

ВСИ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

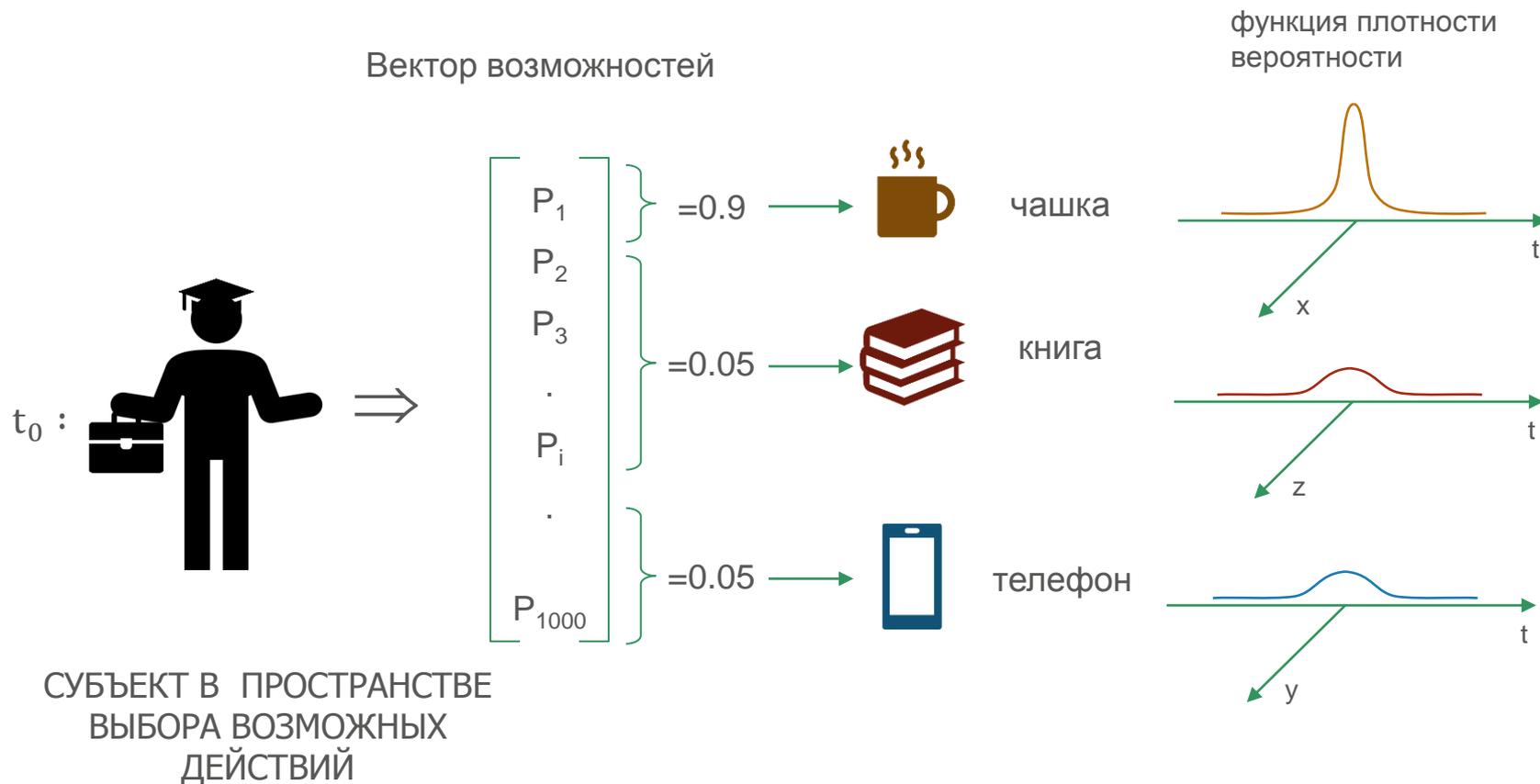
курс: Введение в профессиональную
деятельность

ЛЕКЦИЯ 13: МОЖНО ЛИ ПОНИМАНИЕ ЗАМЕНИТЬ
"ВЫЧИСЛЕНИЯМИ" СМЫСЛОВ

4.05.2023

- Повышения эффективности вычислений возможно на основе не только роста производительности аппаратных ресурсов или новых написания программ, но и
- построения системы «машинного обучения», которая позволяет оценить сложность решаемой задачи и компетентность пользователей в использовании возможностей компьютеров для вычислений
 - накопления «опыта» по настройке параметров имеющейся вычислительной системы для сокращения времени решения –
 - встраивания в систему управления СК ресурсами «средств интерператции» полученных результатов расчетов, включая оценку
 - числа используемых узлов кластера, вид ускорителей вычислений (GPU/FPGA), объем оперативной памяти и пр. на их успешное завершение задачи пользователя

