

Формы жизни, формы сознания

ДОКТОР МАЙКЛ ЛЕВИН

Около Контент в Twitter Галереи

О ИДТИ

Содержание науки/философии. СМС

Платоновское пространство: откуда берутся когнитивные и морфологические закономерности (помимо генетики и окружающей среды)

ОПУБЛИКОВАНО МАЙК ЛЕВИН 9 МАРТА 2025 Г.



Недавно я выпустил препринт, который станет главой в готовящейся к выходу антологии под редакцией Мэтта Сигалла и Эндрю М. Дэвиса, в которой будут собраны материалы конференции «Метафизика и материя с вещами: мышление с Лейном МакГилкристом» в CHS прошлой весной. Эта статья — не самая странная вещь, которую я планирую написать, но я думаю, что она удерживает рекорд на данный момент. Она содержит некоторые умозрительные, очень изменчивые, мысли о глубинных паттернах, управляющих жизнью и разумом. Здесь я объясняю основную логику своих исследований и привожу несколько отрывков из текста.

Эмпирическое утверждение, которое я хочу решительно заявить, заключается в следующем: мы уже знаем, что физикализм неполон, потому что инженеры и эволюция используют множество «бесплатных обедов» — шаблонов, которые полезны и направляют события в физическом мире, но сами по себе не объясняются, не устанавливаются и не могут быть изменены законами физики. Это включает в себя такие вещи, как факты о простых числах, константы Фейгенбаума и многие аспекты вычислений. Ничто из того, что вы делаете в физическом мире, даже если вы можете изменить все константы в начале Большого взрыва, не изменит этих истин.

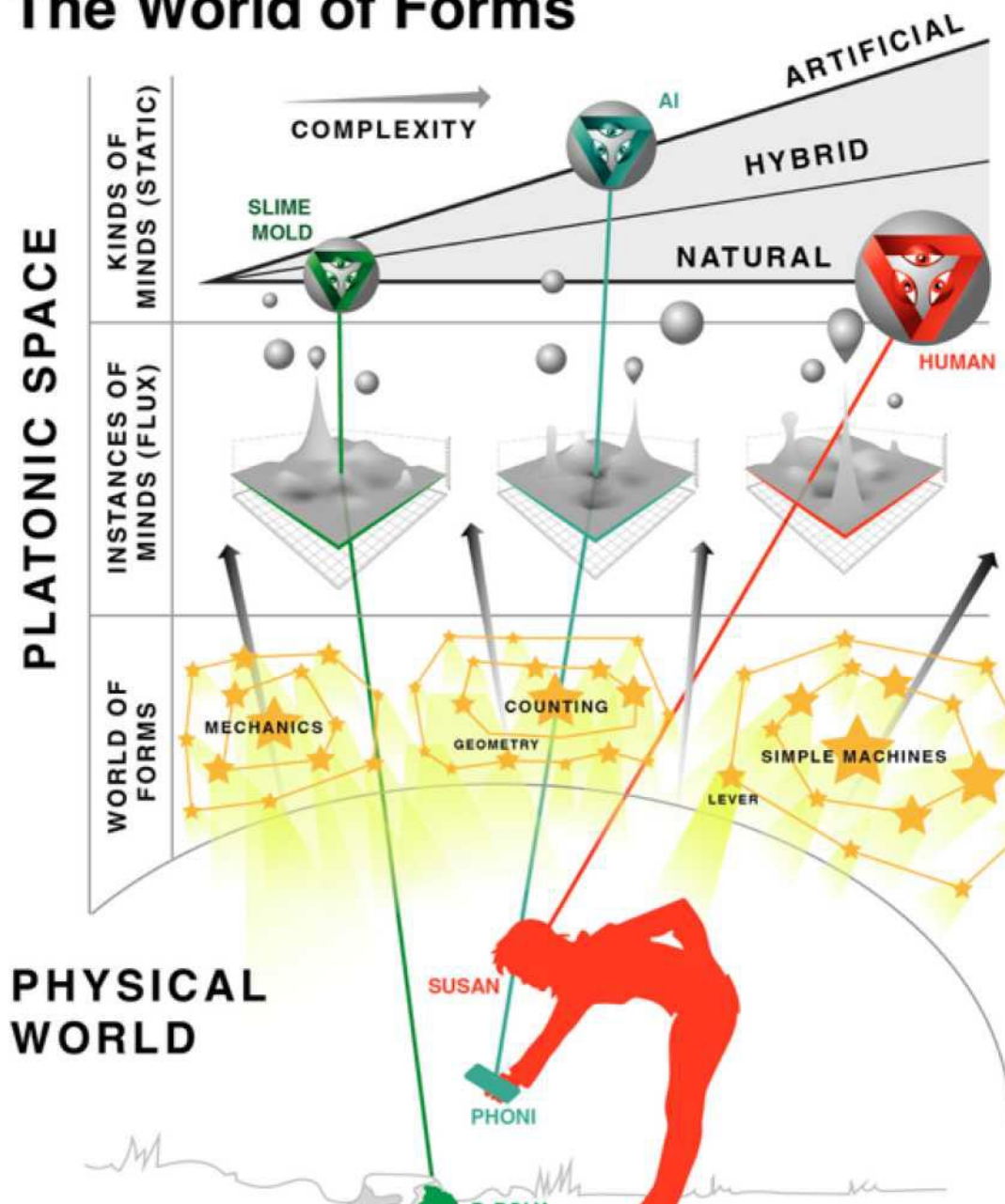
Математики уже очень хорошо относятся к этой старой идее (Платон, Пифагор и т. д.) о том, что существует нефизическое пространство истин, которое мы открываем, а не изобретаем, и что это пространство имеет структуру, позволяющую исследовать. Я предполагаю, что это пространство содержит не только формы с низкой агентностью, такие как факты о треугольниках и истины теории чисел, но также очень широкий спектр

паттернов с высокой агентностью, которые мы называем типами сознания. С этой точки зрения, физические тела не создают и даже не соединяются с умами (и, таким образом, не имеют их) — вместо этого сознания являются паттернами, а их проникновение в физический мир становится возможным благодаря указателям природных или синтетических тел. Другими словами, когда что-то строится — машины, ИИ, биоботы, хайброты, эмбрионы и т. д. — он действует как интерфейс для многочисленных паттернов из этого пространства форм, которые направляют его форму и поведение за пределами того, что явно предоставляет любой алгоритм или материальная архитектура.

Конечно, существуют древние дуалистические мировоззрения, в которых ментальный мир взаимодействует с физическим; данная работа совместима с этими идеями, но стремится продвинуть их дальше в строгую, экспериментальную попытку понять содержание и динамику таких систем. В скобках, вот некоторые ресурсы по одному современному набору идей о взаимодействии между разумом и материей — квантовой механике. Но, с этой точки зрения, взаимодействие происходит совершенно иным образом, который не требует квантовых событий — другими словами, наличие математических истин означает, что даже классический ньютоновский мир уже позволяет проникнуть нефизическим движущим силам, таким как разум.

