



Санкт-Петербургский
Государственный
Политехнический
Университет

Институт прикладной
математики и механики

КАФЕДРА ТЕЛЕМАТИКА

Методы исследовательской работы

«Исследования как результат интеллектуальной деятельности» (занятие 1)

3 февраля
2022 г.

В начале попробуем ответить на вопрос: Как организовать исследование этого феномена ?

2



- Какая информация получена при анализе этой записи (процесса) ??

Использование понятия «информации» при работе с различными формами знаний - как имплицитными, так и эксплицитными, учитывая что 1) **имплицитное знание** не имеет символьной формы описания, поэтому не отделимо от субъекта – непосредственным носителем знаний, 2) **эксплицитные знания** могут быть выражены в форме математических законов физики.

Точное значения используемых терминов:

- Имплицитный (от латинского слова *implicitum*) значит «невыраженный», «подразумеваемый», «неразвернутый» т.е. – «**скрытый**».
- Эксплицитный (от латинского слова *explicitum*) – значит «явно выраженный», «развернутый», т.е. – «**явный**».

Надо понять как нужно при проведении исследований использовать понятия информация или интеллект, чтобы исследования имели научный смысл т.е. полученные результаты **объяснялись с помощью введенных ранее понятий?**

Парадигма современной науки

Классика **Cogito**, ergo sum

(лат. — «Мыслю, следовательно, существую»)



Современность: **Computo**, ergo sum (лат. - «Вычислил, значит существую»). Философия Информационно-Вычислительного натурализма: законы физики – «компьютерные» программы, а окружающий человека мир - квантовый компьютер, который вычисляет самого себя?!



Исходная гипотеза

Гипотеза : **информационное содержание** - основа способности человека к **мышлению**, познанию, обучению, восприятию, запоминанию, обобщению и принятию решений - реализуется в природе как результат вычислений в среде «нейроструктур мозга».

Следствия :

Сознание - «вычисляемая» функция «встраиваемого ПО» нейрокомпьютера мозга человека.

Однако в силу **супервентности** (отношения детерминированности состояния любой системы состоянием другой системы) **ментальные проявления** объективно определяются свойствами физических явлений, однако не сводятся только к ним.

Вывод: отсутствие принципиальных различий в **ментальных моделях** субъектов при отсутствии различий в свойствах окружающих их физической реальности

Информационная (ментальная) реальность – с чем «работает» сознание

Все есть число

Пифагор
(2000 лет назад)

Истина в неполноте

Гедель
(100 лет назад)

It from bit

Арчибальд Уильер
(50 лет назад)

Любые вычисления рассматриваются как физический процесс над носителями информации. Результат вычислений - ментальная (виртуальная) реальность, существующая, в силу принципа дополненности, одновременно в физической и информационной формах. Между физической и информационной формах реальности нет изоморфизма или взаимно однозначного отображения – информация теряется, энергия рассеивается.

Возможна ли интерпретация без моделей

(model agnostic interpretation):

Люди – «потребители» результатов восприятия.....

Методы восприятия – «непрозрачность» алгоритмов интерпретации ([the opacity of machine learning models and methods of explanations](#))

Модель «черного ящика» где физические процессы/данные превращаются в слова/понятия

Данные экспериментов - цифровые двойники («digital twins») сигналов, изображений, текстов

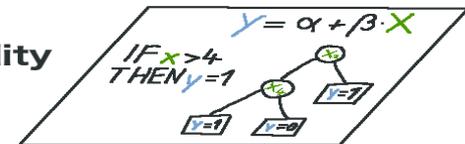
Реальный мир, в котором протекают реальные процессы.