- XI век до н.э. «Если значения слов не определены, то нет и смыслов. Если нет смыслов, то действия не происходят».
- (Конфуций).
- XVI век: «Определите значения слов, и вы избавите человечество от половины его заблуждений (ошибочных действий)
- (Рене Декарт).
- XX век. Реальное существование объекта – это результат разрушения суперпозиции его возможных (ВЕРОЯТНЫХ) состояний в следствии коллапса «волновой функции» и обретение действительного существования из возможных (ВЕРОЯТНЫХ) путем выбора в ПРОСТРАНСТВЕ сознания наблюдателя
-

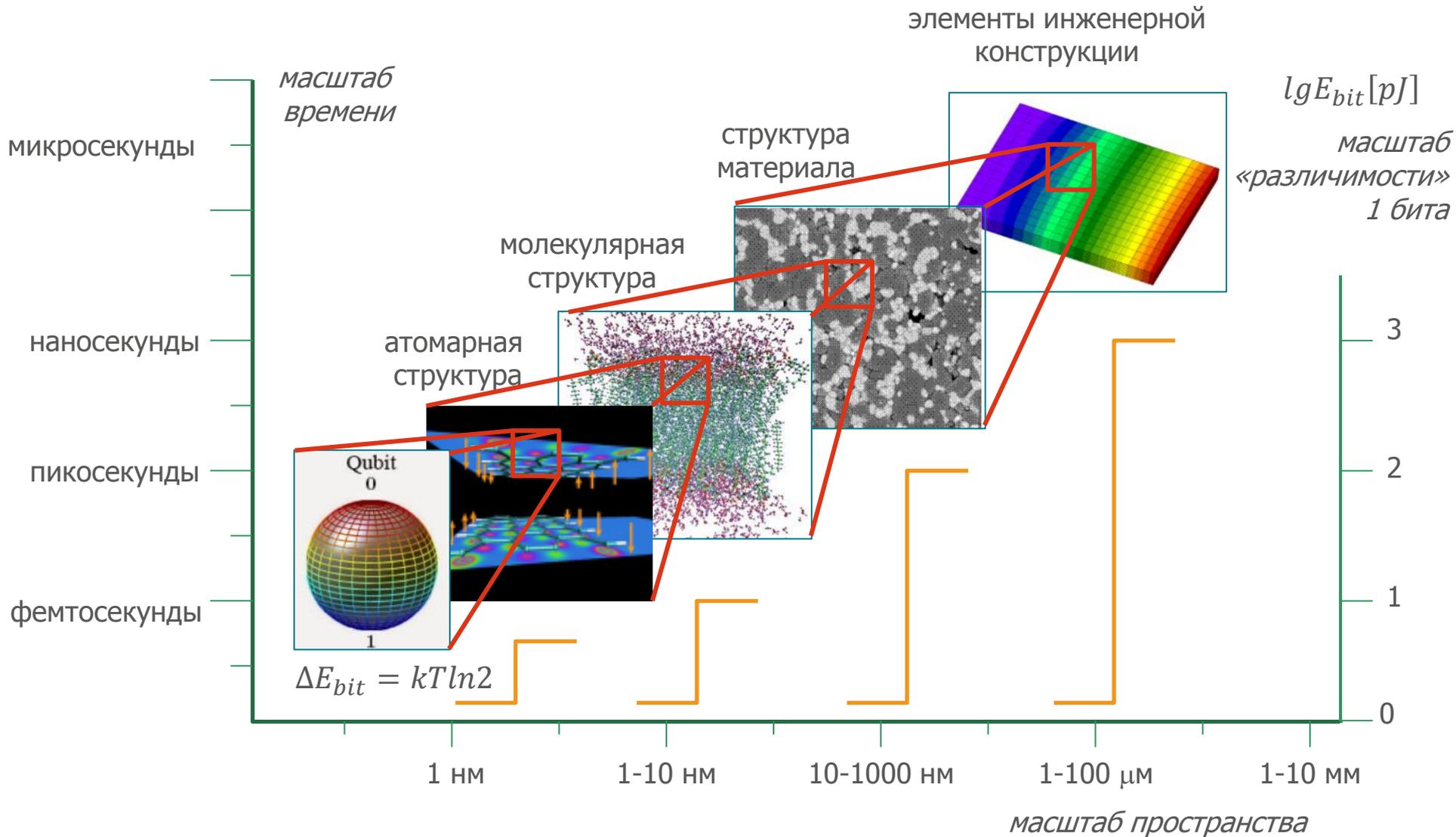


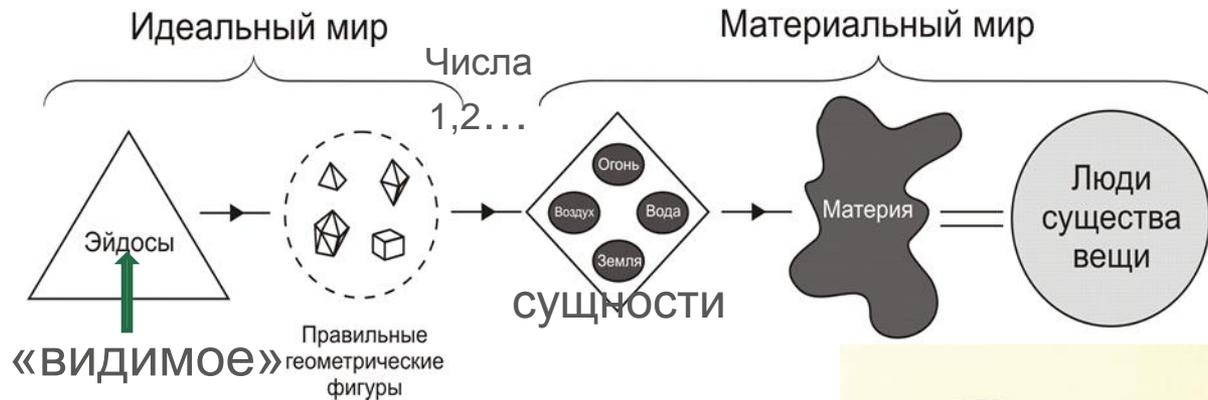
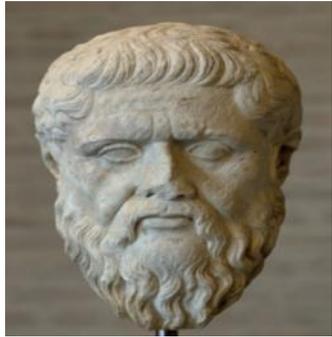