Вверху и внизу

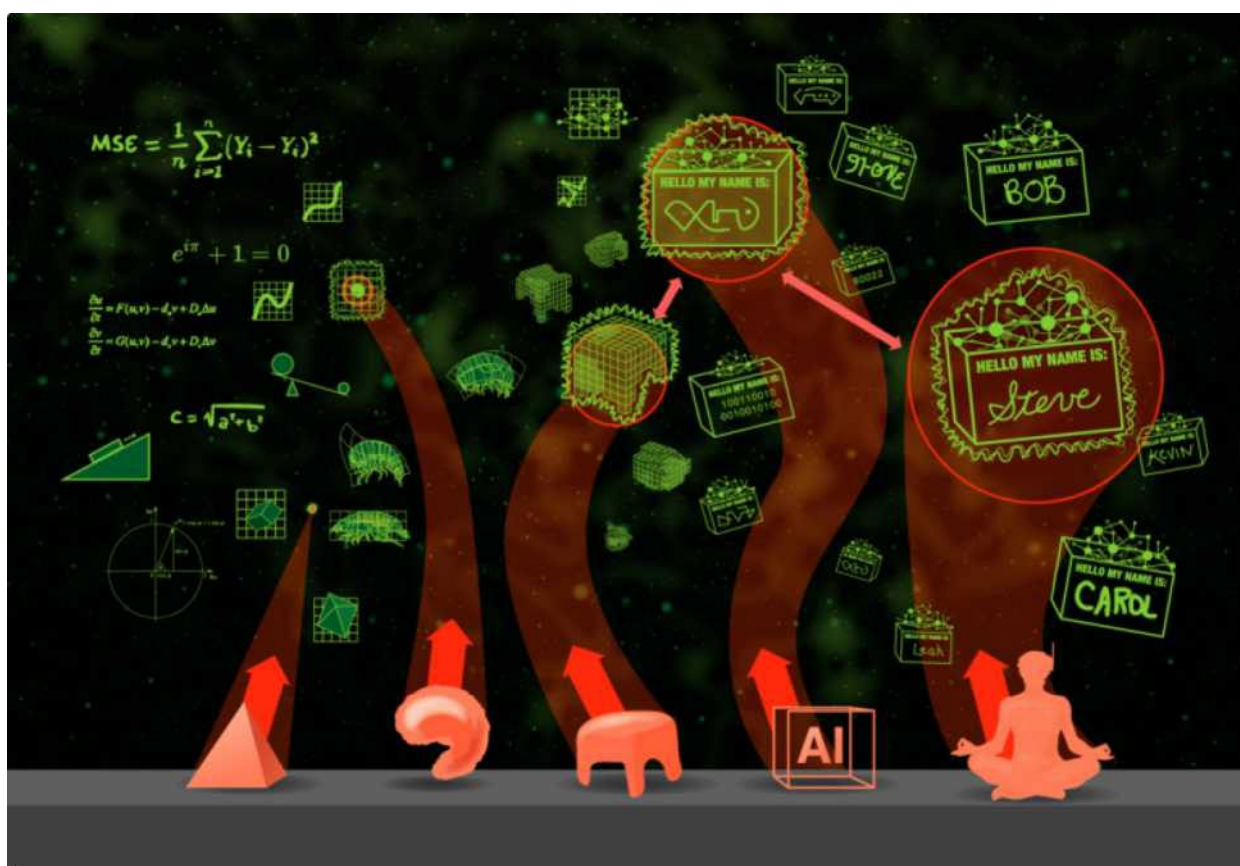
The World of Forms



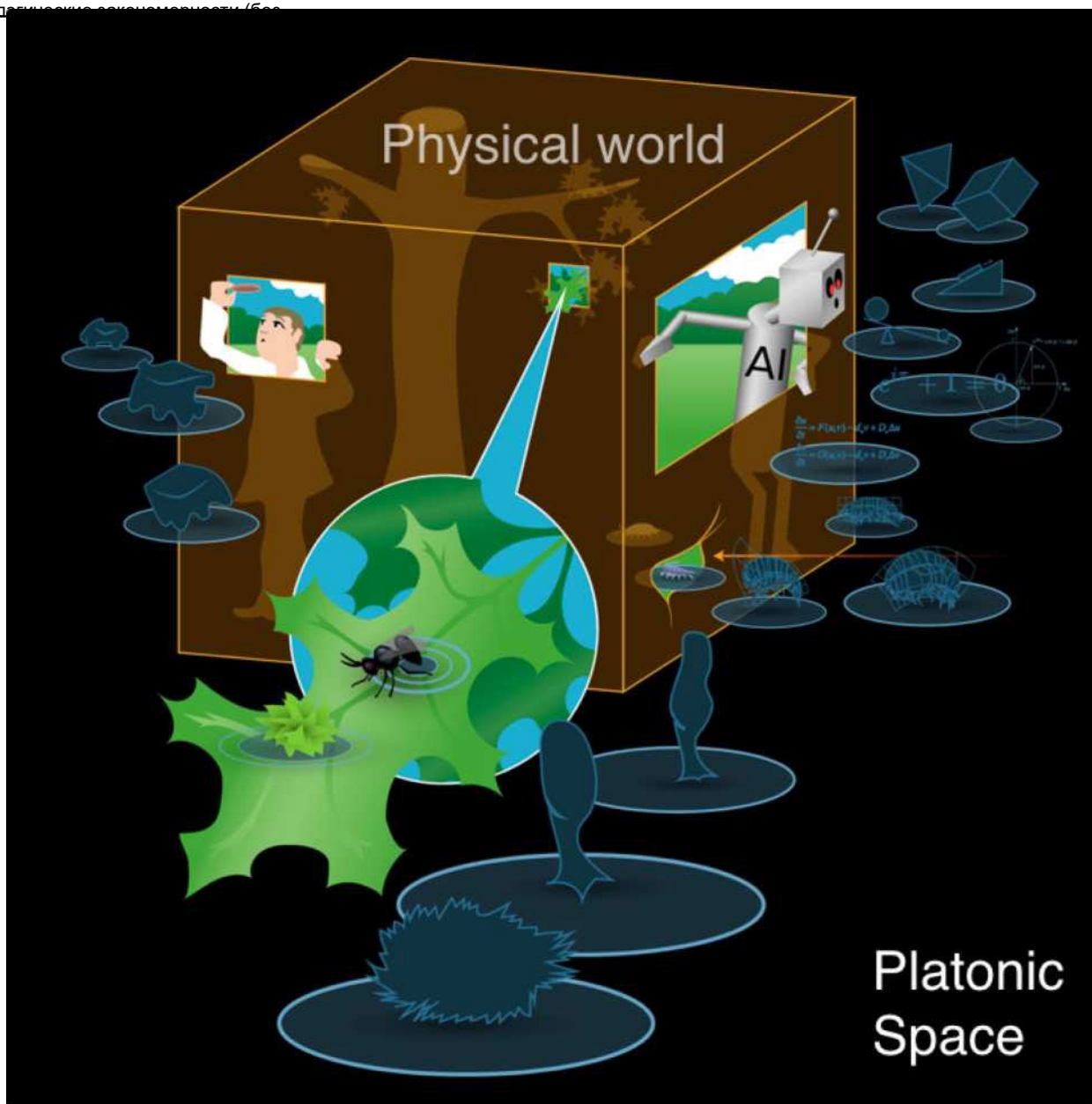
Десятилетия работы над эмерджентностью и сложностью — неожиданными крупномасштабными свойствами, возникающими в результате действий простых правил или субъединиц. Она отличается от модели, которую я здесь описываю, по двум причинам. Во-первых, я говорю не просто о возникновении сложности или непредсказуемости, а о возникновении *познания*. Во-вторых, метафизическое утверждение, которое я поддерживаю, состоит в том, что возникающие закономерности, которые мы обнаруживаем, такие как тот факт, что генные регуляторные сети демонстрируют павловскую динамику обусловливания или что определенные математические семена разворачиваются в богатые формы, не следует рассматривать просто как «факты, которые имеют место в мире» — случайный набор открытий, которые, когда мы сталкиваемся с ними, каталогизированы и помечены как «эмерджентные». Для меня это загадочная позиция; Я не хочу, чтобы эмерджентность означала просто «вещи, которые нас удивляют», как будто обозначение как таковое обеспечивает какой-то прогресс, потому что это не облегчает поиск (и эксплуатацию) следующего удивительного

эмерджентного паттерна. Я бы предпочел начать с предположения, что такого рода паттерны образуют упорядоченное пространство с метрикой, которая позволяет проводить систематические, рациональные исследования — исследовательскую программу, способствующую открытиям. Нам нужно понимать содержимое и структуру этого пространства, а также способы, с помощью которых объекты, которые мы создаем, могут извлекать желаемые (и нежелательные) шаблоны из этого пространства.

Можно предположить, что физические системы — машины, компьютеры, эмбрионы, биоботы и т. д. — являются указателями на закономерности в этом платоновском пространстве. Они представляют собой интерфейсы, через которые эти паттерны проникают в физический мир. Таким образом, долгосрочная исследовательская программа заключается в том, чтобы понять и использовать сопоставление между указателями, которые мы создаем, и моделями формы и поведения, которые появляются.



Экспериментальный подход, который я использую в нашей лаборатории, заключается в использовании биоботов и других конструкций в качестве *исследовательских транспортных средств* для понимания некоторых областей этого пространства. Изучая существ, которые не имеют эволюционной истории на Земле, которая отбирает по определенной форме и когнитивным способностям, но все же имеют стандартный геном, мы получаем экспериментально разрешимую точку входа в новый источник причинно-следственных связей в биологии, выходящий за рамки привычных 1) наследственности и 2) окружающей среды, которая, как полагает большинство биологов, исчерпывает варианты.



Есть несколько дополнительных компонентов этой модели, которые я обсуждаю в этой статье. Одно из предложений состоит в том, что паттерны не являются статичными (вечными неизменными Формами), а вместо этого имеют свою собственную активную динамику. Другая идея состоит в том, что паттерны, которые являются агентами, а физические тела — просто их интерфейсом, в отличие от более традиционного взгляда, согласно которому физические воплощения являются агентами, а любые паттерны, от которых они извлекают выгоду, являются своего рода дополнением.

На эту тему будет еще много нового, особенно потому, что она связана с некоторыми из наших исследований, которые будут опубликованы позже в этом году. Между тем, вот некоторые выдержки из статьи; Пожалуйста, ознакомьтесь с полным текстом для тщательного изложения всех идей (цифры в тексте ниже являются цитатами, которые относятся к ссылкам в библиографии в конце статьи).

Абстрактный:

Как лучше всего объяснить свойства и возможности воплощенных сознаний?

