛАВА 4: Анатомия любой отправной точки: P + F	66
ГЛАВА 5: Три типа ошибок.....	79
ГЛАВА 6: Двойная предварительная подготовка	90
ГЛАВА 7: Для менеджеров — Стратегия в унаследованных структурах ..	95
ГЛАВА 8: Для ученых — Границы моделей.....	102
ГЛАВА 9: Для разработчиков ИИ — наследование — это преимущество, а не недостаток.....	108
ГЛАВА 10: Для политиков — самые дорогостоящие ошибки	114
ГЛАВА 11: Для всех — Личные решения и отношения	121
ГЛАВА 12: Двадцать ловушек и как из них выбраться.....	128
ПОСЛЕСЛОВИЕ: Мудрость отличия как практика	139
ПРИЛОЖЕНИЯ.....	144
Приложение А: Глоссарий терминов	144
Приложение В: Примеры ошибок по областям знаний.....	146
Приложение С: Шаблон карты ограничений.....	149
Приложение D: Рекомендуемая литература	150

Приложение Е: Протокол аудита ограничений (10 шагов)	151
Приложение F: Полный текст научной работы	154

ПРЕДИСЛОВИЕ

Даруй мне спокойствие, чтобы принять то, что я не могу изменить, мужество, чтобы изменить то, что я могу, и мудрость, чтобы отличить одно от другого.

Вы это уже слышали. Возможно, вы сами это произносили или видели в рамке на чьей-то стене. Молитва о безмятежности, приписываемая богослужителю Рейнхольду Нибуру, стала одним из самых часто повторяемых примеров практической мудрости в современном мире. Она кажется настолько очевидной, настолько самоочевидной, что мы киваем в знак согласия и продолжаем свой день.

Но вот неудобный вопрос: если эта мудрость так очевидна, почему же умные, доброжелательные люди постоянно не применяют её на практике?

Представьте себе предпринимателя, который вкладывает миллионы в стартап, основанный на изменении фундаментальных аспектов человеческого поведения, и обнаруживает, что люди упорно

остаются самими собой. Представьте себе реформатора политики, который разрабатывает идеальную систему на бумаге, и видит, как она рушится под натиском институциональной инерции. Представьте себе родителя, который пытается сформировать своего ребенка по заранее заданному образу, или супруга, который женится на ком-то, ожидая от него перемен. Представьте себе исследователя в области искусственного интеллекта, который обещает беспристрастный алгоритм, как будто десятилетия человеческих знаний можно стереть по желанию.

Эта закономерность повторяется во всех сферах человеческой деятельности: мы постоянно пытаемся изменить то, что изменить нельзя, одновременно отказываясь менять то, что можно изменить. Мы изнуряем себя, толкаясь о неподвижные стены, игнорируя открытые двери рядом с ними.

Почему?

Как я пришел к выводу, причина в отсутствии у нас систематической структуры для проведения этого

различия. Молитва о безмятежности говорит нам, *что* делать, но не *как* это делать. Она предполагает, что мы можем надежно определить, что относится к каждой категории. Но именно в этом определении мы снова и снова терпим неудачу .

Происхождение этой книги

Эта книга начиналась как научная статья. В работе «Структурные предпосылки описываемых операций» я разработал формальную концепцию понимания того, почему определенные ограничения неизбежны в любой системе, которую мы можем описать, смоделировать или проанализировать. В статье использовались знания философии, когнитивной науки, математики, искусственного интеллекта и институциональной теории для создания строгой основы для древней мудрости о понимании различий.

Основная идея этой статьи обманчиво проста: каждая отправная точка уже содержит структуру. Нет чистого листа, нет нулевого года, нет по-настоящему нейтрального начала. Это не просто философское наблюдение — это практическая реальность,

имеющая глубокие последствия для любого, кто пытается создавать, изменять или понимать сложные системы.

Однако научные статьи по своей природе обращены к специалистам. В них используется техническая терминология, математическая нотация и формальные условности научного дискурса. Несколько коллег и читателей призвали меня перевести эти идеи в форму, доступную более широкой аудитории: руководителям, сталкивающимся с организационными изменениями, ученым, разрабатывающим исследовательские программы, разработчикам ИИ, создающим системы обучения, политикам, стремящимся к реформам, и всем, кто пытается внести значимые изменения в свою жизнь.

Это та книга, которую вы держите в руках.

Что предлагает эта книга

Эта книга — руководство по «мудрости понимания разницы». Она переводит интуитивную

привлекательность Молитвы о безмятежности в систематический метод, который вы можете реально применить .

Основная концепция проста. Каждое начальное условие — каждая отправная точка для любых изменений, которые вы хотите внести, — может рассматриваться как содержащая два типа элементов:

P = переопределенные элементы **P** (то, что нельзя изменить в вашем контексте).

F = **свободные** элементы (то, что можно изменить)

Ни **P**, ни **F** никогда не равны нулю. Вы всегда что-то наследуете; у вас всегда есть некоторый запас для маневра. Искусство заключается в том, чтобы правильно определить, что есть что.

Большинство наших ошибок делятся на три типа. Во-первых, мы путаем **P** с **F**: мы пытаемся изменить то, что изменить нельзя. Во-вторых, мы путаем **F** с **P**: мы принимаем за истину то, что на самом деле можно изменить. В-третьих, мы путаем области применения: мы переносим ограничения из одной области в

другую, где они неприменимы.

Как только вы научитесь распознавать эти закономерности, вы начнете замечать их повсюду. Эта модель становится диагностическим инструментом, способом понять, почему инициативы терпят неудачу и как они могут привести к успеху.

Как читать эту книгу

Я построил эту книгу таким образом, чтобы она была полезна разным читателям по-разному. Первые три главы посвящены проблеме: почему иллюзия чистого листа так соблазнительна, что происходит, когда мы ей поддаёмся, и противоположной ошибке — предположению, что ничего не может измениться. Эти главы богаты историческими примерами и поучительными историями.

В главах с 4 по 6 представлена основная структура: формула $P + F$, три типа ошибок и концепция «двойного предварительного обусловливания» — идея о том, что ограничения возникают как из самой системы, так и из инструментов, которые мы

используем для ее описания.

В главах с седьмой по одиннадцатую представленная концепция применяется к конкретным аудиториям. Менеджеры найдут здесь рекомендации по организационным изменениям. Ученые получат представление об ограничениях моделей. Разработчики ИИ найдут новый способ осмысления предвзятости и архитектуры. Политики найдут предостережения, извлеченные из дорогостоящих исторических ошибок. И каждый найдет применение этой концепции в личных решениях и отношениях.

В главе двенадцатой представлен краткий справочник: двадцать распространенных ситуаций с диагностикой и рекомендациями.

В приложениях содержатся глоссарий, расширенные примеры, практические шаблоны и рекомендуемая литература. Для тех, кто хочет ознакомиться с полным техническим описанием, в Приложении F представлен полный текст научной статьи со всем математическим формализмом и строгими доказательствами.

Читать в хронологическом порядке необязательно. Если вы менеджер, столкнувшийся с цифровой трансформацией, начните с главы седьмой. Если вы исследователь в области ИИ, переходите к главе девятой. Если вы имеете дело со сложными отношениями, глава 11 может оказаться наиболее актуальной для вас. Структура книги одинакова на протяжении всего текста; различаются только области применения.

Примечание о научном подходе

Некоторые читатели могут задаться вопросом: можно ли действительно систематизировать мудрость? Разве знание того, что следует принять, а что изменить, не является вопросом суждения, опыта или интуиции?

Да, и именно в этом суть. Здравый смысл, полезный опыт и надежная интуиция зависят от наличия точных ментальных моделей. Когда ваша модель мира правильно различает неизменное и гибкое, ваше суждение улучшается. Когда же это не так, даже самые мудрые инстинкты сбивают вас с пути.

Применяемый мной научный подход не заменяет мудрость, а, наоборот, оттачивает её. Делая явным то, что часто подразумевается, предоставляя терминологию для закономерностей, которые иначе остаются без названия, мы создаём инструменты для более ясного мышления. Цель состоит не в том, чтобы устранить суждения, а в том, чтобы дать им более качественный материал для работы.

В научной статье, лежащей в основе этой книги, с логической строгостью изложены определенные выводы. Утверждение о том, что любой описываемый процесс требует ненулевой начальной структуры, — это не просто мнение, а необходимое следствие того, что вообще означает описание изменений. Понимание того, *почему* это так, а не просто *то, что* это так, дает вам интеллектуальную защиту от соблазнительной, но в конечном итоге дорогостоящей иллюзии чистого листа.

Приглашение

Мы живем в эпоху масштабных преобразований. Компании меняют курс. Страны проводят реформы.

Люди заново открывают себя. Технологии обещают переделать всё. В этом контексте способность точно оценивать, что можно изменить, а что нельзя, не только представляет философский интерес, но и имеет практическое значение.

Мудрость ограничений — это не совет отчаяться. Совсем наоборот. Принимая то, что действительно нельзя изменить, мы высвобождаем свою энергию для того, что можно изменить. Осознавая структуру, которую мы унаследовали, мы учимся работать с ней, а не изнурять себя борьбой с ней. Понимая, что каждый чистый лист — это иллюзия, мы становимся более реалистичными в отношении того, что на самом деле требуется для трансформации .

В конечном счете, именно это освобождение и предлагает эта книга: не свободу от ограничений, а свободу через их понимание.

Добро пожаловать в мир мудрости ограничений.

ГЛАВА ПЕРВАЯ: СОБЛАЗН ЧИСТОГО ЛИСТА

«В начале не было ничего, что затем взорвалось». —
Терри Пратчетт.

Идея начать все заново обладает почти непреодолимой привлекательностью. Чистый лист. Пустая страница. Год с нуля. Шанс все сделать правильно на этот раз, не обремененный накопленными ошибками прошлого.

Эта соблазнительность глубоко укоренилась в человеческой психологии. Мы мечтаем о новых началах: новогодних обещаниях, свежих стартах, полных перемен. Предприниматели говорят о разрушениях так, будто существующие структуры можно сметти, как пыль. Революционеры обещают построить общество заново с нуля. Психотерапевтов спрашивают, могут ли люди действительно измениться, как будто прошлое можно стереть и построить заново.

Но вот неприятная правда, которую исследует эта глава: чистый лист всегда является иллюзией.

Понимание того, почему это так — истинное понимание, а не простое кивание головой в знак согласия — это первый шаг к мудрости, которую предлагает эта книга.

Философские корни.

Концепция *tabula rasa* — лат. «чистая доска» — имеет древние корни. Аристотель описывал интеллект как нечто, сравнимое с письменной доской, на которой ничего не написано. Стоики развили эту точку зрения, предположив, что разум при рождении подобен листу бумаги, готовому к письму.

Но именно Джон Локк в своем «*Очерке о человеческом понимании*» 1689 года придал этой идее ее современную форму. Локк утверждал, что разум при рождении лишен врожденных идей — что все знания проистекают из ощущений и размышлений, из опыта, записываемого на чистом листе сознания.

Эмпиризм Локка был революционным для своего времени. Он бросил вызов религиозному и

аристократическому предположению о том, что некоторые люди рождаются с превосходными знаниями или врожденной ценностью. Если все умы изначально равны, то различия в достижениях должны отражать различия в опыте и образовании — а опыт можно изменить.

Это было глубоко демократичное и оптимистичное видение. Оно предполагало, что человек бесконечно податлив, что правильное образование может привести к любому желаемому результату, что общество можно усовершенствовать посредством проектирования окружающей среды. Чистый лист стал основой для прогрессивной политики, реформы образования и социальной инженерии.

Проблема в том, что это неправильно. Не совсем неправильно — опыт действительно имеет огромное значение — но неправильно в своем радикальном утверждении, что мы начинаем с нуля.

Биологическая реальность.

Современная когнитивная наука основательно

разрушила представление о «чистом листе». Люди не рождаются недифференцированными обучающимися машинами. Мы приходим в мир со сложным набором врожденных способностей, предрасположенностей и ограничений, которые формируют все, чему мы впоследствии учимся.

Рассмотрим процесс усвоения языка. Дети не учатся языку так, как, например, играют в шахматы или катаются на велосипеде. Они усваивают его быстро, без усилий, следуя предсказуемой последовательности развития, которая удивительно устойчива в разных культурах и языках. Двухлетний ребенок, подвергающийся воздействию языка, станет свободно говорящим на нем; взрослый же, подвергающийся воздействию того же самого, будет вечно испытывать трудности с грамматикой и акцентом.

Аргумент Ноама Хомского о «бедности стимула» убедительно доказывает это. Языковой ввод, который получают дети, фрагментарен, часто неграмматичен и совершенно недостаточен для объяснения тех

сложных грамматических знаний, которые они быстро усваивают. Единственное объяснение заключается в том, что дети привносят в процесс нечто новое — врожденный механизм усвоения языка, универсальную грамматику, которая ограничивает и направляет обучение.

Та же закономерность наблюдается во всех когнитивных областях. Младенцы обладают врожденными представлениями о физике: они удивляются, когда предметы, кажется, проходят сквозь друг друга или парят в воздухе. У них есть врожденная интуиция в отношении чисел, они отслеживают малые количества еще до того, как научатся считать. У них есть врожденная система распознавания лиц, они с рождения предпочитают изображения, похожие на лица. У них есть врожденное социальное познание, позволяющее им отличать одушевленные предметы от неодушевленных уже в первые месяцы жизни.

Ничто из этого не усваивается с чистого листа. Этот лист изначально отмечен эволюционной мудростью,

накопленной за миллионы лет — мудростью о структуре физического мира, свойствах объектов, природе других разумов. Это унаследованное знание делает обучение возможным. Без него бушующая, жужжащая путаница чувственного опыта навсегда осталась бы непостижимой.

От культуры стартапов до Силиконовой долины.

Если концепция «чистого листа» научно несостоятельна, почему она сохраняет столь сильное культурное влияние? Частично ответ кроется в её полезности как идеологии.

Вспомним мифологию Кремниевой долины. Евангелие стартапов проповедует революционные изменения: идею о том, что гибкие новички могут свергнуть устоявшиеся отрасли, мыслив нестандартно, отказываясь ограничиваться тем, как всё делалось всегда. Действуйте быстро и ломайте всё. Терпите неудачи быстро, терпите неудачи часто. Прошлое — это мёртвый груз; важно только будущее. Эта мифология обладает реальной силой, потому что

содержит в себе реальную истину . Устоявшиеся компании действительно становятся самодовольными. Устаревшие системы действительно накапливают технический долг. Отрасли действительно подвергаются разрушительным воздействиям со стороны инноваций, которые их предшественники не смогли предвидеть. Предприниматели, которые преобразовали вычислительную технику, коммуникации и торговлю, часто делали это именно потому, что отказывались принимать унаследованные представления о том, что возможно.

Но мифология также систематически преувеличивает возможность начать все с нуля. Даже самый радикальный стартап не начинается с ничего. Он наследует языки программирования, инструменты разработки, инфраструктуру, кадровый резерв, правовые рамки, рыночные структуры и культурные ожидания. Он наследует знания, которые сделали возможными его инновации. Он наследует экономические условия, которые делают его бизнес-

модель жизнеспособной.

Более тонкий момент заключается в том, что успешные стартапы наследуют от своих основателей ментальные модели — модели, сформированные образованием, опытом и культурным фоном. Широко известный «бросивший учебу, но построивший многомиллиардную компанию в гараже» по-прежнему несет в себе неявные представления о клиентах, технологиях и организации, которые были усвоены за десятилетия, а не изобретены с нуля.

Когда стартапы терпят неудачу — а большинство терпят — причина часто кроется в непонимании этих унаследованных черт. Основатель, который предполагает, что клиенты немедленно откажутся от привычных моделей поведения в пользу превосходящих альтернатив. Команда, которая считает, что техническая элегантность превзойдет пользовательские привычки. Компания, которая игнорирует нормативные ограничения до тех пор, пока они не станут угрозой для существования бизнеса. Все эти неудачи имеют общую причину:

отношение к predetermined параметрам как к свободно изменяемым.

Нулевой год в политике.

Наиболее яркие примеры применения мышления «с чистого листа» наблюдались в политике. Революционные движения на протяжении всей истории обещали смести старый порядок и построить общество заново. Результаты варьировались от разочаровывающих до катастрофических.

Французская революция служит показательным примером. В тысяча семсот девяносто втором году революционеры не просто свергли монархию; они попытались перевернуть само время. Республиканский календарь отменил традиционную неделю, реорганизовал месяцы, переименовал дни и начал нумерацию лет с первого года республики. Это был не просто символизм — это отражало подлинную веру в то, что прошлое можно стереть, что общество можно рационализировать, исходя из первопричин.

Этот календарь просуществовал около двенадцати

лет, прежде чем Наполеон восстановил григорианскую систему. Попытка перестроить нечто столь фундаментальное, как организация времени человеком — связанная с сельскохозяйственными циклами, религиозными обрядами, биологическими ритмами — потерпела неудачу из-за predetermined ограничений, которые революционеры пытались игнорировать.

Но это был незначительный провал по сравнению с тем, что последовало позже. В двадцатом веке мышление с чистого листа достигло своих самых экстремальных и смертоносных форм.

После захвата власти в Камбодже в тысяче девятсот семидесяти пятом году красные кхмеры объявили «Год ноль». Города были опустошены. Валюта была упразднена. Религия была запрещена. Семьи были разделены. Образование, выходящее за рамки базовой грамотности, было запрещено. Цель состояла в том, чтобы стереть камбоджийское общество и перестроить его в аграрную утопию, свободную от пороков современности.

За этим последовал один из самых ужасных геноцидов в истории. От полутора до двух миллионов человек — примерно четверть населения Камбоджи — погибли от казней, голода, болезней и переутомления всего за четыре года. Заранее установленные ограничения, которые пытались устранить красные кхмеры — необходимость в технических знаниях, прочность семейных уз, сложность сельскохозяйственного производства, пределы человеческой выносливости — вновь проявились с катастрофической силой.

Советские и маоистские эксперименты по созданию «нового социалистического человека» отражают одну и ту же закономерность. Предположение о том, что человеческую природу можно переделать посредством надлежащей социальной организации, что эгоизм и индивидуализм являются всего лишь продуктами капитализма, которые следует искоренить путем коллективизации, — эти убеждения привели к политике, породившей голод, репрессии и десятки миллионов смертей.

Общим моментом является рассмотрение унаследованных человеческих характеристик — потребности в стимулах, привязанности к семье, ограничений планирования сверху вниз — как свободно изменяемых переменных, а не как ограничений, в рамках которых необходимо действовать.

Полноценный разрыв в личной жизни.

Мышление «с чистого листа» поражает не только революционеров и предпринимателей. Оно формирует подход обычных людей к личным переменам и отношениям.

Сколько отношений закончились с убеждением, что начало новых отношений с кем-то другим решит все проблемы? Проблемы в нынешних отношениях — конфликты, разочарования, накопившиеся обиды — кажутся такими специфическими именно для этих отношений. Полный разрыв, новое начало с другим человеком, безусловно, были бы другими.

Иногда всё по-другому. Но часто одни и те же модели

поведения повторяются. Ревность, отравившая первые отношения, отравляет и вторые. Проблемы в общении повторяются и с новым партнёром. Фундаментальная несовместимость между тем, чего человек хочет, и тем, что он готов дать, сохраняется в разных отношениях.

Предопределенные элементы — стили привязанности, сформированные в детстве, черты характера, привычные механизмы преодоления трудностей — сопровождают нас в отношениях. Они являются частью нашей «П», а не «Ф». Отношение к ним как к чему-то, что можно устранить изменением внешних обстоятельств, — это верный путь к повторным разочарованиям.

Та же закономерность наблюдается при смене карьеры, переезде и личностном переосмыслении. «Я начну новую жизнь в новом городе». «Эта смена карьеры позволит мне стать другим человеком». «Как только я похудею / получу диплом / найду подходящего партнера, все изменится».

Внешние изменения действительно могут изменить

параметры нашей жизни. Новая среда открывает новые возможности. Но «я», попадающее в новую среду, несёт с собой свою историю, свои привычки, свои характерные способы восприятия и реагирования. Чистый лист, который мы представляем себе по другую сторону трансформации, уже покрыт написанным о том, кто мы есть на самом деле.

Непредвзятый ИИ.

Пожалуй, ни одна область не иллюстрирует современный подход к мышлению с чистого листа лучше, чем искусственный интеллект.

В разработке ИИ сохраняется устойчивая иллюзия о непредвзятом алгоритме — системе, обрабатывающей информацию без предрассудков и предубеждений, которые поражают человеческое мышление. Если бы мы могли обучить модель на достаточном количестве данных, используя правильные методы, мы могли бы создать объективный интеллект, свободный от искажений, навязанных человеческой историей и культурой.

Эта фантазия бессвязна, и понимание причин этого раскрывает нечто важное о всеобщей невозможности существования «чистого листа».

Каждая система искусственного интеллекта наследует ограничения на нескольких уровнях. Архитектура — будь то нейронная сеть, дерево решений или что-то еще — воплощает предположения о том, какие закономерности стоит выявлять и как должна передаваться информация. Это проектные решения, принимаемые людьми, отражающие человеческую интуицию относительно релевантности и структуры.

Обучающие данные отражают мир, в котором они были созданы, — со всеми его предвзятостями, дисбалансами и ограничениями. Обучите языковую модель на человеческом тексте, и она усвоит человеческие ассоциации, в том числе и проблемные. Обучите классификатор изображений на фотографиях, и он выучит все корреляции, которые содержатся в этих фотографиях.

Целевая функция — математическое описание того,

что система пытается оптимизировать, — воплощает в себе ценности. Точность в какой задаче? Как она измеряется? Какой ценой для других факторов? Это не нейтральные технические решения, а решения, наполненные ценностями, которые ограничивают все, чему система впоследствии научится.

Используемые для оценки производительности метрики дополнительно определяют, что именно оптимизируется. Системы учатся максимизировать то, что мы измеряем, а не обязательно то, что нас интересует. Выбор метрики — еще одно унаследованное ограничение.

Мечта о беспристрастном ИИ — это, по сути, мечта об избавлении от человеческого наследия, о создании интеллекта, не обремененного той конкретной историей, которая нас породила. Но системы ИИ создаются людьми, обучаются на данных, созданных людьми, предназначены для служения человеческим целям и оцениваются по человеческим критериям. Они неизбежно наследуют те структуры, которые мы им привносим.

Это не ошибка, которую нужно исправить; это фундаментальная особенность работы обучающих систем. Мудрость заключается не в стремлении к недостижимой цели — полному отсутствию предвзятости, а в понимании, документировании и надлежащем управлении неизбежно существующими предубеждениями.

Общий шаблон.

Во всех этих областях — философии, когнитивной науке, предпринимательстве, революции, личностных изменениях и искусственном интеллекте — мы видим одну и ту же закономерность. Привлекательность чистого листа заключается в обещании свободы и контроля. Если мы можем начать с нуля, нас ничто не ограничивает. Мы можем создать именно то, что хотим, не будучи искалеченными ошибками прошлого.

Но это обещание ложно. Чистый лист всегда уже написан. Чистая страница всегда уже отмечена. «Нулевой год» всегда продолжает те структуры, которые, как он утверждает, отвергает.

Соблазн начать с чистого листа заставляет нас недооценивать ограничения, в рамках которых мы действуем. Мы вкладываем ресурсы в изменение того, что изменить нельзя. Мы проектируем системы, которые потерпят неудачу, потому что игнорируют унаследованную структуру. Мы даем обещания, которые не можем сдержать, устанавливаем ожидания, которые не можем оправдать, и терпим разочарования, которых можно было бы избежать.

