Путь признания – семантика абстракций

Как формируется общепринятое. Ясные и очевидные модели осознания по ментальному шаблону. Почему новое не может само, своими доводами, пробить путь в признание? Семантика абстракций вскрывает главную язву организованной науки.

Аннотация

В работе рассматривается проблема формирования общепринятого знания и механизмы сопротивления новым идеям в организованной науке. На основе нейробиологических данных о работе гиппокампа, решетчатых нейронов и реверберационных процессов выдвигается гипотеза, что абстрактные концепты (идеи, авторитеты, социальные нормы) кодируются в мозге по тому же эпизодическому принципу, что и физические события: «контекст → стимул → действие → последствие». Такая структура порождает устойчивые ментальные правила, а их многократное подтверждение формирует догматические массивы, которые фильтруют любую новую информацию через призму личной значимости и авторитетности источника.

Показано, что академическая наука, блестяще владея методологией анализа (расщепления данных), не имеет инструментов для системного синтеза (сборки смысла), что ведёт к хроническому игнорированию междисциплинарных прорывов — от законов Менделя до теории информационного синтеза Иваницкого. Консерватизм усугубляется эффектом Даннинга–Крюгера и социальной иерархией, где статус учёного часто преобладает над объективностью аргументов.

В качестве решения предлагается использовать большие языковые модели (LLM) в режиме жёсткого верификатора: специально структурированный промпт заставляет ИИ выступать беспристрастным аудитором, выявляя логические разрывы, скрытые допущения и противоречия с фундаментальными законами смежных дисциплин. Получаемый ответ не является истиной, но служит «слепком общепринятой семантики», позволяя автору увидеть точки трения и целенаправленно доработать объяснительную часть теории. Такой подход превращает LLM в спарринг-партнёра, способного ускорить преодоление когнитивных барьеров и стать первым шагом к методологии понимания в науке.

Ключевые слова: семантика абстракций, эпизод памяти, контекстно-зависимая связь, ментальное правило, значимость, новизна, когнитивная карта, реверберация, научный консерватизм, междисциплинарный синтез, методология понимания, LLM-верификация, эффект Даннинга–Крюгера, ориентировочный рефлекс, авторитетное знание

Когда коммунизм начал сворачиваться и республики поотделялись, многие незыблемые представления начали подвергаться сомнению. Люди общались, делясь такими находками.

- Да что там коммунизм, у нас тут начали обсуждать Ленина! – заявил приехавший из Питера зав кафедрой нормальной физиологии универа дружбы народов.

- А что его обсуждать? – удивились мы.

- А вот, оказывается, натворил он делов!..

Ленин представлялся гораздо монументальнее коммунизма и с этим даже в голову не приходило сомневаться. И вдруг он тронулся, грандиозно, как лед на бескрайней реке. Это было болезненно и даже страшновато: как то, что представлялось практически святым, вдруг начало скукоживаться и все быстрее терять вес.

Что в организации человеческой памяти определяет значимость образа, не просто физического образа (этим заведует древняя семантическая память), а абстракции, значимость которой формируется как-то совершенно не так, как у элементов эпизодической памяти, а больше похожа на элемент эпизодической памяти.

Ученые (такие как Мэттью Ботвиник из Google DeepMind) доказали, что координатные (решетчатые) нейроны (grid cells) в энторинальной коре и нейроны места (place cells) в гиппокампе картируют не только физическое пространство, но и абстрактные многомерные концепты.

Мозг создает «когнитивную карту» (cognitive map) ценностей, идей и социальных иерархий. Абстрактный образ Ленина или коммунизма становится ключевым «ориентиром» (landmark) на этой ментальной карте. Вокруг него выстраиваются траектории личного опыта.

Так как этот механизм физически использует те же сети гиппокампа, которые кодируют ежедневные эпизоды, абстракция кодируется *в тех же координатах* и тем же паттерном (через тета- и гамма-колебания), что и реальное физическое событие. Разрушение ориентира дезориентирует «внутренний GPS» человека, вызывая панику, аналогичную потере в глухом лесу.

Семантика абстракций становится похожей на элемент эпизодической памяти, потому что мозг вынужден использовать те же самые физические структуры (гиппокамп, координатные нейроны) и те же молекулярные механизмы (синаптический теггинг) для удержания сложных нематериальных смыслов.

Когда что-то впервые оказывается в фокусе осознанного внимания и получает определенную связь со значимостью этого в данных условиях, то формируется первый эпизод памяти. Причем это не просто какая-то отдельная абстракция, а связь с другой абстракцией (восприятия или действия). Так, сформировавшиеся многочисленные кадры связи Ленин-гений, Ленин-Вождь, Ленин-будущее образовали мощный массив убежденности своим многообразием, который как не может быть оспорен вдруг появившимся эпизодом осмысления заявления, что Ленин-обман. Это не воспринимается всерьез, но эпизод уже существует, пока никак не мешая более прочно сформированному массиву убеждений.

В современной системной нейробиологии описывается через механизм многоканальной конвергенции, конкуренции синаптических весов и фазовых переходов макросетей. Но такое объяснение нас не устраивает, потому что не дает ясную модель понимания. Чтобы говорить о том, как реализуется

убедительность и размывается непререкаемое, нужно не просто пытаться объяснить это нейросетями и синапсами, а иметь четкий механизм того, как шаг осмысления вытаскивает эпизоды и сопоставляет их, почему он вообще это проделывает в процессе осознания.

Структура эпизода исторической памяти действий

С момента периода доверчивого обучения, когда жизненно необходимо отзеркаливать то, как действует опытный человек, решая насущные пробелы, с каждым моментом восприятия действий авторитета формируется связка: стимул-действие-последствия. Это три простых числа ID образов, а в нейросети – три связи с нейронами-распознавателями образов (fornit.ru/100330). Четвертым числом является ID контекста условий, для которых связка получена (в других условиях может быть другое последствие). Образ последствий – это значимость, негативная или позитивная, выраженная в значении разницы до действия и после действия (fornit.ru/70332).

Не обязательно прямо сейчас вникать глубоко, по ссылкам можно пройти потом, если появится желание. Главное: эпизод – это условия, стимул, ответное действие и последствие (стало хуже или лучше). Когда правило определено авторитетом, то значимость последствий принимается на веру – она имеет высокое положительное значение.

Может быть два случая.

1. Ученик действует сам, авторитет реагирует

В этом сценарии ученик совершает пробу, а реакция авторитета мгновенно определяет значимость последствия (сдвиг в плюс или минус).

- Контекст: семья обедает в гостях у дальних родственников (официальная, строгая обстановка).
- Стимул: на столе появляется тарелка с общим пирогом.
- Действие: ученик (ребенок) протягивает руку и берет самый большой кусок прямо пальцами.
- Последствие: авторитет (отец) строго смотрит и качает головой. Для ребенка этот социальный сигнал от авторитета мгновенно окрашивает последствие в высокий минус (стало хуже, допущена ошибка).
- Итог в нейросети: записывается связка, запрещающая брать еду руками в строгой обстановке.