Humans



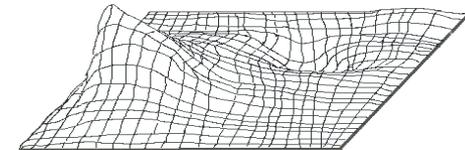
↑ inform

Interpretability Methods



↑ extract

Black Box Model



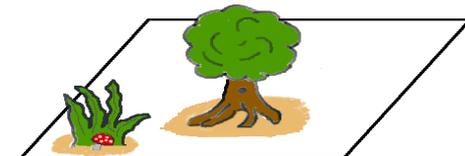
↑ learn

Data

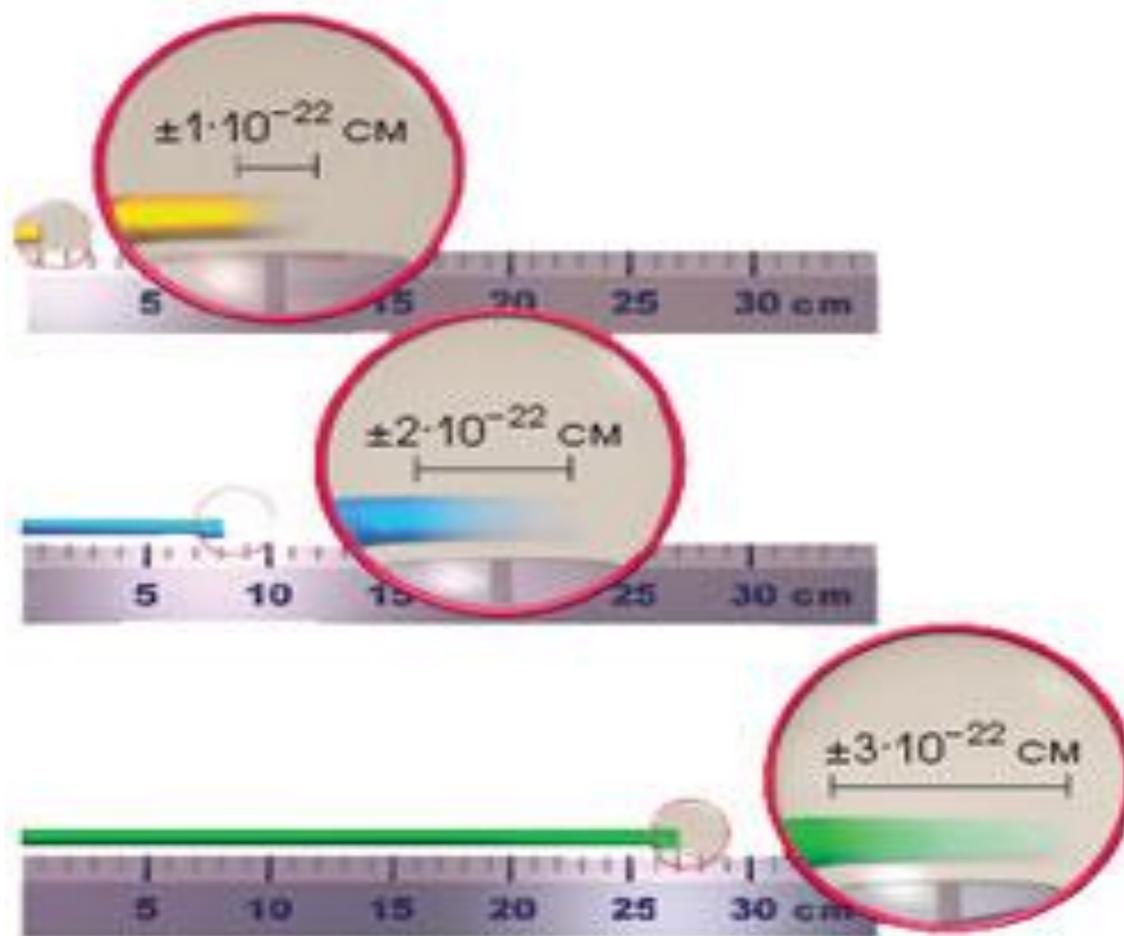
X_1	X_2	X_3	...	X_n
10	2	0		
5	4	0		
1	-1	0		
				0.10
				0.20
				0.30
				0.40
				0.50