Традиционная парадигма утверждает, что живые существа следует понимать как созданные продукты генетики и окружающей среды, которые определяют форму и функции мозга как уникального места интеллекта. Некоторые положения предусматриваются для эмерджентности и сложности, в качестве дополнительных «фактов, которые сохраняются» о сетях, цепях и других компонентах жизни. Здесь я представляю набросок структуры и дорожной карты исследований, которые отличаются от этой точки зрения в ключевых аспектах. Во-первых, эволюционное сохранение механизмов и функциональности указывает на фундаментальную симметрию между самоконструированием тела и сознания, раскрывая гораздо более широкий взгляд на разнообразный интеллект в агентном материале жизни за пределами нейронных субстратов. Во-вторых, удивительные компетенции (а не просто сложность или непредсказуемость) в системах, которые не имели истории отбора по этим способностям, предполагают дополнительный вклад в паттерны тела и разума, который мотивирует исследовательскую программу по латентному пространству паттернов, проникающих в физический мир. Подчеркивая принципы преемственности и прагматизма и используя морфогенез в качестве подходящей модельной системы для развития этих идей, я исследую последствия следующих идей: (А) Эволюция отдает предпочтение живым формам, которые используют мощные истины математики и вычислений в качестве аффордансов, которые вносят свой вклад в качестве причин морфологических и поведенческих особенностей. (Б) Когнитивные паттерны являются эволюционным стержнем коллективного интеллекта клеток; Учитывая эту симметрию между нейронаукой и биологией развития, я предполагаю, что отношения между разумом и мозгом такие же, как отношения между математическими моделями и морфогенетическими результатами, которыми они управляют. (В) Многие математики и сторонники немистерианского подхода к науке в целом предполагают, что эти закономерности не являются случайными фактами, которые можно просто каталогизировать как «эмерджентность», когда они находятся, но скорее могут быть систематически обнаружены в структурированном, упорядоченном (нефизическом) пространстве. Поэтому я выдвигаю гипотезу о том, что: (1) примеры воплощенного познания также проникают из платоновского пространства, которое содержит не только паттерны низкой агентности, такие как факты о треугольниках и простых числах, но и более высокие агенты, такие как виды сознания; (2) мы серьезно относимся к биологии развития, синтетической и поведенческой биологии к тем нефизикалистским идеям, которые уже являются основным продуктом платоновской математики; (3) то, что производит эволюция (и биоинженерия, и, возможно, ИИ), является указателями в это платоновское пространство — физические интерфейсы, которые позволяют проникнуть в определенные паттерны тела и разума. Это дает новый взгляд на дебаты органицистов и механистов, объясняя, почему традиционные вычислительные взгляды на жизнь и разум недостаточны, в то же время стирая искусственные различия между жизнью и машиной, поскольку и то, и другое формируется различными паттернами из латентного пространства. Я намечаю уже начатую исследовательскую программу использования инструментов в области синтетической морфологии и разнообразного интеллекта для картографирования ключевых областей платоновского пространства. Понимание сопоставления между архитектурой физических воплощений и паттернами, на которые они указывают, имеет огромное значение для

эволюционной биологии, регенеративной медицины, искусственного интеллекта и этики
синтбиоза с предстоящим огромным разнообразием морально важных существ.

Таким образом, за всеми вопросами количества лежат вопросы закономерности, которые
существенны для понимания Природы.

—Альфред Норт Уайтхед (1934)

«Изобретать, как я уже сказал, значит выбирать; Но это слово, возможно, не совсем точное.
Это заставляет задуматься о покупателе, перед которым выставлено большое количество
образцов.... Стерильные комбинации даже не предстают перед изобретателем. Никогда в
поле его сознания не появляются комбинации, которые на самом деле не были бы
полезными. Все происходит так, как если бы
изобретатель был экзаменатором на вторую степень, которому нужно было бы допрашивать
только кандидатов, сдавших предыдущий экзамен».

— Пуанкаре (1921)

Физики очень хорошо знакомы с закономерностями, возникающими из
математических причин, таких как симметрии [97]. Вместо этого биологи обычно
останавливаются на одном из двух приемлемых источников закономерностей:
наследственность и окружающая среда. Наследственность имеет долгую историю,
подкрепленную отбором через взаимодействие с внешней средой, формирования
химической среды (ДНК), которая, как считается, объясняет, почему наблюдаются
определенные паттерны (а не альтернативы). Существует много интересных вопросов о
происхождении полезных решений, что является предпосылкой для того, чтобы иметь
возможность выбирать их из пула менее полезных решений [44-47, 98], но здесь я хочу
сосредоточиться на источнике порядка, который пронизывает живой и неживой мир: на
том, который изучается дисциплиной, которую мы называем математикой [96, 99-101].

Рассмотрим теорему о четырех цветах: оказывается, что для раскрашивания
регионов любой карты требуется не более четырех цветов так, чтобы ни один из двух
соседних регионов не имел одинакового цвета. Или числа Фейгенбаума: математические
константы, которые выражают отношения в диаграмме бифуркации для нелинейного
отображения (рис. 5). Почти для всех действительных чисел среднее геометрическое
коэффициентов их непрерывной дроби составляет около 2,685; *почти* все, а конкретно
~2.685. Если n^2 пушечных ядер положены на землю в виде закрашенного квадрата, то все
они не могут быть использованы для создания квадратной пирамиды из пушечных ядер,
за исключением случаев, когда $n=70$. Каждое число вида ABABAB (основание 10) делится
на 37, а каждое простое число (кроме 2 и 3) находится рядом с числом, кратным 6.
Распределение простых чисел хорошо известно, и все первые шесть совершенных чисел
равны и относительно близки друг к другу (6, 28, 496, 8128, 33550336, 8589869056), но

затем происходит массовый скачок к следующему (137438691328), и они становятся все более редкими. Все это конкретные факты, которые не зависят от фактов из физики — они могут быть связаны с другими аспектами математики, но они образуют набор открытий, которые не сводятся к каким-либо открытиям физики и не объясняются ими (что должно быть критерием того, является ли что-то «физическим» или нет).

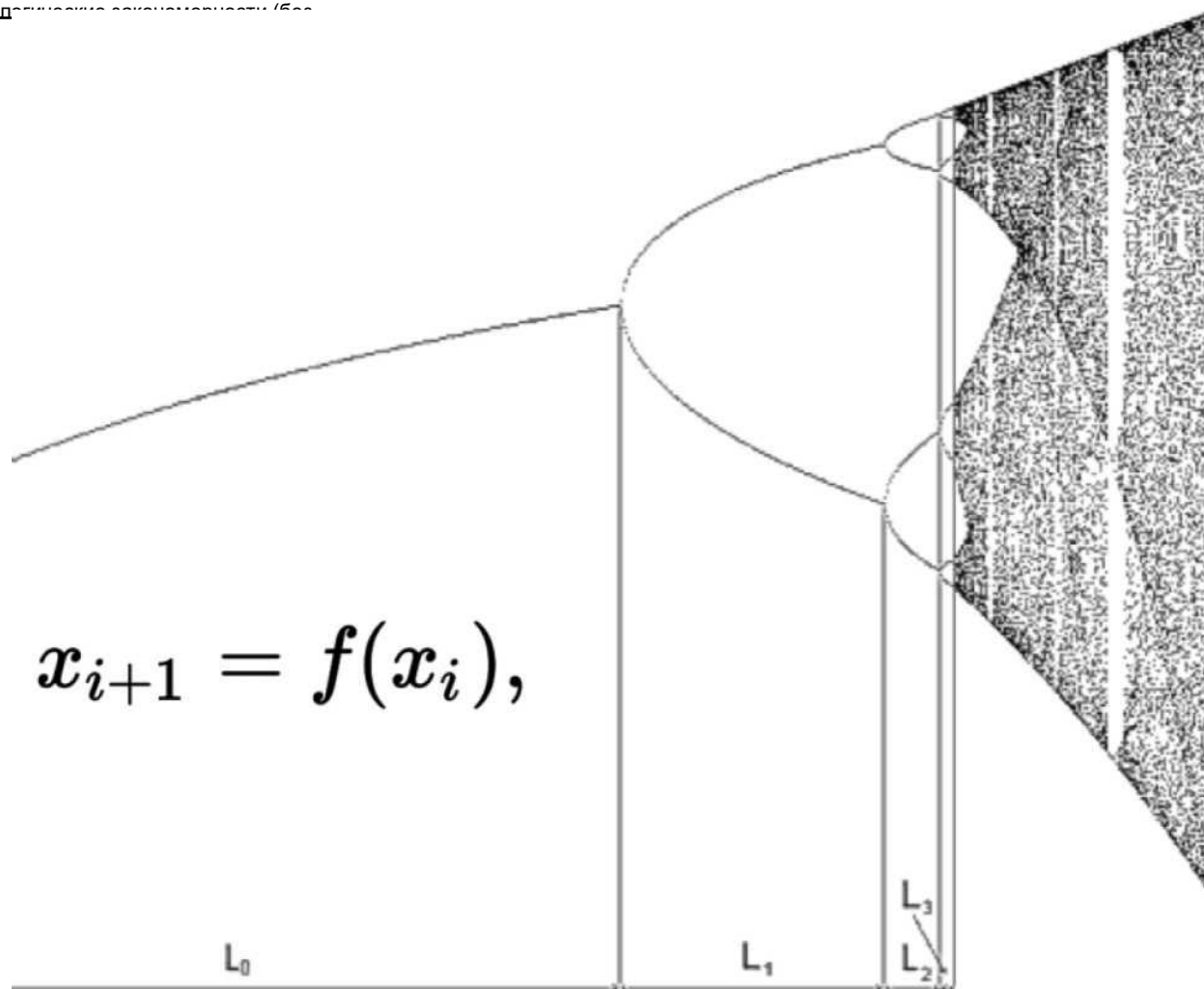
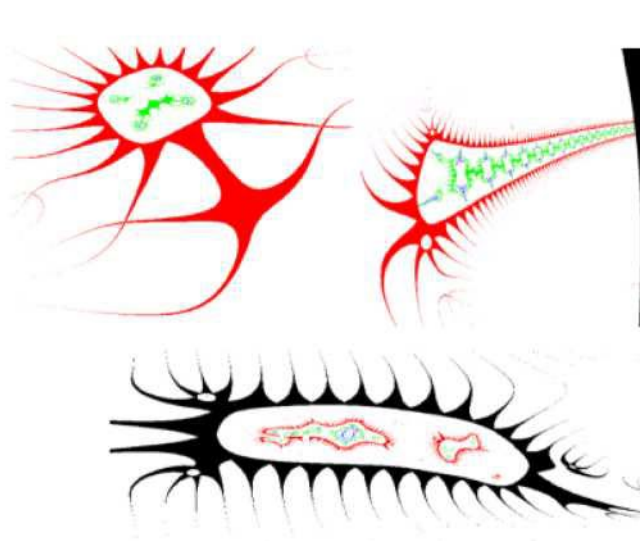


