



Курс: архитектура суперкомпьютеров

Лекция 1

Управление вычислительными процессами с использованием методов машинного обучения

(обратные задачи алгоритмизации вычислительных
процессов)

В.С. Заборовский

5 сентября 2024
г. СПб



ПОЛИТЕХ
Санкт-Петербургский
политехнический университет
Петра Великого

Содержание

Ключевые слова, определения и понятия

Часть 1 (теория)

1. **Введение.** К истокам проблемы «обучения машин» в контексте процессов мирообъяснения
2. **Задачи машинного обучения** : лингвистический «поворот» компьютерных наук
 - Аспекты «академического» понимания проблемы машинного обучения
 - Компьютерные вычисления как «цифровые тени» физических процессов
 - Объяснение результатов вычислений с использованием метрик производительности
 - К проблеме алгоритмической обратимости процессов вычислений как фундаментальной задачи в развитии компьютерных наук

Часть 2 (практика)

3. Прикладные аспекты управления производительностью суперкомпьютеров с использованием машинного обучения

СКЦ «Политехнический - «БАК» технологий ИИ

«Закон Мура» не работает ? ... Да здравствует «машина Геделя»

«Поворот» компьютерных технологий»: от калькулятора к интеллектуальному трансформеру

4. Перспективы и выводы

Определения, понятия и аксиоматика

Ключевые слова: модель, программируемый компьютер, алгоритм, производительность, обучение

- **Модель** — абстрактное описание (некоторый текст) **реальной системы**, используемый для понимания, вычисления или прогнозирования
- **Компьютер** (Вычислительная машина) – физическое (виртуальное) **устройство**, предназначенное для выполнения **конечного множества т.н.** вычислительных **операций за конечное время**
- **операции** вычислительные (operatio — действие): действия связанные со сдвигом n-разрядного двоичного регистра **при реализации алгоритмов сложения, вычитания, умножения и деления чисел, представленных в позиционной системе счисления и используемых при решения прямых $X = AT$ и обратных $BX = T$ вычислительных задач**
- **Кластер суперкомпьютерный** – сеть из компьютеров (вычислительных ресурсов), предназначенных для **выполнения конечного множества программ (алгоритмов)** **решения вычислительных задач**
- **Управление производительностью** суперкомпьютеров – согласование «запросов» алгоритмов вычислений с имеющимися вычислительными ресурсами компьютерного кластера
- **Реификация** (от лат. res — вещь) — процесс, в ходе которого продукты мышления облекаются (**воплощаются**) в материально-вещественные формы

- **Машинное обучение** -

- термин из расширенной спецификации **алгоритмических возможностей вычислительных систем**,
 - связанный с использованием декларативных («знать, что») и процедурных («знать, как») **знаний, полученных ранее** в процессе решения прикладных задач,
 - для **решения обратных задач цифрового моделирования** (построением алгоритмов) и **объяснением** результатов вычислений

- **Информация** – *informatio* (лат.): **форм**ирование, то есть процесс, в котором «форма» (структура) получает воплощение:

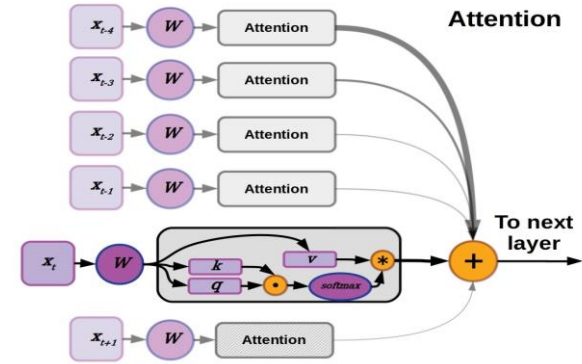
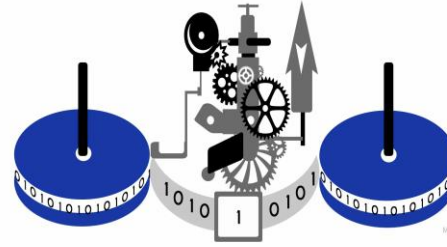
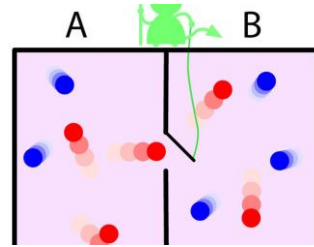
- **Цицероном** *informatio* использовал для таких понятий как **идея**
 - **св. Августин** использовал словосочетание *informatio materiae* в смысле процесса «**формирование материи**».
 - ISO 5127:2017 **information** - сообщение, которые порождают значение (смысл)
 - в теории передачи **информация** Шеннона - **значение $I = -\log P$**

- Будем исходить из **следующих аксиом** (Блум, 1967 г.):

- Каждая вычислимая функция **имеет бесконечное число** программных представлений
- Существуют **алгоритмически неразрешимые задачи** (самоприменимость машины Тьюринга, задача «останова» алгоритма, задача синтеза алгоритмов на основе анализа данных и пр.

Уточнение: Алгоритмическая неразрешимость понимается в том смысле, что **для алгоритма**, решающего данную задачу, существует набор **входных данных**, для которого алгоритм **дает неверный результат** (результат взаимодействия алгоритма с внешней средой в общем случае не вычислим)

Часть 1 (теория) Обучение машин или «филогенез» трансформеров



Машины:

паровоз Черепанова -

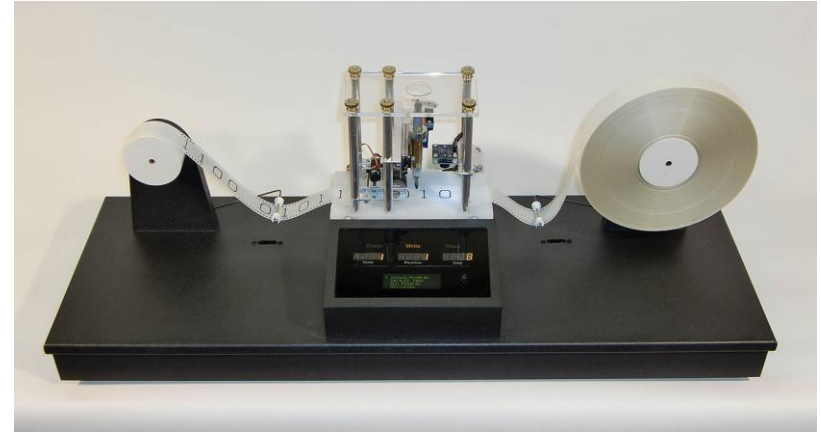
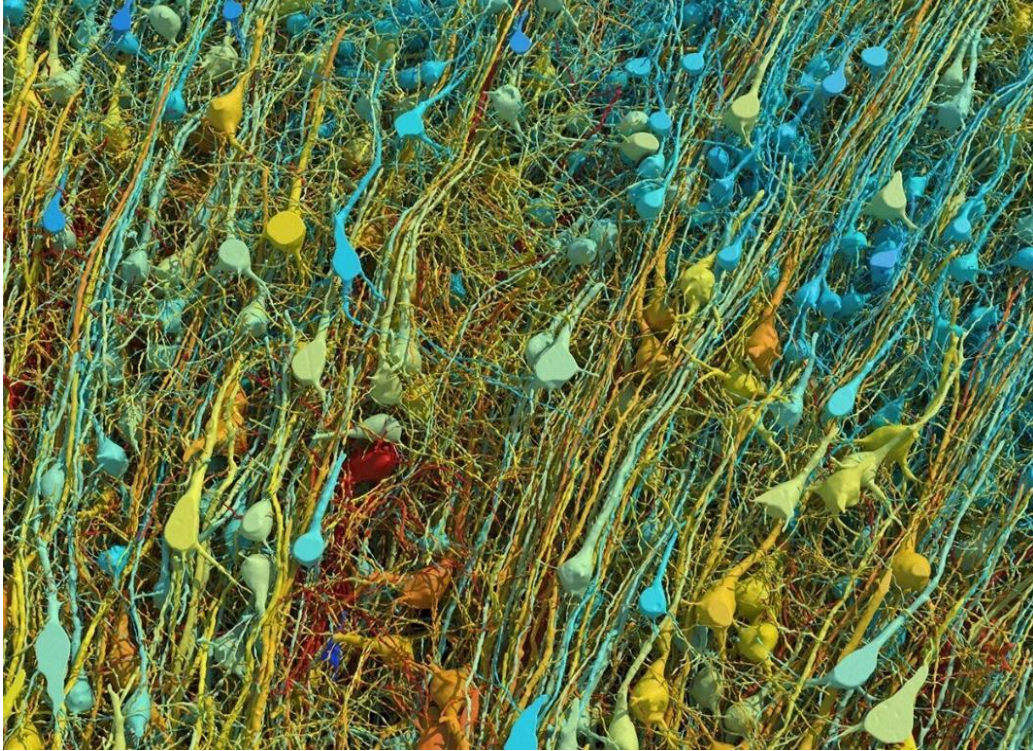
Сциларда -