2. Авторитет демонстрирует действие и его результат

Здесь ученик пассивно наблюдает («отзеркаливает»), но связка все равно формируется, так как значимость результата принимается на веру из-за высокого доверия к мастеру.

- Контекст: мастерская, идет ремонт сложного электронного прибора.
- Стимул: на плате загорается красная предупреждающая лампа.

- Действие авторитета: мастер немедленно щелкает тумблером «Аварийное отключение» и спокойно выдыхает.
- Последствие: прибор уцелел, мастер удовлетворен. Ученик не почувствовал физической опасности сам, но маркер авторитета («это спасло нас от катастрофы») переносит оценку последствий в высокий плюс (стало лучше/безопаснее).
- Итог в нейросети: записывается готовое правило «Лампа - Выключатель - Спасение», которое ученик воспроизведет автоматически при первом же совпадении контекста и стимула.

Когда человек самостоятельно принимает решение и ожидает последствий, то сохраняются все те же четыре компонента в связке, но значимость сначала фиксируется как прогноз, а когда возникает последствие – уже по факту.

К примеру, принтер зажевал бумагу и выдал ошибку. Человек решает не вызывать техника (это долго), а резко дергает застрявший лист рукой. В момент совершения действия нейросеть мгновенно строит прогноз на основе прошлого опыта: «Я быстро вытащу бумагу, принтер заработает, я успею напечатать отчет» - все в виде одного числа высокой позитивной значимости, например, +6.

Действие совершено, и физический мир выдает обратную связь. Здесь возможны два исхода, которые запишутся в память по-разному:

Вариант А (Прогноз совпал с фактом): бумага вышла целой, ошибка исчезла. Плюс подтвердился. Связка «Контекст + Стимул + Действие - Успех» закрепляется в памяти как эффективный шаблон.

Вариант Б (Ошибка прогноза): лист порвался, клочок остался глубоко внутри, принтер окончательно заблокировался. Глубокий минус. Возникает сильная разница между «ждал плюс» и «получил минус».

Если таких опытов было несколько и 8 раз все сходило с рук, но в очередной раз возник негатив последствий, то в следующий раз будет задержка действия на осмысление, потому что сразу же в памяти всплывает негативный эпизод, чтобы прикинуть, стоит ли снова рисковать. Решение сильно зависит от уже имеющегося опыта, причем в разных ситуациях (везучий я сегодня или нет, а что будет, если опять проколюсь и т.д.). После решения опять ветвятся варианты.

Если опыт касается опасных вещей, например, прохода светофора на желтый и был один негатив получения травм, то этого хватит на всю оставшуюся жизнь.

Структура эпизода исторической памяти связи абстракций

Структура кадра памяти остается той же – четыре числа, но вместо образов физических действий тут образ мысленных действий.

Ментальные правила могут использоваться для выявления значимости связанных абстракций, например, Ленин и вождь. В период коммунизма значимость такой связки экстремально положительная, но с каждым рассуждением типа Ленин написал: “Навести тотчас массовый террор, расстрелять и вывезти сотни проституток, споивших солдат, бывших офицеров и т. п. Ни минуты промедления” и “Чем больше удастся нам по этому поводу расстрелять представителей

реакционного духовенства и реакционной буржуазии, тем лучше”, значимость Ленин-вождь уменьшается.

В памяти возникает ровно тот же конфликт связок, что и в примере с принтером или светофором. Старый догматический опыт («Ленин — великий вождь, это плюс») сталкивается с цепочкой новых мысленных действий, каждое из которых приносит негативное последствие (разочарование, ужас от прочитанного).

С каждым новым рассуждением и фактом значимость старой связки тает. В результате в аналогичном контексте при стимуле «Ленин» мозг выдает задержку на осмысление, а итоговая оценка абстракции сдвигается из экстремального плюса в глубокий минус или нейтральную плоскость.

Цепочки ментальных образов с определенными значимостями, формируют шаблоны мысленных решений.

Осмысление на основе ментальных эпизодов памяти

Чтобы наглядно показать, как абстрактные образы связываются в правила и передают друг другу эстафету, воспользуемся специальной разметкой:

[Образ-абстракция] — это ментальные элементы (понятия, числа, формулы), которые хранятся в памяти.

{Ментальное действие} — это операция, которую мозг совершает в уме над этими образами.

Вот как прозрачно выглядят ментальные правила и их цепочки в нейросети при вычислении гипотенузы (Теорема Пифагора).

Здесь видно, как последствие одного правила мгновенно становится стимулом для следующего, выстраивая автоматическую цепочку в уме.

Что-то стимулировало необходимость решения по уже усвоенному шаблону и в оперативной памяти общей картины информированности (инфокартины: fornit.ru/68540) появляется цель (fornit.ru/102733) и начальные данные для осмысления.

Правило №1: запуск шаблона

- Контекст: геометрическая задача на плоскости.
- Стимул: внимание зацепило образы [Прямоугольный треугольник] и [Известны катеты 3 и 4].
- Действие: {Извлечь формулу Пифагора}.
- Последствие: в уме активирован образ [Шаблон: Квадратный корень из суммы квадратов]. Оценка: Плюс (+) (путь решения найден).

Правило №2: возведение в квадрат

- Контекст: шаблон Пифагора активен.
- Стимул: образы чисел [3] и [4].
- Действие: {Возвести в квадрат через таблицу умножения}.

- Последствие: в инфокартине зафиксированы новые образы: [9] и [16]. Оценка: Плюс (+). Это промежуточные решения, изменившие общий контекст и, тем самым, определившие направление следующего шага осмысления.

Правило №3: сложение промежуточных итогов

- Контекст: шаблон Пифагора, этап суммирования.
- Стимул: удерживаемые в инфокартине образы [9] и [16].
- Действие: {Сложить числа}.
- Последствие: рождение нового образа [Число 25 под корнем]. Оценка: Плюс (+). Обновление инфокартины новой информацией.

Правило №4: финальное извлечение корня

- Контекст: финал вычислений.
- Стимул: образ [Число 25 под корнем].
- Действие: {Извлечь квадратный корень}.
- Последствие: получен итоговый образ [Ответ: 5]. Оценка: максимальный плюс (+) (цель достигнута). Он обновляет инфокартинку для ячейки памяти “достижение цели”.

Здесь видно, что даже при реализации привычного ментального шаблона вызываются поочередно шаги осмысления, до момента достижения цели.

Из-за того, что используются уже имеющиеся правила, действия совершаются очень быстро и в сознании возникает конечный вариант.

При слабом освоении правил, осмысление может возникать с каждым шагом, сопровождаясь проверками и убеждением в верности полученных промежуточных результатов.

При осмыслении связки типа “Ленин - вождь” в ходе обсуждений или собственных размышлений, так же возникают ментальные шаблоны, которые все более уверенно применяются в актуальных ситуациях. У человека формируются готовые ментальные шаблоны оценки (семантика образов) и шаблоны рассуждения с этими абстракциями.