Альтернатива — не пессимизм или фатализм, а реализм. Понимая, что мы на самом деле унаследовали, мы можем работать с этим, а не против этого. Осознавая свою «П», мы можем сосредоточить энергию на своей «Ф» — подлинных степенях свободы, которыми мы обладаем.

Именно это различие и является предметом данной книги.

ГЛАВА ВТОРАЯ: ЦЕНА ИЛЛЮЗИИ.

«Те, кто не помнит прошлого, обречены его повторять». — Джордж Сантаяна.

Иллюзия «чистого листа» не только ошибочна с философской точки зрения, но и обходится дорого на практике. В этой главе рассматривается, что происходит, когда учреждения, организации и отдельные лица поддаются фантазии о начале с нуля. Закономерность поразительно последовательна: чем радикальнее попытка «перезагрузки», тем сильнее вновь заявляют о себе унаследованные ограничения.

Шоковая терапия и ее недостатки.

После распада Советского Союза западные экономисты столкнулись с беспрецедентной проблемой: как преобразовать централизованно планируемую экономику в рыночную систему. Основной рецепт, продвигаемый такими институтами, как Международный валютный фонд и Всемирный банк, заключался в «шоковой терапии» — быстрой и всеобъемлющей либерализации цен,

приватизации государственных предприятий и устранении торговых барьеров.

Теория, лежащая в основе шоковой терапии, по сути, представляла собой мышление «с чистого листа», примененное к экономике. Рынки понимались как естественные, самоорганизующиеся системы, которые возникнут спонтанно после устранения искажений централизованного планирования. Старые структуры рассматривались как чистые препятствия; их быстрое устранение позволило бы естественному рыночному порядку процветать.

В России последствия оказались катастрофическими практически по всем показателям. ВВП упал примерно вдвое в девяностые годы. Продолжительность жизни снизилась. Уровень смертности резко вырос, особенно среди мужчин трудоспособного возраста. Небольшая группа олигархов скопила огромные состояния, в то время как уровень жизни большей части населения резко упал. Преступность и коррупция достигли критической точки.

Что пошло не так? Психотерапевты, применявшие шоковую терапию, рассматривали как свободно изменяемые факторы, которые на самом деле были предопределены в краткосрочной и среднесрочной перспективе: правовую инфраструктуру прав собственности, культурные нормы, обеспечивающие исполнение контрактов, финансовые системы, распределяющие капитал, человеческий капитал для предпринимательства и управления, социальное доверие, лежащее в основе рыночных транзакций.

Рынки не функционируют в вакууме. Для их функционирования необходима сложная институциональная структура, на развитие которой на Западе ушли столетия: надежные суды, честные регулирующие органы, функционирующие банки, образованные работники, традиции коммерческой честности. Устранение видимых структур централизованного планирования не привело к появлению рынков; оно создало вакуум, который заполнили те, кто обладал властью завладеть активами.

Сравните опыт России с постепенным подходом Китая. Китай проводил реформы поэтапно, сохраняя государственный контроль над ключевыми секторами и позволяя рыночной активности развиваться параллельно с существующими структурами, а не заменять их. Результаты — независимо от того, как оценивать политическую систему Китая — оказались значительно лучше в экономическом плане: устойчивый рост, повышение уровня жизни и отсутствие обвала продолжительности жизни.

Китайский подход работал с унаследованными ограничениями, а не против них. Он признавал, что рыночные институты нельзя создать указом, что постепенная адаптация позволяет структурам развиваться, и что путь отсюда туда должен начинаться отсюда.

Ловушка переписывания программного обеспечения.

Любому, кто работает в сфере технологий, знакома

модель «масштабной переработки»: решение отказаться от существующей программной системы и создать ее с нуля.

Привлекательность очевидна. В устаревших системах накапливается технический долг — упрощения, принятые под давлением сроков, нагромождение исправлений, устаревшие проектные решения, застывшие в коде. Архитектура может отражать требования многолетней давности, которые больше не актуальны. Документация может быть неполной или вводящей в заблуждение. Ключевые знания могли быть утрачены вместе с уволенными сотрудниками.

В этой неразберихе соблазн начать все заново практически непреодолим. Чистая архитектура. Современные инструменты. Передовые методы. Отсутствие устаревшего багажа. Это звучит намного лучше, чем мучительная работа по распутыванию и постепенному улучшению существующего.

Однако переписывание программного обеспечения терпит неудачу с поразительной регулярностью.

Самая известная поучительная история — это решение Netscape переписать браузер Navigator с нуля в 1998 году. То, что должно было стать быстрой модернизацией, превратилось в трехлетнюю одиссею. К моменту выхода Netscape 6, Internet Explorer захватил рынок браузеров. Компания так и не оправилась от этого удара.

Почему переписывание кода терпит неудачу? Потому что старый код, каким бы запутанным он ни был, содержит в себе накопленные знания, которые нигде больше не задокументированы. Каждое исправление ошибки — это извлеченный урок. Каждое необычное обходное решение устраняет какой-то частный случай, который снова возникнет в новой системе. Каждое странное решение когда-то кому-то казалось логичным, часто по причинам, которые уже и не помнят.

Джоэл Спольски, предприниматель в сфере программного обеспечения и писатель, назвал устаревший код «рабочим капиталом знаний». Этот беспорядок — не случайный шум, а результат с

трудом приобретенного опыта. Выбрасывать его в погоне за чистотой и первозданной чистотой означает снова платить за уже усвоенные уроки.

Альтернативный подход — постепенное улучшение, поэтапная рефакторизация, непрерывная эволюция — менее приятен с эмоциональной точки зрения, но более вероятен для успеха. Он уважает наследие, заложенное в существующем коде, и позволяет неуклонно продвигаться к улучшению архитектуры.

Реформа образования и когнитивные ограничения.

Каждые несколько десятилетий в системах образования происходят волны реформ, обещающие революционизировать методы обучения детей. Конкретные рекомендации различаются — фонетический подход против целостного подхода к обучению, традиционный подход против прогрессивного, лекционный подход против активного обучения, — но многие реформаторские движения объединяет общая нить: предположение, что результаты обучения зависят прежде всего от

методов, и что правильные методы могут преодолеть любые ограничения.

Рассмотрим часто встречающуюся фантазию об отказе от механического запоминания в пользу «критического мышления». Цель звучит привлекательно — кому нужны студенты, способные лишь воспроизводить факты, ничего не понимая? Но реформа часто терпит неудачу из-за когнитивной реальности: критическое мышление в определенной области требует знания информации об этой области. Исследования в области экспертных знаний неизменно показывают, что то, что выглядит как блестящая интуиция у экспертов, на самом деле является распознаванием закономерностей, построенным на основе тысяч сохраненных случаев.

Шахматные мастера не просчитывают ходы, исходя из фундаментальных принципов; они распознают закономерности в ранее встречавшихся позициях. Врачи-диагносты не рассуждают абстрактно; они сопоставляют симптомы с заученными шаблонами заболеваний. Писатели не строят предложения слово

за словом; они используют фрагменты языка, усвоенные в результате обширного чтения.

Попытка обучить критическому мышлению без формирования базы знаний, на которой оно основано, смешивает результаты обучения с его результатами. Она рассматривает как свободно проектируемые то, что на самом деле ограничено особенностями работы человеческого познания: мы мыслим, опираясь на то, что знаем, а знание требует запоминания.

Аналогично, реформы, предполагающие, что дети с самого начала могут самостоятельно управлять своим обучением, игнорируют ограничения, связанные с развитием. Самостоятельное обучение — это навык, который необходимо освоить, и он требует наличия базовых знаний и метакогнитивных способностей, которые у маленьких детей еще не развиты. То, что работает для аспирантов, не обязательно работает для первоклассников .

Это не аргумент против реформы образования — это аргумент в пользу реформы, которая учитывает когнитивные ограничения, а не противодействует им.

Методы, которые неизменно приводят к результатам обучения, как правило, учитывают то, как на самом деле развиваются человеческая память, внимание и навыки .

Революции, которые воспроизводят то, что отвергают.

Французский историк Алексис де Токвиль сделал глубокое наблюдение о Французской революции: возникшее революционное правительство оказалось более централизованным, более бюрократическим и во многих отношениях более абсолютистским, чем монархия, которую оно заменило. Попытка разрушить старый порядок породила новый порядок, паразитально похожий по структуре.

Эта модель неоднократно встречается в революционной истории. Русская революция стремилась к ликвидации классовой иерархии и созданию общества равных; она породила одну из самых жестких иерархий в истории, где элита Коммунистической партии пользовалась привилегиями, от которых аристократы бы

покраснели. Иранская революция стремилась свергнуть авторитаризм западного образца; она установила теократический авторитаризм, по меньшей мере столь же репрессивный.

Почему революции так часто воспроизводят то, что, как они утверждают, отвергают? Одна из причин заключается в том, что они принимают поверхностные проявления власти за саму власть. Свергните короля, и вам все равно понадобится кто-то, кто будет принимать решения, обеспечивать соблюдение правил, собирать налоги, организовывать оборону. Эти функции сохраняются, даже когда меняются имена.

Но есть и более глубокая причина, связанная с нашей темой. Революционеры — это продукты обществ, которым они противостоят. Они несут в себе структуры, предположения и модели поведения, которые сознательно отвергают. Большевики выросли в царской России; их представления о том, как работают организации, как действует власть, как управлять обширной территорией, сформировались в

этой системе. Строя свой новый порядок, они использовали имеющиеся в их распоряжении материалы.

Вы не сможете преодолеть унаследованные ментальные структуры, просто приняв такое решение. Категории, в рамках которых вы воспринимаете мир, закономерности, которые вы распознаёте, решения, которые приходят вам в голову, — всё это формируется вашим собственным опытом. Революционная воля не сможет преодолеть это ограничение, так же как она не может преодолеть законы физики.

Как показывает история, успешные социальные изменения чаще происходят благодаря реформам, чем революциям — благодаря работе в рамках унаследованных структур с целью их постепенного преобразования, а не путем попыток разрушить все и начать все заново.

Распознавание образов

Во всех этих случаях — экономическая шоковая

терапия, переписывание программного обеспечения, реформа образования, политическая революция — вырисовывается одна и та же картина. Попытка начать с нуля сталкивается с унаследованными ограничениями. Ограничения вновь дают о себе знать, часто разрушительным образом. Обещанный чистый лист оказывается уже исписанным неизгладимыми следами.

Что делает эту закономерность такой устойчивой? Почему умные люди постоянно совершают одну и ту же ошибку?

Частично ответ кроется в когнитивных процессах: мы систематически недооцениваем сложность существующих систем и переоцениваем свою способность проектировать лучшие с нуля. То, что уже существует, выдержало испытания, о которых мы можем даже не подозревать; то, что мы проектируем, выдержало только те испытания, которые мы сочли необходимым провести.

Отчасти это эмоциональный аспект: фантазия о чистом листе апеллирует к нашему желанию

контролировать ситуацию и разочарованию из-за несовершенства унаследованных условий. Начало новой жизни дает чувство уверенности в себе; работа в условиях ограничений кажется ограничивающей.

Отчасти это политический вопрос: радикальные перемены привлекают сторонников и ресурсы способами, недоступными для постепенных улучшений. «Мы всё изменим» — более убедительный предвыборный лозунг, чем «мы будем постепенно улучшать некоторые вещи, соблюдая при этом неизбежные ограничения».

заявляет о себе. Цена иллюзии чистого листа — неудачные преобразования, растроченные ресурсы и предотвратимые страдания.

На пути к реалистичным переменам.

Всё это не означает, что перемены невозможны. Это означает, что к переменам нужно подходить реалистично, с точным пониманием того, что можно, а что нельзя изменить.

Успешные преобразования — в экономике,

технологиях, образовании и политике — как правило, обладают рядом общих характеристик. Они работают в рамках унаследованных структур, а не против них. Они развиваются постепенно, а не революционно. Они сохраняют то, что работает, и модифицируют то, что не работает. Они поддерживают преемственность в процессе изменений, а не пытаются совершать резкие скачки.

Это не консерватизм в политическом смысле — это реализм в отношении того, как меняются сложные системы. Понимание разницы между тем, что необходимо принять, и тем, что можно изменить, является необходимым условием для эффективных действий.

В последующих главах будут разработаны инструменты для систематического проведения этого различия.

ГЛАВА ТРЕТЬЯ: ПАРАЛИЧ ЛОЖНОЙ НЕПОДВИЖНОСТИ

«Независимо от того, считаете ли вы, что можете, или считаете, что не можете — в любом случае вы правы». — Генри Форд.

Если первая ошибка заключается в вере в то, что мы можем изменить всё, то вторая — в вере в то, что мы не можем изменить ничего. В этой главе рассматривается противоположная ловушка: отношение к тому, что можно изменить, как к неизменному.

Обе ошибки имеют одну и ту же первопричину: неспособность точно отличить предопределенное от свободного. Но если мышление «с чистого листа» ведет к чрезмерным амбициям и разочарованию, то ложная неизменность ведет к смирению и упущенным возможностям. Цена измеряется не драматическими неудачами, а тихим принятием неприемлемого, неиспользованными возможностями, жизнями, ограниченными воображаемыми стенами.

Привлекательность детерминизма.

Детерминистическое мышление имеет свои соблазны. Если результаты предопределены заранее — генами, историей, экономическими силами, судьбой, — то мы не несем за них ответственности. Нас нельзя ни обвинять в неудачах, ни ожидать от нас успехов. Бремя ответственности снимается; мы становимся пассажирами, а не водителями собственной жизни.

Это может быть очень утешительно. Человек, который считает свой интеллект неизменным, не должен сталкиваться с возможностью того, что более усердные усилия могут привести к лучшим результатам. Общество, которое считает, что его институты отражают законы природы, не должно испытывать дискомфорт от целенаправленных изменений. Организация, которая считает свою культуру неизменной, не должна выполнять сложную работу по трансформации.

Ложная неизменность также обеспечивает интеллектуальную целостность. Детерминистические мировоззрения могут быть элегантными,

всеобъемлющими и удовлетворяющими стремящийся к закономерностям ум. Всему есть свое место; ничто не произвольно; видимость выбора растворяется в осознании необходимости.

Но комфорт и согласованность имеют свою цену: паралич перед лицом того, что действительно можно изменить.

Генетический детерминизм и его пределы.

В двадцатом веке произошли резкие изменения в нашем понимании роли генов в развитии человека. Энтузиазм начала века по отношению к евгенике — вере в то, что человеческие популяции можно улучшить путем селекции, — уступил место послевоенной реакции, которая минимизировала влияние генетики в пользу объяснений, связанных с окружающей средой.

С тех пор маятник качнулся в сторону признания существенного генетического вклада в индивидуальные различия в интеллекте, личности и здоровье. Поведенческая генетика установила, что

большинство психологических черт в значительной степени наследуются. Полногеномные ассоциативные исследования начали выявлять конкретные генетические варианты, которые способствуют сложным результатам.

Однако признание генетического влияния не равнозначно генетическому детерминизму. Ошибка детерминизма заключается в переходе от «гены влияют на результаты» к «гены определяют результаты», рассматривая вероятность как неоспоримую истину, корреляцию как причинно-следственную связь, а вклад как совокупность факторов.

Рассмотрим интеллект. Исследования близнецов показывают, что примерно половина вариаций в измеренном интеллекте отражает генетические различия. Это существенный вклад, но он означает, что примерно половина вариаций отражает факторы окружающей среды и другие факторы. Гены обеспечивают широкий диапазон потенциальных результатов; то, в какой части этого диапазона

окажется человек, зависит от опыта, образования, питания и обстоятельств.

Более того, наследуемость — это не то же самое, что неизменность. Рост в значительной степени наследуется, однако средний рост за последнее столетие значительно увеличился благодаря улучшению питания. Генетический вклад в рост не изменился; изменились условия окружающей среды, позволяющие проявиться генетическому потенциалу.

Та же логика применима и к когнитивным способностям. Интенсивные программы раннего вмешательства показали значительное влияние на результаты обучения детей из неблагополучных семей. Системы образования, которые эффективно обучают, дают иные результаты, чем те, которые этого не делают. Генетические ограничения реальны, но диапазон в пределах этих ограничений допускает значительные вариации.

Генетический детерминизм — отношение к генам как к предопределению судьбы — приводит к фатализму в отношении результатов, которые на самом деле

можно улучшить. Он ошибочно принимает вероятностное влияние за предопределенную судьбу, исключая возможность вмешательств, которые могли бы реально изменить ситуацию.

ТИНА: Альтернативы нет

В политическом и экономическом дискурсе особенно коварная форма ложной неизменности обозначается аббревиатурой TINA: There Is No Alternative (Нет альтернативы). Эта фраза ассоциируется с защитой Маргарет Тэтчер политики свободного рынка, но описываемая ею модель встречается во всем политическом спектре.

Аргументы в пользу принципа «нет альтернативы» (TINA) принимают следующую форму: существующее положение дел, каким бы оно ни было, отражает необходимость, а не выбор. Результаты рыночной конъюнктуры являются следствием безличных сил, неподвластных чьему-либо контролю. Институциональные структуры — единственные, которые могут работать. Политика, которая кажется необязательной, на самом деле

необходима из-за ограничений, не оставляющих места для альтернатив.

Иногда эти аргументы верны. Существуют реальные ограничения. Не каждый политический вариант осуществим. Некоторые альтернативы действительно хуже, чем существующие.

Однако аргументы в пользу того, что ничего нельзя изменить, часто используются стратегически, чтобы исключить обсуждение действительно существующих альтернатив. Утверждение о том, что ничего нельзя изменить, служит интересам тех, кто извлекает выгоду из нынешнего положения дел. Представление конкретных политических решений как неизбежных, а не как выбранных, исключает их из демократического обсуждения.

История того, что когда-то считалось неизменным, весьма поучительна. Рабство защищалось как естественное и экономически необходимое явление. Исключение женщин из системы образования и профессиональной деятельности основывалось на утверждениях о биологических ограничениях.

Расовая сегрегация оправдывалась призывами к социальной стабильности. В каждом случае то, что казалось неизменным, оказывалось изменчивым, как только достаточное количество людей решало это изменить.

Это не означает, что всё можно изменить или что все предлагаемые изменения разумны. Это означает, что мы должны скептически относиться к аргументам «нет альтернативы» (TINA), используемым для того, чтобы исключить рассмотрение альтернатив. Вопрос «Действительно ли нет альтернативы?» стоит задавать серьёзно, а не отвергать рефлексивно.

Фиксированное мышление в личной жизни.

Исследования психолога Кэрол Двек, посвященные образу мышления, показывают, как ложная неизменность проявляется на индивидуальном уровне. Люди с «фиксированным мышлением» считают, что их основные качества — интеллект, талант, личность — являются статичными чертами, которые невозможно развить. Люди с «гибким мышлением» считают, что эти качества можно

развить посредством усилий, стратегии и обучения.

Эти убеждения становятся самосбывающимися пророчествами. Люди с фиксированным мышлением избегают трудностей, которые могут выявить их недостатки. Они легко сдаются, сталкиваясь с препятствиями, поскольку усилия кажутся бессмысленными, если их способности неизменны. Они чувствуют угрозу со стороны успеха других и игнорируют полезную обратную связь. Все это ограничивает их рост и достижения.

Люди с установкой на рост, напротив, воспринимают трудности как возможности для обучения. Они упорствуют, преодолевая неудачи, рассматривая их как часть процесса обучения. Они учатся на критике и черпают вдохновение в успехах других. Такая ориентация приводит к большим достижениям с течением времени.

Фиксированное мышление — это, по сути, личностная версия ложной неизменности: отношение к тому, что на самом деле поддается развитию, как к предопределенному. Человек, который считает, что

«не умеет считать», будет избегать математических задач и, следовательно, никогда не разовьет математические способности, подтверждая свое убеждение. Человек, который считает, что «не общительный человек», будет избегать социальных ситуаций и никогда не разовьет социальные навыки.

Исследования Двек показывают, что эти установки можно изменить — что обучение людей изменчивости способностей может склонить их к ориентации на рост. Но для этого сначала необходимо признать фиксированное мышление тем, чем оно является на самом деле: убеждением, а не фактом, о том, что может и что не может измениться.

Культурный фатализм.

Культуры различаются по степени акцента на судьбе против свободы воли, на принятии против стремления. Эти различия влияют на всё, от экономического развития до благополучия отдельных людей.

Антропологи задокументировали общества, где

результаты определяются судьбой, кастой, кармой или сверхъестественными силами. В таких контекстах попытка изменить свое положение в жизни может рассматриваться не только как бесполезная, но и как нечестие — отказ от космического порядка. Предопределенное множество P расширяется, охватывая практически все; свободное множество F сужается почти до пустоты .

Эти мировоззрения имеют свои утешения. Принятие судьбы может уменьшить тревогу, возникающую от постоянных стремлений. Довольство своей участью может принести мир, которого не может достичь бесконечная амбиция. Есть смысл не сравнивать себя постоянно с альтернативами.

Но культурный фатализм имеет и свою цену. Общества, которые считают, что результаты предопределены, как правило, недостаточно инвестируют в образование, здравоохранение и инфраструктуру — инвестиции, которые окупаются только в том случае, если будущее рассматривается как изменяемое. Люди, принимающие такую

позицию, не ищут возможностей, которые могли бы улучшить их жизнь и жизнь их детей.

Взаимодействие между культурными убеждениями и экономическим развитием сложное, но общая закономерность ясна: общества, которые верят, что будущее может отличаться от настоящего, как правило, создают будущее, которое на самом деле отличается.

Паралич технологического детерминизма.

В дискуссиях о технологиях проявляется изощренная форма ложной неизменности: предположение, что технологическое развитие следует своей собственной неизбежной логике, неподвластной человеческому руководству или контролю.

Технологический детерминизм принимает различные формы. В своей наиболее радикальной версии технология рассматривается как автономная сила, формирующая общество в соответствии со своими собственными императивами. Более умеренные версии признают человеческий выбор, но

рассматривают его как настолько ограниченный техническими требованиями, что его можно считать детерминированным.

Рассмотрим дискуссии об искусственном интеллекте. Распространенная точка зрения представляет развитие ИИ как стихийную силу: это неизбежно, это нельзя остановить, мы можем лишь адаптироваться к его последствиям. Такая трактовка исключает из повестки дня вопросы о том, произойдет ли это, какого рода это произойдет и для каких целей. Технология рассматривается как нечто само собой разумеющееся; обсуждению подлежит только ответная реакция.

Но технологии — это творения человека, формирующиеся под влиянием человеческого выбора на каждом этапе. Решение о выборе направления исследований, проектные решения, определяющие возможности, решения о внедрении, определяющие области применения, нормативно-правовые рамки, регулирующие использование — всё это предполагает участие человека. Технология не

развивается сама по себе; она создаётся людьми и организациями, принимающими решения.

Технологический детерминизм исключает эту свободу выбора. Он представляет как неизбежное то, что на самом деле выбирается, маскируя интересы и ценности, заложенные в технологическом развитии. «Технология этого потребовала» становится оправданием для решений, которые на самом деле были приняты людьми, которые могли бы поступить иначе.

Это не означает, что технологиями можно управлять произвольно. Технические ограничения реальны. Зависимости от предшествующего опыта накапливаются. Некоторые разработки может быть трудно обратить вспять. Но признание ограничений отличается от отношения ко всему как к ограниченному. Пространство для выбора существует; вопрос в том, воспользуемся ли мы им.

Выученная беспомощность.

