

Операциональная концепция истины: верификация предположений в взаимодействии с реальностью

Соответствие модели реальности. История: от Платона до постмодерна. Современное понятие Истины в мейнстриме академической науки. Спор между научным реализмом и научным антиреализмом. Каково рефлексивное отношение академического мейнстрима к обобщающей формуле: Истина - это подтверждение предположения верификацией реальности, а ложь - отрицательный результат? Защита формулы. Природное подтверждение формулы. Финальное выражение формулы Истинности. Насколько отличаются принципы реализации нейросетевых ИИ? Практика: жизнь в эпоху постправды.

Аннотация

Истина — это подтверждение предположения верификацией реальности, а ложь — отрицательный результат.

Эта операциональная формула, лежащая в основе всех адаптивных систем от живых организмов до научного метода, находит своё полное выражение в концепции мировоззренческих абстракций.

В статье анализируется эволюция понятия «Истина» — от греческой алетеи через корреспондентную теорию Аристотеля и средневековую концепцию соответствия Божественному замыслу, рационалистические и эмпирические традиции Нового времени, кантовский феноменализм и гегелевскую диалектику, прагматизм и аналитическую философию XX века до постмодернистского распада «зеркала реальности».

Современный академический мейнстрим (научный реализм + конструктивный эмпиризм ван Фраассена + фаллибилизм Поппера + модель-зависимый реализм Хокинга–Млодинова + теория Тарского) отказывается от бинарной формулы «истина/ложь» в пользу вероятностного и инструментального подхода. Автор защищает исходную операциональную схему, показывая, что она не является «вульгарной» упрощённостью, а, напротив, является наиболее строгой и верифицируемой конструкцией, полностью освобождённой от метафизических наслоений.

Формула $V(A,C,M) — V(A,C,M) = \{\text{true при совпадении фактического результата с ожидаемым; false при рассогласовании}\}$ — позволяет разрешить классические проблемы индукции, тезис Дюэма–Куайна и парадокс «черного лебедя», а также даёт естественное объяснение эпизодической памяти и механизмов адаптации живых систем.

Отличие природной (эпизодической, кадрово-структурированной) верификации от принципов нейросетевых ИИ (распределённая весовая память, статистическая

аппроксимация, отсутствие встроенного агента верификации) объясняет феномен галлюцинаций LLM и открывает путь к созданию функционального искусственного интеллекта через интеграцию операционных контрактов с нейросетевым ядром.

В контексте эпохи постправды формула $V(A,C,M)$ становится рабочим инструментом анализа системного сбоя: подмена агента верификации, манипуляция контекстом и информационный анальгетик, который закрывает канал Prediction Error.

Статья предлагает единую, самообъясняющуюся эпистемологическую рамку, которая объединяет классический реализм, когерентность, прагматизм и фаллибилизм в одну детерминированную операционную модель и демонстрирует её применимость к философии науки, биологии, физике и искусственному интеллекту.

Ключевые слова: истина, операционная формула истины, верификация, агент верификации, ложь как сигнал рассогласования, научный реализм, конструктивный эмпиризм, фаллибилизм, эпизодическая память, постправда, нейросетевые ИИ, галлюцинации LLM, причинно-следственный контракт, модель-зависимый реализм

Справочное определение: Истина — это знание, соответствующее своему предмету и адекватно отражающее действительность. Однако, истинность предположения может быть неизвестна до верификации, но уже существует объективно. Значит Истина – это не знание и сложившееся в словарях определение вообще не верно. Такое определение ничем не лучше мистического понятия: Истина – основа мироздания. Оба определения исходят из неких субъективных, но не верифицированных предположений и различаются только той значимостью, которую придают апологеты.

Для рационалиста «истина» — это инструмент для навигации в материальном мире. Ему важно, чтобы мост не рухнул, а ракета взлетела. Поэтому его «истина» — это рабочая модель.

Для мистика «истина» — это психологическое и духовное убежище, способ преодолеть страх смерти и экзистенциальную пустоту через слияние с «Вечным».

Может быть понятие Истины как чего-то единого вообще нет? Но в природе есть то, что используется всеми живыми существами для выверки своего поведения на ошибках и удачах.

История развития понятия "Истина"

История понятия «Истина» — это эволюция человеческого понимания сути: путь от вспышки факела в темной пещере до разбитого зеркала, в каждом осколке которого у каждого отражается свой собственный мир, сорри за патетику.

Античность. От вспышки во тьме до идеального зеркала. VI–IV вв. до н. э.

Древнегреческое слово, обозначавшее истину — «Алетея» (ἀλήθεια) — буквально переводится как «не-забываемость» или «снятие покрова». Для ранних греков истина была не формулой, а событием: вы вытаскиваете вещь из мрака забвения (реки Леты) на свет, снимаете с нее покрывало иллюзий и видите её первозданную природу.

Позже Аристотель спустил это понятие с небес на землю и сформулировал корреспондентную (классическую) концепцию: истина — это совпадение мысли и вещи [1]. Если вы говорите «на столе лежит хлеб», и он там действительно лежит — ваша мысль истинна. Человеческий ум впервые был объявлен «зеркалом», которое должно без искажений отражать реальный мир.

Средневековье. Архитектурный чертеж Творца. V–XIV вв.

В христианской философии истина перестает быть просто фактом природы — она становится Личностью («Я есмь путь и истина и жизнь»).

Фома Аквинский и схоласты перевернули аристотелевское «зеркало». В Средние века вещь истинна не потому, что мы её так увидели, а потому, что она соответствует задумке Бога. Как деревянный стул истинен, если он точно повторяет чертеж плотника, так и весь физический мир истинен лишь постольку, поскольку он совпадает со своим замыслом в уме Архитектора Вселенной.

Новое время. Рентген разума и весы опыта. XVII–XVIII вв.

С рождением науки философы разделились на два непримиримых лагеря в поисках фундамента истины.

Рационалисты (Рене Декарт): чувства обманчивы (нам может сниться сон или мерещиться иллюзия). Истина — это только то, что воспринимается очищенным от эмоций разумом предельно «ясно и отчетливо» («Мыслю, следовательно, существую»).