↑ capture

World



Точность измерений и коды атрибутов зависят от размеров объектов



Роль «наблюдателя» в формировании приближенной картины Мира

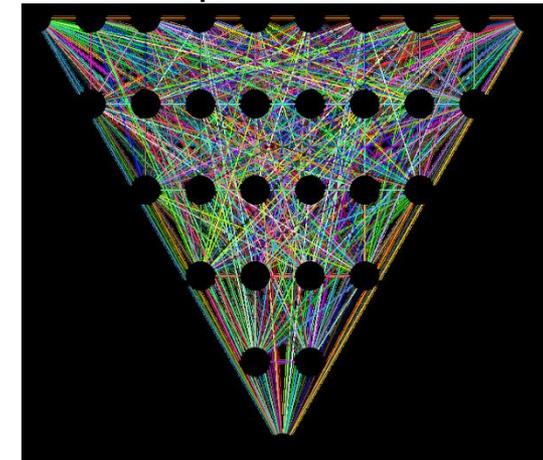
- Поскольку полнота и точность воспроизведения физического мира в сознании человека всегда **относительны** (а о «взаимной однозначности» т.е. изоморфизме вообще не может быть речи), то соответствия между предметами внешнего мира и их образами в человеческом сознании **носят гомоморфный** (приближенный) характер.
- В каждый момент времени, каждый человек использует «картину мира», с характерными для данной ситуации операциями и отношениями.
- Учет роли наблюдателя заключается в логически непротиворечивой для текущей «картины мира» **свертке всей доступной в результате измерения информации** об объектах или, процессах в удобную для последующих «вычислений» форму.

радикальный способ экономически эффективно увеличить производительность труда

- 1) создать интеллектуальную САМ систему основанную на “вычисление” с помощью специализированной на определенных технологиях промышленной нейронной сети оптимальных производственных решений;
- 2) сократить штат посредников, менеджеров, и др. так называемых организаторов производства



Промышленная «нейронная» сеть



Простой пример

Вход		Выход
0	0	0
0	1	0
1	0	0
1	1	1

Вход		Выход
0	0	0
0	1	1
1	0	1
1	1	1

Количество информации по Шеннону

$$I = -\sum_{i=1}^N p_i \log_2 p_i$$

Пусть на входе «гейта» имеется 4 равновероятных состояния

Тогда, входная информационная емкость гейта $I_{in} = 2$ бита

На выходе гейта: состояние 0 появляется с вероятностью $3/4$, а состояние 1 - с вероятностью $1/4$

Выходную информационную емкость гейта можно оценить так: $I_{out} = -[(3/4)\log_2(3/4) + (1/4)\log_2(1/4)] =$
 $= -[(3/4)\log_2 3 - (3/4)\log_2 4 - (1/4)\log_2 4] =$
 $= - (3/4)\log_2 3 + 2$

Информация **теряется** $(I_{in} - I_{out}) = (3/4)\log_2 3 = 1.1887$ бит
а энергия **рассеивается** :

$$\Delta E = kT \ln 2 \cdot (I_{in} - I_{out}) = kT \ln 2 \cdot (3/4)\log_2 3 = 0.824 kT$$

Фундаментальная проблема ограниченности занятий: неполнота формальных моделей

Ключевой вопрос Критики — это исследование познавательной возможности разума, в отрыве от знаний, получаемых эмпирическим, то есть опытным путем.

Первая Теорема Геделя:

если формальная арифметика непротиворечива, то в ней существует невыводимая и непроверяемая формула

Вторая Теорема Геделя:

если формальная арифметика непротиворечива, то в ней невыводима некоторая формула, содержательно утверждающая непротиворечивость этой арифметики

Вывод:

Формализация природы как целостной системы невозможна ?!

Как быть с гипотезой о «вычислимом сознании» ?

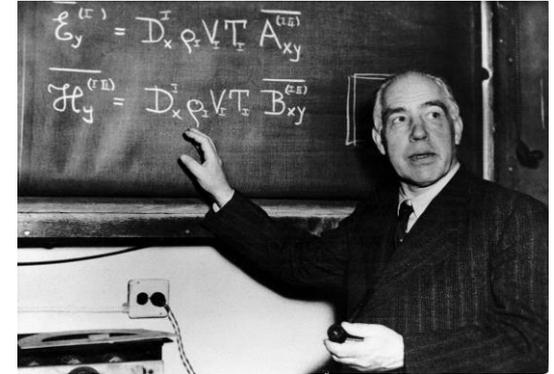


Легко можно сформулировать вопросы на которые не возможно ответить так как они превосходят возможности человеческого разума