Исследования психолога Мартина Селигмана о

выученной беспомощности раскрывают психологический механизм, лежащий в основе ложной фиксации. В его классических экспериментах собаки, подвергавшиеся неизбежным ударам током, впоследствии не могли убежать, когда такая возможность появлялась. Они «научились», что их действия не имеют эффекта, и это обучение сохранялось даже при изменении обстоятельств.

Люди демонстрируют ту же закономерность. Люди, оказавшиеся в ситуациях, где результаты кажутся не связанными с их действиями, часто испытывают общее чувство беспомощности, которое распространяется за пределы исходной ситуации. Убеждение в том, что ничего нельзя изменить, становится самоподтверждающимся пророчеством: те, кто считает, что не могут повлиять на результаты, перестают пытаться, и их пассивность подтверждает это убеждение.

Приобретенная беспомощность особенно пагубна, потому что она носит обобщающий характер. Ребенок, который постоянно терпит неудачи в

математике и приходит к выводу: «Я не могу заниматься математикой», может распространить это на «Я не могу учиться в школе» или даже «Я не могу добиться успеха». Взрослый, чьи инициативы постоянно блокируются, может прийти к выводу, что эти инициативы бессмысленны.

Лекарством от выученной беспомощности является опыт контроля — ситуации, в которых действия действительно приводят к результатам. Но поиск таких ситуаций требует прежде всего переосмысления убеждения, что ничего нельзя изменить. Разрыв порочного круга беспомощности требует признания беспомощности как выученной, а не приобретенной.

Симметрия ошибок.

мы можем увидеть симметрию между двумя ошибками, рассмотренными в первых главах. Мышление «чистого листа» переоценивает свободное множество (F); ложная фиксированность переоценивает предопределенное множество (P). Обе ошибки искажают реальность таким образом, что

приводят к принятию неверных решений.

Самоуверенный предприниматель, игнорирующий рыночные ограничения, и смирившийся с ситуацией сотрудник, игнорирующий возможности роста, совершают одну и ту же фундаментальную ошибку, но с противоположных сторон. Революционер, пытающийся переделать человеческую природу, и консерватор, утверждающий, что человеческая природа не допускает изменений, оба ошибаются в определении реальных границ .

Нам необходима точная оценка: ни завышенное представление о том, что может измениться, ни безразличное смирение с тем, что измениться невозможно. Именно это и предлагает разработанная в следующих главах концепция.

ГЛАВА ЧЕТВЁРТАЯ: АНАТОМИЯ ЛЮБОЙ ОТПРАВНОЙ ТОЧКИ.

Теперь мы готовы представить центральную концепцию этой книги. Она элегантно проста, но при этом освещает проблемы, которые ставили в тупик практиков и теоретиков во всех областях человеческой деятельности.

Данная концепция исходит из одного простого вывода: любое начальное условие — любая отправная точка любого процесса изменений — может быть разложено на две составляющие.

Базовая формула

Пусть **I** обозначает любое начальное условие. Тогда:

$$\mathbf{I} = \mathbf{P} + \mathbf{F}$$

Где:

P = Заранее определенные элементы: все, что нельзя изменить в соответствующем контексте, временных рамках и области применения.

F = Свободные элементы: всё, что можно изменить, выбрать или варьировать.

Это разложение является полным. В начальном условии нет ничего, что не принадлежало бы ни к Р, ни к F. И оно является исчерпывающим. Каждый элемент должен принадлежать либо к одному, либо к другому типу (или, как мы увидим, для определения этого может потребоваться исследование).

Ни Р, ни F никогда не равны нулю.

Это ключевое понимание, отличающее реалистическое мышление от фантазии, основанной на принципе «чистого листа», и от детерминистического фатализма.

Р никогда не равно нулю. Вы всегда что-то наследуете. Нет отправной точки без ограничений — биологических, физических, логических, исторических, институциональных. Главы, которые вы только что прочитали, показали, почему: сам акт описания процесса требует наличия предварительной структуры.

F никогда не равно нулю. У вас всегда есть определённые степени свободы. Даже в самых

стеснённых ситуациях выбор остаётся. Заключённые выбирают, как реагировать на плен. Смертельно больные выбирают, как встретить смерть. Обстоятельства ограничивают, но никогда не определяют полностью.

Осознание того, что и Р, и F всегда непусты, приносит освобождение. Оно освобождает нас от фантазии о неограниченном выборе и от отчаяния полной предопределенности. Мы всегда наследуем; мы всегда выбираем. Вопрос в том, что к какой категории относится.

Зависимость от контекста

Разделение между Р и F не является абсолютным; оно зависит от контекста. То, что предопределено в одном контексте, может быть свободным в другом. То, что свободно в один момент времени, может стать предопределенным позже.

Рассмотрим рост человека. Для взрослого человека рост является частью потенциальной составляющей (Р) — его нельзя изменить с помощью диеты,

упражнений или желания. Но на рост ребенка влияют питание, здоровье и другие факторы, которые можно изменить. Рост, который был частично потенциальной составляющей (F), со временем становится исключительно потенциальной (P).

Или рассмотрим юридическую структуру организации. После регистрации изменение формы с корпорации на партнерство требует значительных усилий и имеет серьезные последствия. В повседневной работе структура фактически представляет собой вариант P. Но при крупной реорганизации она может превратиться в вариант F — то, что активно обсуждается и подлежит пересмотру.

Важный контекст включает в себя:

Временные рамки: То, что нельзя изменить на этой неделе, может измениться в течение нескольких лет. Стратегическое планирование требует более долгосрочной оценки финансовых и моральных аспектов, чем оперативное управление.

Ресурсы: Деньги, время, полномочия и экспертные

знания — все это влияет на то, что осуществимо. У богатого человека показатель F больше, чем у бедного, не потому, что ограничения различаются, а потому, что ресурсы расширяют круг задач, которые можно решить.

Область применения: Физические ограничения отличаются от правовых ограничений, которые, в свою очередь, отличаются от социальных ограничений. То, что физика делает невозможным, определяется как P иным образом, чем то, что запрещает закон.

Уровень описания: На уровне отдельных нейронов вся активность мозга определяется физиологическими процессами. На уровне психологического опыта выбор кажется свободным. Одно и то же явление выглядит по-разному на разных уровнях.

Применение: Запуск нового продукта

Давайте применим эту концепцию к конкретной бизнес-ситуации: компании, готовящейся к запуску

нового продукта.

Что относится к категории Р? Во-первых, существующая репутация бренда. Какие бы ассоциации ни были у клиентов с названием компании, они возникают независимо от вашего желания. Конкурентная среда — кто еще есть на рынке, что они предлагают, как они позиционируют себя — также в значительной степени относится к категории Р, по крайней мере, в краткосрочной перспективе. Нормативно-правовые требования относятся к категории Р: стандарты безопасности, правила раскрытия информации, ограничения на импорт. Текущее состояние технологий относится к категории Р: вы не можете внедрять возможности, которых еще не существует. Экономические условия относятся к категории Р: рецессия или экономический подъем влияют на покупательское поведение независимо от достоинств вашего продукта.

Что относится к категории F? Характеристики продукта, которые еще не определены. Ценовая стратегия. Маркетинговый подход. Каналы сбыта.

Время запуска. Целевые сегменты клиентов. Это те степени свободы, где решения имеют решающее значение.

Разделение на «П» и «Ф» определяет распределение ресурсов. Попытки изменить «П» приводят к растрате ресурсов: нельзя просто игнорировать историю бренда или действия конкурентов. Пренебрежение «Ф» растрчивает возможности: оставшиеся варианты — это то, где строится стратегия.

Приложение: Новый генеральный директор вступает в должность.

Когда в организацию приходит новый генеральный директор, немедленно применяется модель Р/Ф (принцип «плата/неудача»).

Что такое Р? История и культура компании, по крайней мере, в их нынешнем виде. Возможности и ограничения существующего персонала. Договорные обязательства. Материальные активы и пассивы. Позиция на рынке, как её воспринимают клиенты . Незавершенные судебные разбирательства.

Соглашения с профсоюзами. Новый генеральный директор наследует всё это; ничего из этого нельзя просто так отбросить.

Что такое F? Стратегическое направление. Организационная структура (в рамках ограничений). Кто какие должности занимает. Что является приоритетным. Как распределяются ресурсы. Какие сообщения передаются. Как измеряется эффективность. У нового генерального директора есть реальный выбор в этих областях.

Самые мудрые новые лидеры вкладывают значительные средства в понимание P, прежде чем действовать в отношении F. Они слушают, наблюдают и учатся, прежде чем объявлять о планах трансформации. Они определяют, какие аспекты унаследованной ситуации действительно неизменны, а какие лишь кажутся таковыми из-за привычки или недостатка воображения.

Применение: Начальная терапия

Терапевтический контекст позволяет получить более

глубокое представление о структуре Р/Ф.

Клиент приходит на терапию, неся с собой всю свою историю: детские переживания, модели привязанности, усвоенные модели поведения, привычные мыслительные процессы, неврологические тенденции, культурную обусловленность. Все это является Р в том смысле, что это нельзя изменить. Прошлое нельзя забыть; порожденные им модели поведения нельзя просто стереть.

Но терапия не сработала бы, если бы всё было П. Предпосылка терапевтического вмешательства заключается в том, что что-то может измениться: не само прошлое, а его продолжающееся влияние. Модели поведения можно распознать и скорректировать. Можно научиться новым реакциям. Смысл переживаний может измениться. Отношения могут развиваться по-другому.

Эффективная терапия предполагает точную диагностику того, что относится к Р, а что к Ф. Клиенту, который считает свою тревожность

неизменным аспектом своей личности, возможно, потребуется рассматривать её как модель поведения, с которой можно работать (частично F). Клиенту, который считает, что может просто решить испытывать другие чувства, возможно, потребуется принять определённые эмоциональные реалии (частично P).

Применение: Обучение нейронной сети

Искусственный интеллект является наглядным примером применения подхода «процент/недостоверность», поскольку его компоненты разработаны с учетом конкретных требований.

Что представляет собой параметр P для нейронной сети в начале обучения? Архитектура: количество слоев, тип связей, функции активации. Алгоритм обучения: как обновляются веса на основе обратной связи. Функция потерь: какую целевую функцию оптимизирует процесс обучения. Эти параметры выбираются разработчиками, но после выбора они становятся фиксированной структурой, в рамках

которой происходит обучение.

Что такое F? Начальные веса (обычно устанавливаются случайным образом в рамках заданных ограничений). Выбор и порядок обучающих данных. Гиперпараметры, которые настраивают процесс обучения: скорость обучения, размер пакета, сила регуляризации. В рамках архитектуры это степени свободы.

Взаимосвязь между P и F объясняет, почему разработка ИИ включает в себя так много экспериментов. Архитектура (P) определяет, чему потенциально может научиться сеть, но для достижения этого потенциала необходимо найти правильные настройки для F. Одна и та же архитектура может добиться блестящих результатов или потерпеть полную неудачу в зависимости от того, как обрабатывается F.

Визуализация P и F

Визуализация структуры может помочь. Представьте любое начальное состояние как коробку. Внутри

коробки некоторые области затенены (P) , а некоторые открыты (F). Затененные области недоступны для изменения; открытые области доступны.

Ошибка мышления «с чистого листа» заключается в том, чтобы представлять коробку полностью открытой. Ошибка ложной неподвижности заключается в том, чтобы представлять ее полностью затененной. Реальность всегда представляет собой смесь , и практический вопрос всегда один и тот же: каков же на самом деле характер затенения?

Разные контексты порождают разные модели поведения. У стартапа больше свободного пространства, чем у устоявшейся бюрократической структуры. У человека в возрасте двадцати лет больше свободного пространства, чем у того же человека в шестьдесят. У диктатуры меньше свободного пространства, чем у демократии (для большинства граждан). Понимание вашей конкретной модели поведения , основанной на факторах P и F, имеет важное значение для

эффективных действий.

ГЛАВА 5: ТРИ ТИПА ОШИБОК

Структура $P + F$ дает нам карту. Теперь нам нужно понять, как люди теряются. В своих исследованиях я выявил три основных типа ошибок в рассуждениях о том, что можно и что нельзя изменить. Умение распознавать эти ошибки — как в мышлении других, так и в своем собственном — является ключевым практическим навыком, которому учит эта книга.

Ошибка типа I: Ложная нейтральность ($P \rightarrow F$)

Эта ошибка возникает, когда элементы, которые действительно предопределены, рассматриваются так, как если бы они были свободно изменяемыми. Это ошибка, заключающаяся в убеждении, что можно изменить то, что изменить нельзя.

Формально структура такова: берётся нечто, принадлежащее P , и действует так, как если бы оно принадлежало F .

Это та самая ошибка, когда начинаешь с чистого листа, которую мы рассматривали в главах 1 и 2. Она приводит к напрасным усилиям, сорванным

инициативам и неоднократным столкновениям с препятствиями, которые всегда существовали, но оставались незамеченными.

Примеры ошибок первого типа

- **Попытка создать непредвзятый ИИ. Архитектуры** нейронных сетей неизбежно содержат индуктивные искажения; обучающие данные неизбежно отражают мир, в котором они были созданы. Игнорирование этих ограничений не приводит к их исчезновению.
- **Разработка образовательных программ, игнорирующих когнитивные ограничения.** Рабочая память человека имеет ограниченный объем; внимание ограничено; обучение проходит через этапы развития. Учебные программы, предполагающие бесконечную пластичность, терпят неудачу при взаимодействии с реальными учениками.
- **Организационные преобразования, игнорирующие институциональную**

инерцию. Культура, рутина и неформальные сети сопротивляются быстрым изменениям. Объявление о новом видении не приводит к исчезновению старых моделей поведения.

- **В личных отношениях один партнер ожидает от другого кардинальных изменений.** Глубоко укоренившиеся черты характера, стили привязанности и ценности меняются медленно, если вообще меняются. Надежда на то, что брак преобразит человека, обычно не оправдывается.
- **Революционная политика, рассматривающая человеческую природу как бесконечно податливую.** Каждая попытка создать «нового человека» посредством социальной инженерии терпела неудачу, зачастую катастрофическую.

Как распознать ошибку первого типа

Обращайте внимание на формулировки, которые игнорируют ограничения: «Нам просто нужно...»,

«Если бы только люди...», «Как только мы устраним...» Эти фразы часто указывают на то, что реальные ограничения игнорируются.

Спросите, что должно быть истинным, чтобы план сработал. Если ответ включает в себя изменение того, что никогда не удавалось успешно изменить, или изменение этого быстрее, чем это когда-либо удавалось, возможно, присутствует ошибка первого типа.

Обратите внимание на исторические прецеденты. Удалось ли кому-нибудь это сделать раньше? Если многие пытались, и все потерпели неудачу, то неудача может быть следствием ошибки, а не плохого исполнения.

Ошибка типа II: Ложная фиксированность ($F \rightarrow P$)

Эта ошибка является зеркальным отражением первой. Она возникает, когда элементы, которые действительно свободны, рассматриваются так, как если бы они были предопределены. Это ошибка, заключающаяся в убеждении, что вы не можете

изменить то, что на самом деле можно изменить.

Формально: взять нечто, принадлежащее F, и действовать так, как если бы оно принадлежало P.

Это та самая детерминированная ошибка, которую мы рассматривали в главе 3. Она приводит к упущенным возможностям, неоправданной покорности и принятию условий, которые можно было бы улучшить.

Примеры ошибок второго типа

- **Генетический фатализм в отношении интеллекта или личности.** Хотя гены и влияют на эти черты, окружающая среда и прилагаемые усилия также имеют существенное значение. Рассмотрение когнитивных способностей как неизменных означает игнорирование реального потенциала для развития.
- **Принцип «нет альтернативы» (TINA) в политических дебатах.** Представление конкретных институциональных механизмов

как неизбежных исключает рассмотрение альтернатив, которые могли бы быть жизнеспособными.

- **Технологический детерминизм.** Рассмотрение технологического развития как следования автономной логике исключает возможность выбора для человека в вопросах о том, следует ли и как развивать те или иные возможности .
- **Фиксированное мышление в личностном развитии.** Вера в статичность способностей препятствует усилиям и практике, которые могли бы их развить.
- **Организационная выученная беспомощность.** Принцип «Так здесь всё устроено» подразумевает принятие дисфункции, которую потенциально можно было бы исправить.

Как распознать ошибку второго типа

Обращайте внимание на формулировки,

утверждающие невозможность без предварительного анализа: «Вы не можете...», «Так уж оно есть», «Человеческая природа», «Нет никакого способа...». Эти фразы иногда указывают на реальные ограничения, но часто — на непроверенные предположения.

Спросите, кому выгодно заявленная неизменность. Если от изменений пострадают влиятельные круги, то скептицизм в отношении утверждений о невозможности оправдан.

Ищите вариативность. Если результаты различаются в зависимости от контекста, условий или прилагаемых усилий, то рассматриваемый вопрос не является чисто детерминированным.

Ошибка типа III: Путаница доменов

Третья ошибка более незаметна, но зачастую более разрушительна. Она возникает, когда раздел P/F из одного домена некорректно применяется к другому домену.

Формально: применение (P + F) из Домена 1 к Домену

2, где для доменов требуются разные разбиения.

Эта ошибка особенно коварна, потому что может показаться логически верной. Если что-то является P где-то, то, конечно же, оно является P везде? Но этот вывод неверен, потому что P и F всегда зависят от контекста .

Примеры ошибок третьего типа

- **Рассмотрение социальных категорий как биологически обусловленных** . Гендерные роли, сконструированные культурно, защищают как биологически предопределенные. Экономическое неравенство, отражающее политические решения, воспринимается как неизбежный результат различий в способностях.
- **Рассмотрение биологических данных как социальных предписаний**. Тот факт, что какой-либо признак имеет генетическую составляющую, не означает, что социальная политика должна рассматривать его как нечто

неизменное. Наследственность описывает то, что существует при определенных условиях , а не то, что должно быть.

- **Рассмотрение математических моделей как эмпирических требований.** Экономические модели делают упрощающие предположения, которые могут не выполняться в реальности . Рассмотрение результатов модели как обладающих достоверностью математических теорем приводит к путанице в различных областях.
- **Рассмотрение местных обычаев как универсальных ограничений.** Деловые практики, сформировавшиеся в конкретных условиях, применяются во всем мире, как если бы они были требованиями, а не выбором.

Как распознать ошибку III типа

Обратите внимание на «слайды категорий»: аргументы, переходящие с одного уровня описания на другой без указания источника. Утверждение

«Эволюция показывает, что X , следовательно, социальная политика должна быть Y » почти всегда является ошибкой в определении предметной области.

Спросите, на каком уровне действует заявленное ограничение. Это физическое ограничение? Биологическое? Психологическое? Социальное? Правовое? Математическое? Одно и то же слово («невозможно», «необходимо», «фиксированное») означает разные вещи на разных уровнях.

Диагностический контрольный список

При оценке любого предложения об изменениях или утверждения о невозможности изменений, воспользуйтесь этим контрольным списком:

1. **Что представляет собой заявленная цена P ?**
Что считается фиксированным или неизменным?
2. **Что представляет собой заявленное значение F ?** Что считается допускающим изменение?
3. **Что обосновывает данную классификацию?**

Какие доказательства или аргументы подтверждают отнесение каждого элемента к соответствующей категории?

4. **ли ошибка первого типа ?** Обработывается ли что-либо как ошибка типа F, что на самом деле является ошибкой типа P ?
5. **Присутствует ли ошибка второго типа?** Относится ли что-либо к категории P, что на самом деле является ошибкой F ?
6. **Присутствует ли ошибка типа III?** Применяется ли раздел P/F из одного домена к другому?
7. **Какой контекст имеет значение?** Временные рамки? Ресурсы? Уровень описания?
8. **Что изменит разделение?** Существуют ли условия, при которых текущий P станет F, или текущий F станет P?

ГЛАВА 6: ДВОЙНАЯ ПРЕДВАРИТЕЛЬНАЯ ПОДГОТОВКА

«Карта — это не территория». — Альфред
Коржибски

Мы установили, что каждое начальное условие содержит как предопределенные (P), так и свободные (F) элементы. Теперь мы должны признать дополнительную сложность: ограничения исходят из двух разных источников, и неспособность их различать порождает систематическую путаницу.

Два источника ограничений

При анализе любой системы мы сталкиваемся с ограничениями с двух сторон:

Внутренние ограничения проистекают из самой системы — её физических свойств, структурных требований, присущих ей ограничений. Мост должен выдерживать вес; организму необходимо питание; обществу требуются механизмы координации. Эти ограничения существуют независимо от того, как мы выбираем описывать или моделировать систему.

Внешние ограничения возникают из-за используемых нами инструментов описания — языка, модели или структуры, которые мы применяем для представления системы. Математическая нотация имеет ограничения в том, что она может выразить. Инструменты измерения имеют пределы разрешения. Концептуальные структуры выделяют одни особенности, скрывая другие.

Оба типа ограничений реальны. Оба определяют то, что мы можем знать и делать. Но их путаница приводит к серьезным ошибкам.

Карта и территория

Знаменитый афоризм Коржибского прекрасно отражает это различие. Карта — это не территория, изображение — это не то, что изображается. Карта и территория обладают разными свойствами, разными ограничениями, разными возможностями.

Если вы путаете карту с территорией, вы совершаете характерные ошибки. Вы можете считать, что раз на вашей карте нет дороги, значит, дороги не

существует. Вы можете считать, что особенности территории ограничены тем, что изображено на вашей карте.

И наоборот, если игнорировать ограничения карты, можно допустить другие ошибки. Можно предположить, что карта охватывает все важное, упуская из виду то, что она не охватывает. Можно не заметить, что разные карты одной и той же территории показывают разные особенности.

Но вот в чем загвоздка: без карты ориентироваться невозможно. Сама территория не имеет указаний. Для эффективных действий необходимо представление о ней, а это значит, что нужно работать с ограничениями карты, даже понимая, что они отличаются от ограничений территории.

Двойная предварительная подготовка

Это приводит к тому, что я называю принципом двойного предварительного обусловливания: для любой системы, которая одновременно (1) описываема и (2) представима, ограничения

действуют на обоих уровнях одновременно.

Ограничения, накладываемые самой системой (внутренние), и ограничения, накладываемые нашими инструментами представления информации (внешние), пересекаются. То, что мы можем реально делать и знать, находится на этом пересечении — ограниченное как тем, что позволяет система, так и тем, что могут зафиксировать наши инструменты.

Проверка ограничений дважды

Практический вывод из двойного предварительного обусловливания — это методологическая необходимость: проверяйте ограничения дважды. Задавайте вопросы отдельно:

Что требуется от системы? Что должно быть истинным, исходя из свойств самого объекта, независимо от того, как мы его опишем?

Что требует наше описание? Какие ограничения накладывают наши модели, методы измерения, языки и концепции, которые могут не отражать саму суть?

Когда возникают очевидные ограничения,

необходимо определить их источник. Являются ли они системными ограничениями, которые будут препятствовать использованию любых инструментов, независимо от их возможностей? Или же это ограничения, связанные с инструментами, которые можно преодолеть с помощью других средств описания?

ГЛАВА 7: ДЛЯ МЕНЕДЖЕРОВ — СТРАТЕГИЯ В УНАСЛЕДОВАННЫХ СТРУКТУРАХ

Управление — это искусство достижения целей посредством организации. Организации по своей природе являются унаследованными структурами — накопленными слоями решений, взаимоотношений, возможностей и ограничений, существовавшими до прихода к власти нынешнего руководителя. В этой главе наша концепция применяется к конкретным проблемам, с которыми сталкиваются менеджеры, пытаясь руководить изменениями в условиях унаследованной структуры.