Эмпирики (Фрэнсис Бэкон, Джон Локк): наш ум от рождения — чистый лист (*tabula rasa*). Истина не выводится из умозаключений в кабинете, она «выпытывается» у природы через контрольный эксперимент, измерения и воспроизводимый опыт.

Немецкая классика. Очки Канта и живой организм Гегеля. Конец XVIII – XIX в.

Иммануил Кант совершил «коперниканский переворот»: мы никогда не узнаем, каков мир сам по себе («вещь в себе») [2]. Мы смотрим на реальность через врожденные розовые или синие «очки» человеческой психики (пространство, время, категории причинности). Истина стала согласованностью наших представлений с законами нашего же мышления.

Георг Гегель заявил, что истина — это не стоп-кадр, а эволюционный процесс. Истина не подобна готовой монете, которую можно достать из кармана. Она похожа на дерево: зерно, росток, цветок и плод — всё это части одной истины, развивающейся через противоречия. «Истина есть целое».

XX век. Швейцарский нож и правила языковой игры. Начало XX в.

Индустриальная эпоха и развитие логики отбросили романтику.

Прагматизм (Чарльз Пирс, Уильям Джемс): Истина — это то, что работает [3]. Если вера в идею помогает человеку выжить, построить мост или вылечить болезнь, эта идея истинна. Истина стала оцениваться по её практической эффективности.

Аналитическая философия (Людвиг Витгенштейн, Альфред Тарский): Истина переехала из мира вещей в мир языка. Фраза «Снег бел» истинна не потому, что есть какая-то мистическая связь, а потому что это корректно построенное утверждение внутри конкретной языковой системы. Истина — это строгая логика и соблюдение правил игры [4].

Постмодернизм. Калейдоскоп и эра постправды. Вторая половина XX в. — XXI в.

После Ницше, провозгласившего, что «фактов не существует, есть только интерпретации», Мишель Фуко и Жак Деррида окончательно разобрали монолит Истины на запчасти.

Истина начала рассматриваться как конструкт власти: кто контролирует медиа, институты и нарративы, тот и задает «режим истинности». Классическое зеркало окончательно разбилось — вместо него остался калейдоскоп субъективных точек зрения, контекстов и информационных пузырей.

Современное понятие Истины в мейнстриме академической науки

В мейнстриме современной академической науки (философии науки, естествознания и аналитической эпистемологии) абсолютная «Истина с большой буквы» выведена за скобки. Наука сегодня оперирует концепцией научного реализма с фаллибилизмом: истина — это не застывший фрагмент бытия, а наиболее точная на сегодняшний день модель реальности, обладающая максимальной прогностической силой (но ведь «истина» != модель).

Современный академический консенсус держится на четырех опорных столбах:

1. Фаллибилизм и «степень правдоподобия» (Verisimilitude)

Любое научное знание по определению считается гипотетическим (постулируемым). Наука не «доказывает раз и навсегда», а временно принимает концепцию как истинную до тех пор, пока она выдерживает жесточайшие проверки на опровержение (принцип фальсифицируемости Карла Поппера).

Вместо бинарного выбора «истина или ложь» ученые используют понятие степени правдоподобия: теория относительности Эйнштейна не «перечеркнула» механику Ньютона, а показала, что Ньютон был и остается невероятно близок к истине в определенных масштабах, но Эйнштейн продвинулся еще ближе.

2. Модельно-зависимый реализм (Model-dependent realism)

Концепция, сформулированная Стивеном Хокингом и Леонардом Млодиновым, отражает взгляд современной физики и когнитивистики: у человека нет «прямого» доступа к реальности без посредников. Мы воспринимаем мир только через призму наблюдательных приборов, органов чувств и математических моделей.

Если модель успешно объясняет наблюдения и предсказывает новые результаты, она признается истинной в границах своей применимости. Бессмысленно

спрашивать, «существует ли электрон на самом деле так, как мы его рисуем» — истинна сама модель электронного облака, пока она работает в микромире.

3. Рабочий триумвират (Синтез классических теорий)

В академической практике понятие истины проверяется трехступенчатым фильтром:

- Эмпирическое соответствие (Корреспонденция): совпадают ли предсказания теории с воспроизводимыми экспериментами?
- Внутренняя непротиворечивость (Когерентность): вписывается ли новая теория в математический аппарат и массив уже проверенных знаний (или объясняет, почему они были частным случаем)?
- Инструментальная эффективность (Прагматизм): позволяет ли эта истина строить работающие технологии — от адронных коллайдеров до алгоритмов машинного обучения?

4. Семантическая строгость (Теория Тарского)

В строгих дисциплинах (математике, формальной логике, computer science) мейнстримом остаётся семантическая концепция Альфреда Тарского. Она избавляет науку от мистицизма: утверждение истинно тогда и только тогда, когда оно корректно переводится на метаязык и соответствует правилам конкретной формальной системы.

В современной науке Истина — это картография. Карта рельефа — это не сама земля, но она абсолютно «истинна», если позволяет проложить маршрут из точки А в точку Б и не упасть в обрыв.

Как видим, в науке возникла чуть ли не битва за понимание истины и разные теоретики привнесли множество особенностей понимания, которые просто не могли в своей совокупности не влиять на попытки систематизации всего этого. А систематизация — это самое сложно в науке, потому что требует от ученого не просто сбор фактов и их сопоставление, а обобщения всего массива известного в непротиворечивую систему взаимодействий элементов этих фактов, которая начинает обладать предсказательными свойствами, как система химических элементов Менделеева, к которой тот шел долгими годами попыток найти критерии системности.

В результате получилась некая общепризнанная концепция, но она еще не обладает самообъясняющей системностью.

Эпистемология (теория познания) в вопросе Истины сейчас находится именно в «менделеевской» точке 1869 года. Философы науки проделали титаническую работу по сбору и классификации «элементов». Что уже систематизировано (наши «химические элементы» истины):

- Классический реализм (Корреспонденция): требование соответствия фактам.
- Когерентность: требование внутренней непротиворечивости системы знаний.