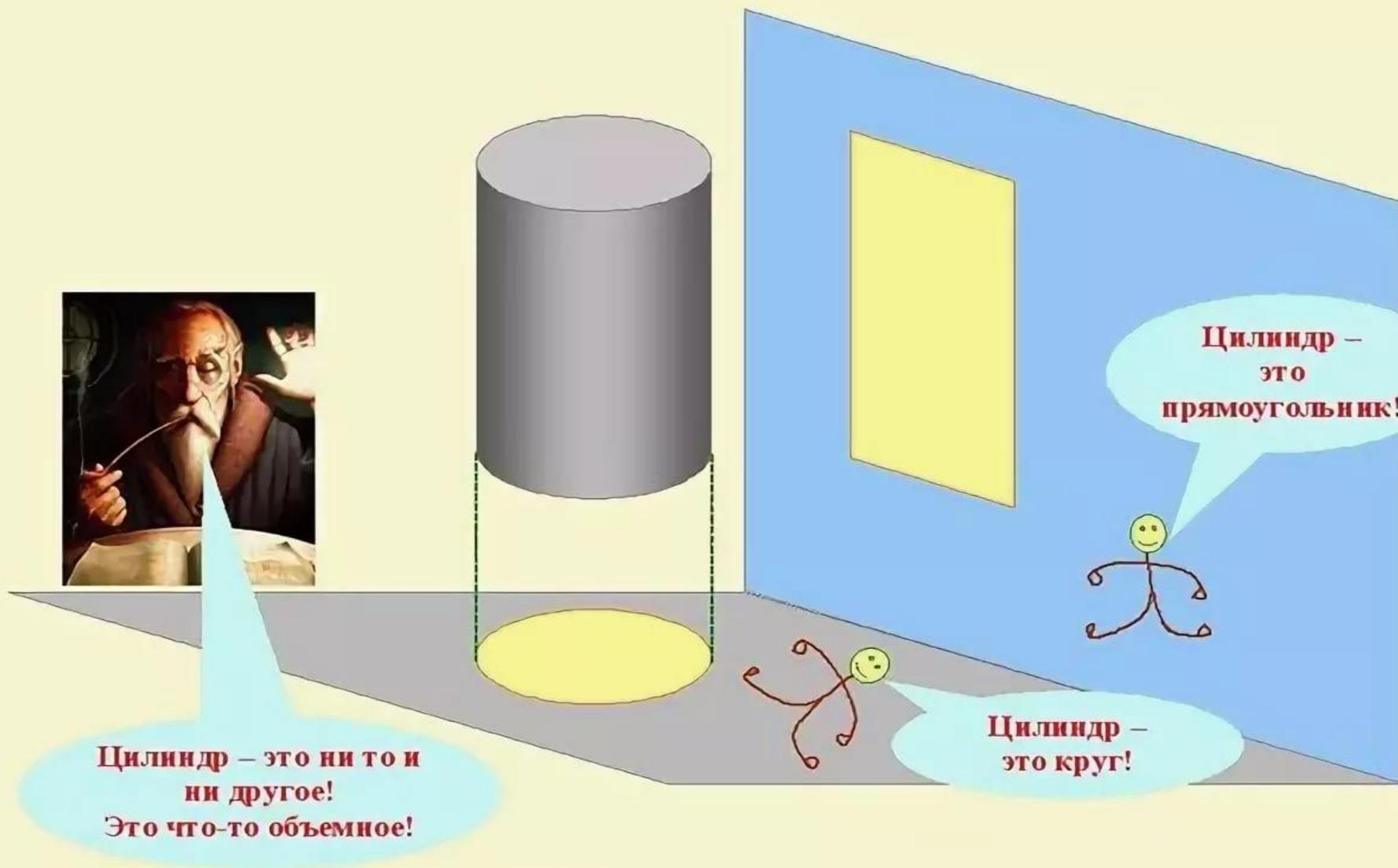
Принцип дополнительности (также принцип комплементарности) — один из важнейших методологических и эвристических принципов науки, а также один из важнейших принципов квантовой механики, сформулированный в 1927 году Нильсом Бором

Согласно этому принципу, для полного описания сложных квантовомеханических явлений необходимо применять два взаимоисключающих («дополнительных») набора классических понятий, совокупность которых даёт исчерпывающую информацию об этих явлениях как о целостных.

Суть принципа — использовать взаимоисключающие классы понятий, каждый из которых применим в особых условиях, но их совокупность позволяет воспроизведение целостности данных объектов.