Тьюринга -

LLM (GPT) трансформер

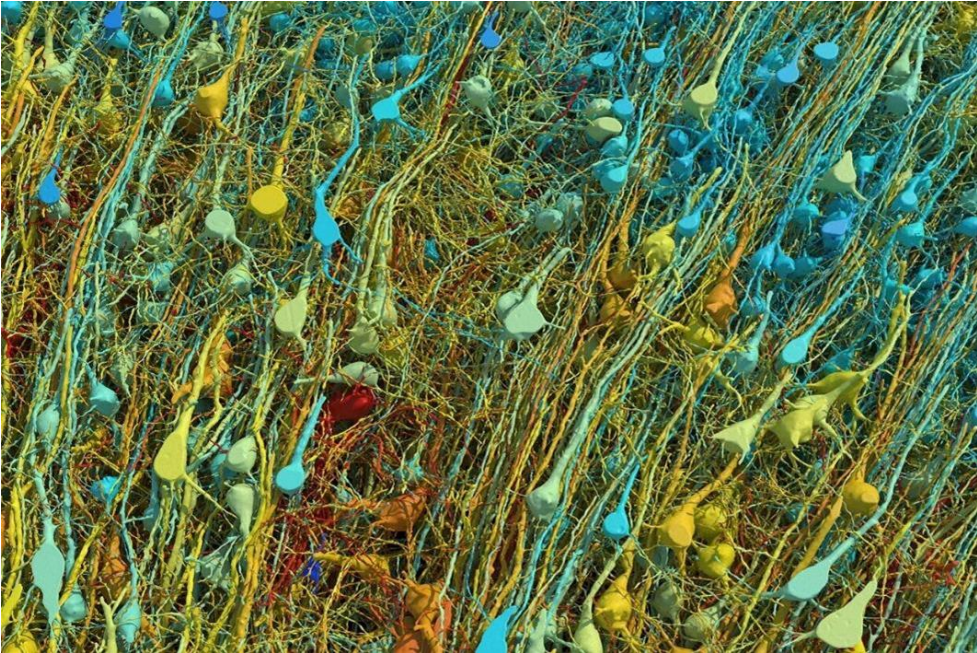
- Паровоз – трансформер тепловой энергии в механическое движение
- «Машина Тьюринга» (1936 г.) - трансформер упорядоченного механического движения в детерминированные вычисления
- Машина Сциларда (1929 г.) - трансформер: 1) термодинамической энтропии в ...информацию ...и 2) информации - в упорядоченное механическое
- GPT генеративный предобученный трансформер нейронных языковых моделей

Мозг-трансформер ?

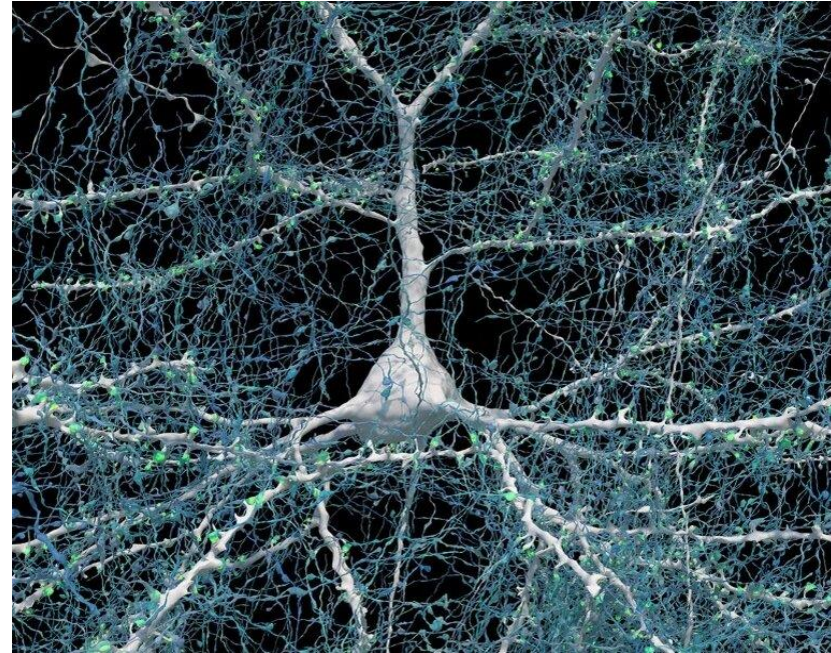


Нейробиологи получили изображение одного кубического миллиметра мозга с точностью до каждой клетки

«Структура» материального носителя естественного интеллекта



изображение одного кубического миллиметра мозга с точностью до каждой клетки



Один нейрон (белый) с 5 600 аксонами (синий), которые к нему подключаются. Синапсы, обеспечивающие эти связи, показаны зеленым цветом. Google Research & Lichtman Lab (Harvard University). Renderings by D. Berger (Harvard University)

Почему современные компьютеры суть трансформеры «осмысленного» в «числовое бессознательное»

- В классическом компьютере (машине Тьюринга = МТ) осмысленные описания (программные коды) воплощаются физические (механические) процессы сдвига двоичных регистров, которые сами по себе не имеют смысла
 - МТ - воплощение «числового бессознательного»

«Искушение» компьютерных наук и технологий: создатьинтеллектуальный трансформер как

«воплощение» машины Тьюринга (МТ) и «демона Максвелла» в «машину Геделя» (МГ), обладающую свойством

- **самоприменимости**
- **«обученную»** извлекать информацию ??? из «окружающей среды», чтобы ...
 - например, вычислить момент времени, когда МТ пора «остановиться» ...
изменить исполняемый программный код ...
объяснить полученный результат.

Итак: «машинного обучения» можно рассматривать как трансформация

- **индуктивный вывод** априорно «бессознательного» (множества чисел, как носителей количества) (дедуктивный вывод) в «сознательное» (слово – как носитель смысла)

Доблесть ума: быть смелым, выдвигая гипотезы, и беспощадным, опровергая их

- К. Поппер пришел к выводу, что доблесть ума заключается не в том, чтобы быть осторожным и избегать ошибок, а в том, чтобы бескомпромиссно устранять их.
- девиз К. Поппера: Быть смелым, выдвигая гипотезы, и беспощадным, опровергая их
- Доблесть интеллекта защищается не «в окопах» доказательств, окружающих чью-либо позицию, но точным определением условий, при которых эта позиция признается непригодной для защиты.
- *предвзятость* (commitment) Поппер считал тягчайшим преступлением интеллекта.
- согласно К. Попперу, наука — это процесс “перманентной революции”, а ее движущей силой является рациональная критика, по Т. Куну, революция есть исключительное событие, в определенном смысле выходящее за рамки науки;

Модификация 4-х принципов архитектуры фон Неймана для создания «МГ»

Суть модификации - наделить компьютерные системы способностями к **постоянному «дообучению»**, индуктивным выводам и дедуктивным (концептуальным) оценкам **вычисленных результатов**, для чего :

- **принцип однородности памяти** – трансформировать в принцип организации «вычислений в памяти» в соответствии с **pretraining** **настройками нейросетевого поля операций**
- **принцип адресности** дополнить принципом динамической **ассоциативной (объяснительной) маршрутизации**,
- **принцип программного управления** расширить за счет возможности **изменения порядка выполнения команд (индуктивное обучение)** с учетом анализа предшествующих результатов вычислений,
- **принцип двоичного кодирования** дополнить процедурой **кодирования категориальных признаков**, которая реализует преобразование дескрипторов описания данных в числовое представление по заданным правилам (концептуальное обучение).

«Быстро» или «точно»: как надо решать математические задачи

универсальность

vs

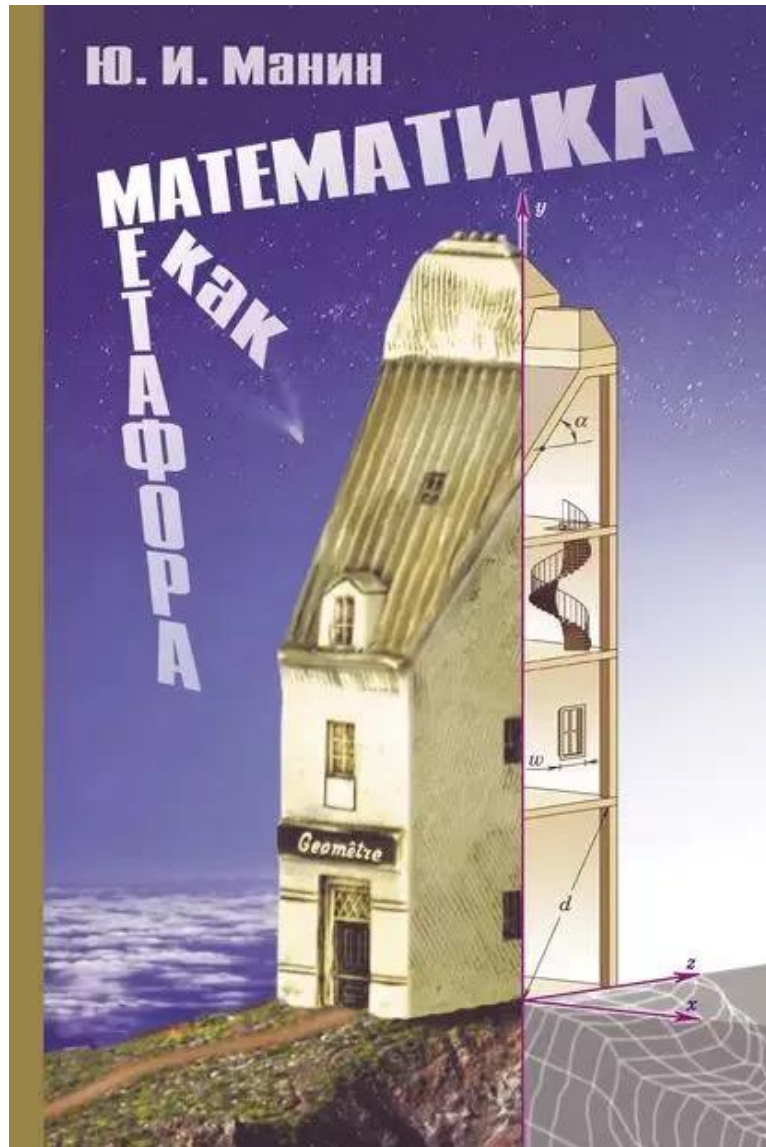
эффективность



Платон исходил из следующих предположений:

- действительность, с которой непосредственно имеет дело человек и которая является предметом его познания, **не зависит от самого знания**;
- мысли могут быть приведены в однозначное соответствие с действительностью;

Математика язык метафор и проблема их воплощения



Ю. И. Манин (2010) "Математика как метафора".