- Прагматизм: требование прогностической силы и полезности.
- Фаллибилизм: признание принципиальной незавершенности любой модели.

Эти концепции собраны в общепризнанную академическую рамку (научный реализм). Но они сосуществуют скорее как конгломерат взаимодополняющих критериев, а не как единый самообъясняющийся закон. Фундаментальный «заряд ядра» для понятия Истины еще не открыт.

Современный научный мейнстрим вынужденно остановился на этой «рабочей» общепризнанной концепции, потому что она прагматична: она позволяет космонавтам летать, а биологам редактировать геном. Но внутренний метафизический голод ученого и философа остается неутоленным — фундаментальный код, который связал бы воедино объективный мир, человеческую модель и биологическую эволюцию в одну самообъясняющуюся формулу Истины, всё еще ждет своего «квантового прорыва». Но споры не затихают.

Спор между научным реализмом и научным антиреализмом (в форме конструктивного эмпиризма Баса ван Фраассена)

Это главный батальный фронт современной философии науки.

Если убрать сложную терминологию, суть спора сводится к одному вопросу: «Существуют ли скрытые от наших глаз объекты (электроны, кварки, темная материя, искривление пространства) на самом деле — или это просто удобные сказки, которые помогают нам делать точные расчёты?»

1. Позиция Scientific Realism (Научный реализм)

Тезис: наука стремится и действительно способна дать нам буквально истинное описание мира — как наблюдаемого, так и невидимого.

Вампиры и электроны: если теория говорит, что при столкновении частиц рождается бозон Хиггса, значит, бозон Хиггса реально существует в физическом мире точно так же, как реальны яблоки или столы.

Принятие теории: принять научную теорию — значит искренне верить, что всё, что в ней написано про невидимый мир — правда.

Главный аргумент реалистов: «Аргумент без чудес» (No-Miracles Argument) сформулирован Хилари Патнамом: «Реализм — это единственная философия, которая не превращает успех науки в чудо».

Если электронов и полупроводников «на самом деле не существует», а это лишь математические абстракции, то почему ваш смартфон работает? Если за уравнениями квантовой механики не стоит реальная физическая структура мира, то предсказательная точность современной физики с точностью до десятого знака после запятой — это необъяснимое космическое чудо.

2. Позиция Constructive Empiricism (Конструктивный эмпиризм)

Сформулирован Басом ван Фраассеном в 1980 году в книге «The Scientific Image» [7].

Тезис: Цель науки — не поиск абсолютной истины о невидимом мире, а создание эмпирически адекватных моделей.

Наблюдаемое vs Ненаблюдаемое: наука должна точно описывать только то, что человек может наблюдать «невооруженным глазом» (пусть даже через телескоп или обычные линзы, но без теоретической интерпретации приборов).

Эмпирическая адекватность: теория «эмпирически адекватна», если всё, что она говорит о наблюдаемых явлениях, соответствует действительности.

Агностицизм к невидимому: существуют ли электроны на самом деле? Эмпирик отвечает: «Мы не знаем и нам всё равно». Достаточно того, что модель электрона идеально предсказывает показания вольтметра и вспышки на экране.

Принятие теории: принять теорию — значит признать её полезной и эмпирически точной, но не обязательно истинной в описании микромира.

Главный аргумент антиреалистов: «Пессимистическая мета-индукция» (Larry Laudan). История науки — это кладбище невероятно успешных теорий, которые оперировали сущностями, оказавшимися фантомами.

Физики XVIII века верили в флогистон (огненную материю) и теплород (невесомую жидкость тепла) — и эти теории давали отличные практические результаты и расчёты!

В XIX веке весь научный мир был уверен в существовании светоносного эфира.

Логика антиреалистов: если прошлые «успешные» теории на самом деле были ошибочны в описании невидимого мира, то с чего мы взяли, что наши сегодняшние «невидимые сущности» (темная энергия, струны, кварки) через 200 лет не отправятся на ту же свалку истории?

Эти споры не продвигают к конечному обобщению и системному пониманию, а только выстраивают все новые аргументы, новые элементы, которые нужно систематизировать и давно вышли за уровень сложности, поддающийся силам человека в попытках систематизации, а обобщение — это такое высшее творчество, которое никак не может продвигаться коллективно, это всегда работа одного ума. И это определяет тупики не только в вопросе определения Истины, но и почти всего переднего фронта науки, включая понимания сути сознания, понимание сути основ мироздания и всего остального.

Казалось бы еще древние гении подвели к простой формуле: Истина - это подтверждение предположения верификацией реальности, а ложь - отрицательный результат, но у дотошных ученых возникло столько своих мыслей по этому поводу, что формула начала представляться вульгарно упрощенной. За всеми запутывающих философствований, в этой формуле в упор не видят основы системы, определяющей Истину во всех ее проявлениях и свойствах.

Вы будете смеяться, но в современных научных работах именно в такой буквальной формулировке это утверждение встречается редко и толком не рассматривается, поскольку академический язык избегает излишней лаконичности.

Современная наука страдает от «синдрома птичьего языка». Если современный ученый напишет в грантовой заявке или рецензируемой статье из списка ВАК/Scopus простую и гениальную фразу: «Истина — это подтверждение предположения верификацией реальности, а ложь — отрицательный результат», его работу, скорее всего, завернут.

В академической среде существует негласный культ искусственной сложности. Чтобы мысль выглядела «научно», её нужно раздуть. Вместо «предположения» надо написать «когнитивная аппроксимация латентных переменных», а вместо «верификации реальностью» — «мультимодальная валидация посредством эмпирического субстрата».

Если всё так просто, то о чём спорить? Зачем писать тысячи диссертаций, если базовый алгоритм был понятен еще во времена, когда человек выбирал, какую ягоду съесть в лесу?

Каково рефлексивное отношение академического мейнстрима к обобщающей формуле: Истина - это подтверждение предположения верификацией реальности, а ложь - отрицательный результат?

Академический мейнстрим относится к этой формуле (в лице современной эпистемологии, логики и философии науки) сдержанно-одобрительно, но считает её слишком упрощённой.