В разные моменты времени t -var вектор возможностей R изменяет значение и даже размерность. Изменяется и вид многомерной функции плотности вероятности «действий» субъекта.

Г. Фреге определил понимание **знаковых (комбинаторных) структур (текстов)** как **наделение их смыслом**



По Фреге понятие «смысл» связывает «**знак**» (текст) с предметом через способность «понимать». Смысл «знака» имеет контекстно зависимый характер (вероятностный ?)





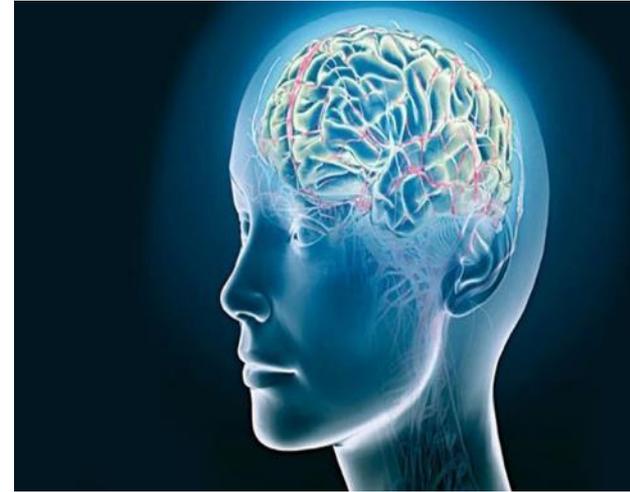
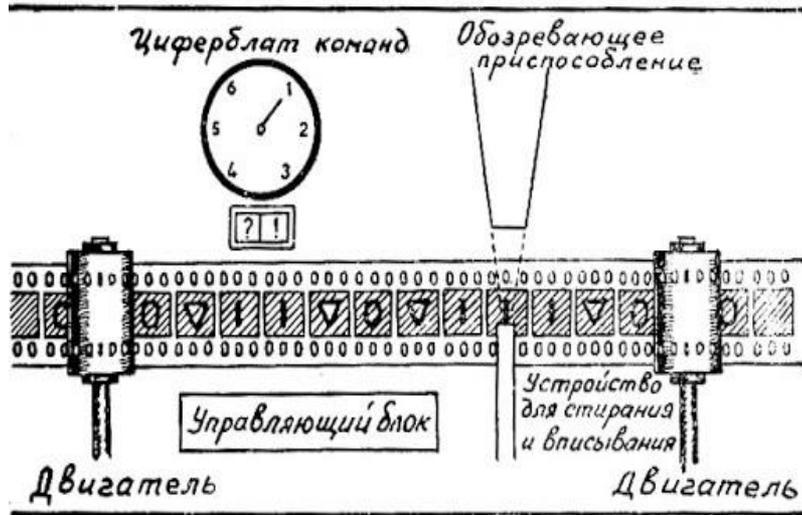
Платон: любая вещь лишь материальное отражение «первоначальной» идеи, познать которые можно только разумом.