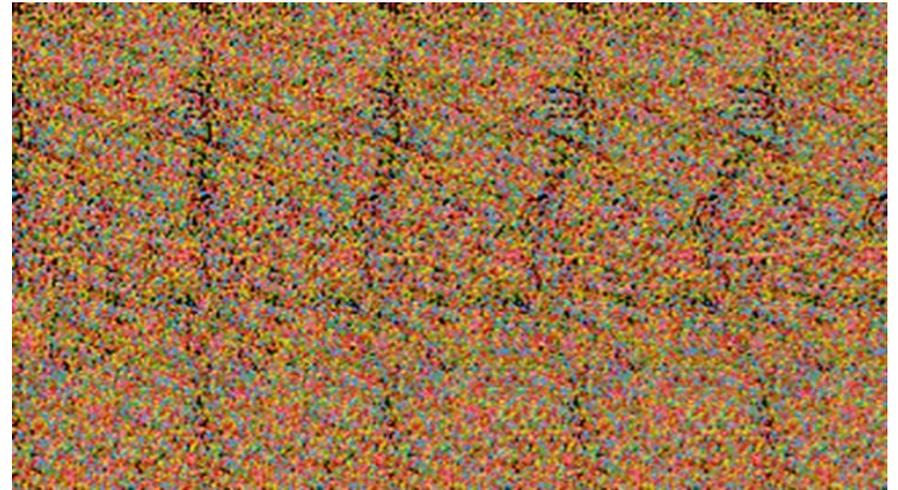
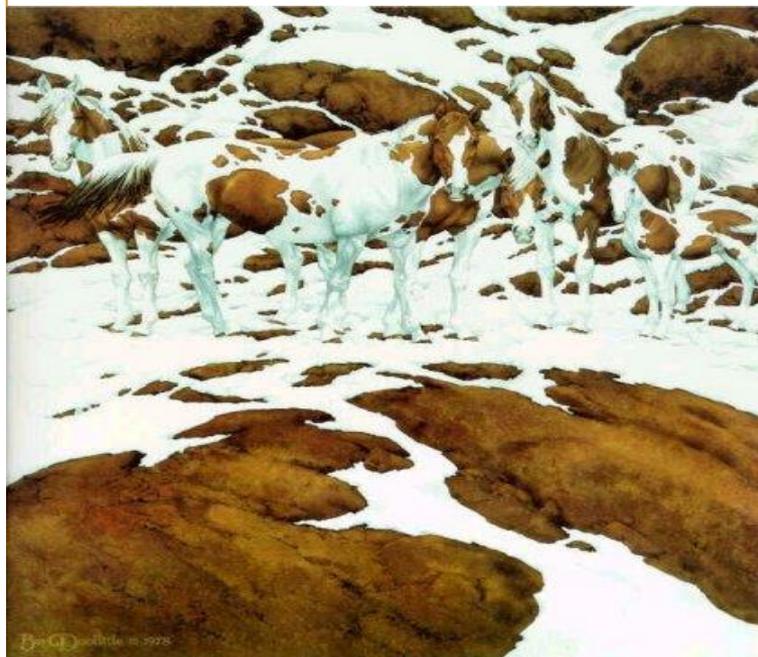


физическая картина явления и его математическое описание дополнительны. Создание физической картины мира требует пренебрежения деталями и уводит от математической точности, а попытка точного математического описания явления затрудняет ясное понимание.



Математика «больших данных» отвечает на вопросы:

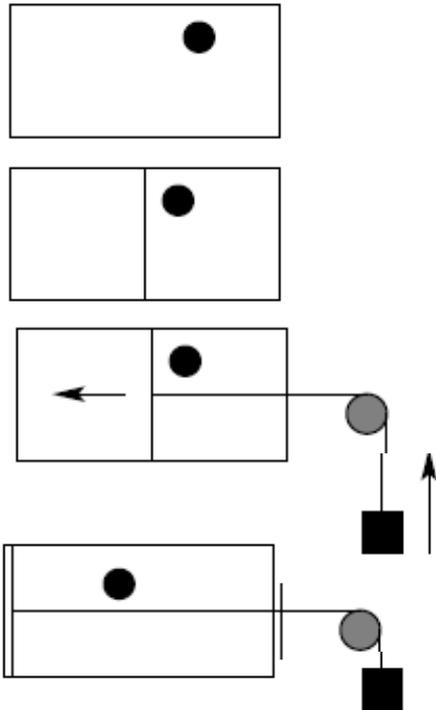
- Как научить вычислительные системы находить «похожее в различном»
- Что надо знать, чтобы выделить «целостное» в наборе фрагментов, образующих поток «больших данных»



Обучение – использование фактов для формулировки принципов ?