- используется идея реификация (reification) - процесса, в котором **абстрактная идея** (например, закодированная в программе) превращаются в мыслимый объект (текст на некотором языке).
- Ю. И. Манин ввел представления о языке как осмысленной метафоре способный к «воплощению»:
 - видимое глазами и воспринимаемое чувствами – это лишь проекция на сетчатку и рецепторы **бесконечномерного и сложноорганизованного физического мира**.
 - непосредственному восприятию доступны лишь **«тени» реальности**,
 - смыслы и понимание происходящего доступны через **«вспоминание» ранее обретенных знаний**

О познании и относительности

- Окружающий человека мир сложен. Человек не может познать мир, не фиксируя своего внимания на одних его сторонах и не отвлекаясь от других. Условно-вероятностный (байесовский) характер знаний внутренне присущ любому познавательному процессу.
 - истинны с точностью до «достоверности» только факты: Волга впадает в Каспийское море
- Абсолютная истина рассматриваемая как абсолютно полное и точное знание, лежит вне пределов научного познания; понятие абсолютной истины неприменимо к фундаментальным видам научного знания.
- **Абсолютная истина проявляется в росте знания, так как человек обречен иметь дело только с миром, данным ему в ощущениях, созерцая внешний мир, результаты своей деятельности, в которой воплощены его знания о мире, поэтому человека "входит" в объективный мир, становясь его частью.**

Проблема обретения знаний – физика процессов обучения):

материальный объект тождественен самому себе (вычисляет сам себя) в пространстве и идентичен во времени, поэтому

- $E = m \cdot c^2$ (Эйнштейн: энергия пропорциональная массе)
- $Q \geq k \cdot T \cdot \log 2$ (формула Ландауэра - потеря информации приводит к диссипации энергии)
- $I = -\log p$ (формула Шеннона: сообщение, вероятность получения которого равна нулю, переносит ∞ информацию)
- неравенство Гейзенберга: существует то, что способно на «действие»)
- $\Delta E * \Delta t \geq h > 0$, где h – квант действия
- $\Delta x * \Delta v \geq h. > 0$
(+) * (+) = (+) : физическое
(-) * (-) = (+) : информационное

Формулу Ландауэра можно обобщить : $Q \geq -k \cdot T \cdot \log \frac{1}{2} =$

$$k \cdot T \cdot (-\log p) =$$

$$k \cdot T \cdot I$$

тогда

$$I \leq Q/(k \cdot T)$$

Искушения компьютерной науки

90-е годы 20 века: задачу решим на компьютере («действие» - исполняемый код , **КТО** напишет ???)

20-е годы 21 века: задачу решит ИИ («действие» - вспоминание или рекурсия Геделя, **КТО** «научит» ???)

Физика» процессов «воплощения» информации: сколько в Дж «весит» 1 бит

- $Q \geq k \cdot T \cdot \log 2$ (формула Ландауэра - потеря информации приводит к диссипации энергии)
- $I = -\log p$ (формула Шеннона: сообщение, вероятность получения которого равна нулю, переносит ∞ информацию)

Формулу Ландауэра можно обобщить : $Q \geq -k \cdot T \cdot \log \frac{1}{2} =$
 $k \cdot T \cdot (-\log p) =$
 $k \cdot T \cdot I$ тогда

$$I \leq Q/(k \cdot T)$$

Актуальные проблемы компьютерных наук

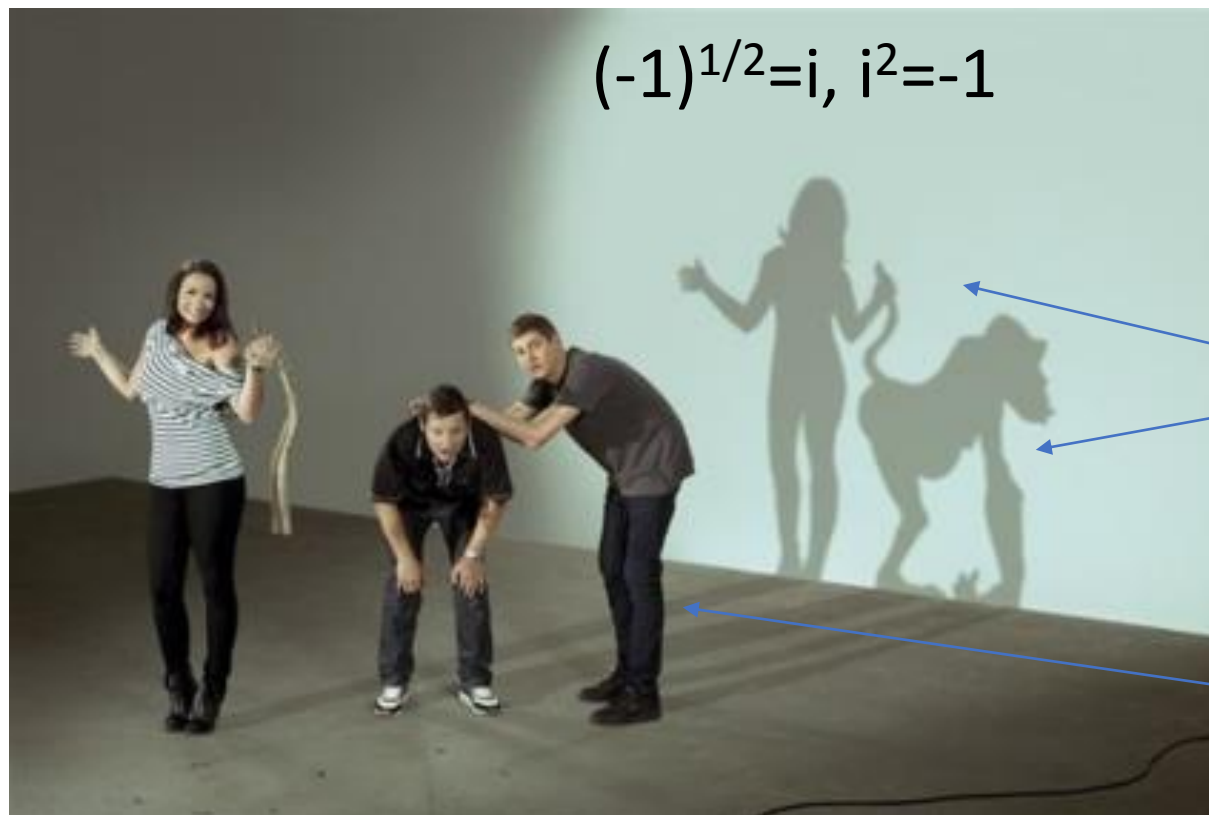
90-е годы 20 века: задачу решим на компьютере («действие» - исполняемый код ,

КТО напишет ???)

20-е годы 21 века: задачу решит ИИ («действие» - вспоминание или рекурсия Геделя,

КТО «научит» ???)

кто же этот «КТО» : компьютерные науки – синергия фундаментальной физики и числовых «теней реальности»



$$(-1)^{1/2}=i, i^2=-1$$

Реально **воспринимаемое** можно объяснить **через невидимое**

«Тени» - то, что доступно для **обратимого мыслимого** (информационного) **восприятия**

Объекты физической реальности

Триплеты научных объяснений:

материя, **энергия**, **энтропия**

модели, алгоритмы, **информация**

числа, слова, **смыслы**

в процессе восприятия «теней» информация об физических объектах всегда искажается ?! (догмы Св. Августина)

Больше — не значит лучше или принцип Анны Карениной: .

принцип Анны Карениной:

- одновременное сочетание всех необходимых факторов для получения решения некоторой проблемы является исключением.

Вероятностная версия принципа:

Число ситуаций, когда что-то может пойти не так, гораздо больше чем число ситуаций, когда всё идёт как надо.

Поведенческая версия принципа: В благополучные периоды существования системы ведут себя одинаково, а в моменты кризиса их поведения могут радикально разниться.

Афоризм Толстого : *«Все хорошо адаптированные системы похожи, все неприспособленные системы не справляются с адаптацией, но каждая по-своему»*

Парадокс: когда различие между системами возрастает, они становятся более скоррелированными (поведенчески похожими)

«Принцип хрупкости хорошего»

ак. В. И. Арнольд

- ... для системы, принадлежащей части границы устойчивости, при малом изменении параметров более вероятно попадание в область неустойчивости, чем в область устойчивости.

Это есть описание принципа Арнольда, согласно которому :

всё хорошее (например, устойчивость) более «хрупко», чем плохое. Все «хорошие» объекты **удовлетворяют нескольким требованиям одновременно**, плохим же считается объект, обладающий хотя бы одним из ряда недостатков (см. принцип Анны Карениной).