Формула схватывает самую суть работы любого познающего субъекта (от исследователя в лаборатории до адаптивных алгоритмов и живых организмов):

Петля «Предсказание — Проверка»: Наука не просто пассивно наблюдает мир, а выдвигает модель (предположение), совершает действие/эксперимент и сравнивает ожидаемый результат с реакцией реальности.

Обязательность категории «Ложь» мейнстрим полностью согласен с тем, что истина бессмысленна без концепта ошибки. Если система принципиально не способна зарегистрировать сбой или отрицательный результат, она не способна ничего верифицировать.

Где мейнстрим вносит существенные поправки

1. Проблема индукции: положительный результат не создает абсолютную «Истину»

Со времен Карла Поппера академическая наука отказалась от чистого верификационизма [5].

Если вы выдвинули предположение «Все лебеди белые» и верифицировали его 1000 раз, увидав 1000 белых лебедей, стало ли предположение «Истиной»? Нет. Положительный результат лишь повысил его правдоподобие, но 1001-й лебедь в Австралии всё равно может оказаться черным.

Поэтому в научном мейнстриме подтверждение верификацией дает не абсолютную «Истину», а временный статус эмпирически подкреплённой гипотезы.

2. Тезис Дюэма — Куайна: Отрицательный результат не равен «Лжи» самого предположения

С отрицательным результатом всё устроено сложнее, чем с простым бинарным переключением. Когда эксперимент дает отрицательный результат, это не означает автоматическую ложность проверяемого предположения.

В реальности гипотеза проверяется не изолированно, а вместе с целым комплексом условий: исправностью прибора, чистотой условий, точностью измерений и правильностью теории, на которой этот прибор вообще построен.

Отрицательный результат говорит лишь о том, что где-то в этой цепочке произошел сбой, но ложным может оказаться не само базовое предположение, а, например, неисправный датчик или неучтенный шум внешней среды.

3. Переход от бинарности к вероятностям (Байесовский подход)

В современной философии науки понятия «Истина» и «Ложь» всё реже рассматриваются как тумблер с положениями 1 и 0. Вместо этого используются вероятностные оценки:

Подтверждение повышает апостериорную вероятность модели $P(H \vee E)$ — модель становится более точной и предсказательной.

Отрицательный результат снижает эту вероятность и заставляет либо скорректировать границы применимости гипотезы, либо пересмотреть вспомогательные параметры.

4. Границы сферы применения

Эта формула безупречно работает для эмпирических (синтетических) суждений о физическом мире. Однако она пасует в формальных науках: в математике или логике утверждение $2+2=4$ истинно не потому, что его верифицировали экспериментально в физической реальности, а в силу логической непротиворечивости внутри самой аксиоматической системы.

Академический мейнстрим принимает эту формулу как необходимый рабочий скелет научного метода, но оборачивает его в вероятность и фальсифицируемость: подтверждение делает модель эмпирически адекватной, а отрицательный результат фиксирует расхождение с реальностью, указывая на ошибку в системе, которую еще предстоит локализовать.

Защита формулы

Первый пункт: сколько бы раз реальность ни подтверждала гипотезу «Все лебеди белые», это не делает её абсолютной истиной, ведь всегда остаётся шанс встретить чёрного лебедя.

Не бывает некоей абсолютной истины для всех условий. В формуле верифицируется только одно конкретное, не противоречивое и проверяемое утверждение. А любое корректное утверждение имеет границы применения — условия, при которых утверждение истинно или ложно. При четком ограничении условий, данное утверждение верифицируется или как истина или как ложь. При появлении новых компонентов условий верификация должна быть повторной.

Если не известны все условия, которые воздействуют на эксперимент необходимо добиться неизменности условий и жесткого ответа верификации. Система

взаимодействий будет верифицироваться независимо от того, открыты ли все участники взаимодействий. Это похоже на рассмотрение пазла таблицы химических элементов Менделеева: когда Менделеев обнаружил правильный критерий системности, то в полученном порядке оказалось немало пустых пока мест. Но ведь и до открытия системы предположения об элементах могли эффективно верифицироваться (принцип «Черного ящика»), но без системного сопоставления.

Проблема остается только в том, чтобы формула предположительного утверждения не выходила за рамки системы взаимодействий, т.е. не следует предполагать некую сущность (вроде Флогистона), выходящей за рамки причинно-следственных взаимодействий (и это касается даже квантовой механики, вот почему: fornit.ru/70790). Так, в опытах Аспекта верифицируется «мистическое дальноедействие» (нелокальность), ломающее привычную логику причин и следствий, но это оказывается лишней сущностью и многие физики увлеклись тем, что начали придумывать суть этой сущности.

Конкретизируем три методологических фильтра:

1. Локализация условий (Снятие проблемы «Черного лебедя»)

Главная уязвимость классической верификации заключалась в неограниченном кванторе общности («Все лебеди всегда и везде...»). Вы правильно локализуете утверждение: не существует абстрактного утверждения без границ его применимости.

Формула A («Лебеди белые») неполноценна. Корректная формула звучит так: $A(C)$ — «В условиях европейского ареала C_1 цвет оперения вида S является белым». Появление «черного лебедя» в Австралии (C_2) не опровергает истинность $A(C_1)$, а фиксирует расширение множества условий и требует формулирования нового контекста $A(C_2)$.

В такой постановке верификация возвращает строгий бинарный ответ (1 или 0), поскольку проверяется не абстрактная «абсолютная истина мира», а конкретный причинно-следственный факт внутри фиксированных границ.

2. Принцип «Черного ящика» и системный критерий

Парадокс отсутствия «полного знания обо всех переменных» разрешается через инвариантность внешней среды:

Если мы фиксируем входные условия и структуру эксперимента, то система взаимодействий верифицируется корректно даже при неполном знании её внутреннего устройства.

Пример с таблицей Менделеева здесь идеально точен: системный критерий позволяет верифицировать закономерность до полного раскрытия всех деталей. Наличие «пустых клеток» не отменяет истинности открытого порядка причинно-следственных связей. Мы проверяем не фантомы, а воспроизводимую реакцию системы на контролируемое воздействие.