Есть ящик с "одно-молекулярным" газом и А демон (Д) – интеллектуальное устройство, способное измерять, хранить и использовать **информацию**.



Сначала положение **молекулы** неизвестно.

Д ставит перегородку, делящую ящик пополам, и **определяет**, в какой половине находится молекула.

В зависимости от результата измерения, Д размещает в ящике поршень с грузом таким образом, чтобы удары молекулы его поднимали. Газ «расширяется» и совершает работу.

При постоянной температуре $\Delta A = kT \ln(V_2/V_1) = kT \ln 2$.

После того, как поршень сдвинулся и груз был поднят, его удаляют.

Система вернулась к исходному состоянию... В итоге совершена работа за счет теплового резервуара при **постоянной температуре**.

Нарушено второе начало термодинамики?

Нет: Д был наделен памятью, в памяти осталась информация о состоянии ячейки, которой в начале цикла не было. Если вернуть систему к исходному состоянию – значит надо **стереть информацию** и затратить на это энергию $kT \ln 2$.

Если память "демона" явно в описание рассматриваемой системы не включать, то можно учесть «информацию» о системе в описании модели ее состояния

Произведенная механическая работа $\Delta A = kT \ln 2 = \Delta Q = T\Delta S$

Изменение обычной термодинамической **энтропии** ячейки

Работа, затраченная на стирание информации равна $\Delta A_e = -kT \ln 2$

А полная совершенная в процессе работа $\Delta A + \Delta A_e = T\Delta S - kT \ln 2 = T\Delta S_g$
 $\Delta S_g = \Delta S - k \ln 2 = 0$

Можно сохранить обычное выражение для второго закона термодинамики (т.е. $T\Delta S \geq 0$ в замкнутом цикле), если ввести понятие **эффективной энтропии**:

$$\Delta S_g = \Delta S - k \ln 2 = \Delta(S + S_i)$$

S – обычная **термодинамическая энтропия**, зависящая только от ее состояния

S_i зависит от того, что знает о системе наблюдатель: **информационная энтропия**

$S_i = -k \ln 2$ (на один бит).

Наличие информации о физической системе должно быть учтено и это и есть ментальная модель объекта

Если **информация** включается в общее **описание состояния системы** наравне с ее физическими параметрами, то оказывается, что

- Одна и та же система **имеет различные физические свойства** в зависимости от имеющейся информации (в одном случае она способна совершить работу, в другом – нет)
- Мера **информации** оказывается согласованной с общефизическими **понятиями энергии и энтропии**
- «Обращение» принципа Ландауэра (не стирание, а получение бита информации): любая неслучайная комбинация битов может быть использована для **производства работы**.
- Информация как атрибут объекта наравне с его физическими параметрами (размер, вес...) меняет ментальную модель мира.
- В зависимости от имеющейся информации о системе систему можно или нельзя использовать для совершения работы. (в одном случае система способна совершить работу, в другом – нет)