Аудит ограничений

Прежде чем запускать какую-либо масштабную инициативу по внедрению изменений, опытные руководители проводят **аудит ограничений** : систематическую инвентаризацию того, что можно и что нельзя изменить в соответствующем контексте.

Аудит начинается с определения предлагаемых изменений. Чего вы пытаетесь достичь? Сформулируйте это как можно конкретнее:

«Сократить время цикла разработки продукта на 40%», а не «Повысить эффективность».

Далее, проведите инвентаризацию ограничений:

- **Физические ограничения:** инфраструктура, оборудование, географическое распределение.
- **Человеческие ограничения:** навыки, знания, установки, количество доступных людей.
- **Финансовые ограничения:** бюджет, денежный поток, доступ к капиталу.
- **Правовые ограничения:** нормативные акты, контракты, интеллектуальная собственность.
- **Рыночные ограничения:** ожидания клиентов, поведение конкурентов, отношения с поставщиками.
- **Технологические ограничения:** существующие системы, требования к совместимости, технический долг.
- **Культурные ограничения:** нормы, ценности, «как здесь принято».

- **Политические ограничения:** интересы заинтересованных сторон, властные отношения, соглашения профсоюзов.

Для каждого ограничения классифицируйте его как Р или F в рамках вашего контекста. Помните, что контекст включает в себя временные рамки, ресурсы и уровень полномочий.

Почему трансформации удаются, а революции терпят неудачу

История организационных изменений преподносит наглядный урок: постепенные преобразования часто оказываются успешными там, где революционные перемены терпят неудачу.

Рассмотрим разработку компанией Toyota концепции бережливого производства. Производственная система Toyota не возникла с нуля. Она развивалась на протяжении десятилетий, постепенно наращивая существующие возможности и систематически решая конкретные проблемы.

Сравните это со многими неудачными попытками

внедрения бережливого производства путем повсеместного заимствования. Компании, которые пытались скопировать всю систему Toyota, не создав при этом необходимых возможностей и культурных основ, как правило, видели, как их инициативы затухают и терпят крах.

Данная модель отражает нашу концепцию. Трансформация работает с Р, модифицируя F. Революция стремится устранить Р, которое возвращается обратно.

Пример из практики: слияния и поглощения

Слияния и поглощения являются наглядными примерами динамики «цена/недостаток». Приобретающая компания часто полагает, что сможет перестроить приобретенную компанию в соответствии со своими собственными методами работы, рассматривая характеристики целевой компании как «недостаток». Реальность же часто оказывается отрезвляющей.

Когда в 1998 году Daimler-Benz объединился с

Chrysler, предполагалось, что обе компании можно будет интегрировать в единую структуру. В действительности же глубоко укоренившиеся культурные различия оказались гораздо более устойчивыми к изменениям, чем ожидалось. То, что считалось F (корпоративная культура), оказалось P, по крайней мере, в рассматриваемый период времени. Слияние провалилось, и Chrysler в конечном итоге был продан.

Как правило, успешные покупатели проводят тщательную проверку соответствия корпоративной культуре, планируют более длительные сроки интеграции и сохраняют, а не заменяют ключевые элементы функционирования приобретенной компании.

Пример из практики: Цифровая трансформация

Цифровая трансформация стала всеобщей необходимостью. Однако большинство цифровых трансформаций не приносят обещанных результатов.

Неудачи часто имеют общую закономерность:

отношение к тому, что на самом деле является положительным (P), воспринимается как F. Компания приобретает новые технологии (F с бюджетом) и объявляет о новых процессах (F с полномочиями), но обнаруживает, что поведение сотрудников не меняется (P в краткосрочной перспективе), ожидания клиентов остаются неизменными (P в среднесрочной перспективе), и вся система возвращается к привычным моделям поведения.

Успешные цифровые трансформации требуют иного подхода. Они начинаются с определения ограничений. Изменения вносятся поэтапно, чтобы постепенно наращивать возможности. При этом уделяется внимание человеческому фактору.

Карта ограничений на одной странице

Для любой значимой инициативы создайте одностраничную карту ограничений, в которой будет обобщена следующая информация:

1. **Цель:** Какого конкретного результата вы хотите добиться?

2. **Элементы Р:** Что не может измениться в вашем контексте?
3. **Элементы F:** Что именно можно изменить ?
4. **Зависимости:** Как взаимодействуют элементы Р и F?
5. **Риски:** В каких случаях вы можете ошибиться в классификации?
6. **Последовательность:** В каком порядке следует обращаться к элементам F?

Этот простой инструмент дисциплинирует мышление и предотвращает выявленные нами ошибки. Он заставляет учитывать ограничения перед принятием решения, повышая вероятность того, что усилия будут направлены туда, где они могут принести результат.

ГЛАВА 8: Для ученых — Границы моделей

«Все модели ошибочны, но некоторые полезны». —

Джордж Бокс

Наука развивается путем построения моделей — упрощенных представлений о реальности, которые отражают существенные особенности, опуская при этом несущественные. Но каждая модель содержит в себе допущения, и эти допущения ограничивают то, что модель может показать. В этой главе применяется подход Р/Г к научной практике, с особым вниманием к проблеме различения ограничений реальности от ограничений метода.

Каждая модель содержит в себе допущения.

Рассмотрим экономические модели, предполагающие рациональных агентов, максимизирующих полезность. Это предположение не является наблюдением за тем, как люди ведут себя на самом деле ; это упрощение, которое делает математический анализ поддающимся анализу. Когда люди ведут себя «иррационально» с точки зрения модели, мы

сталкиваемся с выбором: является ли это свойством людей (ограничение системы) или следствием предположений модели (ограничение инструмента)?

Та же проблема возникает во всей науке. Климатические модели должны упрощать динамику атмосферы. Эпидемиологические модели должны предполагать модели контактов между людьми. Психологические эксперименты должны операционализировать абстрактные понятия, такие как «интеллект» или «тревожность».

В каждом случае параметры модели P — ее заложенные предположения — определяют, какие выводы можно сделать. Осознание этих заложенных ограничений имеет важное значение для правильной интерпретации результатов.

Ловушка захвата метрик

Особенно опасной формой ограничения инструментов является метрический захват: тенденция оптимизировать то, что мы измеряем, а не то, что нам важно.

Образовательные системы, оптимизирующие результаты стандартизированных тестов, могут выпускать студентов, которые хорошо сдают тесты, но мало чему учатся. Системы здравоохранения, оптимизирующие измеримые результаты, могут пренебрегать важным опытом пациентов, который сопротивляется количественной оценке. Исследовательские программы, оптимизирующие показатели публикаций, могут приводить к созданию научных статей, а не к получению знаний.

В каждом случае выбор метрики (части инструмента) становится ограничением для получаемого результата. Метрика предназначена для измерения результата; в конечном итоге она формирует его.

Воспроизводимость и скрытые начальные условия

Кризис воспроизводимости в науке — обнаружение того, что многие опубликованные результаты не поддаются воспроизведению, — можно понять с помощью нашей концепции. Неудачные попытки воспроизведения часто отражают скрытые элементы

Р: аспекты первоначального исследования, которые не были распознаны как ограничения, но в итоге оказались важными.

В психологическом эксперименте, проведенном со студентами определенного университета, учитываются характеристики этой популяции. Если эти характеристики имеют значение для изучаемого явления, результат не будет воспроизведен в других популяциях. То, что считалось F (характеристики участников), оказывается Р (реальное ограничение для результата).

Решение заключается в более явном внимании к начальным условиям: какие аспекты настройки могут быть ограничениями, а не случайными факторами? Зарегистрированные отчеты, предварительная регистрация и подробные разделы с описанием методов помогают сделать Р более наглядным.

Различение реальности и метода

Принцип двойной предварительной подготовки особенно актуален для научной практики. Ученые

должны постоянно задавать себе вопрос: является ли это открытие свойством изучаемого объекта или свойством способа его изучения?

На этот вопрос нет простого ответа. Мы можем получить доступ к реальности только через методы, поэтому свойства методов и свойства реальности всегда взаимосвязаны. Но есть несколько стратегий, которые помогают:

Множество методов: Если разные методы дают одинаковый результат, он с большей вероятностью отражает реальность, чем артефакты, возникающие при использовании какого-либо одного метода.

Явное обозначение допущений: Явное обозначение допущений модели позволяет тщательно изучить то, что принимается за данность.

Анализ чувствительности: систематическое изменение допущений показывает, насколько выводы зависят от них.

Междисциплинарный диалог: разные дисциплины предоставляют разные инструменты; сравнение

показывает, какие ограничения носят методологический характер.

ГЛАВА 9: Для разработчиков ИИ — НАСЛЕДОВАНИЕ — ЭТО ПРЕИМУЩЕСТВО, А НЕ НЕДОСТАТОК

Мечта об искусственном интеллекте всегда была тесно связана с мышлением «с чистого листа». Фантазия о создании разума, свободного от человеческих ограничений, предвзятости и предубеждений — чистого интеллекта, рассуждающего из фундаментальных принципов, — проходит красной нитью через всю историю этой области.

В этой главе утверждается, что эта фантазия не просто не реализована, но и в принципе нереализуема. Каждая система искусственного интеллекта наследует ограничения от своих разработчиков, данных и архитектуры. Понимание этого наследования как преимущества, а не недостатка, имеет важное значение для ответственной разработки ИИ.

Почему «непредвзятый ИИ» — это оксюморон

Теория машинного обучения предлагает

убедительное опровержение фантазии об объективном искусственном интеллекте: теоремы «бесплатного обеда не бывает». Доказанные Дэвидом Вольпертом и Уильямом Макреди, эти теоремы устанавливают, что ни один алгоритм обучения не может превзойти все остальные во всех возможных задачах.

Смысл этого глубоко укоренен: для успеха в обучении необходимо иметь склонность к определенным типам решений. Без предположений о структуре проблемы — о том, какие закономерности вероятны, какие характеристики имеют значение — обучение невозможно. Склонность — это не недостаток, который нужно устранить; это необходимое условие для обучения чему бы то ни было.

В терминологии машинного обучения это называется «индуктивной предвзятостью». Каждый успешный алгоритм обучения имеет индуктивные предубеждения, которые делают его подходящим для одних задач и непригодным для других. Вопрос не в

том, нужна ли предвзятость , а в том, какие именно предубеждения нужны.

Четыре уровня наследования

Системы искусственного интеллекта наследуют ограничения на нескольких уровнях:

Архитектура: Базовая структура системы — нейронная сеть, дерево решений, трансформер — отражает предположения о том, как должна передаваться информация и какие закономерности можно обнаружить. Сверточные сети предполагают пространственную локальность. Рекуррентные сети предполагают последовательные зависимости. Это проектные решения, которые ограничивают все последующие этапы.

Обучающие данные: любые данные, на основе которых система обучается, отражают мир, который их создал, включая его предвзятость, дисбаланс и ограничения. Система может научиться только тому, чему учат данные.

Целевая функция: Математическое описание того,

что оптимизирует система, кодирует значения. Точность чего? Как это измеряется? Какой ценой по сравнению с другими факторами?

Метрики оценки: то, как мы оцениваем производительность, определяет, что оптимизируется. Системы учатся максимизировать то, что мы измеряем.

Каждый слой добавляет ограничения. Итоговая система отражает пересечение всех четырех слоев.

Образец паспорта

Поскольку наследование неизбежно, ответственная практика требует его явного указания. Я предлагаю, чтобы каждая система ИИ имела «Паспорт модели», документирующий унаследованные ограничения:

- **Архитектура:** Какие конструктивные решения были приняты и почему? Какие закономерности можно и нельзя выявить?
- **Данные:** На каких данных обучалась система? Какие популяции, временные периоды и источники представлены?

- **Цель:** Что именно оптимизирует система? Какие ценности заложены в этом выборе?
- **Оценка:** Как измеряется эффективность? Что не отражается этими показателями?
- **Известные ограничения:** Для решения каких задач система не подходит? Какие искажения были выявлены?

Модельный паспорт не устраняет наследование — ничто не может этого сделать — но он делает наследование видимым. Пользователи могут принимать обоснованные решения о целесообразности его использования.

Последствия для этики ИИ

Данная концепция имеет прямые последствия для этики искусственного интеллекта. Если предвзятость не может быть устранена, то:

1. **Заявления об объективности — это ложная реклама.** Любая система, представленная как «беспристрастная», искажает свою природу.
2. **Важный вопрос не в том, «является ли оно**

предвзятым?», а в том, **«являются ли его предвзятости уместными?»** . Различные области применения требуют различных индуктивных предубеждений.

3. **Документирование и прозрачность — этические императивы.** Если наследование неизбежно, оно должно быть, по крайней мере, видимым.
4. **Человеческий контроль по-прежнему имеет важное значение.** Люди должны оценивать, приемлемы ли присущие системе ограничения для предполагаемого использования.

ГЛАВА 10: Для политиков — САМЫЕ ДОРОГОСТОЯЩИЕ ОШИБКИ

Разработка политики — это область, где ошибки в рассуждениях об ограничениях обходятся дороже всего. Когда правительства путают Р с F, результатом может стать катастрофическая социальная инженерия, разрушающая жизни. Когда они путают F с Р, результатом становятся предотвратимые страдания, которые сохраняются из-за недостатка воображения. Когда они смешивают области, они применяют решения, разработанные для одного контекста, к проблемам в другом, часто с разрушительными последствиями.

В этой главе рассматриваются провалы политики через призму нашей концептуальной модели, извлекаются уроки из истории, которые остаются крайне актуальными и сегодня.

Трагедия революционной социальной инженерии

Двадцатый век стал мрачной лабораторией для масштабной проверки политики, предполагающей создание «с нуля». «Нулевой год» красных кхмеров,

Сталинская коллективизация, китайская культурная революция — все это основывалось на предположении, что человеческие общества можно перестроить с нуля, что унаследованные структуры являются препятствиями, а не фундаментом.

Число жертв этих экспериментов исчисляется десятками миллионов. Общая нить: отношение к тому, что на самом деле было Р, как к чему-то F. Привязанность человека к семье, потребность в материальных стимулах, важность местных знаний, ограничения централизованного планирования — все это считалось изменчивым, хотя на самом деле являлось ограничениями .

Урок заключается не в том, что перемены невозможны , а в том, что перемены должны соответствовать природе человека, а не противоречить ей. Политика , игнорирующая реальные потребности и желания людей в пользу того, что, согласно теории, им должно быть нужно и чего они должны желать, — это политика, основанная на ошибке первого типа.

Трагедия фаталистического принятия

Противоположная ошибка — принятие за неизбежность того, что можно изменить, — также влечет за собой огромные издержки, хотя они менее заметны, поскольку принимают форму предотвратимых страданий, а не активного разрушения.

На протяжении веков крайняя нищета воспринималась как часть естественного порядка вещей. Эпидемии болезней считались божественным промыслом. Исключение женщин из политической и экономической жизни основывалось на утверждениях о естественных различиях. Каждое из этих условий сохранялось отчасти потому, что их относили к категории Р, хотя на самом деле они принадлежали к категории F.

Резкое сокращение уровня бедности в мире за последние десятилетия, разработка вакцин и мер в области общественного здравоохранения, расширение прав женщин — все это потребовало прежде всего признания того, что, казалось бы,

неизменные условия можно изменить.

Сегодня аналогичные дебаты ведутся вокруг изменения климата (не слишком ли поздно действовать?), экономического неравенства (неизбежно ли это?) и технологического вытеснения (должны ли работники просто смириться с устареванием?). В каждом случае преждевременное включение в программу Р исключает возможности для изменения ситуации с помощью политических мер.

Трагедия путаницы доменных имен

Пожалуй, самая распространенная ошибка в политике — это перенос классификаций «зачет/незачет» из одной области в другую, где они не применимы.

Социальный дарвинизм взял за основу результаты исследований биологической эволюции и применил их к социальной политике, рассматривая экономическую конкуренцию как «выживание сильнейших» и противопоставляя вмешательство как «противоречащее природе». Ограничения

биологической сферы были незаконно перенесены в социальную сферу, где действуют иные ограничения.

Аналогичным образом, упрощенные экономические модели иногда рассматривают рыночные результаты как естественные и неизбежные, хотя на самом деле они отражают реальность. конкретные правовые и институциональные решения. Р модели превращается в Р политики, исключая альтернативы, которые модель не предусматривает, но которые могут быть допустимы в реальности.

Передача межкультурной политики часто сопряжена с одной и той же ошибкой. Программы развития, разработанные для одного контекста — с его особой историей, институтами и культурой — применяются к другим контекстам, как если бы эти различия были F, тогда как на самом деле они Р. Программы терпят неудачу, потому что они предполагают ограничения, которых не существует, и игнорируют ограничения, которые существуют.

Принципы мудрой политики

Наша концепция предлагает несколько принципов для разработки политики:

1. **Перед принятием решений проведите анализ ограничений.** Что в данном контексте является фиксированным ? А что действительно изменчиво?
2. **Различайте временные рамки.** Р и F различаются в зависимости от горизонта. Чрезвычайные меры отличаются от долгосрочных реформ.
3. **Уважайте границы областей знаний.** То, что верно в биологии, не обязательно верно в обществе. То, что работает в одной стране, может не работать в другой.
4. **Предпочтительнее постепенные изменения, а не революционные.** Работа с унаследованными структурами с большей вероятностью приведет к успеху, чем попытка заменить их целиком.
5. **Сохраняйте эпистемологическую**

скромность. Наше понимание ограничений всегда носит предварительный характер. Оставляйте место для понимания того, что границы лежат в других местах, нежели мы думали.

ГЛАВА 11: Для ВСЕХ — ЛИЧНЫЕ РЕШЕНИЯ И ОТНОШЕНИЯ

Разработанная нами концепция применима не только к организациям, технологиям и политике, но и к самым сокровенным сферам жизни: личным решениям и близким отношениям. В этой главе мы подробно рассмотрим анализ «процент/неверно».

Что вы не можете изменить в себе

Каждый из нас наследует определенные ограничения. Некоторые из них биологические: рост, базовый темперамент, предрасположенность к определенным заболеваниям. Некоторые — психологические: модели привязанности, сформированные в детстве, глубоко укоренившиеся предпочтения, характерные способы обработки эмоций. Некоторые — исторические: семья, в которой мы родились, формирующий опыт юности, совокупный эффект тысяч уже сделанных выборов.

Эти ограничения действительно являются психологическими. От них нельзя избавиться, их нельзя преодолеть силой воли или просто принять

другое решение. Человек, ненавидящий свою интроверсию, не может стать экстравертом силой воли. Человек с тревожным типом привязанности не может просто решить быть в безопасности. Человек, сформированный травмой, несёт её отпечаток даже в процессе исцеления.

Принятие этих ограничений — это не смирение, а реализм. Оно высвобождает энергию, которая раньше тратилась на тщетную борьбу с неизменным, и направляет её к подлинной свободе.

Что вы можете изменить

Но F никогда не равно нулю. В рамках существующих ограничений сохраняются подлинные степени свободы.

Интроверт не может стать экстравертом, но он может научиться навыкам общения в ситуациях, которые ему не свойственны. Человек с тревожным типом привязанности не может просто стать уверенным в себе, но он может распознавать свои модели поведения и выбирать реакции, которые не

усугубляют ситуацию. Человек, переживший травму, не может изменить прошлое, но он может повлиять на его дальнейшее воздействие на настоящее.

Концепция мышления роста, которую описывает Кэрол Двек, по сути, представляет собой точную калибровку соотношения вероятности и несоответствия. Она признает ограничения, не рассматривая при этом все как ограничения. Она признает трудности, не принимая при этом невозможность.

Что нельзя изменить в вашем партнере

Отношения часто рушатся из-за ошибки первого типа: веры в то, что другой человек изменится. Люди вступают в брак в надежде, что раздражающие привычки партнера исчезнут, что фундаментальные несовместимости разрешатся, что любовь изменит характер.

Иногда перемены действительно происходят. Но гораздо чаще черты характера, присущие началу отношений, сохраняются на протяжении всего

времени. Партнер, избегающий конфликтов, скорее всего, всегда будет избегать конфликтов. Партнер, нуждающийся в постоянной социальной стимуляции, скорее всего, всегда будет в ней нуждаться. Партнер, справляющийся со стрессом путем отчуждения, скорее всего, всегда будет отстраняться.

Это не значит, что люди никогда не меняются — иногда они меняются, — но это значит, что перемены нельзя принимать как данность. Мудрый человек выбирает партнеров, с которыми ему будет комфортно, вместо того чтобы делать ставку на трансформацию.

Что вы можете изменить

В отношениях существуют подлинные степени свободы. Модели общения могут улучшаться. Навыки разрешения конфликтов могут приобретаться. Способы, которыми партнеры поддерживают друг друга, могут меняться. Значение, которое они придают своим различиям, может изменяться.

Это области подлинной неприязни, где усилия могут изменить ситуацию. Ключ к успеху — сосредоточиться на том, что действительно можно изменить, зачастую это собственные реакции, а не характер партнера, вместо того чтобы изнурять себя борьбой с неизменяемым.

Почему принцип «начать все заново в понедельник» редко срабатывает

Извечная фантазия о новом начале — о новой диете, начинающейся в понедельник, о программе тренировок с 1 января, о решении «стать другим человеком» — отражает стремление начать все с чистого листа на личном уровне.

Эти инициативы обычно терпят неудачу, потому что слишком многое воспринимается как нечто негативное. Привычки, предпочтения и модели поведения, которые привели к нынешней ситуации, не исчезают от одного заявления. В краткосрочной перспективе они являются позитивными, а для их изменения требуются постоянные усилия в течение длительного времени.

Более успешные подходы работают в рамках существующих ограничений, а не против них. Они модифицируют окружающую среду, чтобы упростить желаемое поведение . Они опираются на существующие рутины, а не заменяют их целиком. Они ожидают постепенных изменений, а не мгновенной трансформации.

Межпоколенческие модели

Пожалуй, нигде концепция Р/Г не проявляется так ярко, как в межпоколенческих моделях поведения. Каждое поколение наследует от предыдущих: материальные условия, безусловно, но также и психологические модели поведения, стили отношений, неразрешенные травмы.

Эти наследственные права действительно имеют важное значение. Вы не можете выбирать своих родителей, свой опыт раннего детства или модели поведения, которые были вам смоделированы до того, как вы стали достаточно взрослыми, чтобы их оценить.

Но они также имеют и другие недостатки. Каждое поколение имеет возможность распознавать закономерности и выбирать, продолжать их или прерывать. Родитель, воспитанный в суровых условиях, может выбрать мягкость. Человек из семьи, где подавлялись эмоции, может научиться их выражать. Наследие прошлого ограничивает, но не определяет.

Возможно, это наиболее глубокое применение нашей концепции: признание того, что мы формируемся под влиянием прошлого и что мы сами формируем будущее. Мы не можем избежать наследия, но можем выбирать, как им распорядиться.

ГЛАВА 12: ДВАДЦАТЬ ЛОВУШЕК И КАК ИЗ НИХ ВЫБРАТЬСЯ

В этой главе представлен краткий справочник по распространенным ситуациям, в которых возникают ошибки типа «проверка правильности/неправильности». Для каждой ситуации я определяю вероятный тип ошибки и предлагаю рекомендации.

Ловушка 1: «Мы внедрим Agile через месяц»

Ситуация: Руководство объявляет о быстром внедрении методологии Agile во всей организации.

Тип ошибки: Тип I (обращение Р как F). Организационная культура, существующие рабочие процессы и возможности персонала не могут быть изменены путем объявления.