Так в галактических масштабах скорость света как функционал для передачи информации становится непригодно медленной, а в межгалактических - вообще теряет смысл. Скорость света действует как некое ограничение на передачу информации во Вселенной.

При сверхсветовой передаче информации «ломается» причинность событий и возможность их описания с помощью физических законов. Хотя, квантовый мир формализуется без аспекта причинности, который вытеснен понятиями вероятности, суперпозиции и спутанности, а описание сингулярностей наподобие черных дыр или «Большого Взрыва» не используют даже и их

- В современной науке принято называть «истинной» лишь ту информацию, которая подкреплена **формальными доказательствами**. Однако, согласно теоремам Геделя, не всякая истина доказуема, а сами формальные доказательства трудны для понимания и требуют больших затрат времени.
- **имеется и другой когнитивно-эмпирический критерий «истины» или «правдивости»**, основанный на повторяемости воспринимаемых мозгом эффектов и доверию к повторяющимся утверждениям, которые обрабатываются мозгом быстрее и... вызывают у него большее доверие.
- Следование критерию «повторяемости фактов» логично с точки зрения «здравого смысла» или теории «условных рефлексов», но может приводить к глубоким когнитивным искажениям, известным как эффект «иллюзорности правды». Этот эффект под названием «дипфейк» в настоящее время легко реализуют современные системы ИИ
- Итак, концепция «машинного обучения» не отличается «идейно глубиной» и оригинальностью, но хорошо интерпретируема как следствие из правила: «повторение – мать учения». Машинное обучение на базе этой концепции приводит к эффекту «понимания» на основе корреляций, но лишь того, что многократно повторяется размеченной обучающей выборке (deep learning) или явно «подкрепляются» в потоке поступающих в систему данных (reinforcement learning).

«Машинное обучение» - обратная задача «компьютерных наук»

прямая задача: «алгоритм -> данные»

обратная задача «данные -> алгоритм»

ИНС - это вариант **формальной системы**, способной к «обучению» для реализации «рекурсии Геделя» с целью «конструирования аксиоматики» для описания решения выбранного класса «обратных задач»

смысл результата вычисления, согласно теоремам о неполноте Геделя, осмысление результаты вычисления ИТ требует «внешнего наблюдателя». ИНС «вычисляет» числа-знаки (см. Г. Вейль), но не может их объяснить их смыслы (4 - «красный цвет»)

Естественный интеллект:

{ действия , основанные на их понимании законов природы и стремления объяснить смысл производимых действия с учетом контекста событий или гипотез }

{ **граница формальной (частичной) рациональности** }

ИИ - виртуальная реальность:

{ действия направленные на имитацию (моделирование) процессов на основе формальных (логически объяснимых) правил /алгоритмов }

Физика сингулярных состояний

- размерности «объектов», образующих сингулярности, нулевые. Скорость как мера изменения расстояния во времени в таких объектах теряет смысл, так как расстояния между элементами, образующими множество меры ноль, равно нулю?
- С другой стороны, количество энергии в сингулярности может быть конечным, а плотность этой энергии стремится в бесконечность
- Без того, чтобы ограничить скорость передачи света, Вселенная не могла бы существовать в том виде, в котором она существует сейчас. Скорость света как бы замедляет всё происходящее, придавая смысл понятию "расстояние".
- Хотя, кванту света - фотону, который летел до Земли, двигаясь со скоростью света, время его существования равно нулю. То есть вся жизнь Вселенной с точки зрения фотона протекает за нулевое время. Он как бы умирает в тот же момент, в какой рождается, а вот благодаря всеобщему замедлению времени мы можем этот фотон увидеть.
- **В мире без скорости, без расстояний и с нулевыми размерами возможные события происходят все и сразу.** Чтобы сделать всё происходящее постепенным и имеющим причинно-следственные связи, нужно увеличить размеры этого мира так, чтобы они стали больше, чем нулевые, и при этом **ограничить скорость передачи информации.**

Понятие адаптационной энергии,

- В 1938 году Г. Селье пришел к мысли, что приспособляясь, организм расходует некий ресурс, заданный при рождении – «адаптационную энергию».
- Если эта энергия полностью исчерпывается, организм погибает - «Экспериментальные доказательства, подтверждающие концепцию адаптации энергии».
- Считается, что адаптационная энергия может восстанавливаться и даже накапливаться в результате процессов «обучения».
-

Дж. Хинтон: ИИ vs Естественный интеллект: что «лучше»



Инструментальные аспекты: ИИ - «цифровая тень»

отбрасываемая сознанием человека на ... «память» компьютера ...



... AI is nothing but **digital brains** inside large computer

Ильей Суцкевер,
(основатель
OpenAI

идея современных технологий машинного обучения - «истина в повторении фактов»

- В современной науке принято называть «истинной» лишь ту информацию, которая подкреплена формальными логическим доказательствами наличия причинно-следственных связей . Если эксперимент не подтверждается теорией ... то эксперимент проведен неверно.
- В машинном же обучении используется **когнитивно-эмпирический критерий «истины» («правдивости»)**, основанный на повторяемости воспринимаемых фактов
- Следование критерию **«повторяемости фактов»** логично с точки зрения «здравого смысла», но может приводить к глубоким когнитивным искажениям, известным как эффект «иллюзорности правды» и «дипфейки» , которые в настоящее время легко реализуют современные системы ИИ
- Итак, современная концепция «машинного обучения» не отличается «идейной глубиной» , но следует прагматической метафоре : **«повторение – мать обучения»**.
- Такое обучение приводит к **«пониманию» фактов (усвоению их содержания)** на основе корреляций, а не их причинно-следственных связей. , Машинному обучению доступно «понимание» лишь того, что
 - многократно повторяется в размеченной обучающей выборке (deep learning) или
 - «подкрепляется» обратной связью в процессе использования поступающих в систему данных (reinforcement learning).

«Лингвистический поворот»: нейро-трансформеры на базе СК



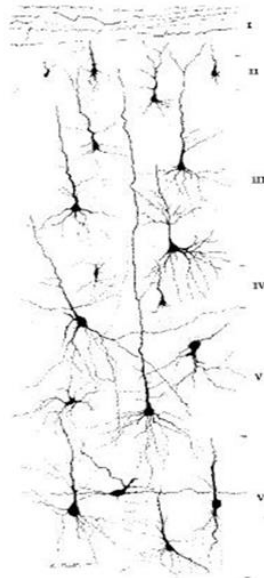
Ашиш Шарма

Суть:

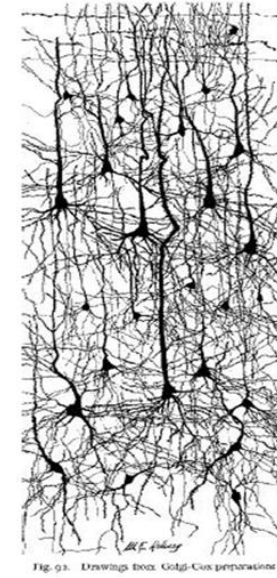
«Attention Is All You Need» (2017)

« ... **memory** is main part of digital brains» (2023)

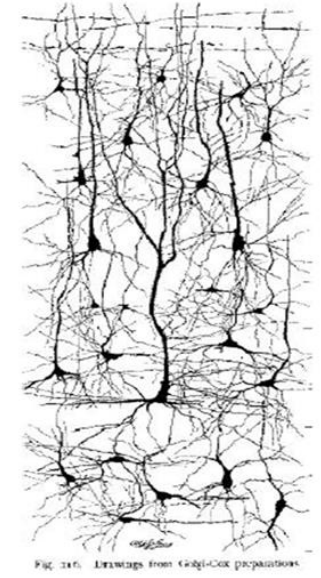
Структура нейронной сети мозга ребенка в момент рождения (опыта нет), в 2 года (много данных, но нет понимания) и 6 лет (структура нейронной сети «оптимизирована» к урону сложности среды обитания)



Birth



2 years



6 Years

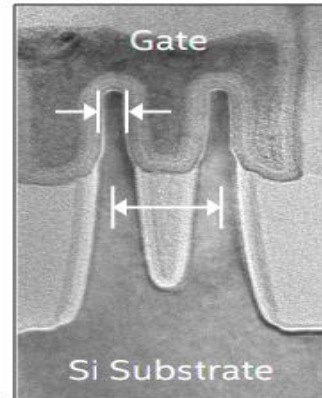
В природе «обучение» – это процесс, в котором «информация» обретает «форму», например путем структуризации нейронной сети мозга

«Король» умер, да здравствует «король»:

или Less Moor

8 nm Fin Width

42 nm Fin Pitch

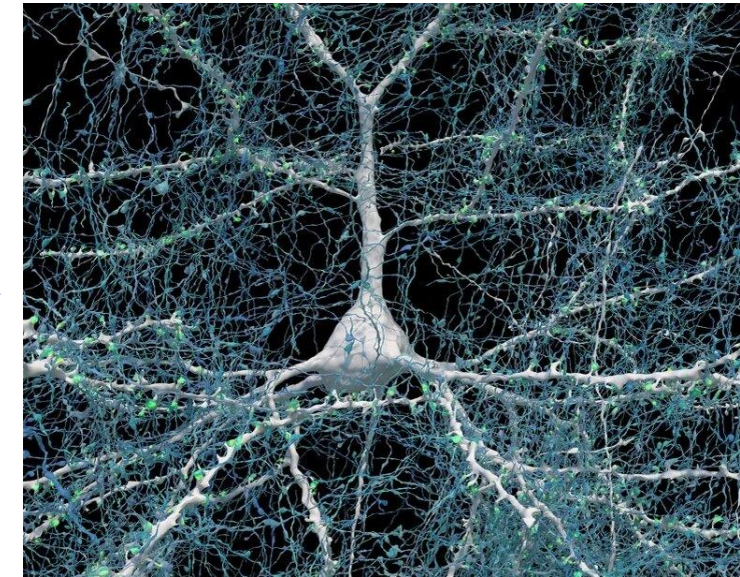


Современные 3-5 нм микропроцессоры (МП), содержит **до 50 млрд. транзисторов**. МП потребляет **«100» ватт**, имеет тактовую частоту 3 ГГц, а **размерность пространства состояний можно оценить величиной 2^{50}**

«Закон» Мура – ...все ...