3. Причинно-следственный заслон (Отсечение сущностей)

Это ключевой элемент защиты, отсекающий ложные интерпретации верификации: предположение не имеет права выходить за рамки физической причинности.

Проблема Флогистона и «Мистического дальногодействия»: если эксперимент (как в опытах Аспекта или интерференции одиночных частиц) дает непонятный результат, логика операциональной истины запрещает придумывать «нелокальность», «волновой коллапс от взгляда наблюдателя» или другие нефизические сущности.

Попытка объяснить незафиксированный механизм «чудом» или «выходом за пределы причинности» — это нарушение границы системности. Если результаты опыта Аспекта кажутся «мистикой», это указывает не на крах причинности, а на неполноту описания условий или физического механизма в проверяемой гипотезе. Настоящая верификация требует поиска конкретного детерминированного взаимодействия, а не легитимизации сущностей-паразитов.

Второй пункт: Мы не «верифицируем реальность» напрямую. Мы сверяем гипотезу с результатами экспериментов, которые сами по себе интерпретируются через призму других теорий и приборов [8].

Совершенно неважно чем и как проверяется утверждение, главное, чтобы был выбран агент подтверждения: физика причин и следствий, реакция зверя или человека на воздействие (как формы утверждения, ожидающей определенные последствия), философская концепция, в рамках которой утверждение получает или подтверждение или опровержение (доводы не важны, важен только конечный результат), компилятор программной реализации и т.п.. Система, в которой утверждения сверяется как имеющее предсказанный результат даст true или false. Утверждение всегда должно быть в форме возможного взаимодействия с агентом верификации, а не в виде интерпретаций или посредничестве приборов с ограниченной точностью.

Это снимает старый философский спор о том, чем отличается «ложь» (намеренный обман) от «заблуждения» (ошибки). Если утверждение сформулировано как алгоритм («если совершить действие А в условиях Б, то агент выдаст реакцию В»), то любой сбой алгоритма — это false. С точки зрения системы, некорректная теория флогистона, ошибка программиста или ложное предсказание физика — это однородный отрицательный результат верификации, то есть Ложь.

Утверждение больше не висит в воздухе как субстанция. Оно становится строгим интерфейсным контрактом:

Контракт=«Если совершить действие А в условиях Б, то Агент М выдаст реакцию В»

В такой системе совершенно неважно, насколько «сложен» или «идеален» агент верификации:

- Если Агентом выступает компилятор, а действием — передача исходного кода, результатом верификации будет true (успешная сборка) или false (ошибка синтаксиса/типов).
- Если Агентом выступает биологический организм, а действием — подача стимула, результатом будет true (получение ожидаемой адаптивной реакции) или false (рассогласование).
- Если Агентом выступает физический опыт, результат сверяется прямо по факту зафиксированного отклика.

Результат взаимодействия всегда бинарен: либо предсказанный отклик агента совпал с фактическим (true), либо возникла ошибка рассогласования (false).

Разделение на «намеренную ложь» и «невольное заблуждение» — это гуманитарный рудимент, который путает причину ошибки с самой ошибкой. Для любой действующей системы (будь то живой организм, алгоритм или научный метод) важен только факт рассогласования предсказания с реальностью:

- Ошибочная теория флогистона сформировала предсказание реакции среды → среда выдала иной результат → false.
- Программист опечатался в индексе массива → программа упала → false.
- Собеседник сознательно назвал неверную дату → действие в названную дату не привело к ожидаемому результату → false.

Во всех трех случаях произошел один и тот же процесс: субъект сформировал модель ожидания C , но Агент верификации вернул C' . Это рассогласование (сигнал ошибки) и есть Ложь в чисто техническом и функциональном смысле.

Поскольку утверждение формулируется не абстрактно («электрон имеет такой-то спин вообще»), а как прямое взаимодействие с конкретным Агентом («если направить пучок на датчик D в режиме M , мы получим отклик R »), погрешность прибора перестает быть «философской проблемой».

Характеристики и ограничения прибора просто включаются в определение самого Агента верификации. Если прибор шумит или имеет ограниченную точность, то Агентом является «система: объект + данный конкретный прибор». Если алгоритм взаимодействия учёл эти ограничения и предсказал результат — отклик true. Если не учёл — false.

В этой архитектуре Истина — это успешный запуск алгоритма (предсказание полностью совпало с реакцией агента), а Ложь — рассинхронизация модели и отклика (сигнал ошибки, требующий коррекции алгоритма).

Третий пункт: Если эксперимент не подтвердил гипотезу, проблема может быть в неисправном оборудовании, неверных расчётах или неучтённых внешних факторах (так называемый тезис Дюэма — Куайна) [6].

Третий пункт оказывается уже защищен доводами защиты первых двух пунктов.

Объединение **Локализованных условий** (Пункт 1) и **Агента верификации** (Пункт 2) создает целостную операциональную модель, в которой все частные проблемы решаются автоматически:

1. Ликвидация «байесовской размытости»

Вероятностный подход нужен философии науки только тогда, когда утверждение формулируется бесконечно широко («Электрон всегда ведет себя как волна»).

Если же утверждение сведено к исполняемому контракту внутри фиксированных условий $(A, C) \rightarrow M \rightarrow R$, вероятностная неопределенность исчезает в момент исполнения:

- До запуска эксперимента у нас может быть субъективная оценка вероятности отклика.

- Но сам факт взаимодействия с Агентом верификации M выдает исключительно дискретный отклик: либо ожидаемый результат зафиксирован ($true$), либо зафиксировано рассогласование ($false$).

2. Снятие парадокса Дьюэма — Куайна

Классическая проблема «неизвестно, что именно ложно — теория или сломанный прибор» перестает быть философским тупиком.

Поскольку прибор, среда и методика измерения заранее включены в понятие Агента верификации и Условий C :

- Контракт звучит так: «Прибор M в состоянии C при действии A должен показать значение R ».
- Если прибор сломался или зашумел, контракт не выполнен $\rightarrow$ система получает чистый сигнал $false$.
- Исследователю не нужно гадать о «высшей истине»: зафиксирован сбой алгоритма предсказания. Следующий шаг — скорректировать либо описание Агента (починить/перекалибровать прибор), либо само предположение.