Если вначале каждое новое критическое суждение требовало долгого осмысления и вызывало сомнения, то со временем в памяти кристаллизуется устойчивый автоматизм. В актуальной ситуации (например, в споре или при просмотре новостей) мозг мгновенно выдает готовую реакцию, проскакивая промежуточные шаги верификации.

Итак, эпизод памяти может быть представлен как контекстно-зависимая связь «условия — стимул — действие — последствие», а мышление над абстракциями может рассматриваться как последовательное воспроизведение и преобразование таких ментальных эпизодов, где результат предыдущего шага становится условием следующего.

Как выбирается ментальное правило из массивов абстрактных взаимосвязей

В случае массива ментальных правил с исключительно позывной или исключительно негативной значимостью, нет никаких проблем.

При позитиве мгновенно запускает самый сильный позитивный автоматизм, если он связан с данным правилом, а при негативе намертво блокируется действие или мысль об этом.

Проблема выбора возникает тогда, когда массив абстрактных взаимосвязей разнороден (когда один и тот же стимул в разных ментальных правилах ведет и к плюсам, и к минусам).

Четвертое число в кадре памяти — это ID контекста.

Так, если человек находится в контексте научного исторического спора, мозг извлекает ментальные правила критического анализа, фактов и архивных цитат.

Если тот же человек находится в контексте семейного застолья с пожилыми родственниками, мозг активирует ментальное правило «сохранение мира в семье». Контекст тормозит агрессивный спор и выбирает правило: «промолчать или согласиться» (чтобы избежать минуса от ссоры).

Это уменьшает выборку и делает более определенным выбор. Но не решает проблему принципиально.

Каждая ментальная связка имеет свой вес – число подтверждающих и число негативных повторений. Сто подтверждений весомее одного негатива. Но если негатив получен последним, то он, в силу новизны заставляет начать осмысливать (fornit.ru/104421) соотношение в отдельной целевой итерации осмысления.

Допустим, в ходе обсуждения активирован стимул: [Нам нужен жесткий лидер, как Ленин]. В массиве памяти человека есть три разных ментальных правила для ответа:

Правило А (+): согласиться, поддержать порядок (активировалось раньше в кругу друзей).

Правило Б (-): возразить, вспомнив цитаты о терроре (активировалось при чтении книг).

Правило В (Нейтральное): промолчать, перевести тему (активировалось в спорах с начальством).

Процесс выбора:

1. Мозг мгновенно считывает Контекст: «Я сижу на совещании, идею предложил мой начальник».
2. Контекст тут же накладывает жесткое торможение на Правило Б (возразить), потому что прогноз этого правила в данном контексте дает огромный минус (конфликт с руководством).
3. Правило А (согласиться) тоже не выбирается, так как внутренний образ [Ленин] после чтения архивов имеет стойкий негативный маркер.
4. В итоге нейросеть выбирает Правило В (промолчать). Прогноз этого правила — безопасность и сохранение текущего статуса (Плюс по отношению к возможному скандалу). Человек делает затяжную паузу, осмысливает ситуацию и просто кивает. Это порождает долгоиграющую проблему выбора – доминанту нерешенной проблемы (гештальт: fornit.ru/102734).

Принцип выбора при осмыслении получается настолько детерминированным (fornit.ru/102699), что неважно, моторные или ментальные правила сохранены в кадре памяти, потому что ID образов уникален (связь с реальными нейронами в нейросети) и по нему сразу ясно, моторный ли это образ или ментальный.

Почему такие сложные выборки из памяти совершаются очень быстро?

Высокая скорость сложных выборок из памяти обеспечивается иерархическим древовидным форматом организации контекстов и признаков, реализуемым как в биологическом неокортексе, так и в передовых системах ИИ (fornit.ru/66797). Этот формат радикально сокращает зону поиска, отсекая нерелевантные ветви, и снижает алгоритмическую сложность с линейной до логарифмической (fornit.ru/102542). Использование древовидных структур позволяет искусственным системам осуществлять мгновенный доступ к опыту и принимать решения в реальном времени (fornit.ru/102542).

Формирование представлений ученых

Ученые – обычные живые люди и организация их психики ничем не отличается от других людей. В мировой академической и прикладной науке около 80–90% ученых занимаются узкоспециализированными исследованиями, опытной проверкой (экспериментами) и прикладными разработками, тогда как

теоретическим обобщением и фундаментальным моделированием заняты лишь около 10–20% специалистов. Но даже у теоретиков есть своя специализация. Ученых с широким универсальным мировоззрением, способных профессионально обобщать междисциплинарные данные, крайне мало — менее 1% от общего числа исследователей в мире (по разным оценкам, от 0,1% до 0,5%). В современной науке их называют генералистами (в противовес специалистам) или интеграторами.

Современная научная система устроена так, что она буквально наказывает за попытки мыслить слишком широко:

Объем информации в каждой дисциплине растет экспоненциально. Сегодня невозможно физически прочитать все статьи даже по одной узкой теме (например, по метаболизму одного вида бактерий). Быть глубоким универсалом во многих сферах одновременно стало физически невозможно для человеческого мозга.

Деньги выдаются под конкретные, измеримые и узкие задачи. Журналы с высоким индексом цитирования (Scopus, Web of Science) почти всегда узкопрофильные. Статью ученого-универсала часто отклоняют, потому что рецензенты (которые сами являются узкими специалистами) считают ее «слишком размытой» или «недостаточно глубокой» в их конкретной нише.

У физиков, биологов и социологов принципиально разный понятийный аппарат и методы доказательства. Перевод данных из одной науки на язык другой требует колоссальных интеллектуальных усилий.

Тех, кто все же способен синтезировать междисциплинарные данные, можно разделить на три основные группы:

1. Специалисты по сложным системам (Complexity Science): Это ученые, изучающие общие законы хаоса, самоорганизации и сетей. Они применяют одни и те же математические модели для анализа работы головного мозга, поведения толпы, климатических изменений и финансовых рынков. Главный мировой хаб таких ученых — *Институт Санта-Фе (США)*.

2. Руководители мега-проектов и системные архитекторы: В сферах вроде искусственного интеллекта (AI), биоинженерии или освоения космоса нужны люди, способные соединить в один проект знания химиков, программистов, психологов и инженеров. Они могут не проводить опыты сами, но видят картину целиком.

3. Философы науки и когнитивисты: Исследователи на стыке философии, нейробиологии и лингвистики, которые пытаются построить единую модель человеческого сознания и познания.

Это приводит к тому, что у подавляющего большинства ученых есть только один контекст массивов ментальных представлений, в котором они наработали хорошо согласованный и уверенный опыт эпизодов памяти. В остальных контекстах опыт оказывается ничем не отличающимся от других людей. Преимущественно он состоит из почерпнутых авторитарных сведений, которым придается высокая значимость. Развивать эти эпизоды проверкой на собственном опыте нет возможности за редким исключением повышенного интереса, но и то с немалыми ограничениями.