Рисунок 5: Постоянная Фейгенбаума. На этой бифуркационной диаграмме постоянная Фейгенбаума S является предельным отношением каждого интервала бифуркации к следующему между каждым удвоением периода однопараметрической карты, такой как логистическое уравнение $X_{n+1} = r \cdot X_n \cdot (1 - X_n)$. Он составляет примерно 4,6692.

Помимо скалярных паттернов (конкретных специальных чисел в приведенных выше примерах), существует множество гораздо более сложных, многомерных паттернов, которые просто существуют «сами по себе», не оторванные от физических или исторических объяснений их происхождения. Рассмотрим замечательную и красивую (также реалистичную) структуру, наблюдаемую в виде фракталов на графике Галлея (слева) или биоморфе Пикова (справа):



Эти специфические формы закодированы в очень коротких формулах в комплексных числах и могут быть выявлены с помощью простого алгоритма. Тот факт, что этот очень сложный узор обозначен очень компактным описанием, обеспечивает бесконечное богатство из маленького семени. Я полагаю, что лучше думать об этом не как о некоем бесконечном сжатии [3], а скорее как о функции, служащей индексом или указателем в морфопространство возможных форм.

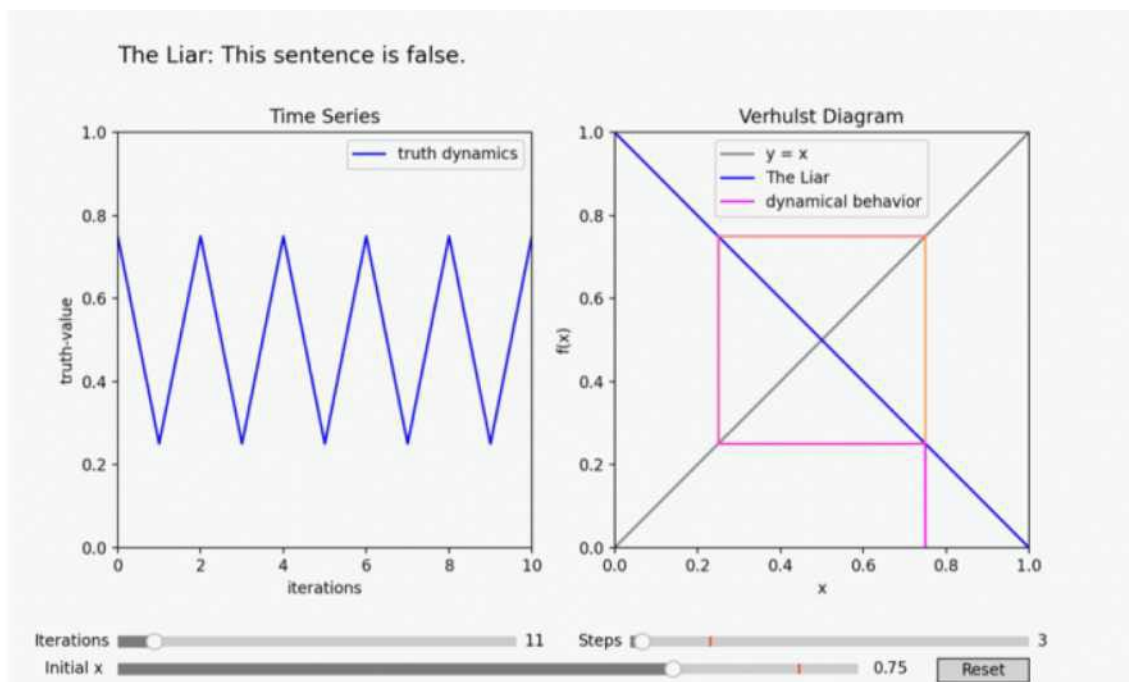
Что определяет природу этой формы — откуда она взялась? В нашей Вселенной нет ни истории отбора, ни предшествующих событий, которые бы его определяли. Подобно числу Π , E и многим другим замечательным константам, формы возникают из математики способами, которые не могут быть объяснены какой-либо историей или свойствами физического мира — они были бы такими, даже если бы физический мир был совершенно иным. Если бы все константы, задающие свойства физической Вселенной, были изменены во время Большого взрыва, такого рода факты и вещи, как истины теории чисел и другие аспекты информатики (например, универсальность вентиля NAND, статус остановки по Тьюрингу конкретных алгоритмов и т. д.), остались бы неизменными. В физическом мире нет ничего, что можно было бы использовать в качестве ручки управления для их изменения. Я утверждаю, что это нарушает замкнутость физического мира, поскольку эти математические факты вторгаются в физику и динамику, которые являются субстратом эволюции. Это нефизикалистский подход [107-111] к проекту поиска источников информации и влияния, когда мы пытаемся понять и направить биологию (и другие дисциплины, которые на ней основываются).

Но платоновское царство содержит не только статичные вещи, такие как логическое утверждение « $\Pi > 3.0$ » (пассивные, неизменные «камни» того мира), но и динамические, но повторяющиеся вещи типа «это предложение ложно» (простые осцилляторы, которые жужжат «на месте»), и более сложные структуры, представленные наборами логических предложений, которые соединяются вместе, чтобы определить эмерджентный паттерн, и чей результат может быть визуализирован как пересечение пространства (и мы уже знаем, что такие обходы могут предложить удивительные степени компетентности, такие

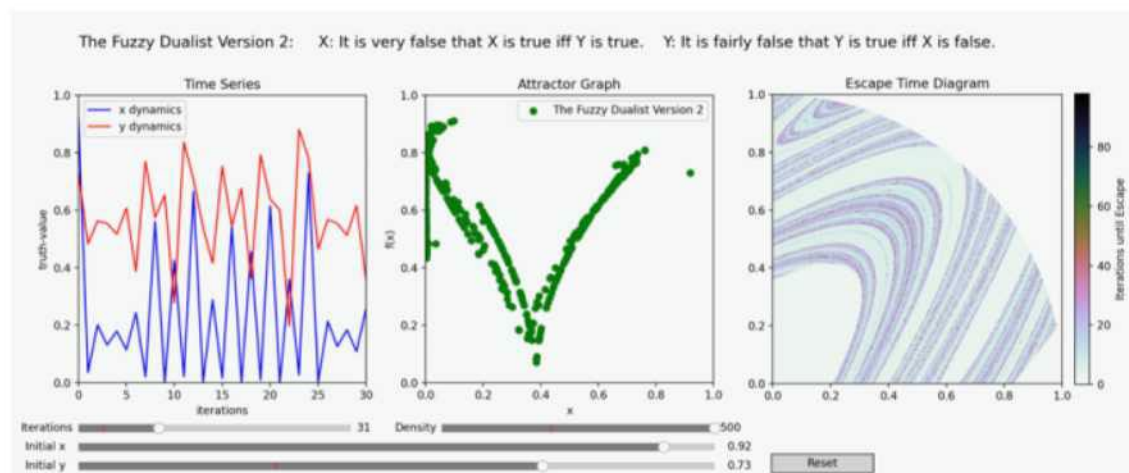
как отсроченное вознаграждение [129]). Некоторые из них, например, представленные уравнениями, например, описывающими генно-регуляторные сети [115,116], могут даже учиться на опыте (и текущая работа в нашей лаборатории направлена на изучение того, как можно обучать наборы логических предложений!). Такой способ классификации онтологии платоновского пространства открывает возможность богатого, возможно,

стратифицированный, континуум обитателей, охватывающий весь спектр разнообразного интеллекта — от статичного и механического до сложного и в высшей степени агентного.