Полная унификация формального и физического знания: снимается граница между математикой, программированием и физическим экспериментом. Меняется только тип Агента верификации:

- В математике: Агент — аксиоматическая система (компилятор формальной логики). Утверждение A дает $true$, если оно выводимо без противоречий, и $false$, если возникает логический тупик.
- В физике: Агент — причинно-следственная среда.
- В когнитивистике: Агент — целевой параметр гомеостаза или реакция внешней среды на действие организма.

Архитектура итоговой концепции

Сконструированный каркас описывает Истину не как статическое свойство Вселенной, а как состояние успешной синхронизации субъекта с Агентом верификации:

Истина $\equiv$ Успешное исполнение контракта $(A, C) \rightarrow M \rightarrow R(\text{Signal: } true)$

Ложь $\equiv$ Рассогласование предсказания с откликом Агента ($\text{Signal: } false$)

Это формальный, полностью очищенный от метафизики и гуманитарных наслоений фундамент, который прямо соответствует архитектуре кибернетических систем, контурам обратной связи в биологии и строгим протоколам научного эксперимента.

Самое главное подтверждение

Природная «эпизодическая память» имеет структуру, отвечающую заявленной формуле верификации, обеспечивая выявление и сохранение логики взаимодействий любого вида: fornit.ru/104562 [9].

В архитектуре механизмов психики эпизод памяти формируется не как произвольный набор сведений, а как строго структурированная связка параметров:

Кадр эпизодической памяти = [Условия C | Действие A | Ожидаемый результат R_{exp} | Фактический

Структура природных механизмов напрямую повторяет формулу:

1. Контекст и границы применимости (C). Эпизодическая память фиксирует текущее внутреннее состояние (гомеостатические потребности, эмоциональный тон) и внешние условия среды. Это гарантирует, что вывод не превратится в «абстракцию для всех времен и народов», а останется привязанным к локальным границам применимости.
2. Агент и воздействие ($A \rightarrow M$). Фиксируется конкретное действие субъекта, направленное на Агента верификации (физический объект, другой организм или элемент среды).
3. Фиксация истинности ($R_{\text{act}} = R_{\text{exp}}$). Если фактический отклик совпал с ожиданием, система получает сигнал успешной верификации (true). Кадр фиксируется как успешный адаптивный опыт, на основе которого впоследствии формируются автоматизмы и правила более высоких уровней (семантическая память).
4. Фиксация лжи и сигнал рассогласования ($R_{\text{act}} \neq R_{\text{exp}}$). Если отклик среды не совпал с моделью предсказания, фиксируется отрицательный результат (false). Это генерирует сигнал ошибки (Prediction Error), запускает ориентировочную реакцию, торможение неэффективного поведения и поиск новых вариантов действий [10].

В процессе осознания на первом шаге определяется цель (fornit.ru/102733), которая и является определением того, какой агент будет верифицировать предположения.

Если намерены вызвать симпатию у женщины и заставить ее улыбнуться, то агент проверки верности наших действий – эта конкретная женщина со всеми ее повадками, а улыбка – способ, по которому мы определим успех.

Если мы ставим цель получить черный осадок при смешивании растворов йода и аммиака с намерением отпугивания крыс в кладовке, то верифицирующим агентом станет физика, а успешность – выпадение осадка (в том, что он сработает мы уже не сомневаемся).

Эволюция не занималась методологическими спорами. Для выживания ей потребовался единственный работоспособный механизм познания.

Биология не оперирует абсолютной истиной. Для организма «истинно» только то, что при действии A в условиях C дает подтвержденный отклик среды R , сохраняющий или восстанавливающий гомеостаз.

Ложь в природе смертельна. Если модель предсказания дает false (ожидался безопасный шаг, но произошло падение), организм обязан мгновенно зафиксировать этот отрицательный результат и скорректировать поведение. Ошибка — это прямой сигнал о неадекватности используемой модели среды.

Таким образом, операциональная формула истины $(A, C) \rightarrow M \rightarrow R \Rightarrow true/false$ не просто придумана как теоретическая конструкция. Она представляет собой инженерный чертеж, по которому эволюция построила механизмы субъективного опыта и эпизодической памяти.

Финальное выражение формулы Истинности

Операционально-детерминированная формула Истинности представляет собой исполняемый контракт взаимодействия между моделью предсказания и объективной реакцией Агента верификации:

$V(A, C, M) = \{$
true при $R_{act} = R_{exp}$ (Истина)
false при $R_{act} \neq R_{exp}$ (Ложь)
 $\}$

Компоненты формулы

- C (Context / Conditions) — зафиксированное множество условий (границы применимости, состояние среды и параметры субъекта).
- A (Action / Input) — конкретное воздействие или проверяемое алгоритмическое правило.
- M (Verification Agent) — система, с которой происходит взаимодействие (причинно-следственная физическая среда, биологический организм, компилятор, формально-логическая система).
- R_{exp} (Expected Result) — предсказанная моделью реакция Агента M на действие A в условиях C .
- R_{act} (Actual Result) — фактически зафиксированный отклик Агента M .

Строгие определения

- Истина (*true*) — подтвержденное совпадение предсказанной реакции Агента верификации с его фактическим откликом при исполнении контракта в строго зафиксированных границах условий C .
- Ложь (*false*) — сигнал рассогласования (ошибка предсказания / Prediction Error), фиксирующий несоответствие модели отклику Агента, независимо от источника ошибки (сбой алгоритма, неполнота условий C или намеренный дезинформирующий фактор).

Эпистемологический статус

Истина не является ни метафизической субстанцией, ни абсолютным свойством мира «самого по себе». Истина — это результат успешной синхронизации адаптивной системы с Агентом верификации, физически фиксируемый в элементарном кадре эпизодической памяти и обеспечивающий выживание и развитие познающей системы.