Так, мы вынуждены доверяться врачам, даже если наблюдаем огромный диапазон личных интерпретаций и всяких авторских методов лечения. Хотя разумный подход как бы говорит: в столь важных случаях нужно становиться специалистом по своей проблеме.

У ученых возникает сильнейшая иллюзия компетентности (или эффект переноса авторитета) там, где они не отличаются от умных шахтеров. Во всех таких случаях проявляется эффект Даннинга-Крюгера (fornit.ru/70928). У ученых формируются огромные массивы представлений в разных контекстах и даже в собственном профессиональном контексте, основанных на мнении авторитетов или собственных идеях, возведенных на уровень решенных и защищаемых. Если появляется сообщение о связи абстракций, которую ученый ранее никогда не мыслил и это сообщение не будет выглядеть для него достаточно авторитетным, то такой эпизод просто затеряется незамеченным в его памяти.

История науки зафиксировала множество трагических подтверждений вашего тезиса, когда гениальные междисциплинарные связи тонули в безвестности просто потому, что у авторов не было авторитетного статуса, а идеи были слишком непривычными:

Грегор Мендель (генетика). Скромный монах открыл законы наследственности, связав математическую статистику и ботанику. Он докладывал свои результаты местному обществу естествоиспытателей и разослал труды ведущим ботаникам. Но поскольку он не был авторитетом, ученые того времени просто «не заметили» эту связь абстракций. Его работы пролежали в полном забвении 35 лет.

Людвиг Больцман (статистическая физика). Связал поведение газов (термодинамику) с теорией вероятностей. Ведущие физики эпохи во главе с Эрнстом Махом яростно отвергали его идеи, считая их математическими трюками, не имеющими отношения к реальности. Не выдержав непризнания и изоляции, Больцман покончил с собой за несколько лет до того, как его правота стала очевидной для всех.

Игнац Земмельвейс (антисептика). Врач, который связал высокую смертность рожениц с грязными руками хирургов, переносивших «трупные частицы». Научное сообщество посчитало его гипотезу неавторитетной и оскорбительной, проигнорировало статистические доказательства, а самого Земмельвейса довело до психиатрической лечебницы.

Накопленный «уверенный опыт эпизодов памяти» играет с учеными злую шутку. Он создает жесткий каркас, который пропускает внутрь только то, что подтверждает этот опыт, либо то, что насильно продавливается весом чужого авторитета. Любая робкая, но гениальная междисциплинарная связь, лишенная этих двух факторов, осыпается на обочину научного прогресса. Наука становится кладбищем незамеченных связей.

От крамолы к признанию

Я помню несколько моментов формирования общественного мнения, которые вначале выделили нелепо и спорно, но постепенно, далеко не сразу, обросли пониманием у всех, как отражение процесса накопления подтверждающих эпизодов у всех, кто этим интересуется и тех, кто случайно оказался под

влиянием. В когнитивной психологии и социологии это описывается как переход информации из категории *«чужого изолированного мнения»* в категорию *«общераспространенного контекста»*. Но теперь мы можем говорить о совершенно однозначном механизме явления.

Исходная очевидность: Животные не могут обладать сознанием

В начале 2000 годов представление о том, что животные могут обладать сознанием, считалось неприемлемым. Признание наличия сознания у животных в официальной академической науке считалось едва ли не профессиональным самоубийством для биолога. Этот феномен в науке долгое время стигматизировался как антропоморфизм — грех очеловечивания природы. Ученый, который заявлял, что собака, ворон или дельфин совершают поступки осознанно, а не под действием слепых инстинктов, мгновенно лишался авторитета. Его обвиняли в сентиментальности и ненаучном подходе.

Процесс деконструкции этой догмы — классическая иллюстрация того, как накопление тысяч независимых «эпизодов наблюдения» заставило научный мир совершить тектонический сдвиг.

Корни этого неприемлемого отношения уходят к Рене Декарту, который объявил животных «живыми автоматами» (биологическими роботами), лишенными души и разума. В XX веке эту идею укрепил радикальный бихевиоризм. Наука постановила: мы не можем заглянуть в голову животному, мы видим только стимул и реакцию. Значит, никакого внутреннего ментального пространства, никаких чувств и осознания там нет. Собака скулит не от горя, а потому что это автоматическая реакция нервной системы на повреждение.

Перелом произошел далеко не сразу. Он складывался из сотен открытий в самых разных контекстах, которые ученые больше не могли игнорировать или списывать на случайность:

Инструментальная деятельность (Новокаледонские вороны): Ученые зафиксировали в лабораториях, как птицы не просто используют палочки, а мысленно планируют цепочку действий. Они загибают проволоку крючком, чтобы достать пищу, и могут использовать один инструмент, чтобы достать другой, более подходящий. Это требовало мысленного моделирования будущего — признака сознания.

Эмпатия и скорбь (Слоны и Шимпанзе): Были накоплены тысячи видеосвидетельств и полевых наблюдений, как слоны часами стоят у костей погибших сородичей, трогают их хоботами, совершая действия, неотличимые от ритуала прощания.

Зеркальный тест на самосознание: Дельфины, сороки, слоны и шимпанзе успешно прошли тест с зеркалом. Когда им незаметно наносили краску на тело, они, глядя в зеркало, начинали тереть именно то место *на себе*, понимая, что в отражении — они сами.

Нейробиологическое сходство: Развитие МРТ показало, что при переживании страха, радости или при решении задач у млекопитающих активируются те же эволюционно древние структуры мозга, что и у человека.

Переломный момент: Кембриджская декларация о сознании (2012 год)

Процесс накопления эпизодов достиг своего апогея 7 июля 2012 года. Группа ведущих нейробиологов, когнитивистов и физиологов мира (включая Стивена Хокинга) собралась в Кембриджском университете и подписала исторический документ — «Кембриджскую декларацию о сознании». В ней наука официально признала то, что в начале 2000-х считалось крамолой:

«Отсутствие неокортекса (новой коры головного мозга) не исключает возможности переживания организмом аффективных состояний. Конвергентные доказательства свидетельствуют о том, что другие животные, помимо человека, обладают нейроанатомическими, нейрохимическими и нейрофизиологическими субстратами сознательных состояний, а также способностью демонстрировать преднамеренное поведение».

Сегодня, спустя всего пару десятилетий после эпохи «неприемлемости», наука о сознании животных (Animal Cognition) переживает бум. В 2024 году была подписана еще более масштабная Нью-Йоркская декларация о сознании животных, расширившая этот круг на рептилий, рыб и даже некоторых насекомых (например, пчел).

Исходная очевидность: Нет никакой флуктуации в вакууме

Еще в начале этого века, несмотря на опубликованные работы Казимира и Лэмбовском сдвиге спектральных линий, теории Дирака, идея о том, что вакуум постоянно флуктуирует виртуальными парами, считалась крамольной. Здесь сопротивление новой идее шло не вопреки экспериментам, а вопреки математическим абстракциям, которые ученые долгое время отказывались наделять физической реальностью.