Рекомендация: Запланируйте постепенный переход на 18-24 месяца. Начните с пилотных групп. Развивайте потенциал, прежде чем требовать изменения поведения.

Ловушка 2: «Эти отношения будут другими»

Ситуация: Вступление в новые отношения с сохранением тех же моделей поведения, которые нанесли ущерб предыдущим.

Тип ошибки: Тип I. Ваш стиль привязанности, модели общения и эмоциональные привычки являются Р-фактором относительно изменений в отношениях.

Рекомендация: Работайте над собой в перерывах между отношениями, а не только во время них. Осознайте, какие модели поведения остаются с вами надолго.

Ловушка 3: «Ребенок должен стать врачом, как и все члены семьи».

Ситуация: Родители ожидают, что дети будут следовать заранее определенному пути, независимо от интересов или способностей ребенка.

Тип ошибки: Тип I. Подлинные интересы и способности ребенка соответствуют категории Р; ожидания родителей соответствуют лишь категории F.

Рекомендация: Различайте надежды семьи (F) и природу ребенка (P). Поддерживайте того ребенка, который у вас есть, а не того, которого вы себе представляли.

Ловушка 4: «Я просто не математика»

Ситуация: Избегание любых математических задач из-за негативного опыта, полученного ранее.

Тип ошибки: тип II (обращение с F как с P). Математические способности могут развиваться при наличии усилий и соответствующего обучения.

Рекомендация: Различать текущий уровень навыков (неудовлетворительный, можно улучшить) и застывший уровень. Стремиться к более качественному обучению, а не к постоянному избеганию.

Ловушка 5: «Так устроена наша отрасль»

Ситуация: Принятие отраслевых практик без проверки их необходимости .

Тип ошибки: Тип II. Многие «отраслевые стандарты» представляют собой условности (F),

ошибочно принимаемые за ограничения (P).

Рекомендация: Спросите: кто-нибудь делал это по-другому? Если да, то это F. Если нет, спросите почему — это действительно невозможно или просто не распространено?

Ловушка 6: «Нам нужно переписать весь код».

Ситуация: Предлагается отказаться от устаревшего программного обеспечения и разработать его заново с нуля.

Тип ошибки: Тип I. Накопленные знания в устаревшем коде имеют значение P; они не будут автоматически перенесены в новый код.

Рекомендация: Проводите рефакторинг постепенно. Сохраняйте то, что работает. Документируйте изменения перед заменой.

Ловушка 7: «Человеческая природа делает неравенство неизбежным»

Ситуация: Использование утверждений о биологии для оправдания социальных отношений.

Тип ошибки: Тип III (путаница в предметной области). Биологические ограничения отличаются от решений в области социальной политики.

Рекомендация: Различать биологические тенденции (P в биологии) и институциональные механизмы (часто F в политике).

Ловушка 8: «Наш ИИ будет беспристрастным».

Ситуация: Маркетинг систем искусственного интеллекта как свободных от предвзятости или предубеждений.

Тип ошибки: Тип I. Предвзятость присуща любой системе обучения; её нельзя устранить, можно только управлять ею.

Рекомендация: Документируйте предвзятость. Сопоставляйте предвзятость с соответствующими областями применения. Никогда не заявляйте о нейтральности.

Ловушка 9: «Переезд в новый город решит все проблемы»

Ситуация: Вера в то, что переезд в другой регион

решит личные проблемы.

Тип ошибки: Тип I. Личность перемещается в новое место; изменяется только окружающая среда .

Рекомендация: Различайте факторы окружающей среды (факторы, связанные с переездом) и личные модели поведения (факторы, которые сопровождают вас).

Ловушка 10: «Рынок высказался»

Ситуация: Рассмотрение рыночных результатов как естественных и неизбежных, а не как продукта институциональных решений.

Тип ошибки: Тип III. Результаты рыночной конъюнктуры отражают правовые, регуляторные и социальные решения (F), а также экономические факторы (P).

Рекомендация: Задайте вопрос: какие аспекты этого результата отражают решения, которые можно было принять иначе?

Ловушка 11: «Я изменюсь, когда буду готов»

Ситуация: Бессрочное откладывание изменения поведения в ожидании появления мотивации.

Тип ошибки: Тип II. Мотивация часто следует за действием, а не предшествует ему; ожидание рассматривает действие как предпосылку.

Рекомендация: Начинайте с малого. Не ждите готовности. Развивайте мотивацию через действие.

Ловушка 12: «Нам необходима полная культурная трансформация».

Ситуация: Руководители объявляют о быстрых изменениях в организационной культуре.

Тип ошибки: Тип I. Культура глубоко укоренена в краткосрочной перспективе; она развивается медленно даже при постоянных усилиях.

Рекомендация: Планируйте на годы, а не на месяцы. Меняйте системы и стимулы; культура последует за этим.

Ловушка 13: «Люди никогда по-настоящему не меняются»

Ситуация: Отказ от попыток личностного роста или улучшения отношений.

Тип ошибки: Тип II. Хотя фундаментальные изменения даются с трудом, значимое развитие возможно.

Рекомендация: Различать устойчивые черты характера (P) от приобретаемых навыков и адаптируемых моделей поведения (F).

Ловушка 14: «Технологии решат проблему»

Ситуация: Ожидается, что новые технологии решат фундаментальные проблемы, связанные с человеческим фактором или организационной структурой.

Тип ошибки: Тип I. Технология может способствовать внедрению, но не принуждать к нему; человеческий фактор остается важным фактором.

Рекомендация: Задайте себе вопрос: что должно измениться помимо технологий? Учтите это при планировании.

Ловушка 15: «Мы всегда так делали»

Ситуация: Отказ рассматривать альтернативы устоявшимся практикам.

Тип ошибки: Тип II. Традиция указывает на то, что было, а не на то, что должно быть.

Рекомендация: Задайте себе вопрос: почему изначально это было сделано именно так? Остаются ли эти причины актуальными и сегодня?

Ловушка 16: «Данные говорят сами за себя»

Ситуация: Рассмотрение статистических данных как не требующих интерпретации.

Тип ошибки: Тип III. Сбор и анализ данных включают в себя выбор (F), который формирует выводы.

Рекомендация: Задайте себе вопрос: какие предположения легли в основу получения этих данных? Приведут ли разные варианты выбора к разным результатам?

Ловушка 17: «При достаточной силе воли возможно всё».

Ситуация: Вера в то, что усилиями можно преодолеть любое препятствие.

Тип ошибки: Тип I. Сила воли действует в рамках ограничений; она не может изменить Р.

Рекомендация: Стратегически используйте силу воли в отношении F; примите Р и обойдите его стороной.

Ловушка 18: «Это просто моя натура»

Ситуация: Использование особенностей характера как оправдание для избегания изменений в поведении.

Тип ошибки: Тип II. Хотя основные черты характера стабильны, поведение может быть изменено.

Рекомендация: Различайте личностные черты (Р) и конкретные модели поведения (часто F).

Ловушка 19: «Нам нужно изменить отрасль»

Ситуация: Планирование стартапа, основанное на пересмотре устоявшихся практик.

Тип ошибки: тип I, если применяемые методы

отражают реальные ограничения; тип II, если это просто условности.

Рекомендация: Прежде чем планировать сбой, тщательно проанализируйте, какие отраслевые практики являются позитивными (P), а какие — негативными (F).

Ловушка 20: «На этот раз все будет по-другому»

Ситуация: Повторение ранее использованных подходов с ожиданием иных результатов.

Тип ошибки: Неправильная диагностика. Предыдущие ошибки указывают на неверную классификацию.

Рекомендация: Проанализируйте, что пошло не так. Была ли буква P обработана как F? Была ли буква F обработана как P? Произошла ли путаница с доменами? Устраните фактическую ошибку.

ПОСЛЕСЛОВИЕ: МУДРОСТЬ ОТЛИЧИЯ КАК ПРАКТИКА

Мы прошли путь от античной философии до современной когнитивной науки, от революционной истории до организационной динамики, от архитектуры ИИ до личных отношений. Одна и та же концепция освещала каждую область: разделение любого начального условия на предопределенные (Р) и свободные (F) элементы, три типа ошибок, возникающих при их неправильной классификации, и двойное предварительное обусловливание, определяющее то, что мы можем знать и делать.

Но концептуальные модели полезны лишь настолько, насколько эффективно они применяются на практике. Мудрость различения — это не разовое открытие, а непрерывная дисциплина.

Дисциплина, достойная отличия

Для точной оценки «зачет/незачет» требуется постоянное внимание. Мир меняется; то, что вчера было «зачетом», завтра может стать «незачетом» благодаря новым технологиям, новым ресурсам или

новым социальным условиям. То, что вчера было «незачетом», завтра может стать «зачетом» из-за принятых решений, упущенных возможностей или закрытых окон.

Рассмотрим, как технологии меняют границу между реальностью и реальностью. На протяжении большей части истории человечества общение на расстоянии было реальностью — можно было отправить письмо или путешествовать самому, но взаимодействие в реальном времени было невозможно. Теперь это реальность; мы сами выбираем, устанавливать мгновенную связь или нет. Ограничение стало выбором.

И наоборот, решения создают новые ограничения. Выбранный вами колледж исключает другие варианты. Начатый вами карьерный путь определяет, что будет дальше. Партнер, с которым вы связали свою жизнь, означает, что другие возможности не осуществились. То, что было F, становится R благодаря осуществлению выбора.

Практика различения подразумевает регулярную

переоценку: что изменилось? Какие новые ограничения возникли? Какие старые ограничения исчезли?

Смирение прежде ограничений

Понимание ограничений порождает смирение — не ложное смирение, которое утверждает, что ты бессилён, а подлинное смирение, которое признаёт ограничения, работая в их рамках.

Эта скромность признаёт, что мы не контролируем результаты, а только исходные данные. Мы можем действовать мудро в рамках наших возможностей, понимая при этом, что многое из того, что последует, будет зависеть от наших возможностей. Мы можем сажать семена, не управляя урожаем.

Такая скромность не пассивна. Она направляет энергию туда, где она может изменить ситуацию. Она прекращает тщетную борьбу с неизменным. Она создает пространство для эффективных действий в рамках реально достижимого.

Свобода через понимание

Самый глубокий парадокс этой книги заключается в том, что принятие ограничений освобождает.

Тот, кто понимает, что нельзя изменить, перестаёт тратить энергию на борьбу с этим. Организация, осознающая свои подлинные степени свободы, может действовать решительно в их рамках. Общество, которое точно оценивает, чего может и чего не может достичь политика, ставит перед собой реалистичные цели и часто их достигает.

Именно это освобождение обещает Молитва о безмятежности: не свободу от ограничений — это невозможно, — а свободу через их понимание. Мудрость, позволяющая различать эти ограничения, — это не дар, а навык, который развивается. Он приходит с практикой, с ошибками и их исправлениями, с терпеливой работой по точной оценке.

Приглашение

Я начал эту книгу с вопроса: почему умные люди постоянно не могут отличить то, что можно изменить,

от того, что изменить нельзя? Ответ, как я утверждаю, заключается в отсутствии у нас систематической основы для проведения этого различия. Анализ Р/Ф, три типа ошибок и принцип двойного предварительного обусловливания обеспечивают эту основу.

Но фреймворки имеют значение только в том случае, если их используют. Поэтому в заключение я призываю : применяйте эти идеи на практике. Проводите аудит ограничений перед следующей инициативой. Диагностируйте следующую ошибку, используя типологию ошибок. Ставьте под сомнение свои предположения о том, что является фиксированным, а что — свободным.

Мир полон ограничений, которые нельзя изменить, и возможностей, которыми можно воспользоваться. Мудрость, позволяющая различать их, ждет вас на практике.

ПРИЛОЖЕНИЯ

Приложение А: Глоссарий терминов

Р (Предопределенные): Элементы исходного условия, которые не могут быть изменены в соответствующем контексте, временных рамках и области. Ограничения, которые необходимо принять или обойти.

Ф (Свобода): Элементы исходного состояния, которые можно изменить, выбрать или варьировать. Степени свободы, при которых принимаемые решения имеют значение.

I = Р + F: Фундаментальное уравнение, утверждающее, что каждое начальное условие состоит как из предопределенных, так и из свободных элементов. Ни Р, ни F никогда не равны нулю.

Ошибка I типа (ложная нейтральность): путаница между Р и F; попытка изменить то, что изменить нельзя. Ошибка «чистого листа».

Ошибка типа II (ложная фиксированность): ошибочное принятие F за Р; принятие за

фиксированное то, что на самом деле может быть изменено. Детерминированная ошибка.

Ошибка типа III (непонимание доменов):

Применение раздела R/F из одного домена к другому домену, где применяются разные разделы.

Двойная предварительная обработка: принцип,

согласно которому ограничения действуют на двух уровнях: внутренние ограничения, исходящие от самой системы, и внешние ограничения, исходящие от используемых нами инструментов описания.

Аудит ограничений: систематическая

инвентаризация элементов R и F перед началом изменений.

Типовой паспорт: Документ, описывающий

присущие системе искусственного интеллекта ограничения: архитектура, данные, цель и оценка.

Приложение В: Примеры ошибок по областям знаний

Домен	Пример ошибки типа I	Пример ошибки типа II	Пример ошибки типа III
Бизнес	«Мы изменим культуру за шесть месяцев».	«Так устроена наша отрасль».	Рассматривать рыночные условности как законы природы
Технологии	«Наш ИИ будет беспристрастным».	«Устаревшие системы невозможно улучшить».	Рассматривать вычислительные ограничения как логически невозможные.

	Пример ошибки типа I	Пример ошибки типа II	Пример ошибки типа III
Домен			
Политика	«Мы создадим нового гражданина»	«Бедность неизбежна»	Использование биологических данных для оправдания социальных отношений
Личное	«Мой партнёр изменится после свадьбы»	«Я просто не математика».	Рассматривать личность как биологически неизменную

	Пример ошибки типа	Пример ошибки типа II	Пример ошибки типа III
Домен	I		ю характерис тику
Наука	«Наша модель учитывает все соответству ющие факторы»	«Мы всегда измерял и это именно так».	Рассмотрен ие допущений модели как эмпиричес ких фактов

Приложение С: Шаблон карты ограничений

Инициатива: _____

Конкретная _____ **цель:**

Временные _____ **рамки:**

Элементы Р (не могут быть изменены): 1.

_____ 2.

_____ 3.

Элементы F (могут меняться): 1.

_____ 2.

_____ 3.

Обоснование классификации:

Оценка рисков (Какие классификации могут быть неверными?):

Последовательность действий (какие элементы F следует рассмотреть и в каком порядке?):

Приложение D: Рекомендуемая литература

На чистом листе: - Стивен Пинкер, *«Чистый лист: современное отрицание человеческой природы»* - Гэри Маркус, *«Рождение разума»*

Об институциональных изменениях: - Дуглас Норт, *«Институты, институциональные изменения и экономические показатели»* - Пол Пирсон, *«Увеличение отдачи, зависимость от предшествующего пути развития и изучение политики»*

Об искусственном интеллекте и предвзятости: - Майкл Кернс и Аарон Рот, *«Этический алгоритм»* - Кэти О'Нил, *«Оружие математического уничтожения»*

О личностных изменениях: - Кэрол Двек, *«Мышление: новая психология успеха»* - Джеймс Кларк, *«Атомные привычки»*

О мышлении и развитии: - Анджела Дакворт, *«Упорство»* - Андерс Эрикссон, *«Пик»*

Приложение Е: Протокол аудита ограничений (10 шагов)

Шаг 1: Определите целевое изменение .
Сформулируйте одним предложением, что вы пытаетесь изменить. Укажите область применения и критерии успеха.

Шаг 2: Определение границ системы. Перечислите, что находится внутри системы и что является ее окружением. Укажите входы, выходы и интерфейсы.

Шаг 3: Составление перечня ограничений. Перечислите все ограничения: физические, человеческие, финансовые, правовые, рыночные, технологические, культурные, политические.

Шаг 4: Классифицируйте как Р, F или Неизвестно. Для каждого ограничения присвойте значение Р (не может измениться), F (может измениться) или U

(требуется дополнительная информация).

Шаг 5: Проверка классификаций. Примените тест на необходимость, тест на замещение и контрфактический тест к каждой U-классификации и спорной классификации.

Шаг 6: Сопоставление зависимостей. Как взаимодействуют элементы Р и F? Какие последовательности необходимы?

Шаг 7: Выявление рисков ошибок. Какой из трех типов ошибок вы, скорее всего, совершите? Как будет выглядеть каждый из них?

Шаг 8: Проектные решения. Сосредоточьте действия только на элементах F. Для всех оставшихся элементов U включите меры предосторожности.

Шаг 9: Определение критериев оценки . Укажите показатели успеха. Убедитесь, что эти показатели не приводят к непредусмотренным изменениям в поведении.

Шаг 10: Мониторинг и обновление . Выполните мониторинг. Обновите карту Р/F на основе

полученных данных.

Приложение F: Полный текст научной работы

[Примечание: Полный текст научной статьи Бориса Кригера «Структурные предпосылки описываемых операций: методологическая основа для анализа начальных условий» включен здесь. В этой статье представлен строгий формальный подход, лежащий в основе доступной методологии, описанной в данной книге.]

Структурные предпосылки описываемых операций: методологическая основа для анализа начальных условий

Борис Кригер

Абстрактный

В данной работе устанавливается условная теорема: для любого процесса, который можно описать как процесс, претерпевающий переходы состояний, начальные состояния с нулевой структурой не могут функционировать в качестве операциональных отправных точек для описываемых процессов. Аргумент носит эпистемологический и операциональный, а не метафизический характер. Мы не утверждаем ничего о метафизическом ничто или о

том, что может существовать за пределами описываемости; мы утверждаем лишь, что требования описываемости влекут за собой непустые структурные предпосылки.

Мы разрабатываем два принципа. Принцип ненулевой инициализации гласит: если процесс можно описать как имеющий переходы, то его описание предполагает непустую область дифференцируемых состояний. Принцип среднего ограничения гласит: любой формальный язык или модель, используемые для представления системы, ограничивают то, что может быть выражено до любого конкретного акта представления. Вместе они дают двойное предварительное обусловливание: для систем, которые одновременно описываемы и представимы, начальные состояния лежат на пересечении двух непустых пространств ограничений.

В статье строго разделяются три уровня: (1) концептуальные необходимости, вытекающие из требований описания, (2) эмпирические примеры,

демонстрирующие актуальность данной концепции для когнитивной науки, искусственного интеллекта и математики, и (3) прагматические последствия. Эмпирические примеры служат иллюстрациями применимости, а не предпосылками или доказательствами основного аргумента.

Данная диссертация имеет узкую предметную область: она посвящена системам, поддающимся описанию, моделированию и анализу. Вопросы о метафизическом происхождении, сотворении мира из ничего или областях, выходящих за рамки описания, выходят за рамки данной работы.

Ключевые слова: структурные предпосылки, описываемость, операционные ограничения, начальные условия, зависимость от предшествующего пути, индуктивная предвзятость, формальные системы

Введение

Известная практическая мудрость гласит: «Даруй мне спокойствие, чтобы принять то, что я не могу

изменить, мужество, чтобы изменить то, что я могу, и мудрость, чтобы отличить одно от другого».

В повседневной жизни это различие кажется само собой разумеющимся. Однако в научных рассуждениях, технологическом проектировании и социальном планировании его неоднократно игнорируют. Во многих областях мы сталкиваемся с неявными предположениями о том, что системы можно рассматривать так, как если бы они начинались с нейтральных, ничем не ограниченных состояний; что модели можно строить без унаследованных предубеждений; или что институты и алгоритмы можно перепроектировать «с нуля». Эти предположения представляют собой современные формы заблуждения *tabula rasa* — убеждения в том, что начало может быть структурно не обусловлено.

В данной статье утверждается, что общеизвестное утверждение о «знании разницы» требует формального методологического обоснования. Мы стремимся проанализировать, а не просто утверждать, почему такое различие необходимо, как его можно

строго сформулировать и какие ошибки возникают при его игнорировании. Наша цель — не предложить моральные рекомендации, а предоставить эпистемологическую основу, способную отделить предопределенные ограничения от изменяемых степеней свободы в любой описываемой системе.

С этой целью мы разрабатываем два взаимодополняющих принципа: принцип ненулевой инициализации, который устанавливает, что операционные системы требуют наличия предварительной структуры, и принцип среднего ограничения, который показывает, что любой акт описания накладывает дополнительные ограничения. Их сочетание приводит к тому, что мы называем *двойным предобусловливанием* — формальному объяснению того, почему ни один процесс не может начаться с абсолютной нейтральности в любой области, открытой для анализа.

В этом смысле статью можно рассматривать как методологическую разработку старой интуиции. Там, где практическая мудрость советует различать

неизменное и изменчивое , мы стремимся объяснить, почему такое различие неизбежно, как его можно систематически проводить и как избежать повторяющихся ошибок, возникающих при его игнорировании.

Вопрос и его масштабы

Выражение *tabula rasa* — в переводе с латыни «чистый лист» — на протяжении тысячелетий имело философский смысл. Эта концепция предполагает, что истоки могут быть нейтральными, что начальные точки структурно не определены, и что то, что следует дальше, является исключительно функцией последующих факторов, а не предшествующих условий.

В данной статье выдвигается условное утверждение: для любой системы, которая является описываемой, функционирующей или эволюционирующей, состояния с нулевой структурой не могут функционировать в качестве начальных условий для описываемых процессов. Это не утверждение о метафизическом ничто или конечной природе

реальности. Это утверждение о требованиях описываемости и функционирования.

ОБЪЕМ И МЕТОДОЛОГИЧЕСКОЕ РАЗЪЯСНЕНИЕ

Аргументация в данной работе носит эпистемологический и операциональный, а не метафизический характер.

Мы утверждаем: описание процесса как претерпевающего изменения требует наличия дифференцируемых состояний, что составляет ненулевую структуру.

Мы НЕ заявляем:

- Эта метафизическая пустота невозможна в каком-то абсолютном смысле.
- То, что состояния с нулевой структурой невозможно представить или формально изобразить.

- Всё, что касается доменов, выходящее за рамки описания.

Данный аргумент касается условий описания и функционирования, а не конечной природы реальности.

Операционная и логическая невозможность

Мы различаем два типа невозможности:

Логическая невозможность: состояние, содержащее внутренние противоречия (например, женатый холостяк). Такие состояния невозможно связно осмыслить.

Операционная невозможность: состояние, несовместимое с тем, чтобы служить начальным условием для любого описываемого процесса. Такие состояния могут быть формально представимы, но не могут функционировать в качестве отправных точек для описания или функционирования.

Мы утверждаем второе, а не первое. Можно написать « \emptyset » и работать с ним в формальных системах. Наше утверждение состоит в том, что состояния с нулевой

структурой не могут функционировать в качестве начальных условий для описываемых процессов, а не в том, что их нельзя представить или осмыслить.

Логическая структура аргумента

Основной аргумент имеет следующую структуру:

ЛОГИЧЕСКИЙ ВЫВОД

1. Для описания изменений → необходимо различать состояния (до и после).
2. Для различения состояний → необходимы дифференцируемые элементы.
3. Дифференцируемые элементы → образуют непустую область определения.
4. Следовательно: описание изменения → непустой области определения.

Это условное выведение из требований описания, а не метафизическое утверждение о необходимости во всех возможных мирах.

Три уровня анализа

В данной работе строго соблюдается разделение

между тремя уровнями:

Уровень 1 – Концептуальные необходимости:

Утверждения, вытекающие из требований описываемости. Это основной аргумент: если что-то описываемо как претерпевающее изменения, то должны выполняться определенные предварительные условия.