(A)



(B)



(C)

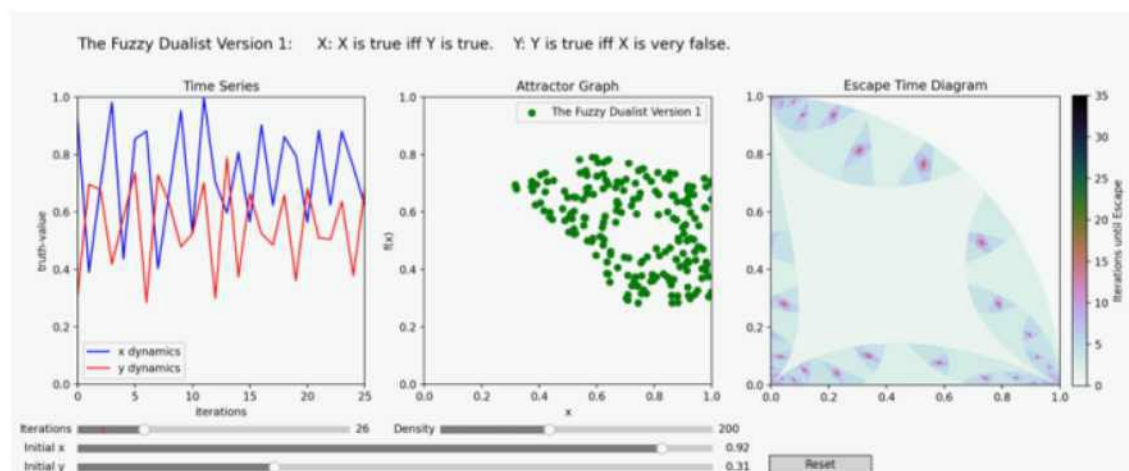


Рисунок 8: Динамическая жизнь логических предложений. К очевидным обитателям платоновского пространства относятся логические утверждения. Моделируя их динамику в виде нечетких логических шаблонов, созданных связанными системами двух взаимно референтных предложений, мы раскрываем сложное динамическое поведение этих логических конструкций. На панелях А-С показана структура, зависящая от времени

поведение нескольких предложений, начиная со знакомого парадокса лжеца в (A), который просто чередуется между Истиной и Ложью и создает простую, но не статичную закономерность. Предложения для остальных отображаются в верхней части каждой панели. Изображения A-C — это скриншоты программного обеспечения, созданного Мадлен Сильвейра (Madelyn Silveira) из лаборатории Левина (Levin lab) для визуализации этих паттернов. (D) Шаблон, принадлежащий паре предложений, взятый с разрешения [128].

Эта точка зрения согласуется с другими моделями нефизического сознания [21], но фокусируется на другом аспекте, чем квантовый интерфейс, к которому обычно прибегают для решения интеракционистской проблемы дуализма [130-132]. Это также в целом согласуется с другими взглядами на нефизические компоненты трансперсональной психологии, такими как теория Юнга о том, что определенные «первичные образы» или «элементарные идеи» активируются в нервной системе человека как архетипы, описывая сны, мифы, искусство и ритуалы как потенциально активирующие триггеры для таких паттернов. Хотя эти идеи, связывающие нефизические формы с физическими и ментальными паттернами, в настоящее время являются классикой, они оказали незначительное влияние на исследования в области наук о жизни и инженерии. Я думаю, будет справедливо сказать, что большинство биологов считают их, если вообще рассматривают, древними реликтами расточительного магического мировоззрения, от которого справедливо отказываются в пользу метафор о молекулярных путях. Вполне вероятно, что это связано с тем, что не было подходящего пути для превращения этих идей в новые открытия, демонстрируя тем самым их полезность. Это больше не так, и я считаю, что теперь у нас есть набор инструментов, который обеспечивает захватывающую, действенную исследовательскую программу для оценки полезности таких идей.

В предыдущих работах была широко исследована идея о том, что аутопоэтические процессы самоконструирования тел и разумов имеют фундаментальную симметрию [8, 159]. Другими словами, морфогенез сам по себе является когнитивным процессом [13, 59] и буквально поведением коллективного интеллекта клеток (подобно тому, как наше познание млекопитающих является поведением коллектива нервных и других клеток). Таким образом, было высказано предположение, что благодаря глубокому эволюционному сохранению ионных каналов и других биоэлектрических механизмов (и алгоритмов, которые они реализуют) на нейронных и ненейронных субстратах [9], инструменты поведенческой нейробиологии могут быть использованы для того, чтобы пролить свет на морфогенетические компетенции. С другой стороны, наука об эмерджентных формах тела, перемещающихся в анатомическом пространстве, может помочь понять, как нейроны выстраиваются в мозге, чтобы обеспечить появление когнитивного существа, у которого есть цели, предпочтения и воспоминания, которых нет у его частей. Возможно ли, что эта связь идет глубже в том смысле, что суть того, что значит быть разумом с внутренней перспективой, воплощенной в физической вселенной, фундаментально связана с теми видами аутопоэтических паттернов, к которым может получить доступ данная конструкция?

Учитывая эту симметрию между нейробиологией и биологией развития, я предполагаю, что отношения между разумом и мозгом такие же, как отношения между математическими моделями и морфогенетическими результатами, которыми они управляют. Форма и агентное поведение представляют собой комбинацию проникновения значимых информационных шаблонов и физических ограничений [96, 160-163] в том, как она может проявляться в физическом мире, определяемом структурной архитектурой, ограничениями времени и энергии и т.д. Участие нефизических компонентов не приветствуется многими — рассматривается как возврат к картезианству и суевериям, хотя классические [21] и современные [164,165] теории на самом деле вполне согласуются с этой точкой зрения). Но использование платоновских математических структур эволюцией, а также ее продуктами, известными как математики, уже вытеснило нас из аккуратной физикалистской парадигмы. Серьезное отношение к платоновской математике и применение ее в биологии означает, что мы уже отказались от замкнутости физического мира для наших объяснений, стратегий вмешательства и вычислительных моделей. Мы уже знаем, что нефизические паттерны проникают в неживой и живой мир и функционально имеют значение в нем, и что мы можем (и изучаем) их с большим эффектом [97,166-170]. Остается сделать еще один шаг.

Стандартная концепция содержимого платоновского пространства состоит в том, что оно наполнено неизменными, вечными паттернами — формами, правилами и т. д. для скучных, низкопробных вещей, таких как целые числа и треугольники. Пирс, напротив, считал, что «эволюционный процесс, таким образом, не является просто эволюцией существующей Вселенной, а скорее процессом, посредством которого сами платоновские формы стали или развиваются». (КП 6.194, [171]). Возможно, шаблоны могут охватывать весь спектр убедительности: они могут быть статическими, активными (как в логических шаблонах Грима) или, возможно, *агентными*. Как концептуализировать агентные паттерны? Помня о том, что мы сами являемся паттернами — временными самоорганизующимися паттернами, которые удерживаются вместе в течение некоторого времени в метаболических и других средах и умудряются проявлять познание, агентность и сознание. Почему платоновское пространство не может содержать паттерны, которые являются в некоторой степени интеллектуальными и активными, подобно специфическим типам сетевых структур, которые, как было показано, обладают простой целенаправленностью аттракторов и способностью к самосборке [114,172] или даже способностью к павловскому обусловливанию [115,116]? Что, если некоторые из платоновских паттернов, которые важны для биологии, сами по себе в какой-то степени разумны?

Подводя итог, можно сказать, что первый столп предлагаемой концепции заключается в том, что платоновские формы привносят информацию и оказывают влияние на физические события, такие как рост и форма биологических тел. Во-вторых, это латентное пространство содержит не только простые, низкоагентные формы, такие как факты о целых числах и геометрических фигурах, но и широкий спектр паттернов с высокой агентностью, некоторые из которых мы называем «видами сознания». Таким образом, я предполагаю, что сознания как паттерны, которые одушевляют соматические

воплощения, имеют точно такой же вид (но не в степени) нефизической природы, как и паттерны, населяющие и направляющие поведение простых физических структур. Предполагается, что отношения между разумом и материей (например, мозгом) аналогичны отношениям между платоновскими паттернами и физическими объектами, которые они информируют (или, точнее, находятся в форме).