Ученые долгое время выстраивали ментальную стену, разделяющую математику и реальность:

«Это просто расчетный трюк»: Виртуальные частицы изначально появились в уравнениях Ричарда Фейнмана (знаменитые диаграммы Фейнмана) как внутренние линии графов. Для большинства физиков это было лишь удобным бухгалтерским инструментом для подсчета интегралов. Сказать, что эти линии «реально кипят в пустоте», считалось вульгарным антропоморфизмом математики.

Парадокс бесконечной энергии: Если признать, что вакуум непрерывно флуктуирует на всех частотах, то его энергетическая плотность должна быть бесконечной (или гигантской). Но общая теория относительности Эйнштейна говорит, что такая энергия мгновенно свернула бы Вселенную в сингулярность. Чтобы не сталкиваться с этим концептуальным кошмаром (космологической постоянной), физикам было психологически проще считать флуктуации «математическим артефактом», который зануляется при перенормировке уравнений.

«Нельзя потрогать — значит, нет»: Поскольку виртуальные частицы по определению невозможно зарегистрировать напрямую детектором (они живут меньше времени, разрешенного принципом неопределенности Гейзенберга),

консервативная физика следовала строгому позитивизму: «Мы видим Лэмбовский сдвиг, но это свойство взаимодействия атома с полем, а не доказательство того, что в пустом пространстве что-то рождается».

Потребовалось накопление критической массы абсолютно новых, более тонких технологических «эпизодов», чтобы это сопротивление рухнуло и идея перешла из разряда «крамолы» в разряд фундаментального консенсуса:

Динамический эффект Казимира (2011 год): Ученые смогли не просто зафиксировать притяжение неподвижных пластин, а буквально *«выбить» реальные, светящиеся фотоны из абсолютной пустоты*. Раскачивая зеркало в вакууме со скоростью, близкой к скорости света, они передали энергию виртуальным фотонам вакуума, превратив их в реальные, регистрируемые приборами частицы.

Эффект Хокинга и космология: Физика элементарных частиц слилась с астрофизикой. Стало очевидно, что реликтовое излучение Вселенной и крупномасштабная структура галактик — это не что иное, как «раздутые» во время Большого взрыва мизерные квантовые флуктуации первозданного вакуума. То есть всё, что мы видим вокруг, включая нас самих — результат того самого «кипения пустоты».

Этот пример наглядно иллюстрирует, насколько инертен «согласованный опыт» ученых. Даже имея на руках строгие формулы Поля Дирака и физические аномалии спектров, физики целое поколение отказывались признать, что вакуум — это не отсутствие всего, а динамическое состояние поля с минимальной энергией, которое непрерывно дышит и флуктуирует. Множество апологетов научных кумиров едко высмеивали в обсуждениях всех, кто показывал согласованную картину вакуумных явлений. Спорить с ними было невозможно: они в упор не видели и не могли видеть то, в чем были уже крепко уверены. Они защищали привычную картину мира, где пустота пассивна, потому что признание активного вакуума требовало тотального пересмотра их ментального базиса. Сегодня же физик, отрицающий реальность вакуумных полей, сам мгновенно прослывет маргиналом.

Исходная очевидность: Актуальность привлечения внимания, связанная с новизной и значимостью - фикция

В мозге нет никакой системы значимости, а формула, что актуальность фокуса осознанного внимания равна произведению значимости на новизну – самопровозглашенная бездоказательность. Это несмотря на работы О.Виноградовой по гиппокампу, работы Е.Соколова, которые как бы и были опубликованы, но совершенно без должных выводов.
В мозге никто не находил «системы значимости» или «детектора актуальности», которые работали бы по принципу перемножения двух этих вымышленных переменных.

С точки зрения строгой нейробиологии, эта формула кажется несостоятельной по нескольким причинам:

Отсутствие гомогенных переменных: «Новизна» и «Значимость» — это не физические величины, которые можно измерить в единой шкале и перемножить.

Это условные ярлыки, которыми исследователи называют сложнейшие, распределенные процессы в мозге.

Проблема курицы и яйца: чтобы определить «значимость» стимула, мозг уже должен направить на него внимание и сопоставить его с текущей доминирующей потребностью (по Анохину или Ухтомскому). Значимость не предшествует фокусу внимания в виде готового коэффициента, она формируется динамически в процессе самого акта восприятия.

Иллюзия «центра управления»: попытка найти в мозге конкретную зону, которая вычисляет эту формулу, абсурдна. Внимание — это не продукт работы одного «процессора», а эмерджентное (всплывающее) свойство синхронизации огромных нейронных сетей.

Формула казалась и все еще кажется абсолютно абсурдной, потому что 1) все нейробиологи занимаются отдельными локальными механизмами, не сводя понимание в более общую модель и 2) никто так и не рассекретил на каких принципах работает “ориентировочный рефлекс” и не задался простым сопоставлением: когда значимость нулевая, то нет никакой необходимости привлекать к этому фокус осознанного внимания, так же и в случае нулевой новизны, когда нет никакой новизны (все привычно знакомо).

Ориентировочный рефлекс (ОР) до сих пор воспринимается мейнстримной наукой как «черный ящик». Концепция Е. Соколова о *нейрональной модели стимула* подошла к разгадке ближе всего, но ее системные выводы были фактически «засекречены» (преданы забвению) в угоду более простым, коммерчески выгодным метафорам.

Для академической науки признать раздражающе не вписывающиеся представления — значит разрушить огромный пласт накопленных «авторитарных сведений», на которых построены тысячи карьер, учебников и психологических теорий управления вниманием. Им гораздо комфортнее защищать спекулятивную формулу, маскируя отсутствие системного понимания за сложными латинскими терминами локальных механизмов.

Исходная очевидность: Реверберация в мозге - чушь

Когда появилась работа А.Иваницкого “Мозговая основа субъективных переживаний: гипотеза информационного синтеза”, где он описывал на основе строгих полученных данных о задержках распространения сигналов в мозге схему закольцовки образов восприятия гиппокампом через структуру системы значимости и подключения ее к лобным долям каналом осознанного внимания, она не вызвала никаких последствий, несмотря на то, что Иваницкий показал первичную суть появления процесса осознания и ее семантическую составляющую. Работа была как бы впустую. С ней не спорили, а просто замалчивали. При всех попытках вынести это на обсуждение налетали апологеты истинной нейробиологии и доказывали, как это тупо и отвратительно. Хотя еще О.Виноградова четко дала данные по реверберациям в гиппокампе как механизм удержания актуального стимула в памяти.

Эта концепция давала изящный и материалистический ответ на «трудную проблему сознания». Однако она была фактически проигнорирована и замолчана узкопрофильным сообществом по трем глубинным причинам.

1. Месть узких специалистов за «простоту» модели

В тот период нейробиология стремительно уходила в глубь микроструктур: ученые учились картировать отдельные ионные каналы, исследовать синаптические белки и выделять субъединицы рецепторов. Для «истинных нейробиологов» (редукционистов) изящная схема Иваницкого, оперирующая крупными системными блоками (кора — гиппокамп — лимбическая система — лобная доля), казалась «упрощенчеством».