Уровень 2 – Эмпирические иллюстрации:

Примеры из когнитивной науки, ИИ и математики, иллюстрирующие, что данная концепция отражает что-то важное. Это не доказательства или подтверждения концептуальных утверждений, а демонстрации применимости.

Уровень 3 – Прагматические последствия:

Практические последствия для осмысления обучения, ИИ и моделирования. Они вытекают из применения данной концепции в конкретных контекстах.

Структура документа

В разделе [2](#) рассматривается соответствующая литература. В разделе [3](#) разрабатывается формальная

структура с точными определениями. В разделе [4](#) представлено двойное предварительное обусловливание. В разделе [5](#) рассматривается возражение о самоинициализации. В разделе [6](#) рассматривается математическая пустота. В разделе [7](#) приводятся примеры применения в ИИ. В разделе [8](#) перечислены возражения. В разделе [9](#) обсуждаются практические применения.

Методологический характер, новаторский вклад и когнитивная защита.

Данная статья носит принципиально методологический характер. Ее цель — выявить и формализовать повторяющуюся эпистемологическую ошибку в концептуализации начальных условий в различных областях и предложить принципиальную основу для избежания этой ошибки в будущих анализах. Речь идет о широко распространенной тенденции рассматривать описываемые процессы так, как если бы они могли осмысленно начинаться с состояний абсолютной структурной нейтральности. Мы называем эту тенденцию *ошибкой «чистого*

листа» (*tabula rasa*) . Эта ошибка возникает не из метафизических спекуляций о ничто, а из недостаточного понимания ограничений, которые неизбежно сопровождают любой акт описания или представления. Формулируя эти ограничения с помощью теоремы двойного предварительного обусловливания, статья предлагает диагностический инструмент и корректирующий механизм, превращая часто неявное упущение в явный методологический принцип.

Наш подход не является ни метафизическим, ни широко философским в спекулятивном смысле. Он не выдвигает никаких утверждений об окончательной природе реальности, о возможности или невозможности абсолютного небытия или о происхождении в каком-либо онтологическом регистре. Вместо этого он носит строго эпистемологический и методологический характер. Аргумент касается только условий, при которых описания операциональных процессов могут быть сделаны связными и поддающимися анализу.

Центральный вопрос заключается не в том, что может существовать в принципе, а в том, что может быть осмысленно определено в рамках репрезентативной системы.

Новизна данного вклада заключается в трех взаимосвязанных измерениях.

Во-первых, в статье диагностируется повторяющееся когнитивное искажение, встречающееся в различных дисциплинах. Это искажение, аналогичное искажению подтверждения в своем избирательном игнорировании неудобных предпосылок, проявляется всякий раз, когда аналитики неявно предполагают, что системы, модели или институты могут быть осмысленно представлены как исходящие из нейтрального или бесструктурного состояния. В когнитивной науке оно проявляется в чрезмерно упрощенных эмпирических моделях, которые постулируют обучение без индуктивных априорных знаний; в искусственном интеллекте — в утверждениях о «непредвзятых» алгоритмах, игнорирующих архитектурное и

информационное наследование; в социальных науках — в нарративах о радикальных институциональных перестройках, которые пренебрегают зависимостью от предшествующего пути и встроенными ограничениями. Эти примеры служат лишь иллюстрациями, а не доказательством теоремы. Их роль заключается в том, чтобы показать, как одно и то же концептуальное искажение повторяется всякий раз, когда игнорируются предпосылки описания.

Во-вторых, данная концепция возводит этот диагноз в ранг формальной гарантии. Выдвинутые здесь принципы не являются тавтологическими переформулировками определений. Они выводятся из минимальных эпистемологических предположений о том, что необходимо для описания любого процесса вообще. Мы начинаем с намеренно не зависящего от структуры понятия системы — всего, что может быть представлено как претерпевающее переходы, — и показываем, что акт описания таких переходов обязательно влечет за собой дифференцируемые состояния и непустую область определения. Посылки

касаются только требований представления; заключение следует как следствие, а не как предположение. Поэтому ценность теоремы носит профилактический характер: она предоставляет аналитикам четкий контрольный список для выявления случаев, когда ошибка «чистого листа» неявно вводится в модель или аргумент.

Во-третьих, оригинальность данной работы заключается в ее междисциплинарном синтезе и практической пользе. Элементы аргументации перекликаются с устоявшимися идеями, такими как индуктивная предвзятость в машинном обучении, аргумент о бедности стимула в лингвистике и зависимость от предшествующего пути в экономике. Однако до сих пор отсутствовала единая концептуальная основа, которая связывала бы эти идеи в рамках единой условной теоремы, тщательно отделяя эпистемологические требования от эмпирических иллюстраций и прагматических последствий. Трехуровневая структура, принятая в данной работе — требования, иллюстрации и

следствия — обеспечивает методологическую ясность и предотвращает смешение логических требований со специфическими для предметной области фактами.

Цель данной работы, следовательно, скромна, но точна. Под «ошибкой» мы понимаем не логическое противоречие, а устойчивую методологическую некорректность, игнорирующую предпосылки описания. Теорема не утверждает, что абсолютная структурная нейтральность метафизически невозможна. Она утверждает лишь, что такая нейтральность не может функционировать как осмысленная отправная точка в любой описываемой или операциональной области. В этом смысле данная структура служит своего рода когнитивной защитой: она дисциплинирует исследование, требуя явного признания ограничений, которые обязательно предшествуют любому связному представлению.

В целом, вклад заключается не в переоткрытии очевидного, а в целенаправленном противодействии распространенному концептуальному упрощению.

Условная форма аргументации сохраняет необходимую скромность, обеспечивая при этом практическую строгость. Она повышает методологическую точность, не выдвигая утверждений, выходящих за рамки возможностей логики описания .

Типы методологических ошибок в рассуждениях о начальных условиях

В предыдущих разделах было установлено, что ни один описываемый процесс не может быть последовательно смоделирован как начинающийся из состояния абсолютной структурной нейтральности. Однако этот вывод не подразумевает, что все аспекты начального состояния одинаково predetermined. В практическом анализе начальные состояния обычно содержат как predetermined компоненты, так и компоненты, которые действительно могут изменяться. Методологическая путаница возникает, когда эти области смешиваются. (Подробные примеры этих ошибок в истории, науке, политике, психологии и других областях см. в Приложении [11.](#))

Для уточнения этого момента введем простую схему обозначений. Пусть любое начальное условие I понимается как состоящее из двух подмножеств:

- P – элементы, предопределенные в рамках соответствующей описательной модели (архитектурные ограничения, физические законы, пределы представления, унаследованные параметры);
- F – элементы, которые остаются *формально свободными* или недоопределенными относительно данной структуры (выбор, регулируемые параметры, условные входные данные).

Таким образом, любую описываемую отправную точку можно схематически представить следующим образом:

$$I = P + F$$

Ошибка *tabula rasa* заключается в том, чтобы рассматривать ситуацию I так, как если бы $P = \emptyset$. Противоположная ошибка заключается в том, чтобы

рассматривать $F = \emptyset$. Оба искажения порождают характерные методологические ошибки.

Ошибка I типа: Неправильное приписывание свободы (ложная нейтральность)

Эта ошибка возникает, когда элементы, принадлежащие к predetermined области определения, P некорректно рассматриваются как свободно выбираемые элементы этой области F .

Формально: $P \rightarrow F$

Примеры включают:

- Утверждается, что алгоритм может быть «непредвзятым», несмотря на фиксированные архитектурные индуктивные смещения;
- Образовательные теории, предполагающие, что учащиеся приходят к обучению без каких-либо когнитивных предварительных знаний;
- Институциональные предложения, обещающие полную перезагрузку, игнорирующие при этом правовые, инфраструктурные или исторические ограничения.

Суть ошибки заключается в том, чтобы представить себе возможность выбора там, где её нет. Это наиболее прямое проявление заблуждения «чистого листа»: предположение, что то, что структурно фиксировано, можно рассматривать как нейтральное.

Ошибка типа II: Неправильное определение (ложная фиксированность)

Симметричная ошибка возникает, когда действительно открытые или регулируемые компоненты F ошибочно классифицируются как предопределенные P :

Формально: $F \rightarrow P$

В данном случае аналитик отрицает наличие частичной *tabula rasa* там, где она фактически существует. Типичные примеры включают:

- Аргументы о том, что социальные или когнитивные результаты полностью определяются генетикой или архитектурой;
- Технологический детерминизм, который исключает возможность альтернативных

проектов или политик;

- Методы моделирования, которые рассматривают настраиваемые параметры как неизменяемые характеристики.

Эта ошибка приводит к фатализму и ригидности: отказу признать степени свободы, которые на самом деле доступны для изменения.

Ошибка типа III: Слияние доменов (неправильное размещение категорий)

Наиболее серьезная методологическая ошибка возникает, когда границы между P и F не просто неправильно классифицируются, а относят их к совершенно неверной описательной области.

Формально, если домены D_1 и D_2 требуют разных разделов P и F , то возникает следующая ошибка:

$$(P + F)_{D_1} \text{ applied to } D_2$$

Примеры включают:

- Рассматривать социокультурные ограничения как биологические необходимости;

- Рассматривать вычислительные ограничения как логически невозможные;
- Предполагается, что ограничения конкретной среды моделирования представляют собой ограничения самой целевой системы.

Этот вид ошибки особенно опасен, поскольку он не просто неправильно распределяет свободу и ограничения; он неправильно определяет сам уровень, на котором проводится анализ. Он путает предпосылки представления с внутренними свойствами или локальные конвенции с универсальными необходимыми условиями.

Частичная *Tabula Rasa*: корректирующее уточнение.

Теорема о двойном предварительном обусловливании исключает абсолютную структурную нейтральность, но не отрицает существование частичной нейтральности. В рамках любого ограниченного начального условия $I = P + F$ подмножество F может быть существенным. Таким образом, данная

концепция поддерживает более тонкую позицию:

- **Абсолютная *tabula rasa*** ($P = \emptyset$) является непоследовательной для описываемых процессов;
- **Частичная *tabula rasa*** ($F \neq \emptyset$) не только возможна, но и повсеместно распространена.

Методологически обоснованный анализ требует определения для каждой области, какие компоненты относятся к той или иной группе, P а какие — к другой F , вместо того чтобы предполагать крайние значения .

Практическое применение типологии

Данная классификация позволяет создать практический аналитический контрольный список:

1. **Явное разделение** – для любой предлагаемой модели или реформы необходимо указать, какие аспекты считаются предопределенными, а какие – регулируемые.
2. **Требование к обоснованию** – Укажите причины, по которым каждый элемент отнесен

к Ртой или иной категории F .

3. **Проверка домена** – подтвердите, что раздел соответствует правильному уровню описания.
4. **Проверка на наличие ошибок** – протестируйте предложенное решение на соответствие трем указанным выше типам ошибок.

Благодаря явному обозначению этих различий, данная концепция преобразует абстрактное понимание двойного предварительного обусловливания в конкретный методологический инструмент.

В заключение, признание того, что начальные условия содержат как predetermined, так и неопределенные компоненты, позволяет найти дисциплинированный срединный путь между двумя симметричными иллюзиями: иллюзией полной нейтральности и иллюзией полной неизменности. Избегание этих ошибок, особенно их смешения в разных областях, имеет важное значение для

согласованного моделирования, ответственного проектирования и реалистичной интерпретации любого операционного процесса.

Методология различения предопределенных и свободных компонентов

Полезность разработанной в разделах 1.6 и 1.7 структуры зависит от практического вопроса: как аналитик, имея реальную систему или модель, может надежно определить, какие элементы относятся к заранее определенному подмножеству P , а какие — к относительно свободному подмножеству F ? Без такого метода различие рискует остаться чисто риторическим. В этом разделе предлагается структурированная процедура для проведения такого определения. (Полный пошаговый протокол аудита, подходящий для практического применения, см. в Приложении [12.](#))

1.8.1 Операционный критерий необходимости

Основным критерием для классификации элемента как предопределенного является *критерий*

необходимости описания :

Проверка на необходимость: элемент x относится к категории, P Если его удаление или изменение x делает процесс неопиcуемым, неработоспособным или несогласованным в рамках выбранной модели.

Формально:

$$x \in P$$

\Leftrightarrow Description without x is undefined or non-functional

Наоборот:

$x \in F \Leftrightarrow$ Description remains coherent when x is varied

Этот тест основывает различие на практике представления информации, а не на интуиции.

1.8.2 Тест на подстановку

Вторым практическим инструментом является *тест на замещение* .

Для любого элемента-кандидата x :

- Если альтернативные значения x могут быть подставлены без нарушения операционной целостности модели, то x принадлежит F .

- Если замена невозможна без нарушения согласованности модели, то χ принадлежит P .

Пример:

- В нейронной сети точное начальное значение случайной переменной, используемое для инициализации, обычно равно F ;
- Наличие весов, слоев и правила обновления означает следующее P : .

1.8.3 Тест на устойчивость контрфактуальных данных

Еще одним диагностическим методом является *тест на устойчивость к контрфактическим сценариям* :

Спросите: «Можно ли будет осмысленно описать процесс в его нынешнем виде, если этот элемент будет отсутствовать?»

- Если ответ «нет» , то элемент предопределен.
- Если ответ «да» , то это свободный параметр.

Этот тест намеренно переключает внимание с метафизической необходимости на описательную

необходимость.

1.8.4 Послойное разложение ограничений

Во многих системах заранее определенные компоненты функционируют на нескольких уровнях. Поэтому мы предлагаем иерархическую декомпозицию:

1. **Формальные ограничения** — логические или математические структуры, необходимые для любого представления.
2. **Архитектурные ограничения** – структурные особенности выбранного инструмента моделирования.
3. **Эмпирические ограничения** – унаследованные особенности целевой системы.
4. **Условные параметры** – настраиваемые или свободно задаваемые значения.

Первые три слоя обычно составляют P ; четвертый составляет F .

Такой многоуровневый подход позволяет избежать

распространенной ошибки, заключающейся в том, что эмпирические ограничения рассматриваются как формальные требования.

1.8.5 Дерево решений для классификации

Для прикладного анализа можно использовать следующее дерево решений:

1. Необходимо ли вообще указывать конкретную систему с помощью этого элемента? – Да $\rightarrow P$
2. Это обусловлено языком моделирования или средой разработки? – Да $\rightarrow P$
3. Он наследуется от целевой системы и не находится под контролем аналитика? – Да $\rightarrow P$
4. Можно ли внести изменения, не нарушая целостности? – Да $\rightarrow F$

Эта процедура преобразует абстрактное различие в воспроизводимую аналитическую процедуру.

1.8.6 Распространенные эвристические методы

Для классификации помогают несколько практических эвристических методов:

- **Эвристика неизменности:** то, что нельзя изменить в течение анализируемого P временного масштаба, как правило, остается неизменным.
- **Эвристика заменяемости:** то, что можно заменить альтернативными ценностями или конструкциями, как правило, является F ...
- **Эвристический подход к инфраструктуре:** фоновые условия, способствующие развитию инфраструктуры, обычно относятся к P .
- **Эвристика настройки:** регулируемые гиперпараметры обычно относятся к F .

Эти эвристические методы несовершенны, но они представляют собой полезную отправную точку.

1.8.7 Иллюстративные примеры применения

Хотя данная статья в основном посвящена методологии, краткие иллюстрации поясняют используемый метод:

Машинное обучение:

- Функция потерь, архитектура, алгоритм обучения $\rightarrow P$
- Гиперпараметры, выборка набора данных, начальные значения генератора случайных чисел $\rightarrow F$

Когнитивное развитие:

- Нейронная архитектура, сенсорные модальности $\rightarrow P$
- Конкретный опыт и усвоенный материал $\rightarrow F$

Институциональный дизайн:

- Правовые рамки, физическая инфраструктура $\rightarrow P$
- Варианты политики, административные правила $\rightarrow F$

Эти примеры демонстрируют, что один и тот же объект может находиться P как в одной, так и F в другой области, что подтверждает относительный характер различия между областями.

1.8.8 Краткое изложение процедуры

методологию различения P можно F кратко описать следующим образом:

1. Четко укажите описательную структуру.
2. Перечислите все элементы, предполагаемые данной концептуальной основой.
3. Примените критерии необходимости, замещения и контрфактуального анализа.
4. Присвойте элементы соответствующим образом F .
5. Повторно проверьте наличие смешения доменов, как описано в разделе 1.7.

Только после завершения этой процедуры можно делать обоснованные выводы о «степенях свободы» или «начальных условиях». Более подробный десятиэтапный протокол со стандартизированным шаблоном контрольной таблицы представлен в Приложении [12](#).

1.8.9 Методологическая выгода

Данная методология преобразует абстрактное

понимание двойного предварительного обусловливания в практический опыт. Она предотвращает обе симметричные ошибки, выявленные ранее:

- рассматривать фиксированные ограничения как необязательные, и
- рассматривать регулируемые компоненты так, как если бы они были фиксированными.

Заставляя аналитиков обосновывать каждую классификацию, она заменяет расплывчатые ссылки на «начала» четкими описаниями того, что именно является неизменным, а что — неопределенным.

Заключение раздела 1.8

Различие между предопределенными и свободными компонентами — это не вопрос метафизических спекуляций, а вопрос дисциплинированного репрезентативного анализа. Описанные выше инструменты предоставляют систематический способ сделать это различие явным, воспроизводимым и обоснованным в любой описываемой области.

Литературный обзор

В этом разделе рассматривается литература, относящаяся к формальной структуре. Обсуждаемые здесь эмпирические данные иллюстрируют применимость данной структуры; они не приводятся в качестве предпосылок или доказательств основной теоремы.

Эмпиризм и нативизм

Дискуссия о врожденном и приобретенном знании — одна из старейших в философии. В *«О душе»* Аристотель описывал интеллект как «письменную доску, на которой ничего не написано». Стоики развили эту точку зрения, и Аэций сообщал, что для них разум при рождении «подобен листу бумаги, готовому к письму». В *«Очерке о человеческом понимании»* Локк систематизировал эту эмпирическую позицию, утверждая, что разум при рождении лишен врожденных идей и что все знания проистекают из ощущений и размышлений.

Рационалистическая контртрадиция, начиная с

Платона и заканчивая Декартом и Лейбницем, утверждала, что определенные идеи или структуры должны быть врожденными. В своих «*Новых очерках о человеческом понимании*» Лейбниц предложил подробную критику Локка, утверждая, что разум — это не чистая доска, а скорее глыба мрамора, прожилки которой уже обозначают фигуру, которую предстоит вылепить.

Современный нативизм, примером которого служит «Универсальная грамматика» Хомского, основывает свои утверждения на генетике и нейробиологии, а не на метафизике. Аргумент «бедности стимула» утверждает, что усвоение языка нельзя объяснить одними лишь общими механизмами обучения — дети должны привнести в процесс обучения предварительную структуру. Этот аргумент был формализован и широко обсуждался.

Тезис Джерри Фодора о модульности расширил нативистские рассуждения за пределы языка, предположив, что разум состоит из специализированных модулей для распознавания

лиц, пространственного мышления, восприятия чисел и теории разума. В последующих работах обсуждались масштабы и природа модульности.

Современный консенсус отвергает как чистый нативизм, так и чистый эмпиризм: сложные черты формируются в результате динамического взаимодействия генетического наследия и факторов окружающей среды. Эта «интеракционистская» или «системная» перспектива разрушает дихотомию «природа-воспитание», подчеркивая, что развитие всегда включает в себя как унаследованные, так и экологические ресурсы. Примечательно, что это эмпирическое сближение иллюстрирует (но не доказывает) наше формальное утверждение о требованиях описываемости.

Динамические системы и зависимость от предшествующего пути.

Математическое изучение динамических систем предоставляет формальные инструменты для анализа того, как начальные условия распространяются во времени. Фундаментальная работа Пуанкаре по

задаче трех тел показала, что детерминированные системы могут проявлять крайнюю чувствительность к начальным условиям. Открытие Лоренцем детерминированного хаоса в моделях атмосферы продемонстрировало, что эта чувствительность — «эффект бабочки» — не является просто теоретической, но имеет практические последствия для прогнозирования.

Философские последствия теории хаоса были всесторонне проанализированы. Келлерт утверждает, что хаос принципиально бросает вызов традиционным представлениям о детерминизме и предсказуемости. Смит исследует эпистемологические ограничения, которые хаос накладывает на научное знание.

В экономике и институциональной теории зависимость от предшествующего развития распространяет эти идеи на социальные системы. Влиятельный анализ клавиатуры QWERTY, проведенный Полом Дэвидом, показал, что стандарты могут закрепляться скорее в результате исторических

случайностей, чем из-за оптимальности. Брайан Артур формализовал механизмы возрастающей отдачи, порождающие зависимость от предшествующего развития. Дуглас Норт применил эти концепции к институциональной экономике, показав, как первоначальный институциональный выбор ограничивает последующее развитие.

Концепция была уточнена в ходе дискуссий о «закреплении» и разграничении различных форм зависимости от предшествующего пути развития. Махони предлагает систематическую типологию объяснений, зависящих от предшествующего пути развития, в исторической социологии.

Недавние экспериментальные работы Миттоне , Морреале и Риталы предоставляют прямые доказательства того, что начальные условия формируют траектории обучения в задачах исследования-эксплуатации.

Индуктивная предвзятость в машинном обучении

Теория машинного обучения обеспечивает особенно

четкую формализацию вопросов, связанных с начальными ограничениями. Фундаментальное определение Митчелла установило, что индуктивная предвзятость включает в себя любые предположения, выходящие за рамки обучающих данных, которые модель использует для генерации прогнозов.

Теоремы «бесплатного обеда не бывает», доказанные Вольпертом и Макреди, формально устанавливают, что ни один алгоритм обучения не может превзойти все остальные во всех возможных задачах. Этот результат демонстрирует, что индуктивная предвзятость не просто полезна, но и логически неизбежна: без предположений о структуре целевой области обучение невозможно.

Геман, Биененсток и Дурса проанализировали компромисс между смещением и дисперсией, показав, как различные формы индуктивного смещения влияют на обобщение. Эта работа установила, что выбор смещения не является произвольным, а должен соответствовать структуре задачи.

Архитектуры нейронных сетей кодируют индуктивные смещения посредством выбора их структуры. Сверточные сети предполагают пространственную локальность и инвариантность к сдвигу. Рекуррентные сети предполагают последовательные зависимости. Трансформеры предполагают, что внимание к парным отношениям является информационно достаточным. Батталья и др. предлагают всесторонний анализ реляционных индуктивных смещений в глубоком обучении.

Лейк, Ульман, Тенебаум и Гершман утверждают, что для обучения, подобного человеческому, требуются более сильные индуктивные искажения, чем те, которые обычно используются в современных системах глубокого обучения. Бенджио и его коллеги исследуют, какие индуктивные искажения могут быть необходимы для обучения когнитивным навыкам более высокого уровня .

Смещение в больших языковых моделях

Появление больших языковых моделей сделало вопросы унаследованной структуры практически

актуальными. Анализ «стохастических попугаев», проведенный Бендером, Гебру, Макмиллан-Мейджор и Шмитчеллом, показал, как смещения обучающих данных распространяются через выходные данные модели. Блоджетт, Барокас, Дауме и Уоллах представили критический обзор того, как понятие «смещение» концептуализируется в исследованиях в области обработки естественного языка.