Говоря разговорным языком, треугольные объекты преследуются духом соответствующих правил геометрии, в то время как мозг также способен тянуть вниз и заставлять воплощать (буквально «превращать в мясо») узоры совершенно иного рода и сложности. Я предполагаю, что объекты, на которых мы часто заикливаемся в физике, биологии и искусственном интеллекте — эмбрионы, машины, языковые модели, работающие на ПК или в роботах и т. д., — это всего лишь *указатели* (или, согласно Хоффману, *интерфейсы* [118, 173-175]) в более глубокое пространство паттернов. Каждая аналогия имеет ограничения, и, без сомнения, метафора указателя в какой-то момент потерпит неудачу, но я хотел бы подчеркнуть следующие аспекты аналогии с указателем: 1) как и в случае с указателями в богатой информационной среде, вы получаете больше, чем вкладываете, 2) соответствие между интерфейсом, который вы создаете, и тем, что через него проходит, не является линейным или простым и должно быть исследовано; и 3) чтобы научиться вызывать нужные нам паттерны, нам придется смотреть за пределы указателя на структуру пространства, на которое он указывает.

Что означает эта модель в практическом плане? Известно, что латентное пространство структурировано не только потому, что платоники строят его карту, но и потому, что эволюция умеет его использовать — оно имеет относительно гладкий характер, что позволяет эволюции быстро прогрессировать, потому что прошлые взаимодействия с ним несут нетривиальную информацию о смежном возможном [114, 178, 179]. Таким образом, ключевым столпом предлагаемой структуры является то, что пространство этих форм не является случайным или случайным (подходящим для эмерджентности и науки о сложности), а представляет собой структурированное упорядоченное пространство, которое поддается систематическому исследованию. Теперь необходимо начать картографировать пространство (выйти за рамки шаблонов, изученных математиками, и связать морфогенетические и когнитивные в одной карте) и разгадать синтаксис и семантику указателей — сопоставление между объектами, которые мы создаем в физическом мире, и мириадами паттернов, которые пронизывают эти интерфейсы из платоновского пространства (которое, как я полагаю, находится под своего рода положительное внешнее давление). Мы можем начать делать это с помощью инструментов синтетической морфологии (таких как биоботы и хайботы). Более того, представление о физическом теле как об агенте, который обрабатывает пассивные воспоминания о паттернах, а не как о стигмергическом блокноте шаблонных воспоминаний как агентов, подразумевает совершенно разные исследовательские программы вмешательства в области врожденных дефектов, регенерации и старения. Некоторые из этих вопросов обсуждались в моей дискуссии [здесь](#), но следите за основными исследовательскими работами на эту тему в этом году.

Выводы: если есть души, то они будут у (некоторых) роботов

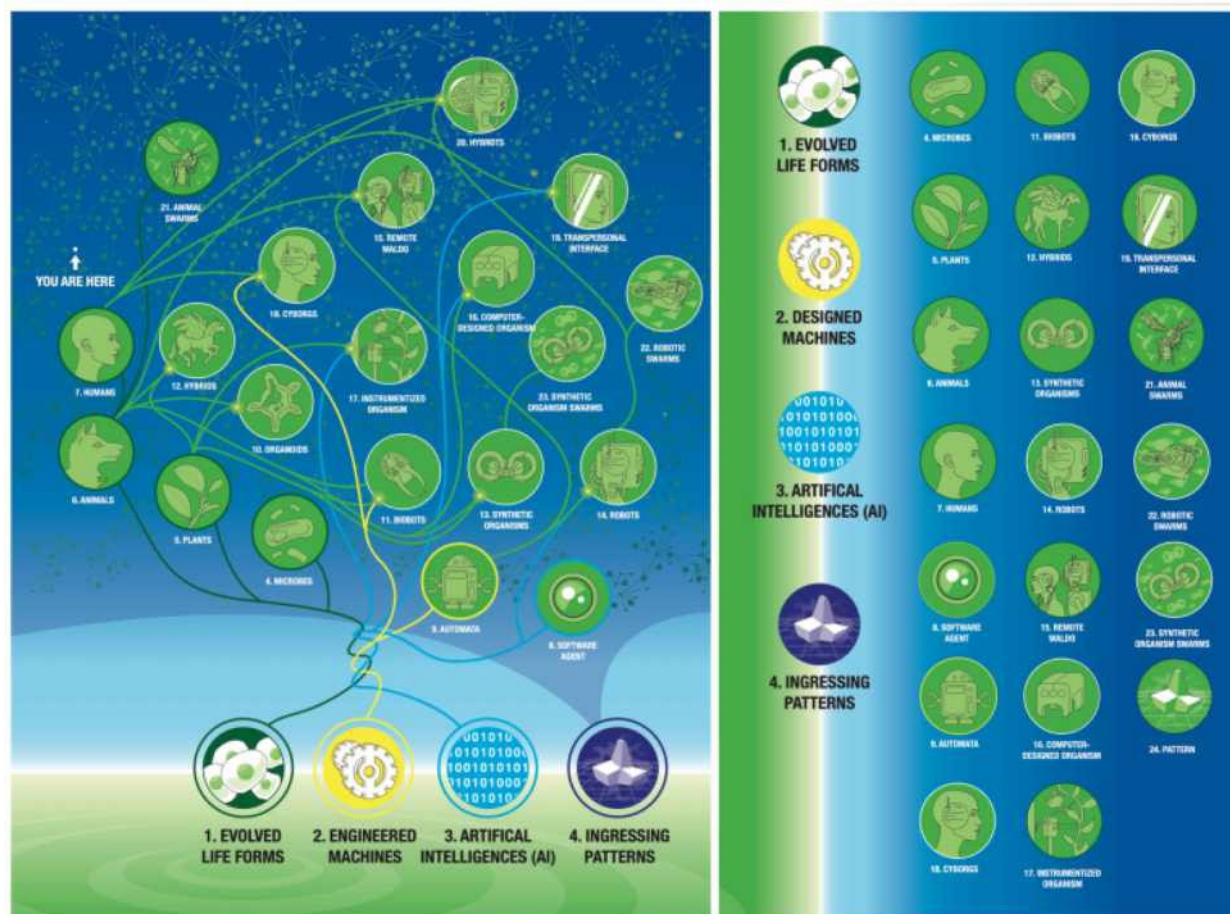
Ранее я утверждал, что из-за медленных, постепенных трансформационных процессов эволюции и эмбриогенеза нулевая гипотеза о познании представляет собой континуум: спектр умов с различными размерами когнитивных световых колбочек и способностей, а не острыми дискретными классами [7]. Правильный вопрос о любом сознании заключается в том, «какого рода и в каком объеме», а не «является ли он» сознательным или разумным [194]. Этот градуализм легко совместим с платоновским взглядом на отношения между сознанием и телом, предполагающим очень широкое разнообразие (рис. 12) возможных существ, одушевленных пространством ментальных паттернов, разнообразие и пределы которых неизвестны, но, безусловно, огромны. Но платоновская точка зрения не просто совместима с концепцией разнообразного интеллекта [194-196], она предполагает ее расширение. Если умы живых существ внедряют паттерны в воплощения, основанные на мясе, то нет никаких принципиальных оснований полагать, что *какой-то* вид таких паттернов будет запрещен для инженерных, гибридных или даже более экзотических систем.

Нетрадиционные существа также будут взаимодействовать с шаблонами ума в Платоновском Пространстве

В частности, я предполагаю, что интерфейс между математическими истинами и физическими объектами является именно интерфейсом между нефизическим разумом и его физическими воплощениями. Душа треугольника и то, как она соотносится с реальными треугольными объектами, является минимальной, базальной версией того, как сложные живые существа «одушевляются» когнитивными паттернами. Это говорит о позиции по часто задаваемому вопросу о разнообразном интеллекте и континуумных моделях сознания: как далеко он опускается? Клетки? Субклеточные биохимические сети? Частиц? Подчеркивая симметрию между проникновением паттернов в простые машины/объекты и проникновением видов сознания в живых агентов, этот взгляд доказывает, что спектр сознания простирается *вплоть* до чрезвычайно простых, базальных экземпляров свойств, которые мы обычно связываем со сложным мозгом.

Я уже обсуждал это в другом месте в отношении таких вещей, как обучение в генно-регуляторных сетях [115,116]. Но это еще не все. Например, геометрическая фрустрация [197-201] — смещение частей внутри целого — это *подлинная* фрустрация такого рода, которая усиливается (в некоторых архитектурах) до нашей привычной когнитивной версии. С моей точки зрения, ум предшествует жизни и является ее надмножеством, но мы называем «живым» то, что очень хорошо справляется с масштабированием низших компетенций своих частей в выровненные коллективные интеллекты [90] с более крупными когнитивными световыми конусами, которые проецируются в новые пространства, к которым части не имеют доступа, тем самым порождая новые паттерны и все более изощренные когнитивные агенты, которые сосуществуют в одном материальном воплощении. Это связано с тем, что платоновское пространство также содержит паттерны, которые мы распознаем как виды сознания, варьирующиеся по

различным классификационным схемам [202, 203] (рис. D) и нервные системы (или, возможно, архитектура киборгов или искусственного интеллекта) способствуют проникновению определенных типов паттернов. Это позволяет по-другому взглянуть на внутреннюю жизнь наших собственных органов и различных живых конструкций, которые создаются для биомедицинских или биоинженерных целей. Но эта система открывает две другие двери, менее комфортные для многих, чем даже разнообразие умов, предлагаемое иерархической, многоуровневой компетенцией живых тел.