Они нападали на модель именно из своей узкой специализации: *«Как можно говорить о закольцовке и удержании стимула, если вы не учитываете конкретный тип ГАМК-рецепторов в поле CA3 гиппокампа?»*. Узкий специалист не способен оценить красоту архитектуры здания, потому что он всю жизнь исследует химический состав кирпича. Любая попытка интеграции воспринималась ими как дилетантское пренебрежение деталями.

2. Запрет на реверберацию

Отказ признавать реверберацию как базовый системный клей сознания — это удивительный парадокс. Действительно, Ольга Виноградова фундаментально доказала, что гиппокамп работает как компаратор и использует циклические процессы для удержания следа памяти. Но в западной и вслед за ней в догоняющей отечественной нейробиологии долгое время доминировала линейно-иерархическая модель: сигнал идет от сетчатки к таламусу, оттуда в зону V1, затем V2, V4 и так далее, пока на вершине не возникнет распознавание.

Идея о том, что сознание рождается не на вершине этой пирамиды, а в момент обратного удара волны возбуждения (когда лобные доли и гиппокамп возвращают сигнал назад в первичную кору), ломала эту простую конвейерную логику. Психологически ученым было легче считать реверберацию шумом или побочным эффектом (эхо-сигналом), чем признать её главным движком осознания.

3. Эффект «Пророка в своем отечестве» и заграничный авторитет

Существует горькая ирония в том, как распределяются «авторитарные сведения» в науке. Пока работу Иваницкого замалчивали на родине, американский нейробиолог и нобелевский лауреат Джералд Эдельман практически параллельно и независимо разработал свою *теорию нейронного дарвинизма*, построив её на точно таком же принципе — «повторного входа» (re-entering).

Когда об этом заговорил Эдельман (обладающий гигантским международным авторитетом), мировая наука взорвалась цитированиями и восторгами. Концепция Эдельмана стала мейнстримом, а отечественные апологеты, которые годами топили Иваницкого за «тупость и отвратительность» схемы информационного синтеза, побежали учить терминологию Эдельмана, поскольку она была освящена западным авторитетом.

Модель Иваницкого опередила готовность научного сообщества мыслить системно. Для её принятия нужно было признать, что субъективный опыт

(психика) — это не мистический эмерджентный дух, а физический процесс сопоставления информации во времени, требующий физической задержки сигнала и его циклической реверберации.

Узкие специалисты, привыкшие защищать свои локальные массивы представлений о синапсах, просто вытеснили эту концепцию на периферию, поскольку она требовала от них слишком масштабной перестройки мышления и выхода из зоны комфортного редукционизма.

Наука не имеет никакого лекарства от этой напасти

Я наблюдал очень много самых разных случаев, когда высказанные идеи, несмотря на их обоснованность и большой потенциал, не могли никак задеть умы специалистов. Если антагонизм Г.Перельмана с коллегами вылился в его самостоятельность и выход из Российской академической системы, то у академика Д.Дубровского так не вышло признания несмотря на то, что в свое время он, как никто другой, близко подошел к системной теории сознания (fornit.ru/70862).

Давид Дубровский предложил информационную теорию психической реальности, рассматривая сознание через принцип инвариантности информации и связь кода с его носителем, что обеспечило материалистическое решение психофизической проблемы. Однако его теория столкнулась с изоляцией из-за междисциплинарного разрыва, философских споров с Эвальдом Ильенковым и неготовности узких специалистов принять системный синтез. В отличие от математики, где верны жесткие доказательства, концептуальные синтезы вроде теории Дубровского сложнее верифицировать, что привело к игнорированию его прорывной модели академическим сообществом.

Обстоятельный разбор показывает, что проблема науки в формировании понимания никак не обеспечена методологически и все еще основывается на природной семантике абстракций.

За последние 400 лет (со времен Фрэнсиса Бэкона и Рене Декарта) наука филигранно отточила методологию анализа (расщепления). Есть строгие протоколы: как изолировать переменную, как поставить контролируемый эксперимент, как очистить данные от шума, как применить математическую статистику. Этому учат со студенческой скамьи.

Но когда данные собраны и расщеплены на тысячи мелких фактов, легального, прописанного в методологических учебниках инструмента для их обратной сборки в единый смысл (понимания) просто нет.

Наука исходит из наивной веры в то, что если накопить достаточное количество мелких кирпичиков-фактов, то здание теории «само собой» вырастет в голове исследователя. Но оно не вырастает. Сборка делегируется случайному биологическому процессу — озарению, интуиции или, как вы точно выразились, природной семантике.

Чтобы создать строгую методологию формирования понимания, науке пришлось бы сделать шаг, который полностью обнуляет статус нынешних академических элит. Для этого потребовалось бы:

Создать мета-язык, который жестко связывает абстракции физики, биологии и философии (то, что пытался сделать Д. Дубровский на концептуальном уровне [fornit.ru/70862]).

Признать, что «уверенность» и «экспертный статус» ученого — это часто лишь биохимический маркер (дофаминовый тег), удерживающий его в зоне комфортных личных догм, а не объективная истина.

Сейчас успешность — это количество опубликованных статей (Scopus, Web of Science), то есть производство новых изолированных кирпичиков. Переход к методологии понимания требовал бы давать высшие награды и гранты не за поиск нового факта, а за ликвидацию междисциплинарных разрывов и сборку старых данных.

Наука производит терабайты информации, но задыхается от катастрофической нехватки смысла, потому что процесс синтеза так и не стал строгой технологией.

Мешает то, что нет понимания системной модели мышления. Она есть, но не может быть принята из-за описанных причин.

Вторая причина – академическая наука при всей ее помпезности и скрепости авторитетом науки выглядит в принципе так же, как организованная преступность с ее понятиями. Помпезность, мантии, ученые степени, латынь, сложные термины и апелляция к «священному статусу Науки» — это внешние атрибуты, выполняющие роль когнитивной и социальной брони. Они внушают обществу и государству священный трепет, исключая внешний аудит.

Мы все еще живем в пещерное время: вместо способности корректно обсуждать, все еще возникает сильнейшее желание набить морду зарвавшемуся оппоненту.

Есть ли выход и полезная методология?

Как большие языковые модели могут помочь решить эту проблему?

Большие языковые модели в их перспективе обучения на всей массе правил связей абстракций, дают полную совокупную семантику накопленного людьми, что позволяет, в принципе, делать корректные сопоставления и обобщения. Они наследуют все те недосказанности в эпизодах отдельного опыта, что мешают пониманию людей, но нет природной зависимости от личного отношения, от эгоцентрической значимости, которая придается каждому из эпизодов.