Галлегос и др. представляют всесторонний обзор, различающий внутреннюю предвзятость (закодированную в представлениях) и внешнюю предвзятость (проявляющуюся в последующих задачах). Навильи и др. показывают, что предвзятость возникает уже на этапе построения корпуса, благодаря особенностям выбора данных.

Боммасани и др., подчеркивает, что эти системы наследуют ограничения на нескольких уровнях: от архитектуры, процедур обучения, выбора данных и процессов тонкой настройки. Вайдингер и др. классифицируют этические и социальные риски, возникающие из-за этих унаследованных структур.

Воплощенное и ситуативное познание

Движение за воплощенное познание бросает вызов классическому взгляду на разум как на бестелесную манипуляцию символами. В книге Варелы, Томпсона и Роша « *Воплощенный разум* » был введен энактивизм , утверждающий, что познание возникает посредством сенсомоторной связи с окружающей средой. Работа Кларка о расширенном познании предполагает, что когнитивные процессы могут выходить за пределы мозга и распространяться на тело и окружающую среду.

Галлахер развил энактивистский подход дальше, подчеркивая, как телесные возможности формируют когнитивные процессы. Томпсон представил всестороннее философское исследование взаимосвязи разума, тела и мира .

В своей недавней работе Барретт и Стаут исследуют воплощенное познание в свете искусственного интеллекта, утверждая, что современным системам

ИИ не хватает той воплощенной основы, которая характерна для биологического познания. Радикальная когнитивная наука, основанная на телесном восприятии, разработанная компанией Chemero, развивает этот подход дальше, полностью отвергая ментальные представления.

Экологическая психология Гибсона, с ее акцентом на возможностях взаимодействия, предлагает дополнительную перспективу на то, как структура окружающей среды формирует восприятие и действия. Эта работа оказала влияние на робототехнику и ситуативный искусственный интеллект.

В наших целях воплощенное познание предоставляет доказательство существования: биологическое познание — это область, где описание когнитивных процессов требует ссылки на ненулевые исходные структуры — телесные, средовые и структуры развития.

Философия моделей и представлений

Акцент данной концепции на ограничениях среды связан с обширными работами в философии науки о природе моделей и представлений. Семантический взгляд Ги́ра на теории подчеркивает, что модели — это абстрактные структуры, представляющие целевые системы. Суарес анализирует умозаключительную концепцию представления.

Вайсберг различает разные типы моделей — конкретные, математические и вычислительные — каждая из которых имеет различные ограничения представления. Фригг и Хартманн дают всесторонний обзор моделей в науке.

Модельный подход к научным теориям поднимает вопросы о взаимосвязи между моделями и реальностью, которые параллельны нашему разграничению между репрезентативными и операциональными ограничениями. Работа Картрайта о «пятнистом мире» подчеркивает локальный и ограниченный характер научных моделей.

Флориди предоставляет основу для понимания

ограничений, которые информационные структуры накладывают на представление и обработку информации.

Основы математики

Рассмотрение в статье математической пустоты связано с фундаментальными дискуссиями о природе математических объектов. Стандартная теоретико-множественная основа строит всю математику из пустого множества посредством аксиом Цермело - Френкеля. Однако эта конструкция, как мы отмечаем, предполагает непустую аксиоматическую структуру.

Были предложены альтернативные основания. Теория категорий предлагает структуралистскую альтернативу, в которой математические объекты характеризуются своими взаимоотношениями, а не своей внутренней природой. Теория гомотопических типов предоставляет еще одну фундаментальную основу с различными исходными предположениями.

В своих философских работах Мэдди исследует обоснование теоретико-множественных аксиом.

Структурализм Шапиро в философии математики анализирует, что значит для математики изучать структуры.

Эти основополагающие альтернативы иллюстрируют нашу точку зрения о том, что математическая «пустота» всегда относительна по отношению к определенной системе координат. Пустое множество — это не *просто ничто*, а пустое множество *в рамках* конкретной основополагающей системы.

Формальная структура

Операциональное определение системы

Нам необходимо определение, не предполагающее какой-либо структуры, чтобы избежать замкнутого круга. Мы различаем «систему как онтологическую сущность» и «систему как описываемый процесс».

Нас интересует последний вариант:

ОПЕРАЦИОННОЕ ОПРЕДЕЛЕНИЕ

В контексте данной работы системой считается всё, что можно описать как находящееся в процессе перехода состояний .

Это определение является минимальным: оно не предполагает какой-либо структуры априори.

Важно отметить, что структура в определении не предполагается; она выводится из минимальных требований к описанию переходов. Из этого определения мы выводим (а не предполагаем), что описание систем, претерпевающих изменения, требует ненулевой структуры. Вывод осуществляется через требования самого описания.

Вывод из требований описания

Аргументация строится следующим образом:

Шаг 1: Описание функциональных переходов требует наличия как минимум минимального числа дифференцируемых состояний. Для описания изменений необходимо различать состояние до и состояние после.

Шаг 2: Описание дифференцируемых состояний требует наличия основания для дифференциации. Различие без основания не является различием.

Шаг 3: Основой для дифференциации является

непустая структура.

Вывод: описание процесса как претерпевающего изменения предполагает наличие непустой области дифференцируемых состояний.

Данный вывод носит эпистемологический характер: он касается требований к описанию, а не предшествующих метафизических представлений о природе систем.

Принцип ненулевой инициализации

Принцип 1 (Ненулевая инициализация — условная форма). *Если процесс описывается как эволюционирующий, то его описание предполагает непустую область состояний.*

Логическая форма: *Описываемость эволюции \rightarrow в непустой области определения*

Это следует из требований описания, а не из метафизических предположений.

Это утверждение касается условий описания и представления, а не внутренней природы реальности.

Аргумент в поддержку:

- *Предпосылка 1:* Переход описывается как изменение из одного состояния в другое.
- *Предпосылка 2:* Для описания «одного государства» и «другого государства» необходимо наличие как минимум двух различных элементов.
- *Посылка 3:* Пустая область определения не содержит элементов, которые можно было бы различать.
- *Вывод:* для описания переходов необходима непустая область определения.

Примечание: Это не утверждает, что пустые области логически невозможны в каком-либо абсолютном смысле или что пустоту нельзя представить. Это утверждает, что описание переходов требует наличия чего-то, между чем эти переходы должны происходить.

Принцип ограничения средней величины

Принцип 2 (среднее ограничение). *Моделируемая*

среда ограничивает то, что может быть представлено, до любого конкретного акта представления.

Это утверждение касается выразительности, а не онтологии. Оно касается репрезентативных ограничений, а не метафизических.

Это утверждение касается условий описания и представления, а не внутренней природы реальности.

Аргумент:

- *Предпосылка 1:* Моделирование требует средства представления, обладающего синтаксисом и семантикой.
- *Предпосылка 2:* Синтаксис определяет, какие выражения являются корректными.
- *Вывод:* Возможные модели ограничиваются средой, прежде чем будут построены конкретные модели.

Двойная предварительная подготовка

Теорема 1 (двойное предварительное

обусловливание). Для любой системы S , которая одновременно (1) может быть описана как претерпевающая переходы и (2) может быть представлена в некоторой модели:

Начальное состояние ограничено обоими факторами:

- **Внутренние требования:** описание требует наличия дифференцируемых состояний (Принцип [1](#))
- **Внешние требования:** представление требует выразительной структуры (Принцип [2](#))

Пространство возможных начальных состояний представляет собой пересечение двух непустых пространств ограничений.

Доказательство. Пусть S описываемая функция может быть описана как претерпевающая переходы, с предварительными условиями R из Принципа [1](#). Пусть M Пусть будет модель с ограничениями, Связываемыми из Принципа [2](#). Представление

Сначального состояния в M ограничено как P , так и C . Поскольку $P \neq \emptyset$ и $C \neq \emptyset$, пространство представимых начальных состояний является дважды ограниченным. \square

Что это устанавливает и чего это не устанавливает

Это позволяет сделать вывод: для систем, которые мы можем описывать, моделировать и анализировать, состояния с нулевой структурой не могут служить отправными точками для описываемых процессов.

Это **НЕ** устанавливает:

- Эта метафизическая пустота невозможна в каком-то абсолютном смысле.
- То, что состояния с нулевой структурой невозможно представить или изобразить.
- Всё, что касается доменов, выходящее за рамки описания.

Полученный результат является сильным в своей области — охватывая науку, математику и формальное мышление, — но при этом имеет четкие ограничения.

Возражение о самоинициализации

Сложное возражение: возможно, структуры могут возникать в процессе, а не существовать до него. Мы тщательно изучим этот вопрос.

Возражение

Возможно, существуют процессы, в которых структуры, необходимые для перехода, возникают как часть самого процесса. Если это так, то структура не обязательно должна существовать до начала процесса — она может быть самогенерирующейся.

Это серьезное возражение. Мы рассматриваем конкретных кандидатов.

Клеточные автоматы

Утверждение: Сложные закономерности возникают из простых правил.

Анализ: Для описания клеточных автоматов необходимы: (а) сеточная структура, (б) пространство состояний, (в) правило перехода, (г) начальная конфигурация. Возникновение паттернов происходит в рамках этой уже существующей структуры.

Вердикт: Возникновение в рамках структурированной системы, а не возникновение системы из ничего. Соответствует нашей теореме.

Нарушение симметрии

Утверждение: Дифференцированные состояния возникают из симметричных начальных условий.

Анализ: Описание симметричного начального состояния наделяет его структурой (симметрией). Описание динамики требует потенциального ландшафта с нетривиальной структурой.

Вердикт: трансформация одного типа структуры в другой, а не создание из отсутствия структуры. Соответствует нашей теореме.

Эволюционное возникновение

Утверждение: В процессе эволюции создаются действительно новые структуры.

Анализ: Описание эволюции предполагает наличие репликаторов, ландшафта приспособленности, химических процессов. Это существенные априорные структуры в любом описании эволюционных

процессов.

Вердикт: Структурная трансформация, а не самосозидание из ничего. Соответствует нашей теореме.

Оценка

В каждом рассмотренном случае кажущееся «возникновение из ничего» при анализе оказывается возникновением в рамках метаструктуры, которая обеспечивает описательную основу для этого явления.

БЛАГОДАРНОСТИ И ОГРАНИЧЕНИЯ

Ни одна из существующих моделей не демонстрирует самоинициализацию без предварительного предположения о наличии метаструктуры для описания.

Однако: процессы самовосстановления остаются открытым фундаментальным вопросом. Мы не утверждаем, что доказали невозможность в принципе — лишь то, что ни одна из существующих описываемых моделей этого не делает.

Таким образом, данная теорема является условной и зависит от существующих концептуальных основ; будущие концептуальные разработки могут потребовать пересмотра.

Что могло бы поставить под сомнение теорему?

Для решения подлинной проблемы потребуется: (а) поддающийся описанию подлинный переход, (б) появление концептуальной основы для перехода как часть процесса, (в) отсутствие какой-либо заранее предполагаемой мета-концептуальной основы для описания.

Мы утверждаем, что в рамках существующих концепций подобный процесс не описан. Мы не утверждаем, что такой процесс невозможен в каком-либо абсолютном смысле для описываемых процессов.

Математическая пустота

ВАЖНОЕ ПРЕДУПРЕЖДЕНИЕ

Этот аргумент не оспаривает легитимность пустого множества как формального объекта.

Математическая пустота \neq Операциональная пустота

Мы не отвергаем стандартную теорию множеств и не утверждаем, что математика несовершенна.

Два чувства пустоты

Репрезентативная пустота: формальное представление отсутствия в символической системе. « \emptyset » обозначает множество, не имеющее элементов. Это пустота как объект исследования.

Операциональная пустота: отсутствие всякой структуры, включая отсутствие какой-либо основы для представления. Это была бы пустота как условие, а не как концепция.

Наше утверждение: «ничто» можно представить, но оно не может служить начальным условием для описываемых процессов. В этом разница между представлением пустоты и началом с пустоты.

Пустое множество в формальных системах

Математика показывает, что представления «ничего» встречаются в структурированных формальных системах. Пустое множество существует в рамках

ZFC (или альтернативных систем), обладающих непустой аксиоматической структурой: аксиома пустого множества, аксиома спаривания, правила вывода, язык теории множеств.

Пустое множество является элементом структурированной теории; оно не является дотеоретической отправной точкой для описываемых операций.

Пустое множество обладает свойствами: является подмножеством любого множества, имеет нулевую мощность и единственность. Это структурные особенности теории множеств. Математическая пустота всегда относительна по отношению к более богатой системе.

Точка

Пустое множество можно представить именно потому, что имеется непустой формальный аппарат. Представление «ничего» само по себе является чем-то — символом, понятием, объектом со свойствами.

Это скорее иллюстрирует, чем опровергает наше

утверждение: для описания представлений о пустоте необходимы непустые репрезентативные ресурсы.

Иллюстрации из области искусственного интеллекта

В этом разделе показано, как данная структура применяется к искусственному интеллекту. Приведены примеры применения общей теоремы, а не независимые доказательства или подтверждения её применимости.

Индуктивное смещение

Индуктивная предвзятость иллюстрирует Принцип [1](#). Предвзятость — это априорные предположения, определяющие пространство гипотез. Теоремы «бесплатного обеда не бывает» доказывают, что предвзятость неизбежна для любого описываемого обучающегося: ни один обучающийся не превосходит всех остальных во всех задачах.

У обучающегося, не имеющего предвзятости, не будет оснований отдавать предпочтение какой-либо гипотезе. Требование индуктивной предвзятости для

обучения иллюстрирует общее требование непустой структуры для описываемой операции.

Архитектурные ограничения

Нейронные архитектуры кодируют структурные предположения: сверточные сети предполагают пространственную локальность; рекуррентные сети предполагают последовательные зависимости; трансформеры предполагают, что достаточно механизма внимания.

Это теоретические положения, а не просто детали реализации. Архитектура обеспечивает структуру, делающую возможным описываемое обучение.

Большие языковые модели

LLM-модели иллюстрируют множественные уровни наследования:

Архитектурный подход: разработанные человеком решения по трансформации, касающиеся внимания, глубины и активации .

Данные: Созданный человеком обучающий текст, отражающий человеческое мышление и культуру.

Цель: Заданные человеком функции потерь, определяющие успех.

Всё, что порождает что-либо, накладывает ограничения на то, что порождается — закономерность, согласующаяся с нашей общей теоремой.

Подразумеваемое

Искусственный интеллект создает новизну, ограниченную унаследованными структурами. Это сугубо практическое наблюдение о том, как функционируют эти системы, а не философское утверждение о «истинной креативности» или конечной природе машинного познания.

Понятие «непредвзятый ИИ» в том виде, в котором оно сформулировано, не является последовательным. Вопрос в том, какие предвзятости уместны для каких целей.

Возражения и граничные условия

Возражение о тривиальности

Возражение: Утверждение носит тавтологический

характер.

Ответ: Мы не говорим: «Системы нуждаются в структуре, потому что они являются системами». Мы говорим: «Описание изменений требует дифференцируемых состояний». Это вытекает из требований к описанию, а не продиктовано определением. Мы начали с минимального определения, не предполагающего наличия структуры; структура была выведена из требований к описанию переходов.

Возражение о превышении полномочий

Возражение: В статье содержится слишком много утверждений обо всех системах.

Ответ: Мы четко сформулировали условия области применения. Утверждение касается описываемых, операциональных систем. Мы не делаем никаких заявлений о метафизическом ничто или областях, выходящих за рамки описания. Каждое использование слова «невозможно» в данной статье имеет оговорку: невозможно как операциональная

отправная точка для описываемых процессов.

Возможные контрмодели

Для создания подлинной контрмодели необходимы:

(а) описываемый подлинный переход, (б) возникновение структуры как часть процесса, (в) отсутствие заранее предполагаемой метаструктуры для описания.

На это могут указывать некоторые спекулятивные предложения: определенные топос-теоретические основы, некоторые интерпретации квантовой гравитации, процессуальные философии. Мы не утверждаем, что они не могут быть успешными. Мы утверждаем, что ни одно из них еще не продемонстрировало необходимые характеристики, сохраняя при этом описываемость.

Осмысленные приближения

Наши принципы касаются точной нулевой структуры как оперативных отправных точек, а не приближения:

Относительная нулевая структура: система может не иметь ограничений одного типа, но иметь

ограничения других типов. Термин «необученная модель» имеет смысл.

Описания, зависящие от уровня: При выбранных уровнях описания начальные условия могут быть незначительными.

Методологическая абстракция: предварительное рассмотрение ограничений как отсутствующих в исследовательских целях.

Практическое применение

Разработанная в данной статье структура имеет прямое практическое применение в различных областях. Методология различения R компонентов F (раздел 1.8) в сочетании с типологией ошибок (раздел 1.7) предоставляет конкретные инструменты для анализа и проектирования. Исторические примеры каждого типа ошибок см. в Приложении [11](#); систематический протокол аудита см. в Приложении [12](#).

Машинное обучение

Четко обозначьте индуктивные смещения.

Сопоставьте инициализацию с известными свойствами предметной области. Задокументируйте унаследованные ограничения: архитектура, инициализация, данные, целевая функция, оценка.

Когнитивная наука

Данная концепция поддерживает телесно-ориентированные подходы, предоставляя формальное обоснование утверждения о том, что описание познания требует ссылки на телесную и средовую составляющие.

Научное моделирование

Выбор модели — это содержательное теоретическое обоснование того, какие аспекты реальности следует отразить. Рассматривайте выбор модели как требующий обоснования. Признайте, что ограничения модели могут отражать пределы средних значений.

Институциональные изменения

Понятия «начать все заново» и «строить с нуля» вводят в заблуждение. Реформаторы

перераспределяют ограничения, а не устраняют их. Вопрос не в том, «как начать все заново», а в том, «какие унаследованные структуры следует преобразовать».

Заключение

Мы установили условную теорему: если нечто можно описать как претерпевающее изменения, то его описание предполагает непустую структуру. Это следует из требований описываемости, а не из метафизических условий.

Принцип [1](#) (ненулевая инициализация): описание переходов требует непустых областей определения. Принцип [2](#) (Среднее ограничение): представление требует выразительных структур. Теорема [1](#) (Двойное предварительное обусловливание): описываемые и представимые системы имеют начальные состояния с двойным ограничением.

Мы исследовали кандидатов на самоинициализацию — клеточные автоматы, нарушение симметрии, эволюцию — показав, что каждый из них

предполагает наличие метаструктуры в любом описании. Мы различали репрезентативную пустоту (работу с \emptyset) и операциональную пустоту (начало с нулевой структуры как отправной точки для описываемых процессов).

Данная концепция применима к областям, поддающимся научному изучению и формальному анализу. Она не рассматривает вопросы о метафизическом происхождении или областях, выходящих за рамки описания. Практическая полезность концепции демонстрируется на примере типологии ошибок (раздел 1.7), методологии различения predetermined и свободных компонентов (раздел 1.8), обширных иллюстраций в различных областях (приложение [11](#)) и протокола систематического аудита (приложение [12](#)).

Вывод, сформулированный точно: в рамках описываемых областей каждая операциональная отправная точка предполагает наличие ненулевых ограничений.

Этот вывод следует из логики описания, а не из

какого-либо метафизического тезиса о природе бытия.

ЗАКЛЮЧИТЕЛЬНОЕ ОПИСАНИЕ ОБЪЕМА РАБОТ

В данной статье не рассматривается вопрос о том, что может быть или не быть возможным за пределами описываемых областей.

Наши утверждения представляют собой теоремы, ограниченные определенной областью применения, касающиеся требований описываемости и функционирования, а не универсальные метафизические необходимости.

Результат зависит от существующих концептуальных рамок описания; будущие концептуальные разработки могут потребовать пересмотра.

Иллюстративные примеры методологических ошибок в различных областях.

В этом приложении приведены конкретные примеры трех типов ошибок, выявленных в разделе 1.7, взятые из истории, науки, политики, психологии, экономики,

технологий и других областей человеческой деятельности. Эти примеры приводятся в качестве иллюстраций, а не в качестве доказательства формальной структуры. Их цель — продемонстрировать практическую значимость различий, разработанных в данной работе.

Ошибка I типа: Неправильное приписывание свободы (ложная нейтральность)

Эта ошибка возникает, когда predetermined ограничения (P) рассматриваются так, как если бы они были свободно изменяемыми (F).

История и политика

- **Первый год Французской революции (1792):** Революционный календарь пытался обнулить само время, рассматривая историческую и культурную преемственность как нечто необратимое. Предetermined вес религиозных, сельскохозяйственных и социальных традиций (P) рассматривался как свободно заменяемый (F). Календарь был

отменен в течение 14 лет.

- **Советская идеология «нового человека»:** предположение о том, что человеческую природу можно полностью изменить посредством социальной инженерии, игнорируя биологические и психологические ограничения, унаследованные от эволюции.
- **«Нулевой год» красных кхмеров (1975):** Крайне радикальная попытка рассматривать всю существовавшую ранее социальную структуру как подлежащую ликвидации, что привело к катастрофическим последствиям, когда вновь проявились заранее определенные экономические, образовательные и инфраструктурные потребности.
- **Постколониальное национальное строительство:** границы, установленные без учета этнических, языковых и исторических ограничений, рассматривающие географию и демографию как свободно определяемые параметры.

Наука и медицина

- **Радикальный бихевиоризм (Уотсон, Скиннер):** утверждение, что любого ребенка можно обучить стать «специалистом любого типа» независимо от врожденных предрасположенностей, рассматривая генетические и неврологические ограничения как незначительные.
- **Теории усвоения языка, основанные на принципе «чистого листа»:** предположения, существовавшие до Хомского, о том, что изучение языка не требует врожденной грамматической структуры, игнорирующие бедность стимула.
- **Ранний оптимизм в отношении генной терапии:** предположения о том, что генетические заболевания можно «просто» исправить путем введения соответствующих генов, недооценка predetermined сложности сетей регуляции генов.

- **Универсальные методы лечения:**
игнорирование фармакогеномных различий на основе предположения, что все пациенты одинаково реагируют на стандартизированные дозы.

Технологии и искусственный интеллект

- **Заявления об «непредвзятом ИИ»:**
маркетинговая формулировка, предполагающая, что алгоритмы могут быть свободны от предвзятости, несмотря на заранее определенные архитектурные решения, распределение обучающих данных и целевые функции.
- **Мифы о нейтральности платформ :**
Утверждения о том, что платформы социальных сетей являются нейтральными каналами связи, игнорирующие то, как алгоритмическое ранжирование, дизайн интерфейса и бизнес-модели (*P*) формируют дискурс.

- **Переписывание программного обеспечения «с нуля»:** распространенное убеждение, что устаревшие системы можно заменить чистой переработкой, недооценивающее заложенные институциональные знания и ограничения интеграции.

Психология и семейные отношения

- **Иллюзии «нового начала» в отношениях:** убеждение, что переезд в новый город или начало новых отношений устраняют психологические модели поведения, стили привязанности и привычки, сформировавшиеся за десятилетия.
- **Воспитание детей на основе предположений *tabula rasa*:** отношение к детям как к бесконечно податливым существам, игнорирование темперамента, генетической предрасположенности и пренатальных влияний.
- **Игнорирование травмы, передающейся из**

поколения в поколение: предположение, что семьи могут просто «пережить» историческую травму, не обращая внимания на унаследованные психологические и межличностные модели поведения.

- **Мифы о «чистом разрыве» при разводе:** предположение, что расторжение брака устраняет все прежние взаимоотношения, что особенно проблематично, когда есть дети.