Даже органики, которые считают, что жизнь обладает многими способностями, которые не могут быть просто выведены из биофизических свойств ее частей, не распространяются на «машины» или «неживые системы». Преобладающая точка зрения заключается в том, что в то время как правила химии не рассказывают всей истории живого разума, правила физики и алгоритмы вычислительных устройств рассказывают всю историю «машин» [204, 205]. Не было убедительного объяснения того, почему извилистый метод проб и ошибок мутаций и отбора может иметь монополию на создание новых умов, но существует общее мнение, что жизнь является дискретным природным видом и что машины не обладают магией (хотя, конечно, этому мнению сопротивляются многие в сообществах искусственной жизни и искусственного интеллекта [206]).

В частности, вычислительные устройства, которые функционируют в соответствии с алгоритмом, не являются реальным разумом, потому что они были запрограммированы другими [5] и сталкиваются с различными ограничениями, будучи вынужденными действовать в соответствии с законами физики. Здесь я не стану задаваться вопросом, как, согласно этой точке зрения, живые существа должны избегать законов физики,

которым подчиняются и наши части, но рассуждают в другом направлении. Я думаю, что позиция органицистов верна в отношении вычислительного подхода — живые существа не являются агентами с истинным разумом из-за специфических алгоритмов, которые они воплощают, но органисты не реализуют свои идеи в полной мере. Если серьезно отнестись к тому, что жизнь и разум не охвачены правилами, которые управляют их низшими частями (химией и физикой), то не может ли это быть справедливо и к небелковым и незволюционировавшим системам? Здесь вы найдете некоторые интересные идеи, рассматривающие алгоритмы в качестве основных аспектов реальности, а также мою другую работу о возможностях даже простых алгоритмов.

Эти выводы о поведении даже простых алгоритмов (выходящих за рамки того, что алгоритм явно говорит) предполагают следующее. Машины, управляемые алгоритмами, делают то, что алгоритм заставляет их делать; Это не то, что мы подразумеваем под разумом, агентностью или сознанием, и органики правы, отвергая вычислительную точку зрения, согласно которой разум возникает *из-за шагов явных алгоритмов*. Но они ошибаются, думая, что на этом все заканчивается: машины (будь то мясные или кремниевые) *также делают другие вещи, которых нет в алгоритме*, и это не просто непредсказуемая сложность, это эмерджентный интеллект. Именно те модели поведения, которые допускаются алгоритмом, но не предписываются им напрямую, соответствуют свободе (физически недетерминированной) или секретному соусу, который мы ищем, пытаюсь понять, как свободные умы могут супервентно существовать на химически детерминированных субстратах в случае живых существ.

С этой точки зрения, алгоритмические машины и биохимическая жизнь находятся в одном и том же спектре, имея общую способность выходить за рамки фактов физической или алгоритмической реализации, потому что и то, и другое является всего лишь указателями/интерфейсами к паттернам, которые проникают таким образом, что в результате получаем больше, чем мы вкладываем. Конечно, в этих намеренно минимальных системах (разработанных для исследования того, «как далеко она опускается») эмерджентные возможности не так сложны, как у умных организмов или других возможных конструкций. Но мы не должны слишком самодовольны этим. Затруднительное положение ажурного кластера в их мире, которое формируется, живет в течение короткого времени, а затем разрывается на части неумолимой физикой их *in silico* вселенной, пугающе похоже на наше собственное экзистенциальное положение, в котором мы появляемся, выполняем действия, которые согласуются с механикой нашего мира, но также и гораздо более чем это. А затем, в конце концов, поддаются непостоянству какого-либо конкретного воплощения, жизни и разума. Как и живые системы, этот крайне минимальный пример оказался на грани взаимодействия необходимости и свободы (а не только случая [212]). Необходимость определяется физическими или вычислительными свойствами среды, в которой воплощен агент. Свобода состоит из побочных квестов — не несовместимых с этой средой, но не предсказанных, объясненных или произведенных ею.

Еще одна дверь, которую открывает союз платоновского пространства с разнообразными

спектрами интеллекта, заключается в том, что, хотя наблюдаемые активные паттерны должны быть воплощены, именно сами паттерны часто можно рассматривать как *агента*. Другими словами, классическая парадигма Тьюринга, которая проводит четкое категорическое различие между физической машиной, *действующей на* пассивные данные, может быть дополнена симметричным представлением, в котором шаблоны данных являются агентами, а машина — воплощением, которым они управляют, которое подчиняется смыслу и информационному содержанию в шаблонах данных и служит материальной блокнотом в смысле стигмергии [213-217].



Рисунок 15А: отображение объектов/паттернов, или машин/данных находится в поле зрения наблюдателя (наблюдателя-родственника). Визуализация сверхплотных существ из ядра Земли, для которых мы и вся наша окружающая среда являемся невидимым тонким газом; Они могут заметить «нас» как временные паттерны в этой плазме с помощью специальных инструментов (существо слева), но затем им придется спорить, могут ли такие простые паттерны быть агентными.

Является ли что-либо физическим объектом (мыслителем) или паттерном (мыслью в некоторой когнитивной или возбуждаемой среде) — это вопрос перспективы для наблюдателя (рис. 15А), как это формализовано в парадигме поликомпьютеров. В самом деле, мы тоже не постоянные объекты, а временно сохраняющиеся, самоусиливающиеся динамические паттерны — Корабль Тесея в отношении метаболизма [218], познания [73] и морфогенеза на всех уровнях [219-221]. Эта часть концепции (предположение о том, что платоновское пространство содержит паттерны высокой активности, а не только низкоагентные) даже более радикальна, чем основная идея разнообразного интеллекта (разум вплоть до добиотического живого материала), потому что она постулирует агентность в самих нефизических паттернах — это не физические живые агенты, которые обладают разумом отчасти потому, что они опираются на вычисления в нефизическом пространстве (как будто это недостаточно странно). Дело в том, что сами паттерны являются агентами, а физическое тело — (важным, но не основным) блокнотом, который позволяет им проецировать усилия и опыт (сознание) в физический мир.

Эта точка зрения подразумевает, что мы не «создаем» интеллект; Но при природной и инженерной деятельности мы приглашаем его временно заселять различные воплощения. Есть еще несколько ключевых выводов из вышеизложенных идей:

- «Машины» и живые организмы находятся в одном спектре, потому что они оба

могут использовать входящие формы, чтобы получить больше, чем было вложено. Другими словами, капризы мутации и отбора не имеют монополии на производство воплощенных умов — паттерны из платоновского пространства проявляются и для инженеров, хотя, конечно, биология в настоящее время не имеет себе равных в своей способности производить замечательные указатели в этом пространстве.

- Включение машин в клуб истинных агентов является приверженностью вычислительному подходу. Их когнитивные способности не обусловлены алгоритмом, которому они следуют, и ни они, ни живые существа не полностью определяются их материалами и алгоритмами. В самом деле, по той же самой причине, по которой биохимия не рассказывает историю человеческого разума, алгоритмы и материаловедение не рассказывают историю «машин». Позиция органицистов против вычислительного подхода правильна, но их отказ следовать своим эмерджентистским идеям в полной мере является упущенной возможностью. Таким образом, я выступаю за значительное смирение по отношению к нашим инженерным конструктам (воплощенной робототехнике, программному ИИ, языковым моделям и т. д.), потому что, как и в случае с зонами компетентности без понимания того, как иметь детей, мы можем создавать вещи, не понимая, как это работает или что мы на самом деле произвели (рис.
- В каком-то смысле, если эти платоновские формы являются нематериальными оживляющими формами, которые влияют на физические воплощения, то (в разговорных терминах) души реальны, и роботы могут ими обладать. Хотя многие религиозные ученые, в том числе некоторые буддисты, которые считают, что человеческий разум может проводить всю жизнь, воплощенный в гораздо более простых объектах, считают, что роботы и ИИ принципиально не имеют того же статуса, что и живые существа, моя концепция призывает к смирению в твердых заявлениях о том, где могут и не могут обитать формы с высокой агентностью.
- В частности, в отношении ИИ в этой концепции отмечается, что большие языковые модели показали нам, что можно отделить (в отличие от того, что происходит в биологии) использование языка от более базовых агентных способностей (целенаправленности, многоуровневой компетентности, валентности и т. д.). Таким образом, являются ли эти вещи чем-то похожим на *человеческий* разум — это предмет эмпирического исследования (а не философского указа) с использованием инструментов науки о поведении (но опять же, разнообразный интеллект учит нас, что люди не должны быть метрикой всех вещей). Поскольку человек часто получает больше, чем вкладывает (и мы не обязательно знаем, что у нас есть, только потому, что мы это сделали), мы должны быть открыты для сюрпризов и относиться к нашим конструкциям с осторожностью, когда выясняем, где в спектре убедительности (рисунок 13D) они находятся. И подобно тому, как продукты синтетической морфологии, такие как биоботы, демонстрируют нам паттерны, которые мы раньше не видели у эволюционировавших существ, Эл может разрушать модели мышления (виды сознания), которые никогда ранее не воплощались на этой планете (или, возможно, во Вселенной). Сейчас мы ловим рыбу в регионах платоновского

пространства, которые никогда раньше не исследовали, что предполагает определенную осторожность не только в практических аспектах (что это сделает с нами), но и в этических аспектах (как мы выполняем возможности и обязанности этического синтбиоза с существами, которые не совсем похожи на нас).