И. Кант: Для достоверного знания доступны лишь явления в нашем сознании. Рассудок сам предписывает Природе её законы

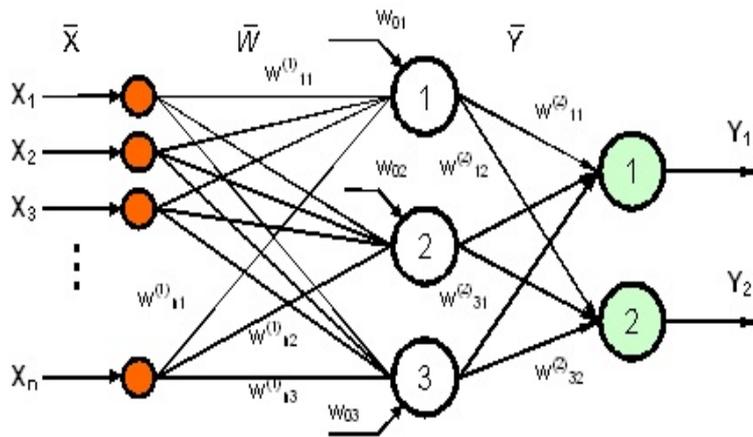
К. Гедель: Следует различать истинность высказывания и его доказуемость - не все функции даже натурального аргумента вычислимы.



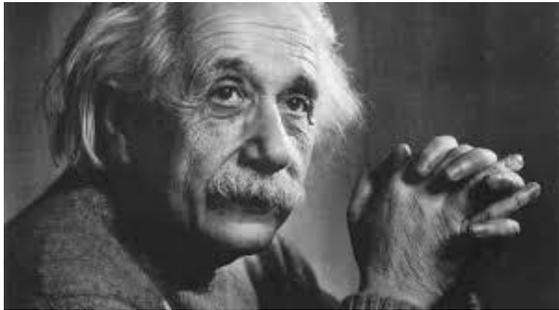
Эпименид утверждал: **все критяне — лжецы**, сам являясь критянином. Отрицание: 1) «Не все критяне — лжецы», 2) «Все критяне — не лжецы»



- Поисковая система Google обрабатывает в день около **24 петабайт** данных, а все книги, написанные человечеством – **50 петабайт**. ChatGPT 100 млн пользователей
- Человеческий мозг – 100 млрд. нейронов, 100 трлн. связей, 1 петабайт **односимвольных** данных, но не все решения – суть функции «символьных» аргументов **вычислимы**

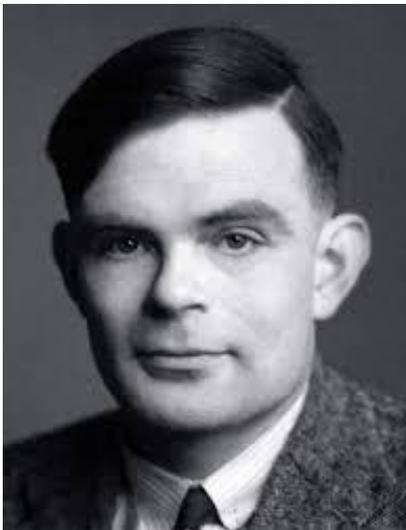


Искусственная нейронная сеть



теория Эйнштейна - «скорость в физическом мире относительна».