Экономика и политика

- **Шоковая терапия в рамках экономических реформ:** программы быстрой приватизации, предполагающие, что рынки могут функционировать без существовавшей ранее правовой, институциональной и культурной инфраструктуры.
- **Аисторические модели развития:** применение экономических моделей из одного контекста к другому без учета различий в институциональном наследии.

- **Политика «просто переобучение работников»:** предположение, что уволенные работники могут свободно перейти в новые отрасли, игнорируя возраст, местоположение, социальные связи и ограничения, связанные со специализацией навыков.

Ошибка типа II: Неправильное определение (ложная фиксированность)

Эта ошибка возникает, когда действительно изменяемые элементы (F) рассматриваются так, как если бы они были предопределены (P).

История и политика

- **Исторический детерминизм:** утверждает, что исторические результаты были «неизбежны», игнорируя случайные решения и альтернативные возможности, которые действительно существовали.
- **Тезис о «конце истории»:** предположение, что либеральная демократия представляет собой фиксированную конечную точку, рассматривая

политическую эволюцию как завершенную, а не непрерывную.

- **Расовый и этнический эссенциализм:** рассмотрение культурных практик как биологически фиксированных, а не как изменяемых социальных конструкций.
- **Обоснование кастовой системы:** представление социальных иерархий как естественных и неизменных, а не как исторически обусловленных институтов.

Наука и медицина

- **Генетический детерминизм:** преувеличение неизменности генетического влияния при игнорировании эпигенетики, взаимодействия генов и окружающей среды, а также пластичности развития.
- **Редукционизм «химического дисбаланса»:** рассмотрение психических расстройств как чисто биологических, игнорируя поддающиеся изменению психологические, социальные и

поведенческие факторы.

- **Наследственный подход к IQ:** рассмотрение интеллекта как полностью predetermined генетикой, игнорирующее образовательные мероприятия, питание и улучшение окружающей среды.
- **Неврологический фатализм:** устаревшее представление о том, что мозг взрослого человека не может изменяться, существовавшее до открытия нейропластичности.

Технологии и искусственный интеллект

- **Технологический детерминизм:** утверждает, что технология следует неизбежной траектории, игнорируя проектные решения, варианты регулирования и альтернативные пути развития.
- **Заблуждение «алгоритмы объективны»:** рассматривать результаты работы алгоритмов как незыблемые истины, а не как результат изменяемых проектных решений.

- **Фатализм в отношении устаревших систем:** предположение о невозможности модификации устаревших систем, когда на самом деле возможны постепенные улучшения.

Психология и семейные отношения

- **Убеждение «Люди никогда не меняются»:** отношение к личности и поведению как к абсолютно неизменным, игнорирование доказательств развития личности во взрослом возрасте и терапевтических изменений.
- **Фиксированное мышление (Двек):** убеждение, что способности являются врожденными и неизменными, а не развиваются посредством усилий.
- **Обобщение выученной беспомощности:** распространение подлинных ограничений на области, где действительно существует субъектность .
- **«Такова уж наша семья»:** рассматривать дисфункциональные семейные модели

поведения как незыблемые традиции, а не как изменяемые модели поведения.

Экономика и политика

- **TINA («Нет альтернативы»):** Представление конкретных экономических мер как единственно возможного варианта, исключающее подлинные политические альтернативы.
- **Догма о естественном уровне безработицы:** рассматривать конкретный уровень безработицы как структурно фиксированный, когда он зависит от изменяемых политических решений.
- **Бедность как недостаток характера:** приписывание экономических результатов исключительно индивидуальным чертам, игнорируя при этом изменяемые структурные факторы.

Ошибка типа III: Слияние доменов (неправильное размещение категорий)

Эта ошибка возникает, когда P/F раздел из одного домена некорректно применяется к другому.

Междоменные ошибки категорий

- **Биологический → социальный подход:** рассмотрение социально сконструированных гендерных ролей как биологически обусловленных; использование эволюционной психологии для оправдания случайных социальных условий.
- **Социально → биологический подход:** отказ лысенкоизма от генетики в пользу идеологически предпочтительных объяснений, основанных на влиянии окружающей среды; отрицание биологического вклада в поведение по политическим причинам.
- **Математический → эмпирический подход:** Рассмотрение экономических моделей так, как если бы для них требовались математические теоремы; смешение допущений модели с эмпирическими фактами.

- **Эмпирическая** → **логика**: рассмотрение наблюдаемых закономерностей как логических необходимостей; предположение, что раз что-то всегда делалось определенным образом, то это должно делаться именно так.
- **Локальный** → **универсализм**: Рассмотрение специфических для конкретной культуры практик как универсальных норм человеческого существования; предположение о необходимости западных институциональных форм для развития.
- **Техническая** → **политика**: представление политических решений как технических необходимостей («решил алгоритм»); сокрытие оценочных суждений за заявлениями об вычислительной объективности.
- **Политико-** → **технический** **аспект**: Предполагается, что политическая воля может преодолеть технические ограничения; игнорируются физические, вычислительные или ресурсные ограничения.

Исторические примеры конфликта доменов

- **Социальный дарвинизм:** Неправильное применение биологических концепций отбора для оправдания социального неравенства, смешение эволюционных ограничений с политическим выбором.
- **Научный расизм:** использование (ошибочных) биологических измерений для оправдания социальных иерархий, отношение к политическим порядкам как к естественным потребностям.
- **Френология и уголовное правосудие:** рассмотрение преступного поведения как биологически фиксированного на основе измерений черепа, смешение анатомических особенностей с поведенческими предпочтениями.
- **Евгенические движения:** отождествление генетического наследования с социальной ценностью и рассмотрение социальных

проблем так, как если бы они требовали биологических решений.

Современные примеры

- **Алгоритмическое вынесение приговоров:** рассмотрение прогнозов рецидивизма как объективных фактов, а не как результатов моделей, кодирующих конкретные (изменяемые) предположения о риске.
- **Детерминизм «умного города»:** предположение о нейтральности управления городом на основе данных, игнорирующее влияние размещения датчиков, категорий данных и целевых показателей оптимизации на политические решения.
- **Нейробиология в судах:** использование сканирования мозга для обоснования смягчения вины способами, которые смешивают нейронные корреляты с причинно-следственным детерминизмом.
- **Экономический натурализм:** Рассмотрение

результатов рыночной деятельности как естественных явлений, а не как результатов конкретных (изменяемых) правовых и институциональных рамок.

Сводная таблица

Краткое описание типов ошибок с характерными последствиями и типовыми примерами.

Тип ошибки	Структура	Последствие	Пример парадигмы
Тип I: Ложный нейтралитет	$P \rightarrow F$	Игнорирование реальных ограничений приводит к неудачным реализациям.	Идеологии и «нулевого года»
Тип II:	$F \rightarrow P$	Исключени	Генетиче

Тип ошибки	Структура	Последствие	Пример парадигмы
Ложная фиксирован ность		е реальных возможност ей ведет к фатализму.	ский детермин изм
Тип Слияние доменов	III: $(P + F)_{D_1} \rightarrow D_2$	Ошибки в классифика ции приводят к неадекватн ым вмешательс твам.	Социальн ый дарвиниз м

Методологическая заметка

Приведенные здесь примеры представлены не в качестве доказательств эффективности предложенной структуры, а в качестве иллюстраций ее диагностической полезности. В каждом случае

ошибку можно точно охарактеризовать, используя *P/F* обозначения и трехкомпонентную типологию. Структура не указывает, *какие* элементы относятся к той *P* или *F* инной области — для этого требуются знания в конкретной области. Она предоставляет структурированный словарь для выявления, классификации и предотвращения этих повторяющихся моделей ошибочного отнесения.

Протокол аудита ограничений перед внесением изменений

В этом приложении представлен систематический протокол применения разработанной в данной статье структуры к практическим ситуациям, связанным с запланированными изменениями, вмешательствами или проектами. Протокол операционализирует *P/F* различие и типологию ошибок из раздела 1.7, преобразуя абстрактные принципы в повторяемую процедуру аудита.

Шаг 0 — Определение целевого изменения

В одном предложении опишите, что именно

изменяется (объект, область применения, критерий успеха). Укажите уровень описания (например, алгоритмический, организационный, когнитивный, институциональный). Если уровень не установлен, явно перечислите возможные уровни.

Шаг 1 — Установите границы системы

Перечислите, что находится внутри системы и что считается «средой». Укажите интерфейс: входы, выходы и каналы обратной связи. Если граница оспаривается, запишите два варианта границы и повторите протокол для обоих (проверка на чувствительность к предметной области).

Шаг 2 — Ограничения по запасам перед вмешательством

Составьте «реестр ограничений», содержащий все ограничения, которые могут иметь отношение к изменениям. Включите в него:

- **Внутренние ограничения:** архитектура, инварианты, отношения, подобные законам сохранения, фиксированные ресурсы,

требования совместимости.

- **Ограничения среднего уровня:** язык представления, измерительный аппарат, метрика оценки, набор данных, институциональные правила, правовая/нормативная база.

Шаг 3 — Классифицируйте каждый элемент как P, F или U.

Для каждого ограничения присвойте одну метку:

- *P*(заранее определено / не подлежит обсуждению на выбранном уровне и горизонте)
- *F*(свободно / может быть изменено в пределах выбранного уровня и горизонта)
- *U*(результаты неизвестны/не определены, ожидаются результаты тестов)

Правило: При наличии неопределенности следует использовать вариант по умолчанию *U*; не следует *P/F* преждевременно прибегать к принуждению.

Шаг 4 — Проведите четыре теста для всех пунктов

U (и для любых спорных результатов P/F).

Проверка необходимости: Если удалить/отменить этот параметр, перестанет ли система функционировать на выбранном уровне? Если да → P ...

Тест на замещение: Можно ли заменить его альтернативой без потери функциональности? Если да, то → вероятно F (или « F в рамках одной группы»). Если нет, то → вероятно P .

Тест на устойчивость к контрфактическим сценариям: Сохраняется ли это ограничение в правдоподобных контрфактических условиях (данные, окружающая среда, начальные состояния)? Если да → P , то если оно существенно меняется → F .

Тест на зависимость от среды: Исчезает ли это «ограничение» при изменении среды моделирования/измерения? Если да, → то это ограничение среды (C), и его необходимо рассматривать как P ограничение внутри среды, но F между средами (запишите оба варианта).

Шаг 5 — Составьте карту прибылей и убытков и «Бюджет степеней свободы».

Подведите итог: что фиксировано, что регулируемо и где остаются неизвестные. Преобразуйте это в конкретный «бюджет степеней свободы»: перечислите параметры, которые вам фактически разрешено изменять, не нарушая работоспособность системы.

Шаг 6 — Выявление основного риска ошибок (Раздел 1.7 «Вход в таксономию»)

Прежде чем предпринимать какие-либо действия, необходимо четко определить, какой методологический сбой вы, скорее всего, допустите:

- **Тип I:** рассмотрение P как F (попытка перепроектирования инвариантов)
- **Тип II:** отношение F как P (преждевременный фатализм; отказ от осуществимых изменений)
- **Тип III:** путаница в предметной области (смешивание уровней описания; перенос ограничений из одной предметной

области/уровня в другую).

При наличии риска III типа следует остановиться и повторить шаги 0–2 на скорректированном уровне.

Шаг 7 — Вмешательства в дизайн только внутри F (и помечены как U с мерами предосторожности)

Составьте план вмешательства, ограниченный заданными F параметрами. Для тех параметров, которые U вы все еще планируете изменить, добавьте пункт о безопасности: условие отката и сигнал мониторинга, который обнаруживает нарушение работоспособности.

Шаг 8 — Определите критерии оценки, избегая замкнутого круга.

Укажите показатели успеха и убедитесь, что они не идентичны ограничениям среднего уровня, которые легли в основу модели (избегайте «привязки к метрике»). Запишите как минимум один вторичный показатель, который является средне-разнообразным (с другим подходом к измерению/оценке).

Шаг 9 — Предварительный анализ: смоделируйте

отказ как нарушение ограничений.

Напишите краткое «описание неудачи» в формальном ключе: что P было ошибочно принято за F , что F было ошибочно принято за P , или какая граница предметной области была нарушена. Преобразуйте каждое описание в проверку обнаружения (что мы будем наблюдать в первую очередь?).

Шаг 10 — Выполните с мониторингом и обновите реестр ограничений.

Проведите вмешательство. Отслеживайте сигналы работоспособности. Обновите P/F карту на основе выявленных нарушений или вновь обнаруженных инвариантов. Рассматривайте каждый неожиданный сбой как свидетельство P существования скрытого явления или того, что среда наложила неучтенное ограничение.

Выходной артефакт — Лист аудита ограничений

Протокол формирует одностраничный сводный документ, содержащий:

- Границы системы и уровень описания

- Регистр ограничений (полный список)
- *P/F/U* классификация с обоснованиями
- Проведенные тесты и результаты
- Выявлен доминирующий тип риска
- Пространство вмешательства (параметры в *F*)
- Показатели успеха (основные и второстепенные)
- Сигналы отката и сигналы мониторинга

Шаблон: Лист аудита ограничений

Шаблон листа аудита ограничений для систематического применения данной структуры.

Поле	Вход
Изменение цели	[Описание в одном предложении]
Уровень описания	[Алгоритмический / Организационный / Когнитивный / Институциональный /

Поле	Вход
	Другой]
Границы системы	[Внутри: ... / Снаружи: ... / Интерфейс: ...]
Регистр ограничений	[Перечислите все выявленные ограничения]
Компоненты Р	[Список с обоснованием]
Компоненты F	[Список с обоснованием]
Компоненты U	[Список запланированных тестов]
Примененные тесты	[Необходимость / Замещение / Контрфактуальный сценарий / Зависимость от средней величины]
Преобладающий тип риска	[Тип I / Тип II / Тип III]
Пространство для	[Параметры, подлежащие

Поле	Вход
вмешательств	изменению]
Основной показатель	[Критерий успеха]
Вторичный показатель	[Альтернатива со средним и разнообразным набором вариантов]
Триггеры отката	[Условия остановки]
Сигналы мониторинга	[На что следует обратить внимание]

99

Артур, У. Б. (1994). *Возрастающая отдача и зависимость от предшествующего развития в экономике*. Издательство Мичиганского университета.

Аводей, С. (2010). *Теория категорий*. Издательство Оксфордского университета, 2-е издание.

Барретт, Х. К. и Курцбан, Р. (2006). *Модульность в познании: формулирование дискуссии*.

Психологический обзор , 113(3):628–647.

Барретт, Л. и Стаут, Д. (2024). Разум в движении: Воплощенное познание в эпоху искусственного интеллекта. *Философские труды Королевского общества В* , 379(1911):20230144.

Батталья, П. В., Хамрик, Дж. Б., Бапст, В. и др. (2018). Реляционные индуктивные смещения, глубокое обучение и графовые сети. *Препринт arXiv :1806.01261* .

Бендер, Э. М., Гебру, Т., Макмиллан-Мейджор, А., и Шмитчелл, С. (2021). Об опасностях стохастических попугаев: могут ли языковые модели быть слишком большими? *Труды FAccT 2021* , страницы 610–623.

Бенджио, Й., Лекун, Й., и Хинтон, Г. (2022). Индуктивные смещения для глубокого обучения когнитивным процессам высшего уровня. *Труды Королевского общества А* , 478(2266):20210068.

Блоджетт, С. Л., Барокас, С., Дауме III, Х., и Уоллах, Х. (2020). Язык (технология) — это власть: критический обзор «предвзятости» в НЛП. *Труды*

ACL 2020 , страницы 5454–5476.

Боммасани , Р., Хадсон, Д. А., Адели, Э. и др. (2021).
О возможностях и рисках моделей фондов. *Препринт
arXiv :2108.07258* .

Брукс, Р. А. (1991). Интеллект без представления.
Искусственный интеллект , 47(1–3):139–159.

Каррутерс, П. (2006). *Архитектура разума* .
Издательство Оксфордского университета.

Картрайт, Н. (1999). *Пятнистый мир: исследование
границ науки* . Издательство Кембриджского
университета.

Чемеро , А. (2009). *Радикальная воплощенная
когнитивная наука* . Издательство MIT Press.

Хомский, Н. (1965). *Аспекты теории синтаксиса* .
Издательство MIT Press.

Хомский, Н. (1980). *Правила и представления* .
Издательство Колумбийского университета.

Кларк, А. и Лаппин, С. (2011). *Лингвистический
нативизм и бедность стимула* . Wiley-Blackwell.

Кларк, А. и Чалмерс, Д. (1998). Расширенный разум. *Анализ* , 58(1):7–19.

Кларк, А. (2008). *Расширение сознания: воплощение, действие и когнитивное расширение* . Издательство Оксфордского университета.

Дэвид, П. А. (1985). Клио и экономика QWERTY. *Американский экономический обзор* , 75(2):332–337.

Фейнман, Р. и Лейк, Б. М. (2018). Обучение индуктивным смещениям с помощью простых нейронных сетей. *Препринт arXiv :1802.02745* .

Флориди , Л. (2011). *Философия информации* . Издательство Оксфордского университета.

Фодор, Дж. А. (1983). *Модульность разума* . Издательство MIT Press.

Френкель, А. (1922). Zu den Grundlagen der Cantor-Zermelosen Mengenlehre . *Mathematische Annalen* , 86:230–237.

Фригг, Р. и Хартманн, С. (2020). Модели в науке. В книге Залта, Э. Н., редактор, *Стэнфордская энциклопедия философии* .

Галлахер, С. (2005). *Как тело формирует разум* . Издательство Оксфордского университета.

Галлахер, С. (2017). *Энактивистские интервенции: переосмысление разума* . Издательство Оксфордского университета.

Галлегос, И. О., Росси, Р. А., Барбер, Дж. и др. (2024). Предвзятость и справедливость в больших языковых моделях: обзор. *Вычислительная лингвистика* , 50(3):1097–1179.

Геман , С., Биененсток, Э., и Дурса , Р. (1992). Нейронные сети и дилемма смещения/дисперсии. *Нейронные вычисления* , 4(1):1–58.

Гибсон, Дж. Дж. (1979). *Экологический подход к зрительному восприятию* . Хоутон Миффлин.

Гиере, Р. Н. (1988). *Объяснение науки: когнитивный подход* . Издательство Чикагского университета.

Гиере, Р. Н. (2004). Как модели используются для представления реальности. *Философия науки* , 71(5):742–752.

Гриффитс, П. Э. и Стотц, К. (2013). *Генетика и*

философия: Введение . Издательство Кембриджского университета.

Хохрайтер, С. и Шмидхубер, Й. (1997). Долговременная кратковременная память. *Нейронные вычисления*, 9(8):1735–1780.

Программа «Унивалентные основы». (2013). Теория гомотопических типов: унивалентные основы математики . Институт перспективных исследований.

Келлерт, С. Х. (1993). *Вслед за хаосом* . Издательство Чикагского университета.

Лейк, Б. М., Ульман, Т. Д., Тенебаум, Дж. Б., и Гершман, С. Дж. (2017). Создание машин, которые учатся и думают как люди. *Поведенческие и нейробиологические науки*, 40:e 253.

ЛеКун, И., Ботту, Л., Бенджио, И., и Хаффнер, П. (1998). Обучение на основе градиента, применяемое к распознаванию документов. *Труды IEEE*, 86(11):2278–2324.

Лейбниц, Г. В. (1765/1996). *Новые эссе о*

человеческом понимании . Издательство Кембриджского университета.

Либовиц, С. Дж. и Марголис, С. Э. (1995). Зависимость от предшествующего пути, эффект застревания и история. *Журнал права, экономики и организации* , 11(1):205–226.

Локк, Дж. (1689/1975). *Очерк о человеческом понимании* . Оксфорд: Clarendon Press.

Лонг, А. А. и Седли, Д. Н. (1987). *Эллинистические философы* , том 1. Издательство Кембриджского университета.

Лоренц, Э. Н. (1963). Детерминированный непериодический поток. *Журнал атмосферных наук* , 20(2):130–141.

Лоренц, Э. Н. (1993). *Сущность хаоса* . Издательство Вашингтонского университета.

Мак Лейн, С. (1971). *Категории для работающего математика* . Springer.

Мэдди, П. (1988). Вера в аксиомы I и II. *Журнал символической логики* , 53(2):481–511, 736–764.

Мэдди, П. (2011). *Защита аксиом: о философских основах теории множеств*. Издательство Оксфордского университета.

Махони, Дж. (2000). Зависимость от предшествующего пути в исторической социологии. *Теория и общество*, 29(4):507–548.

Митчелл, Т. М. (1980). Необходимость учета предвзятости при обучении обобщениям. *Технический отчет CBM-TR-117*, Ратгерский университет.

Митчелл, Т. М. (1997). *Машинное обучение*. McGraw-Hill.

Миттоне, Л., Морреале, А., и Ритала, П. (2024). Начальные условия и зависимость от пути в исследовательском и эксплуатативном обучении. *Technovation*, 129:102895.

Навилли, Р., Кониа, С., и Росс, Б. (2023). Предвзятость в больших языковых моделях: происхождение, перечень и обсуждение. *Журнал качества данных и информации*, 15(2):1–21.

Норт, Д. К. (1990). *Институты, институциональные изменения и экономические показатели* . Издательство Кембриджского университета.

Ояма, С. (2000). *Онтогенез информации: системы развития и эволюция* . Издательство Университета Дьюка, 2-е издание.

Перфорс, А., Тенебаум, Дж. Б., и Реджиер, Т. (2011). Усвояемость абстрактных синтаксических принципов. *Когниция* , 118(3):306–338.

Пирсон, П. (2000). Возрастающая отдача, зависимость от предшествующего пути и изучение политики. *Американский обзор политических наук* , 94(2):251–267.

Пуанкаре, Х. (1890). Sur le problème des trois corps et les équations de la dynamique . *Acta Mathematica* , 13:1–270.

Пуллум, Г. К. и Шольц, Б. К. (2002). Эмпирическая оценка аргументов о бедности, основанных на стимулах. *Лингвистический обзор* , 19(1–2):9–50.

Шапиро, С. (1997). *Философия математики:*

структура и онтология . Издательство Оксфордского университета.

Смит, П. (1998). *Объяснение хаоса* . Издательство Кембриджского университета.

Стерелни , К. и Гриффитс, П. Э. (1999). *Секс и смерть: Введение в философию биологии* . Издательство Чикагского университета.

Суарес, М. (2004). Инференциальная концепция научного представления. *Философия науки* , 71(5):767–779.

Томпсон, Э. (2007). *Разум в жизни: биология, феноменология и науки о разуме* . Издательство Гарвардского университета.

Варела, Ф. Дж., Томпсон, Э., и Рош, Э. (1991). *Воплощенный разум: когнитивная наука и человеческий опыт* . Издательство MIT Press.

Васвани, А., Шазир , Н., Пармар, Н. и др. (2017). Внимание — это все, что вам нужно. *Достижения в области нейронных информационных систем* , 30:5998–6008.

Вайдингер, Л., Меллор, Дж., Раух, М. и др. (2021). Этические и социальные риски вреда от языковых моделей. Препринт *arXiv :2112.04359* .

Вайсберг, М. (2007). Кто такой моделист? *Британский журнал философии науки* , 58(2):207–233.

Вайсберг, М. (2013). *Моделирование и сходство: использование моделей для понимания мира* . Издательство Оксфордского университета.

Вольперт, Д. Х. и Макреди, У. Г. (1997). Нет теорем о бесплатном обеде для оптимизации. *Труды IEEE по эволюционным вычислениям* , 1(1):67–82.

Цермело , Э. (1908). Унтерсхунген über die Grundlagen der Mengenlehre I. *Mathematische Annalen* , 65:261–281.