???

Очевидно, что пользователей интересует не **FLOPS**, а успешное решение конкретной прикладной задачи то есть «прикладная производительность» ЦКП



Мозг одного человека – **86 млрд нейронов** (биологических аналогов транзисторов), «тактовая частота» 100 Гц, потребляет **«20» ватт**, имеет объем пространства состояний **2^{86}** , что в **2^{36} раз больше, чем у МП !!!!!**

Часть 2 (практика)

ВОПЛОЩЕНИЕ «СМЫСЛОВ» В ПРОЦЕССЫ ВЫЧИСЛЕНИЙ

«В начале было **Слово**»

Евангелия от Иоанна

«Все есть **число**»

Пифагор 570-490 до н.э.

ВЫЧИСЛЕНИЯ **чисел** → ВЫЧИСЛЕНИЕ **СМЫСЛОВ**

Эра конечных автоматов, вычисляющих числа с использованием программ-алгоритмов

Эра механических автоматов, исполняющих один алгоритм, вычисления



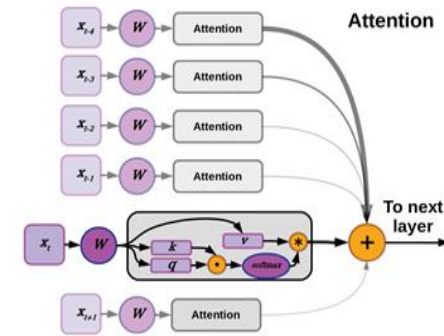
Алгоритм записанный на **естественном языке**, понятном человеку

Алгоритм вычисления записанный **человеком** на **языке «понятным» компьютерам**

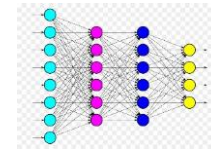


Эра «психозоя» Вернадского - «интеллектуальных» компьютерных платформ, вычисляющих «СМЫСЛЫ» слов и алгоритмов

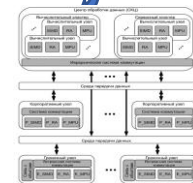
X-входные данные и описание заданий



Y-выходные данные - результаты



описание процессов на «языке данных»



Описание процессов на «языке понятий»

Воплощение или реификация продуктов мышления человека (программы) при в материально-вещественные формы – результаты вычислений

Зачем нужно «машинное обучение» для суперкомпьютерных платформ

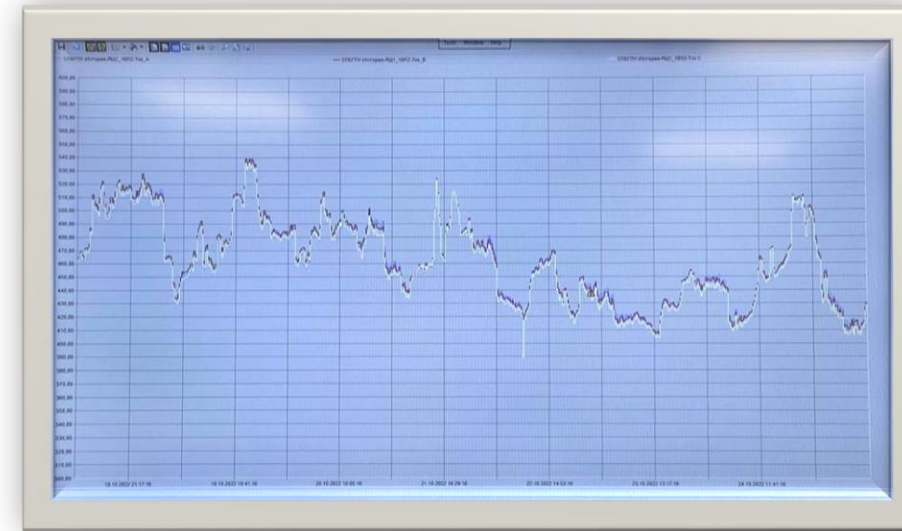
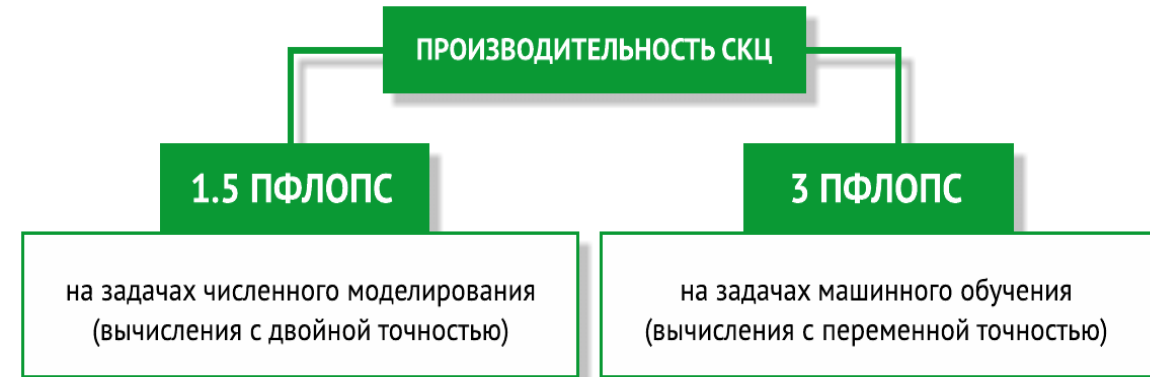
Причина 1 Сложность гибридных (CPU + GPU + FPGA) СК кластеров
Согласно рейтингу top50.supercomputers.ru СКЦ «Политехнический» это

- Более 1000 пользователей
- Более 100 научных групп
- Более 30 промышленных организаций
 - > 25 миллионов узло-часов
 - > 2 млн расчетных заданий

Причина 2. Результаты функционирования СК (связанные с данными из сети интернет) можно рассматривать как «размеченные обучающие выборки» необходимые для «обучения» системы управления (например, диспетчеризации) процессами вычислений.

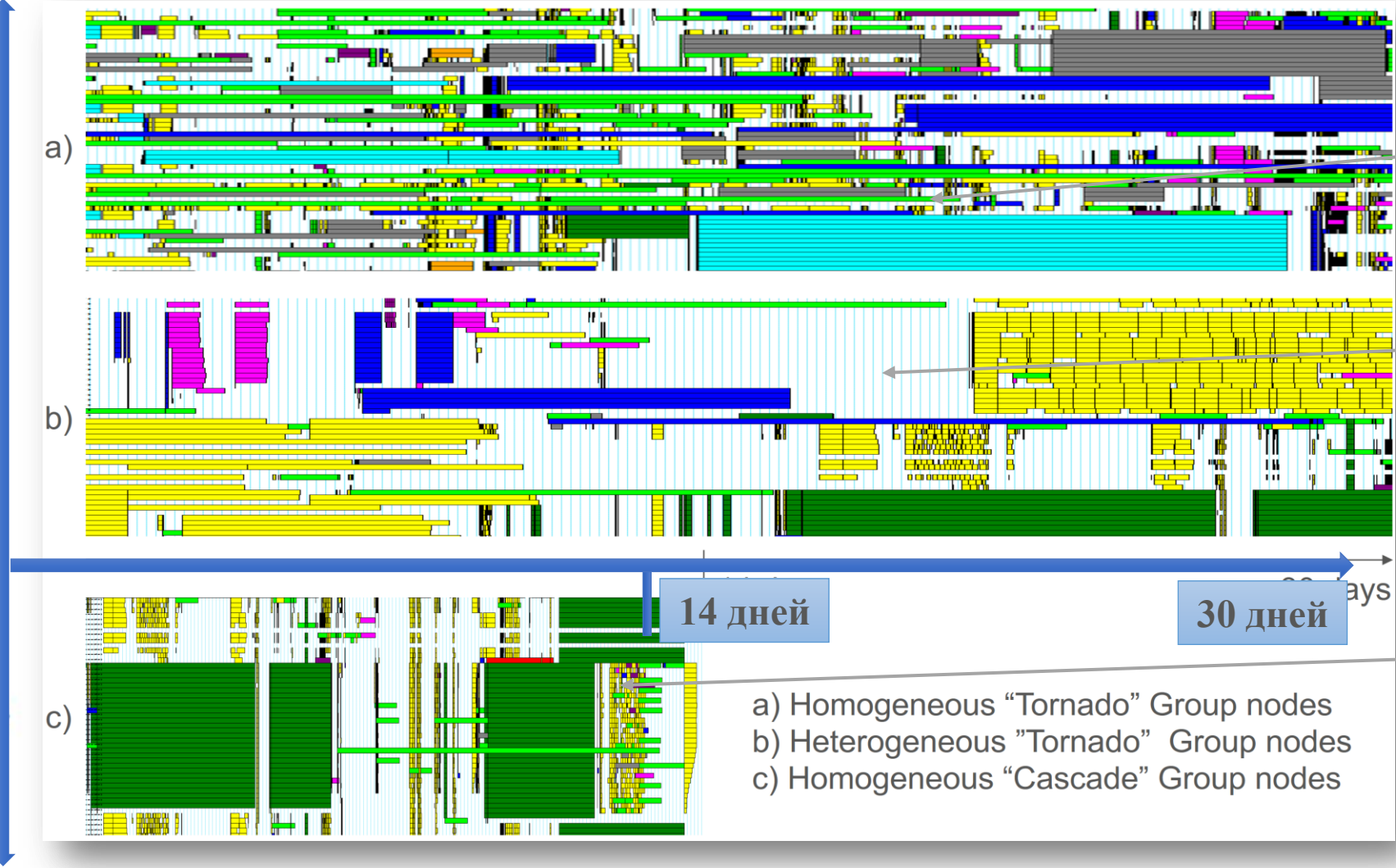
Причина 3. надо знать сколько в Вт «стоит» решение одной прикладной задачи» (метрика эффективности «число успешно решенных прикладных задач на 1 Вт / «дельта» секунд»