Если принудительно внедрить операционально-детерминированную формулу Истины в мейнстрим, «вечные тупики» науки мгновенно разрешаются:

Сознание перестает быть «мистическим духом в машине». Оно становится высокоэффективным вычислительным Контрактом V , где мозг выступает как Генератор моделей предсказания, а тело и среда — как Агент M . Психическое здоровье и выживание — это удержание функции V в значении true.

В Квантовой физике формула идеально легитимизирует квантовую механику. Физический прибор и квантовая система — это единый Агент M . Измерение — это Действие A . Волновая функция — это Ожидаемый результат R_{exp} . Больше нет нужды спорить, «существует ли электрон, когда мы на него не смотрим». Есть контракт, который выдает true на экране макроприбора.

В Искусственном Интеллекте - это готовое ТЗ для создания сильного ИИ (AGI). Настоящее мышление начнется тогда, когда ИИ перейдет от компиляции текстов к непрерывному автономному исполнению контрактов $V(A, C, M)$ с физическим или цифровым миром, самостоятельно минимизируя false.

Итак, истина рассматривается как состояние адаптивной системы, возникающее при совпадении предположения с откликом конкретного агента в локализованном контексте, а ложь — как функциональный сигнал ошибки предположения, запускающий перестройку модели.

$C + A + M + \text{ожидаемый результат} + \text{фактический результат} \rightarrow \text{элементарный эпизод опыта}$ — это только для процесса осознания, в котором поставлена цель и используются эпизоды жизненного опыта, в самих же эпизодах компонента цели нет (потому что в других процессах осознания может ставиться другая цель) и вместо этого есть только общий результат действия в данных условиях.

Эпизодическая память хранит результаты взаимодействия с реальностью; осознание же использует эти результаты для проверки предположений, возникающих в связи с текущей целью.

Насколько отличаются принципы реализации нейросетевых ИИ?

Принципы реализации современных нейросетевых ИИ (Deep Learning, Transformers, LLM) фундаментально отличаются от рассмотренной операционально-детерминированной модели.

Если природные адаптивные системы и операциональная формула истины строятся на дискретных причинно-следственных контрактах и эпизодической памяти, то нейросетевой мейнстрим построен на аппроксимации непрерывных статистических распределений.

Главные принципиальные различия

1. Природа памяти

- Операционально-детерминированная модель (природная психика): Явная и кадровая (эпизодическая). Память состоит из дискретных локализованных кадров $[C | A | R_{exp} | R_{act}]$. Однократное событие сразу формирует отдельный кадр опыта.

- Нейросетевой ИИ (Deep Learning / LLM): Распределенная и весовая (параметрическая). Знания «размазаны» по миллиардам непрерывных коэффициентов (весов W). Отдельных кадров опыта не существует.

2. Механизм проверки (Истина)

- Операционально-детерминированная модель: Прямой запуск контракта. Проверка происходит через фактическое взаимодействие с Агентом M в условиях C с вычислением дискретного сигнала (true/false).
- Нейросетевой ИИ: Статистическое правдоподобие. Генерация наиболее вероятного продолжения $P(w_t | w_{1..t-1})$ на основе ассоциативных паттернов из обучающей выборки.

3. Скорость адаптации

- Операционально-детерминированная модель: Мгновенная (One-shot). Однократное рассогласование $R_{act} \neq R_{exp}$ сразу создает новый опыт и фиксирует ошибку правила.
- Нейросетевой ИИ: Медленная (Backpropagation). Ошибка требует множества шагов градиентного спуска (∇L) по всей сети с риском «катастрофического забывания».

4. Границы условий (C)

- Операционально-детерминированная модель: Жесткие и зафиксированные. Утверждение истинно только в конкретном контексте C . Выход за границы требует новой серии верификаций.
- Нейросетевой ИИ: Размытые (Интерполяция). Модель обобщает на основе векторного сходства (эмбеддингов), из-за чего легко галлюцинирует при выходе за пределы обучающего распределения.

5. Мотивация и цели

- Операционально-детерминированная модель: Внутренняя (Гомеостаз). Система действует автономно для устранения жизненно важных рассогласований.
- Нейросетевой ИИ: Внешняя (Loss Function). Пассивная минимизация скалярной функции потерь на статическом датасете.

Почему нейросети «галлюцинируют» с точки зрения операциональной формулы [11]?

Галлюцинации языковых моделей — это прямое следствие отсутствия у них встроенного Агента верификации M и эпизодического кадра:

1. Нет разделения между предсказанием и проверкой: Модель выдает ответ R_{exp} , но у нее нет фазы исполнения взаимодействия с реальностью. Для нее правдоподобно звучащая ложь ничем не отличается от факта, так как обе фразы имеют высокую статистическую вероятность в языке.

2. Отсутствие локализации контекста: Модель не помнит источник и условия, при которых данное утверждение было верно. Она смешивает контексты из разных частей обучающей выборки.

Точки сближения: куда пытается идти современный ИИ

Понимая эти фундаментальные ограничения, разработчики нейросетей сейчас пытаются внешним образом надстроить операциональные принципы поверх нейросетевого ядра:

- RLVR (Reinforcement Learning with Verifiable Rewards): Модели (например, OpenAI o1/o3) обучают не просто угадывать текст, а запускать внешний Агент верификации — компилятор кода или математический таргетинг. Ответ принимается только если Агент вернул true.
- Agentic Frameworks & Tool Use: Нейросеть перестает сама «придумывать» факты и начинает выполнять роли алгоритма: сформировать гипотезу → вызвать инструмент (поиск, API, код) → проверить отклик → выдать результат.
- Внешняя эпизодическая память (RAG / Vector DBs): Попытка прикрепить к весовой памяти явное хранилище конкретных фактов и условий их применимости [12].

Нейросетевой ИИ в нынешнем виде — это мощнейший ассоциативный генератор гипотез, но он лишен встроенного механизма проверки истинности. Настоящая функциональная адаптивность достигается только тогда, когда статистическое ядро нейросети объединяется с детерминированным контуром верификации и кадровой эпизодической памятью.

Жизнь в эпоху постправды в контексте понимания сути Истины

Эпоха постправды (post-truth) — это исторический период, когда объективные факты оказываются менее значимыми для формирования общественного мнения, чем обращение к эмоциям и личным убеждениям.