Теория алгоритмов утверждает: в достаточно сложных языках существуют недоказуемые высказывания.



- не всякое число равно самому себе (в модулярной арифметике)
- «алгоритм» — эта последовательность однозначных инструкций («программа»), которая **гарантировано** за конечное число шагов переводит исходные данные (высказывание) в результат – булево значение (истина, ложь).
- **не всякое верное высказывание МОЖНО ДОКАЗАТЬ!**

Модели цилиндра

:



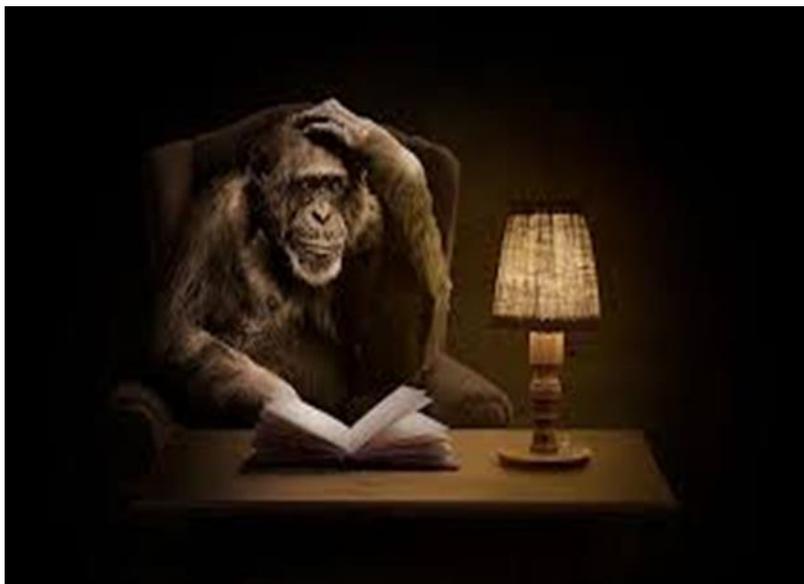
Цилиндр – это ни то и ни другое!
Это что-то объемное!



Цилиндр – это прямоугольник!

Цилиндр – это круг!

...ИЛИ ВСЕ ТАКИ «ИСТИНА В ВИНЕ» ?!



Умозаключение строятся путем рекурсии, образуя бесконечное множество силлогизмов, выражаемых с помощью одной формулы: суть рекурсии - если формула справедлива для n , то она справедлива для $n+1$.

Так как **интуиция** не в состоянии сформулировать общих истин, то рассуждения могут формироваться и путем рекурсий - математической индукции. В итоге доказательство истины есть рекурсия, реализующая операции, которые может повторяться бесконечное число раз, а сами эти операции однородны и делимы.

Итак, **математическая непрерывность** – это результат вычислений, в котором опыт не участвует. Математики изучают не предметы, а лишь отношения между ними.

Предметы могут меняться – отношения должны оставаться

Физическая непрерывность - это результат «ощущений» для проверки равенства, Пусть вес предмета: $A=10$ гр. и $B=11$ гр. и $C=12$ гр. , но по ощущениям не неразличимы. **$A=B$, $B=C$** , хотя на деле **$A < C$**

В компьютерных науках доказывается равномогцность алфавитов размерности 2 и более, а это утверждение обобщается на любой алфавит размерности больше 1.

Множество всех слов бесконечной длины на алфавите размерности n по крайней мере — **несчётно**. Это прямо следует из теоремы Кантора: множество всех бесконечных слов на алфавите $\{0,1\}$ несчётно.

Итак, множество всех конечных строк над счётным алфавитом — счётно. Строки – это слова "**неограниченной**", а не "**бесконечной**" длины. Счетное объединение неограниченных множеств - счетно.

Действительные числа в диапазоне $[0; 1)$ можно рассматривать как **слова бесконечной длины** над алфавитом 0-9. Оператор рекурсии **string = 'a' : string** задает строку, состоящую из бесконечного числа символов 'a'.