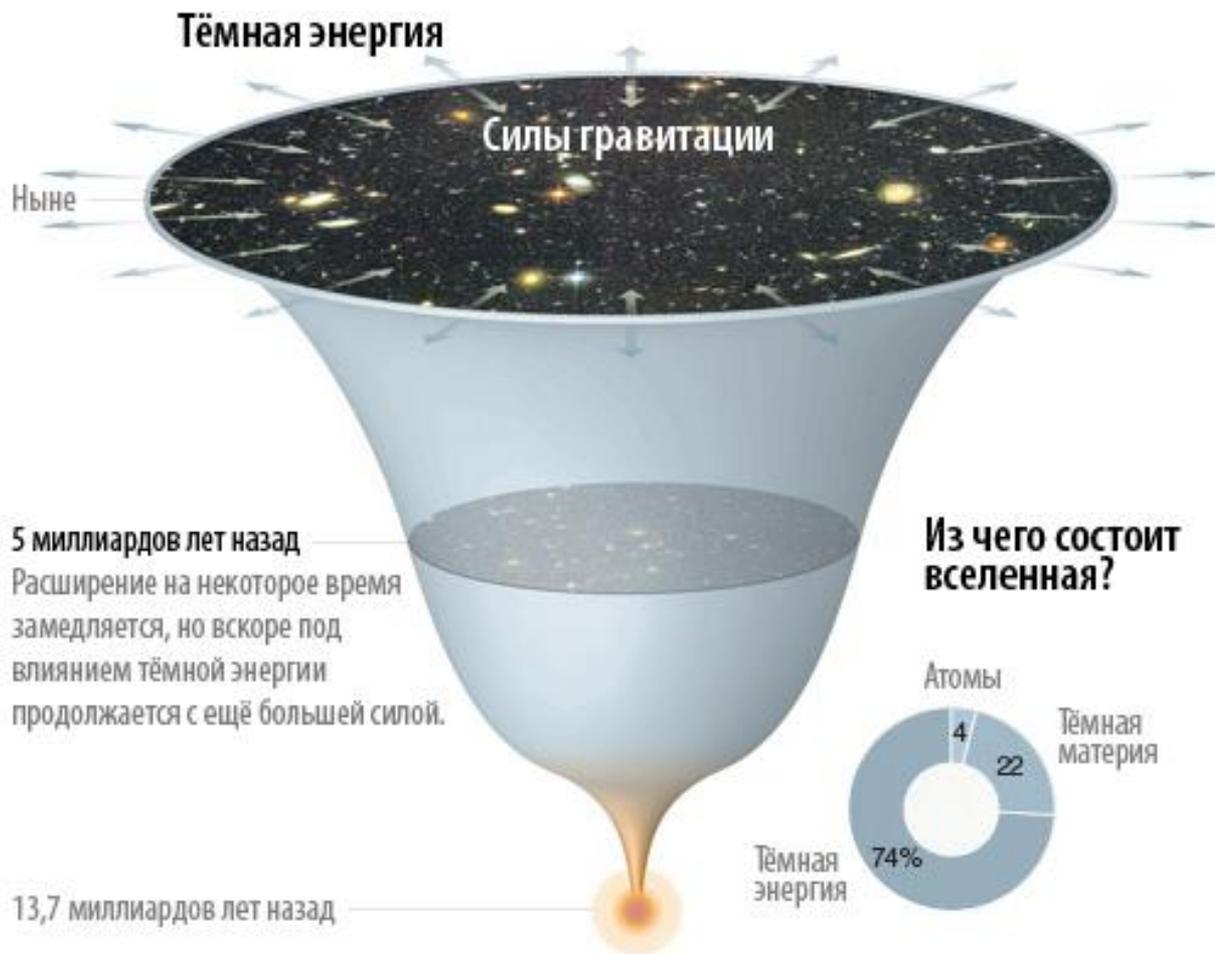
Проблема «материализации» знаний в структурах мозга и интуитивное понимание мира

19

- «Наше» прошлое «записано» в нейросетях мозга, которые формируют то, как мы воспринимаем мир в целом и его конкретные объекты в частности.
- «Наши» реакции не спонтанны - большинство из них запрограммировано устойчивыми нейронными связями. Каждый объект (стимул) активирует ту или иную нейронную сеть, которая в свою очередь вызывает набор определенных химических реакций в организме.
- Эти химические реакции заставляют человека действовать или чувствовать себя определенным образом. Эмоциональные реакции – не более чем результат химических процессов, обусловленных сложившимися нейросетями, и основываются они на прошлом опыте.

Итого: в 99% случаев мы воспринимаем реальность не такой, какая она есть, а интерпретируем ее на основе готовых образов из прошлого.

Физическая vs информационная модель реальности



Сколько информации содержится во Вселенной ?



Обработка информации, проводимая в режиме «мягких вычислений», например, «1 полюс два=3 миллиона рублей» требует создания принципиально новых технических устройств, воплощающих возможности работы информационно открытых систем.

Итого:

- Интеллектуализация вычислителя на основе обучения приводит к изменениям в его физической структуре и программном обеспечении
- Обученная «Машина» обретет способность моделировать процессы мышления, если реализует свойства «процессора управляемого данными», решающего «обратные алгоритмические задачи» с использованием методов интеллектуальной регуляризации