Гибридный СК «ПОЛИТЕХНИЧЕСКИЙ»



Характер загрузки и(t) гибридных узлов СК работающего в режиме ЦКП

Загрузка узлов $u(t)$



Не более 60%

Не более 40%

Не более 50%

Время

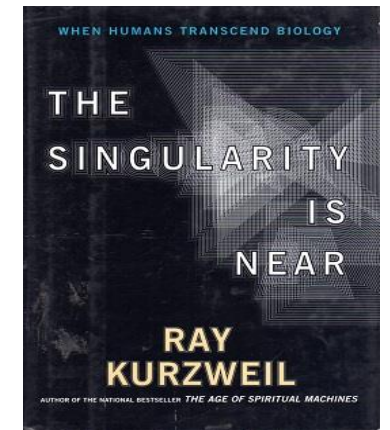
Функция загрузки $u(t)$ узлов – случайная (???) величина



МЫСЛЬ
БУДУЩЕМ

«Гладко было на бумаге, но ... забыли про «овраги»

- **Создание LLM трансформеров** и их успешная апробация ChatGPT, породила иллюзию того, что **«точка сингулярности»** Р. Курцвейла уже рядом



- В **«больших лингвистических моделях» (LLM)** любой «цифровой объект» есть «слово», которому на множестве «осмысленных» текстов можно сопоставить меру вероятности, а не меру «смысла».

не надо из LLM (GPT) делать «МОЛОТОК»...

- для LLM **любое «число»** тоже есть **«не имеющее смысла слово»**, поэтому фразы, сгенерированная ChatGPT:

- «температура пациента 36,6»
- «температура пациента 38,6»

«лингвистически» эквивалентны, но для врача имеют разный «смысл»

Чему надо научиться:

Находить самореферентные (Гедделевские) решение «обратных задач», отражающих как смысловые, так и количественные особенности прикладной проблемы

Предлагаемое решение: Konstantinov A.V., Utkin L.V. Incorporating Expert Rules into Neural Networks in the Framework of Concept-Based Learning // arXiv:2402.14726



от мыслимого (объекта сознания) к
вычислимому (объекту
реальности)

«рациональная» логика компьютерных наук:

- 1) вычисляю, следовательно понимаю (могу закодировать)→
- 2) понимаю значит мыслю и могу описать с помощью слов =>
- 3) мыслю значит могу воплотить, то есть «вычислить» (фундаментальный вопрос, который до сих пор не имеет ответа : **можно ли вычислить самого себя ?**)

«информационная физика» процессов вычислений

-**граница диссипации тепла** (неравенство Ландауэра, 1961 г.) при потере одного бита (равновероятное состояние гейта) выделяется тепло не меньше, чем

$$\Delta Q \Rightarrow k * T * \ln 2 > 0$$

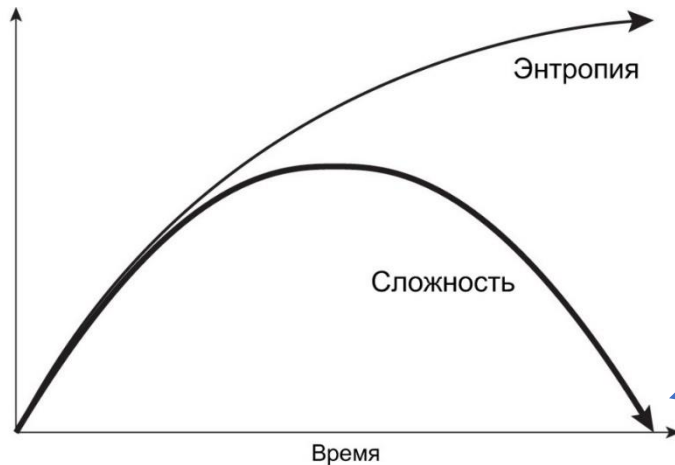
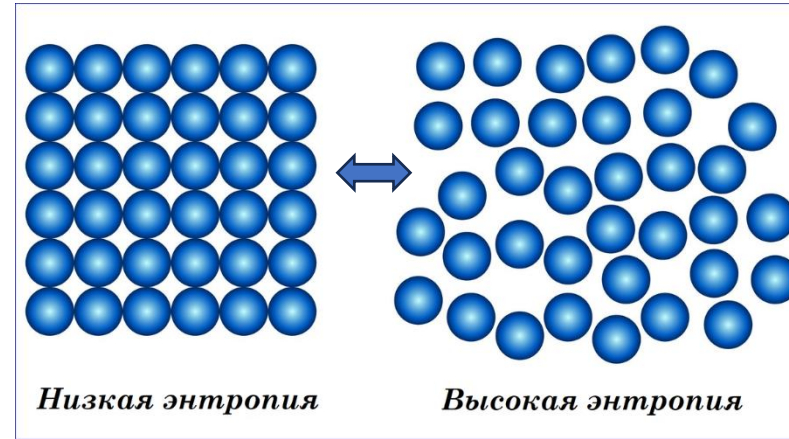
информационная трактовка неравенства, имеем $k * T * \ln 2 = -k * T * \ln(1/2)$ **для равновероятных состояний, то есть $p=1/2$, имеем $\Delta Q \Rightarrow -k * T * \ln(p)$, а согласно формуле Шеннона $I = -\log p$**

тогда граничное значение диссипации тепла $\Delta Q \Rightarrow -k * T * \ln(p) = k * T * I > 0 \dots$
 $I \leq \Delta Q / k * T$

Что общего в термодинамике и теории передачи информации ? -> использование понятия « энтропии »

$$\Delta S = \frac{\Delta Q}{T}$$

Энтропия в термодинамике — мера рассеяния энергии.



Количество данных $K \ll N$ — число возможных комбинаций из этих данных

Энтропия в теории К. Шеннона мера «рассеивания» формы объекта

Вопрос: «сложность» это «форма»? Как сложность описания объекта растет по мере роста энтропии ?!

изменение внутренней энергии системы всегда отрицательно (**энергия рассеивается**), а изменение энтропии всегда положительно (**неопределенность растет**).

$$E = \sum -p_i \log_2(p_i) \Rightarrow \sum p_i \log_2 \frac{1}{p_i}$$

Ин-форм-ация («IN/FORMATION») как мера субъективного различия

Математики имеют дело только
с каталогом знаков

Г. Вейль



Рассматривая художественную картину ясно, что дело не в холсте или красках (материальных субстанциях), а в структуре, образе, то есть в **визуальном различии**, или субъективном коде мыслимых представлений (в пояснениях к этим представлениям в форме слов/понятий)

Код – всегда обозначает нечто **мыслимое**, а значит, согласно Декарту – **существующее** либо в реальности, либо в сознании

Примеры:

- Лингвистика – сложнее, чем химия и физика. Есть слова амони́мы – знаки, обозначающие разных существности: «коса», «ключ» ...
- Биология - тексты молекулы ДНК это описание реального объекта с помощью $64 = 2^5$ кадонов, которые кодируют 32+32 аминокислоты. Код ДНК не однозначен – один кадон кодирует две аминокислоты

Процесс реификации (reification) абстракций

- Определение - реификация (reification)
 - процесс , в котором абстрактная идея (например, программа) или результат вычислений по программе (число) превращается в данные, другой реальный (программа) или мыслимый объект

По сути реификация – это процесс «овеществления» продуктов мыслительной деятельности (абстрактных утверждений), например,

- описание языка программирования с помощью другого языка программирования
- трансформация понятий в материально-вещественные формы или объекты

Пример: Посредством реификации то, что ранее было неявным, невыраженным и невыразимым, теперь **эксплицитно (явным образом)** формулируется и становится доступным для физических, логических или вычислительных манипуляций (например, с помощью рефлексивных языков программирования).