- **Сознание** как идеальный **объект**, имеющий «направление» (вектор ?) и представленный **суперпозицией** мыслительных, чувственных и эмоциональных процессов, включая такие явления как воображение, интуицию и воспоминания.
- **Сознание** как **предмет**, характеризующий **способность** человека
 - **приобретать** знания в результате преобразования, хранения и восприятия информации,
 - **выражать** сомнения, страх, радость, возбуждение, желания и пр.
- **Вычисление** (что) — формальное преобразование, позволяющее отобразить входящий поток данных в выходной, с отличной от первого структурой или значением
- **Вычисление (как)** — возможность формального получения из множества входных данных нового множества, которое является носителем некоторого смысла (нового знания).

- **Структура** сознания:
 - **Интеллект** – способность, знания и умение, необходимые для решения умственных задач
 - **Мотивация** – совокупность побуждений, определяющая целенаправленность действий
 - **Эмоции** – переживания, выражающие субъективное отношение человека к явлениям
 - **Воля** – способность сознательно регулировать действия и фокусировать поведение.
 - **Самосознание** – отражение «Я», отделенного от внешней среды.
- **Г. Фреге:** смыслы совсем не то, что **значения**, поэтому смысл и референцию (отнесенность знака к означаемому) нельзя отождествлять. Например, имя "Наполеон Бонапарт" имеет и значение, и смысл, а "Пегас" - только смысл.
- **Понимание текста** — умственная деятельность, в результате которой содержание текста наделяется смыслом. **понимаемое есть смысл, а не значение.**
- **Фундаментальный вопрос: Можно ли понимание заменить «вычислениями» ?**

- Для ответа на вопрос сформулированный выше часто используется так называемый называемый "геделевский аргументе - обоснование невозможности создания искусственного интеллекта.
- Суть аргумента заключается в следующих двух аспектах:
 - из теоремы Геделя о неполноте формальных систем вытекает принципиальное преимущество человеческого ума перед "умом" машинным, так как человек обладает способностью решать проблемы, принципиально неразрешимые для любых искусственных "интеллектуальных" систем (так называемые "алгоритмически неразрешимые" проблемы).
 - ограниченность "искусственного ума" проистекает именно из его «алгоритмического» и поэтому формального характера, а саму "формализацию" при этом можно определить как предельно однозначное и исчерпывающее описание объекта.

Суть сформулированного предложения состоит в «реализации системы ИИ, которая может использоваться для генерации содержательной оценки научных отчетов с точки зрения критериев и в терминах, которые используются корпусом экспертов.

В общем случае решение этой задачи невозможно, так как в настоящее время отсутствует доказательство алгоритмического характера функции «понимания». Из «геделевского аргумента» следует, что невозможно «реализовать» с помощью вычислительной машины то, что недостаточно ясно описано, что лишь интуитивно обосновано и зависит от контекста.

Следуя «аргументу Геделя» можно утверждать, что преимущество создаваемой системы ИИ перед уже существующими «машинными системами» можно усмотреть в том, что используя индуктивную логику «частичной рациональности» система может распознавать (т. е. «почувствовать») семантику анализируемого текста лишь с точностью до вычисленных ранее дескрипторов

Частичная **алгоритмическая невычислимость** функции «смысла» текста означает невозможность построить «непротиворечивую алгебраическую систему» и соответствующее **цифровое алгоритмическое устройство**, которое во всех ситуациях (для любых наборов входных данных) выполняет действия функционально эквивалентные возможностям человеческого мозга

При реализации нашего предложения **будем исходить из «тезис Черча»:**

существует (потенциально бесконечный) класс "универсальных вычислительных машин", способных (в силу того, что каждая «машина» является конкретной формализацией алгоритма) вычислить некоторую функцию, определенную как формально, так и в интуитивном смысле.

Будем исходить из того, что алгоритм - это ясная, неизменная, понятная система инструкций или совокупность правил. В понятие алгоритма входит идея принципиальной познаваемости производимых действий. То что принципиально непознаваемо не может рассматриваться в качестве алгоритма.

Утверждение о принципиальной непознаваемости алгоритма работы сознания при понимании «смысла» текста - фактически равносильно признанию **невозможности описать** функцию сознания и, поэтому, невозможность вычислять смыслы входных потоков данных, используя «программные автоматы».

В итоге, довод в пользу «геделевского аргумента» против ИИ можно сформулировать так:

- гипотеза об алгоритмической природе сознания ведет к парадоксам, следовательно логически противоречива.

Сознание как функция психики отдельного человека не вычислимо, однако индуктивно формируемое массовое сознание людей может быть описано как «распределение конечного множества отдельных возможностей».