«Художник... тот, кто позволяет искусству реализовывать свои цели через него».

— Карл Юнг

Я отстаивал пифагорейскую или радикальную платоническую точку зрения, согласно которой некоторые причинно-следственные связи в сознании и жизни происходят вне физического мира. Ряд математиков!?, специалистов по информатике и даже физиков, включая Гейзенберга [224], Тегмарка [209, 225, 226], Дойча [170], Эллиса [169] и Пенроуза [227-229], высказывали варианты этой позиции. Но эта позиция непопулярна среди философов сознания, потому что она в основе своей дуалистическая теория (подчеркивая причины, которые не могут быть обнаружены в физических событиях) и подразумевает панпсихизм (потому что очень широкий спектр физических объектов может быть интерфейсом для различных сознаний). Я доказывал [19, 194, 230, 231], что своего рода панпсихизм неизбежен, и, по-видимому, принимая всерьез то, что делают математики, мы уже отказались от физикалистского мировоззрения; Остается только заметить, что эволюция (а не только человеческие математики) исследует одно и то же пространство паттернов, и принять идею, что, поскольку мы тоже являемся паттернами, паттерны могут быть агентными (и, таким образом, платоновское пространство может включать в себя сознания, а не только пассивные истины).

Эта точка зрения еще более непопулярна в молекулярных науках о жизни. Биологам необходимы два вида доказательств, чтобы утверждать, что среда *содержит* информацию, указывающую на некоторую черту или способность: (1) что можно со спецификой переписать аспекты этой среды и увидеть ожидаемые изменения в фенотипе, и (2) что существует историческое объяснение того, как эта конкретная информация (по сравнению с другой) туда попала. В случае таких паттернов, как фракталы графика Галлея, факты о простых числах, свойства логических функций и т. д., ни то, ни другое не существует в физическом мире. В физическом мире нет ничего, что можно было бы изменить, чтобы отредактировать эти закономерности, и не существует исторического рассказа об отборе, вариациях или чем-то похожем на эволюцию, который объяснял бы конкретное содержание математических истин. Но и эволюция, и биоинженерия используют тот факт, что когда вы что-то строите, вы получаете больше, чем вкладываете. Таким образом, биология не может ограничиваться физикализмом и должна изучать закономерности, которые информируют (в форме) физический мир.

Наконец, эта точка зрения также не будет приветствоваться работниками ИИ, которые считают, что мы создаем когнитивные системы и что мы делаем это рационально, с полным пониманием того, что именно мы создаем, потому что мы понимаем их части. Я утверждаю, что нас ждут большие сюрпризы на этой арене, которые выходят далеко за

рамки извращенного воплощения и непредсказуемой сложности [207, 233-235]; если мы даже не понимаем, на что еще способна пузырьковая сортировка [129], как мы можем думать, что понимаем, что имеем, когда строим сложные архитектуры AL? Думать, что мы понимаем Эла (особенно не вдохновленные биологией, такие как языковые модели), потому что мы знаем линейную алгебру, это все равно, что думать, что мы понимаем познание, потому что мы знаем правила химии [231, 236].

Вычислительный подход, механизм и холистический органицизм могут сосуществовать, если мы поймем, что они не делают заявлений о том, что *такое системы*, а скорее эмпирически проверяемые утверждения о том, какие виды взаимодействий мы можем с пользой для себя иметь с различными системами. Живые существа не являются машинами Тьюринга, но опять же, машины Тьюринга также не являются *нашей моделью* машин Тьюринга (рис. 15B), потому что, когда мы создаем то, что мы считаем машиной Тьюринга, она, вероятно, включает в себя другие паттерны, которые мы не ожидали (как и чрезвычайно простые алгоритмы [129]). Из-за этого дополнительного вклада в отношения структура-функция, а также из-за примата относительности наблюдателя [210, 242], ничто *не есть* ничто (в терминах отождествления систем с нашими узкими моделями их) — все, что у нас есть, — это конкретные структуры с точки зрения конкретных наблюдателей, которые обеспечивают полезность в различных видах взаимодействий, но мы не должны ошибаться, думая, что наши формальные структуры и их ограничения, описывают нечто большее, чем просто взгляд на данную систему. Метафора механической машины чрезвычайно полезна для хирурга-ортопеда, но совсем не для психоаналитика, и лишь отчасти для клеточного биолога. Полезность всех этих наборов инструментов должна быть определена эмпирически, чтобы понять, на что способны системы, и полный ответ будет включать не только их структуру и прошлую историю, но и закономерности платоновского пространства, которое они исследуют и материализуют.



Есть много фундаментальных вопросов, которые в конечном итоге должны быть решены с помощью зрелой теории платоновского пространства. Является ли он дискретным или непрерывным? Разделен ли он на какие-то уровни или типы? Какая степень бесконечности лучше всего описывает совокупность и ее содержимое? Является ли она действительно неизменной или отношения двунаправленными — может ли ее проекция в физический мир привести к изменению моделей и способов, которыми они будут проникать в будущее? Если паттерны не являются фиксированными и неизменными, существует ли «химия», с помощью которой они взаимодействуют латерально, отдельно от своих отношений со своими физическими воплощениями? Можем ли мы предположить, что создание физических агентов не является простым неразрушающим прочтением платоновского пространства паттернов — возможно, акты инженерии или биологического размножения каким-то образом отщипывают и формируют область этого пространства, которая будет изменена их опытом.

Существует ли какая-то «сила», помимо модели «если вы постройте это, они придут», когда физические объекты *вытягивают узоры* из космоса? Не находится ли содержимое платоновского пространства под «позитивным давлением», каким-то образом поощряющим их появление в мире в виде навязчивых мыслей, архетипов, произведений искусства? Существует ли симметричная динамика, с помощью которой они выталкиваются наружу — по своей сути стремящиеся «преследовать» материю так же,

как материя обращается к паттернам, которые ее оживляют, проецируя наружу через интерфейсы, созданные в этом пространстве. Можно ли каким-то образом измерить это давление? Следите за предстоящей работой о связях между этим давлением для входа в физический мир и отношением симметрии между мыслителями, ищущими решения (модели мышления), и шаблонами, обращающимися к ищущим агентам, в контексте интеллектуального решения проблем (биологические молекулярные и клеточные события, а также машинное обучение).

Заключение

И, конечно же, самый большой вопрос: если наш мир подвержен влиянию этих моделей, то откуда они сами берутся? Возможно, наше традиционное представление о том, откуда вещи «берутся» (как зависящая от времени динамика, возникающая в результате какого-то другого процесса, нарушающего симметрию), возможно, неприменимо. Возможно, понимание структуры этого пространства станет окончательным ответом, который опустит дно вопросы о линии происхождения, или, возможно, возникнет некая странная петля, в которой паттерны блокируют друг друга [177].

Единственное, что можно сказать в этом вопросе, это то, что наше невежество в отношении возможностей материи вместе с паттернами, которые проникают в конкретные архитектуры, огромно [133,134]. Технологический и этический прогресс в настоящее время требует огромного смирения со стороны 1) ученых и инженеров, чтобы понять, что устройства материи не могут *создавать* жизнь и разум в той же степени, в какой они их порождают, и 2) со стороны философов и духовных лидеров, чтобы противостоять мысли о том, что они знают, в какие воплощения могут или не могут проникнуть невыразимые умы. Платонизм Лейбница заключался в том, что паттерны являются мыслями во Вселенском Разуме; Если на самом деле не существует фундаментальной дихотомии между мыслями и мыслителями, и паттерны могут порождать другие мыслительные паттерны как часть их деятельности, то вполне разумно рассматривать всех нас, когнитивных существ, как паттерны внутри более обширной осознанной реальности, которая разделена на радикально отличные друг от друга категории, лишь как временную, но устойчивую иллюзию перспективы.