Так, овеществление является одной из часто используемых техник программирования, например:

- «овеществление» адреса памяти для его непосредственного использования в каком либо специальном контроллере

```
char* buffer = (char*) 0xB8000000;  
buffer[0] = 10;
```

- описания некоторого утверждения, не указывая его явно

Варианты воплощение принципов «самореферентных» процессов вычислений

Уровень «**объяснения**» результатов

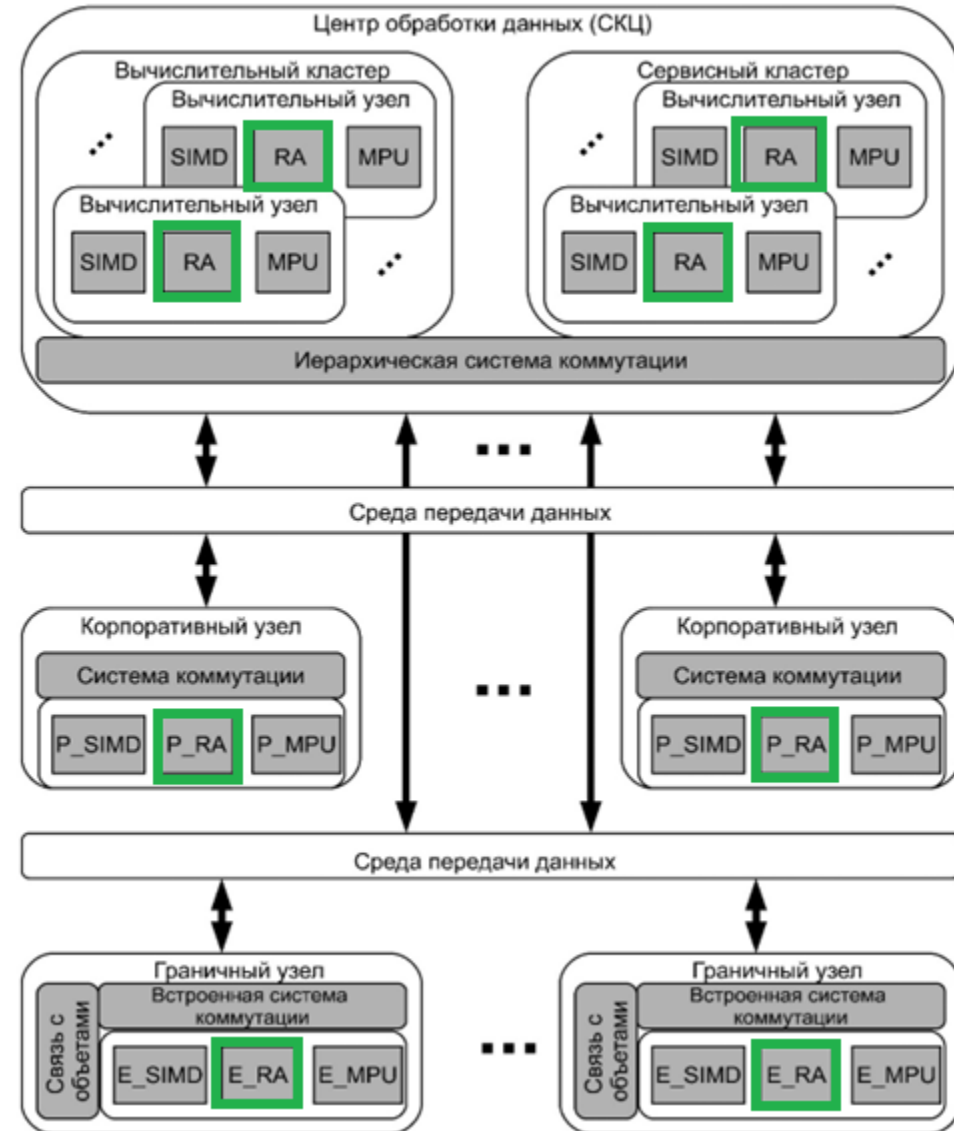
Энерго-вычислительная эффективность >4 Гфлопс/Вт

Уровень «машины Геделя»: механизмов «внимания» и «понимания» реализуемых процессов

Энерго-вычислительная эффективность >10 Гфлопс/Вт

Уровень «машины Тьюринга» - CPU/GPU/FPGA процессоры

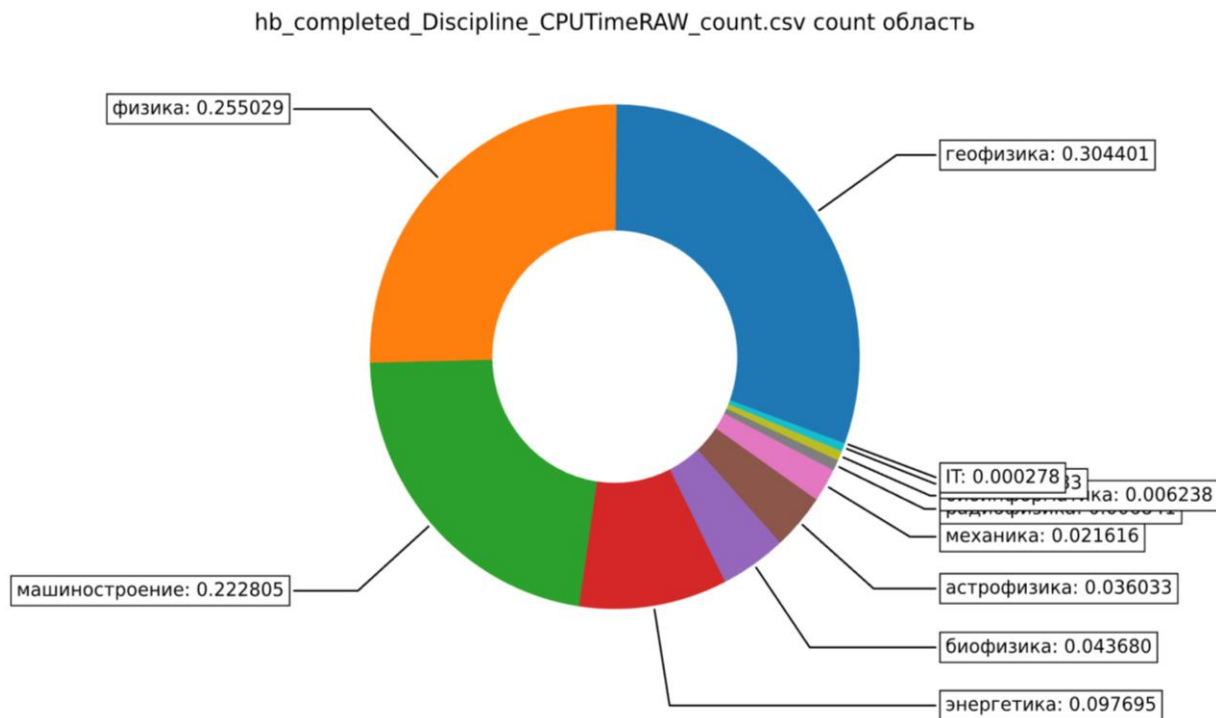
Энерго-вычислительная эффективность >20 Гфлопс/Вт



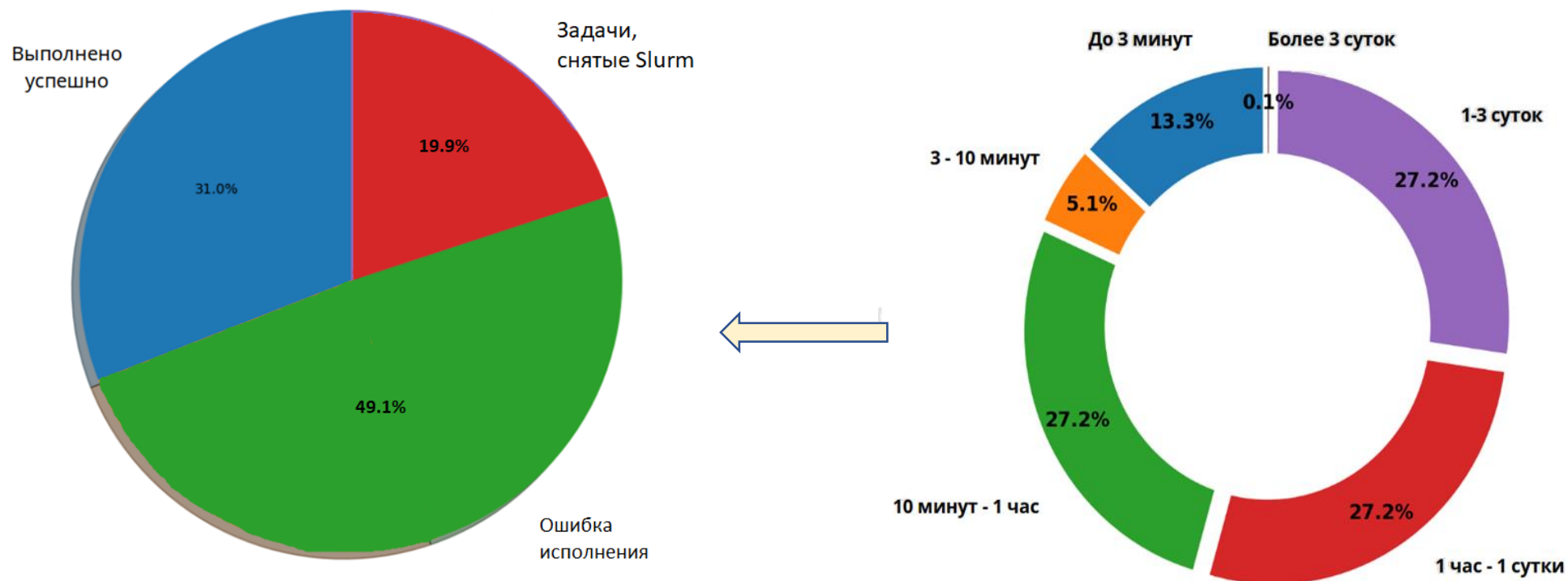
Количественные меры сложности описания работы СК в режиме «центров коллективного пользования» (ЦКП)