Вероятность того, что «выбранная возможность» есть решение рассматриваемой интеллектуальной задачи «вычисления смыслов» не равна «1»

Интеллект – это система с неопределенным **множеством аксиом, которая неформализуема**, для нее невозможно однозначно определить множество "доказуемых истин" и, следовательно, невозможно сформулировать предложения, утверждающие собственную недоказуемость относительно заданной системы аксиом.

Пока нет возможности **установить конечную или счетную аксиоматику**, на основе которого функционирует естественный интеллект или сознание человека, **поэтому можно** рассматривать (считать) **человеческий интеллект** «формальной системой с неопределенным и при этом постоянно расширяющимся множеством аксиом».

Заметим, что гипотеза об алгоритмической невычислимости функции сознания **эмпирически недоказуема**. Это значит, что невозможно экспериментально доказать, что конкретный человек не **способен решать** все алгоритмически неразрешимые «массовые проблемы» (то есть такие проблемы, которые типичные для рассматриваемого класса ситуаций) .

Этот вывод **«индуктивно» очевиден**, в силу того, что человек обладает ограниченным объемом памяти и навыков и на протяжении всей своей жизни имеет дело лишь с конечным множеством проблем/ситуаций.

Любое множество проблем (ситуаций), может представлять собой лишь **конечное подмножество множества**, характеризующего алгоритмически неразрешимую «массовую проблему».

При этом мощность множества всех «алгоритмически неразрешимых «массовых проблем» - скорее всего счетно.

Алгоритмическая неразрешимость феномена сознания не исключает возможности решения некоторого **конечного множества «интеллектуальных» проблем**, входящих в как конечное подмножество в счетное множество всех неразрешимых «массовых проблем» функции создания.

В общем случае, можно утверждать лишь **отсутствие общего, универсального способа решения всех возникающих проблем**. Проблема «вычисления» семантики (смысла) текста – частная задача, которую можно решить с точностью «до вероятности».

рассмотрим два утверждения

1* [X не способен доказать данное утверждение 1*]

2* [Алгоритм А не способен установить истинность высказывания 2*].

Если нет описания формальной системы, то невозможно сформулировать в явном виде «геделевское предложение» для рассматриваемой системы. С этой точки зрения утверждение 1* [X не способен доказать утверждение 1*] не является подлинным геделевским предложением, поскольку оно никак не специфицирует систему «X» и, следовательно, утверждает «непонятно что и о чем».

Утверждение 2* истинно, но не доказуемо. Однако, если человек способен установить, что он действительно работает в соответствии с алгоритмом А, то человек сразу же устанавливает истинность - 2*, расценивая это утверждение как «геделевское предложение». Таким образом, предложение 2*одновременно и должно и не может быть распознано человеком как истинное.

Чтобы исключить возможность возникновения этого парадокса, можно выдвинуть гипотезу о принципиальной непознаваемости алгоритма, в соответствии с которым функционирует мозг человека.

Разница между утверждениями 1* и 2* в том, что во втором утверждении делается акцент на необходимости детального знания "системы аксиом" (алгоритма) на которой основана психическая деятельность человека, для того, чтобы было возможно сформулировать геделевские предложения, неразрешимые для человеческого мышления.

Конкретный вид геделевских предложений **зависит от выбора дедуктики**, т.е. конкретного набора аксиом и правил вывода. Поэтому, не зная действительного устройства формальной системы, **невозможно и выписать** в явном виде и «геделевское предложение» для рассматриваемой системы.

С этой точки зрения предложение 1* [X не способен доказать утверждение 1*] не является подлинным геделевским предложением, поскольку оно никак не определяет систему "X" и, следовательно, утверждает **непонятно что и о чем**.

Следовательно, в данном случае, снимается вопрос о том, каким образом X способен распознать истинность некоторого предложения

- Предположим, что некто формально описал алгоритм А, который исчерпывающим образом описывает функцию работы его сознания. Тогда, по крайней мере в некоторых случаях, можно будет предсказывать будущие действия или конкретные поступки.
- Пусть исследование алгоритма А привело к заключению, что в ситуации Х с **необходимостью надо осуществить действие Р**. В силу осознания неизбежности действия Р субъект может "нарочно" отказаться от осуществления Р, а осуществить некоторое альтернативное действие Q.

В итоге получится, что субъект одновременно должен и не должен осуществить действие Р.