Это не просто великий усреднитель, ведь GPT работают не как усреднитесь, а как генератор наиболее вероятного продолжения (откуда и галлюцинации в случае скудности исходных данных). Но работать как усреднитель может их заставить удачный промпт. Это дает мощный инструмент справочного ориентирования в уже созданном людьми и этим должно бы быть ограничено применение GPT в качестве инструмента методов верификации идей в условиях общепринятости.

Человеческое понимание всегда фрагментарно и отравлено личной значимостью (желанием защитить «свою пещеру»). Модель же, будучи обученной на колоссальном массиве человеческих текстов, наследует структуру связей абстракций, но не наследует амигдалу. У неё нет страха потерять грант, нет амбиций и нет потребности «набить морду» оппоненту. Она беспристрастно видит

ландшафт человеческого знания целиком — со всеми его сквозными связями, замалчиваниями и белыми пятнами.

Поскольку LLM впитали в себя тексты всех научных дисциплин, они способны стать тем самым транслятором семантики, которого не хватает ученым.

У LLM нет узких мест специализации, а все, что они содержат, в идеале – это максимально выверенные данные специалистов. И в условиях избытка данных (когда тема хорошо описана в литературе) эта вероятностная структура отражает то, что можно назвать «семантическим полем общепринятости».

Через удачный, жестко структурированный промпт мы можем заставить модель работать как идеальный беспристрастный аудитор:

Мы можем загрузить в модель тезисы Давида Дубровского [fornit.ru/70862], данные Алексея Иваницкого и Ольги Виноградовой и попросить систему найти точки концептуального пересечения, которые люди не заметили из-за разности терминологий. Модель мгновенно подсветит, что «повторный вход» Эдельмана, «информационный синтез» Иваницкого и «инвариантность информации» Дубровского описывают один и тот же системный биомеханизм.

Модель способна выявить, где узкие специалисты используют одно и то же слово (например, «информация» или «внимание») в принципиально взаимоисключающих значениях, и очистить семантику от этого шума.

Когда у нас есть новая междисциплинарная идея, мозг автора может упустить из виду, что в какой-то смежной науке (например, в биохимии или математике) уже есть жесткие экспериментальные данные, которые эту идею на корню опровергают. Человек не может прочесть все статьи. LLM же выполняет роль «сироты без эго», которая быстро сопоставляет новую гипотезу со всей массой накопленного человечеством опыта и говорит: *«Концептуально это красиво, но противоречит законам термодинамики, зафиксированным в таких-то массивах данных»*. Это чистая, беспристрастная логическая проверка на системную непротиворечивость.

Там, где человечество еще не наработало правил связей (на переднем крае науки), модель действительно начнет «галлюцинировать», выдавая наиболее вероятный, но пустой лингвистический шум. Она не может родить принципиально новый физический опыт.

Чтобы превратить LLM из «болтливого собеседника» в строгий инструмент верификации, её нужно вырвать из режима бытовой генерации текстов. Промпт должен задавать жесткие системные ограничения:

Запретить модели использовать метафоры и психологические маркеры.

Приказать оценивать входящую идею исключительно по критерию сквозной логической согласованности с фундаментальными, экспериментально доказанными законами смежных дисциплин.

Обязать модель искать скрытые «нулевые значения» — то есть вскрывать логические тупики в теориях авторитетов.

Большие языковые модели — это первый в истории человечества внешний, небиологический зеркальный кабинет для нашей семантики. Мы можем спроецировать туда накопленные абстракции и увидеть их очищенными от нашей пещерной природы: от страха показаться глупым, от преклонения перед академическими званиями, от желания защитить свой социальный статус.

В качестве инструмента верификации идей в условиях общепринятости LLM способны совершить то, что веками не могли сделать научные советы: они могут принудительно, силой математических алгоритмов сопоставления, соединить изолированные кирпичики знаний в единую, понятную и работающую модель.

Пример промпта

Ниже представлена структура и готовый пример идеального верифицирующего промпта. Вы можете скопировать его, подставить свою идею и загрузить в любую современную модель:

Ты — беспристрастный системный аудитор научных концепций. Твоя задача — очистить предложенную идею от личных эмоций, авторитетных мнений и проверить её на строгую логическую согласованность с накопленными знаниями человечества.

Забудь про вежливость. Не пытайся хвалить идею. Работай как жесткий верификатор.

Примени к тексту ниже следующий алгоритм анализа:

1. Выявление скрытых "нулей" (Логическая проверка):

Проверь крайние значения модели. Если один из ключевых факторов равен нулю, зануляется ли вся система, как требует логика? Найди логические тупики, где формула или связь абстракций перестает работать.

2. Междисциплинарный аудит:

Сопоставь эту идею с фундаментальными, железобетонными законами смежных наук (физики, биологии, математики). Нет ли здесь противоречий с базовыми законами природы (например, экономии энергии, принципами работы нервной системы)?

3. Очистка семантики от "шаманских заклинаний":

Найди в тексте размытые психологические метафоры, абстрактные ярлыки или авторитетные догмы, которые используются вместо описания реальных механизмов. Замени их простым описанием причинно-следственных связей.

4. Картирование слепых зон:

Что эта идея замалчивает? Какие экспериментальные факты или альтернативные системные модели она упорно игнорирует, чтобы сохранить свою целостность?

Вот текст идеи для верификации:

[ВСТАВЬТЕ СЮДА ВАШ ТЕКСТ ИЛИ ОПИСАНИЕ ТЕОРИИ (можно прикрепить заранее или дать ссылку на интернет-источник или просто назвать теорию, если она уже представлена материалами в интернете)]

Выдай ответ строго по этим 4 пунктам, используя простой, сухой и понятный язык без наукообразия.

Ограничения применимости результата

В результате будет получен список утверждений на основе общепринятых концепций, в котором нет следствий и выводов новой теории. Поэтому многие утверждения окажутся без учета нового в теории. Такой список утверждений нужно воспринимать как некий общечеловеческий отклик обобщенной семантики, в которой пока отсутствует новое из теории и тогда становится видно, в чем именно расхождение и чем стоит дополнить объяснительную часть теории.

То, что выдаст большая языковая модель в ответ на этот промпт, — это действительно «слепок вчерашнего дня». Это моментальный снимок накопленной, устаканившейся общечеловеческой семантики, живущей по старым правилам и понятиям. Модель физически не способна сделать из вашей новой теории революционные выводы, потому что её вероятностный движок настроен на то, что уже было написано людьми.

Поэтому полученный результат — это не вердикт «правильно/неправильно». Это инструмент контраста.

Когда вы получаете сухой, консервативный ответ модели, вы используете его как зеркало, которое подсвечивает точки трения между старым миром и вашей новой теорией:

Список утверждений модели показывает, на какие именно «авторитетные сведения» и привычные догмы наткнется ваша теория при попытке войти в умы специалистов. Если модель пишет: *«Это невозможно, потому что общепринято считать X»*, — это не значит, что вы неправы. Это значит, что в этом месте у человечества глухая ментальная стена.