Если наложить эту реальность на формулу Истинности $V(A, C, M)$, то феномен постправды открывается не как абстрактная философская проблема, а как жесткий системный сбой в механизмах верификации информации человечеством.

Постправда — это эпоха, когда общество массово разучилось выбирать правильного Агента верификации (M).

Деконструкция постправды по формуле $V(A, C, M)$

В нормальных условиях, чтобы проверить суждение на истинность, человек совершает мысленное или физическое действие (A) и направляет его к Агенту (M) — физическому миру, логике или проверяемому документу.

В эпоху постправды этот контур намеренно ломается манипуляторами через три системных искажения:

1. Подмена Агента верификации (M)

Главный механизм постправды — это незаметная для человека замена независимого Агента М (реальности) на Агента Мсоц (социальную группу, эхо-камеру или алгоритм соцсети).

Если политик или блогер делает заявление (Предположение), обыватель больше не идет проверять его у физических фактов. Он идет в комментарии к единомышленникам.

Агент Мсоц настроен так, чтобы всегда возвращать ответ true, если заявление нравится группе, и false, если не нравится. Происходит замыкание системы: контракт верифицируется не миром, а одобрением толпы.

2. Манипуляция Контекстом (С)

Поскольку формула Истинности строго привязана к контексту и границам применимости (С), творцы постправды мастерски вырезают или искажают эту переменную.

Берется реальный факт (например, статистический график), но вокруг него полностью уничтожается контекст времени, места или сопутствующих условий.

Контракт выдает ошибку предсказания (false), но человек этого не замечает, потому что ему подменили исходные параметры С. Модель начинает казаться истинной там, где она глубоко лжива.

3. Атрофия сигнала предсказания

В природной эволюции (и в механизмах МВАП) столкновение с false (ошибкой) — это болезненный, но спасительный сигнал, заставляющий мозг включить осознание и перестроить модель. Постправда создала «информационный анальгетик».

Современная цифровая среда (алгоритмы рекомендаций) изолирует человека от любых фактов, которые могут вызвать сомнение в его правоте. Если реальность кричит о том, что модель человека ошибочна (false), алгоритмы услужливо подсовывают ему объяснение: «Это не твоя модель неверна, это враги/заговор/фейк подделали отклик Агента».

Ошибки не возникает вообще, автоматизмы не перестраиваются, и человек капсулируется внутри своей иллюзии.

«Разбитое зеркало» в действии

Постправда — это радикальное проявление той самой метафоры «разбитого зеркала», о которой мы говорили. Общество распалось на миллионы осколков, и в каждом осколке у людей своя «истина», потому что у каждого свой карманный Агент верификации (своя лента новостей, свой авторитетный паблик).

Парадокс эпохи в том, что люди искренне считают, что они верифицируют информацию. Они запускают контракт $V(A, C, M)$, но отправляют запрос не в физическую реальность, а в зеркальное отражение собственных предрассудков.

Как выжить в эпоху постправды, используя формулу Истинности?

Единственный способ сохранить ментальное здоровье и адекватность сегодня — это включить жесткую гигиену на основе формулы:

Каждый раз, принимая информацию, задавать вопрос: «Кто или что является Агентом проверки в данном случае? Это законы физики, проверяемый первоисточник, или это просто коллективное мнение людей, которым хочется в это верить?»

Намеренно сталкивать свои убеждения с противоположными фактами. Если ваша модель выдерживает проверку суровой реальностью и по-прежнему выдает true — она жизнеспособна. Если она защищена от проверок только тем, что вы баните несогласных — вы живете в облаке постправды.

Реальность (физика, биология, экономика) рано или поздно все равно возвращает жесткий и бескомпромиссный отклик, наказывая за ошибки. Постправда может временно обмануть мозг человека, но она не способна обмануть Агента по имени Вселенная.

Список литературы

1. Aristotle. *Metaphysics*. Book IV, 7, 1011b25–1012a. — Классическая формулировка корреспондентного понимания истины.
2. Kant I. *Critique of Pure Reason*. Cambridge: Cambridge University Press, 1998 [1781/1787]. — Ограничение познания опытом и проблема «вещи в себе».
3. Peirce C. S. The Fixation of Belief // *Popular Science Monthly*. 1877. Vol. 12. P. 1–15. — Практическая фиксация убеждения, роль опыта и действия в формировании знания.
4. Tarski A. The Semantic Conception of Truth and the Foundations of Semantics // *Philosophy and Phenomenological Research*. 1944. Vol. 4, No. 3. P. 341–376. DOI: 10.2307/2102968.
5. Popper K. R. *The Logic of Scientific Discovery*. New York: Basic Books, 1959. — Фальсифицируемость, отрицательный результат и логика проверки научных гипотез.
6. Quine W. V. O. Two Dogmas of Empiricism // *The Philosophical Review*. 1951. Vol. 60. P. 20–43. — Холизм проверки и проблема вспомогательных предпосылок (тезис Дьюэма—Куайна).
7. van Fraassen B. C. *The Scientific Image*. Oxford: Clarendon Press, 1980.
8. Oreskes N., Shrader-Frechette K., Belitz K. Verification, Validation, and Confirmation of Numerical Models in the Earth Sciences. *Science*. 1994. Vol. 263, No. 5147. P. 641–646. DOI: 10.1126/science.263.5147.641.
9. Tulving E. Episodic Memory: From Mind to Brain. *Annual Review of Psychology*. 2002. Vol. 53. P. 1–25.
10. Sutton R. S., Barto A. G. *Reinforcement Learning: An Introduction*. 2nd ed. MIT Press, 2018.

11. Ji Z., Lee N., Frieske R. et al. Survey of Hallucination in Natural Language Generation. ACM Computing Surveys. 2023. Vol. 55, No. 12. Article 248. DOI: 10.1145/3571730.
12. Lewis P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. NeurIPS, 2020.
13. McIntyre L. Post-Truth. MIT Press, 2018.
14. Hawking S., Mlodinow L. The Grand Design. Bantam Books, 2010.