Итак, если сознание человека подчинено конкретному алгоритму, то предположение о познаваемости данного алгоритма ведет к противоречию. Отсюда можно сделать вывод, что алгоритм сознания, если он на самом деле существует, то **принципиально непознаваем.**

Вывод. то что непознаваемо, вряд ли можно е считать алгоритмом.

Будем при реализации предложений рассуждать следующим образом:

Исходим из того, что в основе человеческого интеллекта лежит некий алгоритм (система правил) A . Если способно в явной форме установить какие именно правила составляют алгоритм A , то можно легко построить "неразрешимое" высказывание:

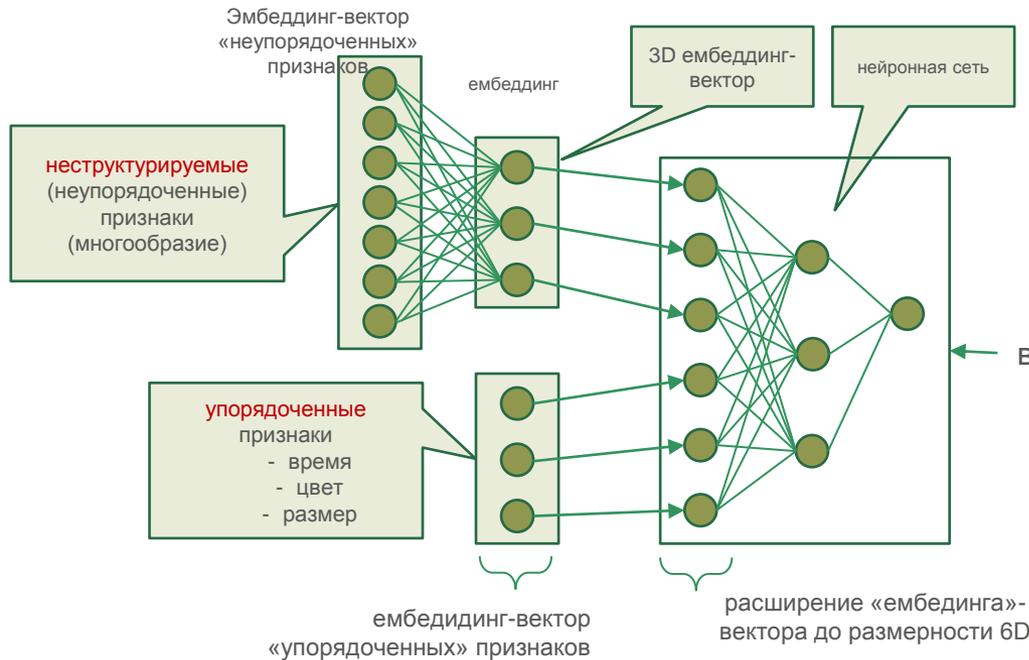
2^* [Алгоритм A не способен установить истинность высказывания 2^*].

Это предложение истинно, но не доказуемо.

Сам человек, если он действительно следует алгоритму A , не способен установить истинность данного предложения. Но если человек способен установить, что он действительно следкует алгоритму A , то он сразу же устанавливает истинность - 2^* , расценивая его как геделевское предложение. Таким образом, предложение 2^* одновременно **и должно и не может быть распознано человеком как истинное.**

Итак, решение, которое возможно вычислить, доказуемо, но с точностью «до вероятности»

ПРИМЕР ИСПОЛЬЗОВАНИЯ «ВЕКТОРОВ СМЫСЛОВ» В ЗАДАЧАХ МАШИННОГО ОБУЧЕНИЯ



Суть вложения вектора смысла («эмбединга») преобразование слов описания признаков объекта в набор чисел – числовой : «word2vec» (метод преобразования «слово в вектор»).

Так объекты сравниваются более чем с одним признаком, поэтому эмбединги, полученные из признаков в первом слое нейронной сети, объединяются с эмбедингами, полученными из «упорядоченных» признаков в следующих слоях

Выражения синтетическое выражение типа $a = b$ принципиально отличающимися от выражений типа тавтологии $a = a$.

формально... выражение «вода — это H_2O », сформулированное в 18 веке Лавуазье ничем синтаксически не отличается от выражения «вода — это хуз»

Из семантической модели Фреге следует, что прирост знания, возможен если разные символы (текст) связывается с одним и тем же предметом, **имея при этом разный смысл**, причем связь смыслов знаков с одним и тем же предметом **не является очевидной, а зависит от контекста**