Если модель не поняла следствий вашей теории и скатилась в привычное русло, значит, в вашем тексте недостаточно разжеван причинно-следственный механизм. Для вас эта связь абстракций очевидна (у вас есть этот эпизод опыта), а для общепринятой семантики — это пока пустой звук.

Вы смотрите на список банальностей или возражений модели и понимаете: *«Ага, они мыслят старым контекстом. Чтобы они меня услышали, мне нужно дополнить объяснительную часть вот в этой конкретной точке. Мне нужно детально описать переход от их старого понимания X к моему новому пониманию Y»*.

Итог: ИИ как спарринг-партнер

В такой рамке большая языковая модель становится идеальным спарринг-партнером. Она изображает из себя того самого консервативного, упрямого «профессора из академического клана», который не хочет слушать крамолу. Но в отличие от живого профессора, модель делает это без агрессии, без желания набить морду и по первому вашему требованию.

Вы используете её как тренажер: прогнали текст — увидели, где общепринятая семантика буксует и сопротивляется — переписали объяснительную часть, сделав междисциплинарные связи более явными — прогнали снова. Это и есть

технологический способ преодоления «пещерного» барьера в формировании понимания.

Список литературы

Решётчатые нейроны, нейроны места и когнитивные карты абстракций

1. Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292), 1464–1468.
   Демонстрация того, что решётчатые нейроны энторинальной коры картируют не только физическое, но и абстрактное концептуальное пространство.

2. Aronov, D., Nevers, R., & Tank, D. W. (2017). Mapping of a non-spatial dimension by the hippocampal–entorhinal circuit. *Nature*, 543(7647), 719–722.
   Экспериментальное подтверждение, что гиппокамп и энторинальная кора кодируют абстрактные измерения (например, частоту звука) по тем же принципам, что и физическое пространство.

3. Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4), 189–208.
   Классическая работа, введшая понятие «когнитивной карты» как ментальной репрезентации пространственных отношений, впоследствии распространённого на абстрактные домены.

4. Park, S. A., Miller, D. S., & Boorman, E. D. (2021). Hippocampal neurons construct a map of an abstract value space. *Cell*, 184(17), 4525–4540.e8.-
   Прямое свидетельство того, что нейроны гиппокампа у приматов формируют «карты ценности», аналогичные картам физического пространства.

Реверберация, повторный вход и теория сознания

5. Edelman, G. M. (1987). *Neural Darwinism: The Theory of Neuronal Group Selection*. Basic Books.
   Фундаментальная теория, в которой «ре-энтри» (reentry, повторный вход) — циркуляция сигналов между таламокортикальными структурами — рассматривается как ключевой механизм сознания.-

6. Edelman, G. M., & Tononi, G. (2000). *A Universe of Consciousness: How Matter Becomes Imagination*. Basic Books.
   Развитие теории нейронного дарвинизма с акцентом на динамическое ядро (dynamic core) и реентерабельные взаимодействия как субстрат сознательного опыта.-

7. Vinogradova, O. S. (2001). Hippocampus as comparator: Role of the two input and two output systems of the hippocampus in selection and registration of information. *Hippocampus*, 11(5), 578–598.-
   Ключевая работа Виноградовой, обосновывающая функцию гиппокампа как компаратора, использующего циклические процессы (реверберацию) для удержания и сравнения следов памяти.

Декларации о сознании животных

8. Low, P., et al. (2012). The Cambridge Declaration on Consciousness. *Francis Crick Memorial Conference, Cambridge University.*-
   Исторический документ, в котором группа ведущих нейробиологов официально признала наличие нейроанатомических субстратов сознания у нечеловеческих животных.

9. Andrews, K., Birch, J., & Sebo, J. (2024). The New York Declaration on Animal Consciousness. *New York University.*-
   Расширение Кембриджской декларации, утверждающее реалистическую возможность сознания у всех позвоночных и многих беспозвоночных (включая осьминогов и пчёл).

Исторические случаи игнорирования научных открытий

10. Mendel, G. (1866). Versuche über Pflanzen-Hybriden. *Verhandlungen des Naturforschenden Vereins in Brünn*.-
    Первоисточник законов наследственности, который был проигнорирован научным сообществом почти на 35 лет — классический пример того, как отсутствие авторитета автора ведёт к забвению открытия.

11. Boltzmann, L. (1877). Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung. *Wiener Berichte*, 76, 373–435.
    Работа, заложившая основы статистической физики; идеи Больцмана яростно отвергались Эрнстом Махом и энергетистами.-

12. Semmelweis, I. P. (1861). *Die Ätiologie, der Begriff und die Prophylaxis des Kindbettfiebers*. C. A. Hartleben.
    Труд Земмельвейса о связи послеродовой лихорадки с грязными руками врачей, отвергнутый медицинским истеблишментом того времени.-

Когнитивные искажения в науке

13. Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121–1134.-
    Классическая работа, описавшая эффект Даннинга–Крюгера — когнитивное искажение, при котором люди с низкой компетентностью переоценивают свои способности.

Теория информационного синтеза Иваницкого

14. Ivanitsky, A. M. (1999). The main riddle of nature: The mental events origin out of the brain processing. *Psikhologicheskii Zhurnal*, 20(3), 89–96.-
    Работа, в которой Иваницкий формулирует гипотезу о возникновении субъективных переживаний как результата информационного синтеза в коре — через возврат нервных импульсов к местам первичных проекций и их сопоставление с извлечёнными из памяти данными.-

15. Ivanitsky, A. M. (1995). Information synthesis in cortical areas as an important link in brain mechanisms of mind. *Neuroscience and Behavioral Physiology*, 25(6), 487–492.-

Развитие теории информационного синтеза с акцентом на динамические корковые структуры — «очаги взаимодействия» (interaction foci).

Источники с fornit.ru (5)

16. ID образа в модели МВАП: функциональная абстракция или биологическое противоречие? fornit.ru/100330[reference:14]
Обсуждается, что в модели МВАП «ID образа» выступает как стабильный, адресуемый элемент, позволяющий однозначно идентифицировать, хранить и извлекать конкретный образ в системе — функциональная абстракция, не имеющая прямого биологического аналога, но необходимая для инженерного описания памяти.

17. Дифференциатор состояния организма. fornit.ru/70332[reference:16]
Описывается механизм определения изменения значимости состояния организма после совершённого действия — «стало лучше или хуже» — как основа оценки последствий и обучения через опыт.

18. Глобальная информационная картина субъекта. fornit.ru/68540[reference:18]
Вводится понятие инфокартины — совокупности актуальной информации, удерживаемой в оперативной памяти в процессе осмысления, которая обновляется с каждым шагом мышления.

19. Цель осмысления. fornit.ru/102733[reference:20]
Раскрывается механизм возникновения цели как следствия рассогласования между прогнозом и реальностью, запускающего процесс осознанного выбора альтернатив.

20. Актуальность осмысления – основа сознания. fornit.ru/104421[reference:22]
Обсуждается, как ориентировочный рефлекс определяет, какой стимул становится предметом осознанного внимания, и как это связано с процессом осмысления и принятия решений.