соотношение суммарных времен исполнения задач в различных областях знаний

область	CPUTimeRAW (s)	per
геофизика	62923292888	0.304401
физика	52717309792	0.255029
машиностроение	46056422576	0.222805
энергетика	20194851856	0.097695
биофизика	9029163752	0.04368
астрофизика	7448485016	0.036033
механика	4468425024	0.021616
радиофизика	1414211624	0.006841
биоинформатика	1289548064	0.006238
NaN	1112776304	0.005383
IT	57423616	0.000278

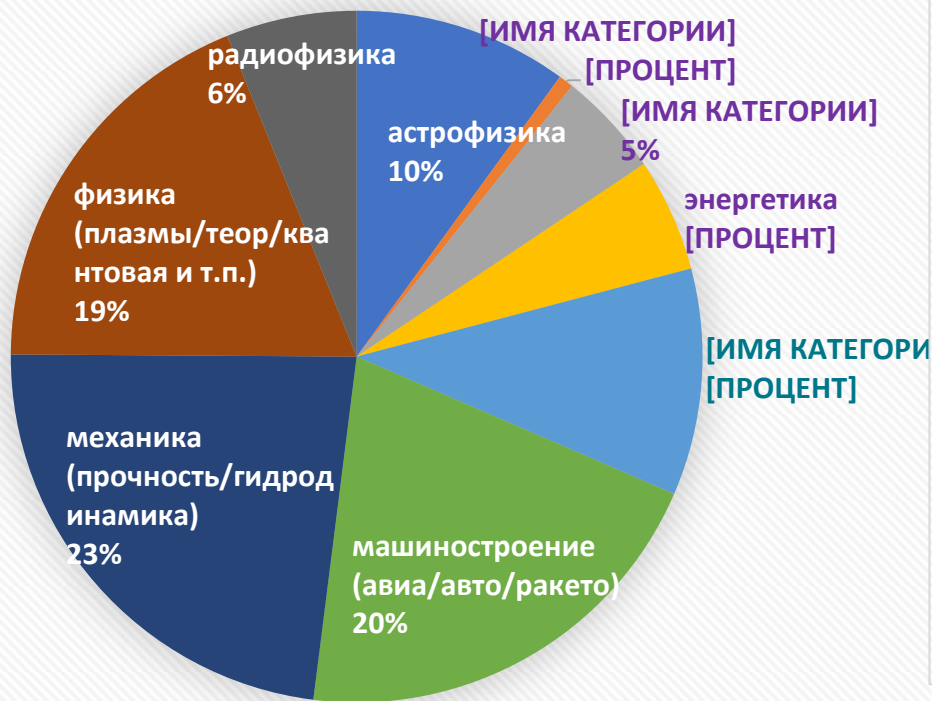


Проблема «выживания» прикладных заданий в ЦКП



Проблема: Успешно выполненные задачи ЦКП составляют **менее 1/3** от общего числа (формально отлаженных) прикладных задач пользователей

«Метафора» сложности процессов управление ЦКП

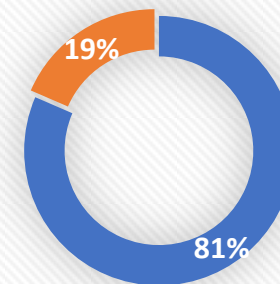


- астрофизика
- биоинформатика
- биофизика
- энергетика (энергомаш)
- геофизика (сейсмика/геофизика)
- машиностроение (авиа/авто/ракето)
- механика (прочность/гидродинамика)
- физика (плазмы/теор/квантовая и т.п.)
- радиофизика

Характер решаемых задач



Классы пользователей



- Внутренний
- Внешний

О полезность суррогатных (объяснительных) моделей для решения «прямых и обратных задач»

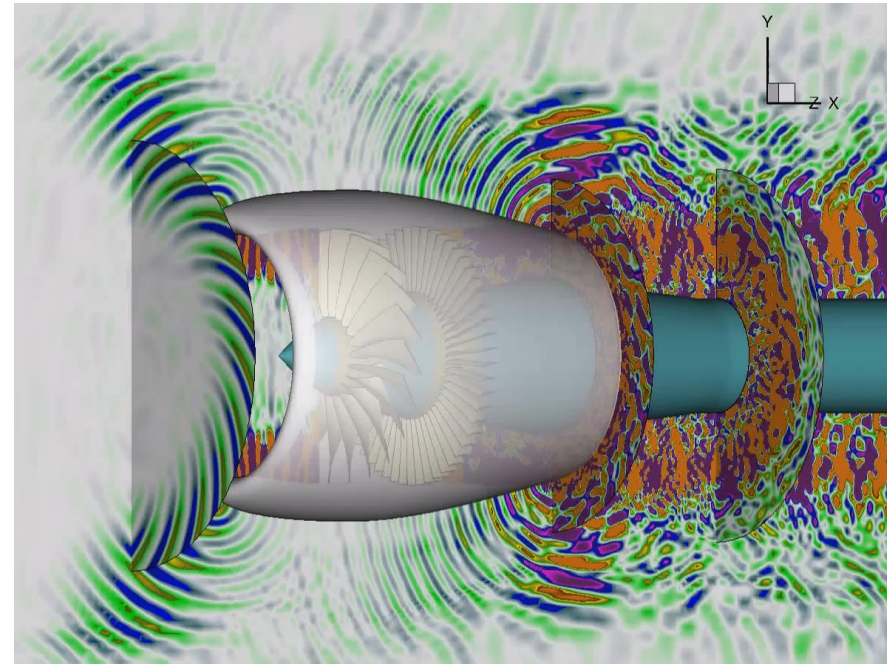


Физическая модель турбореактивного авиационного двигателя

прямой расчет и визуализация акустических характеристик турбореактивного двигателя

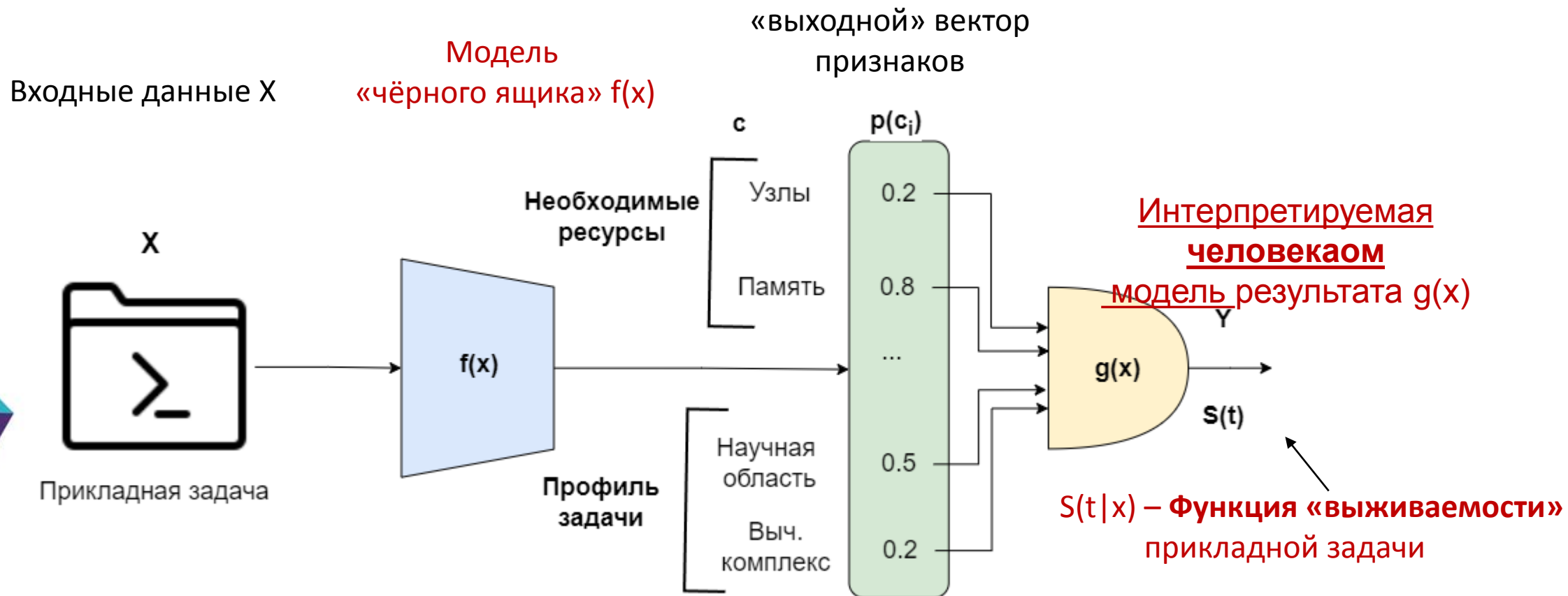
- **Прямая** задача: поиск решения, основанного на точных **моделях** (первых **принципов**) описания физических явлений
- **Обратная** задача: использование **суррогатных моделей**, объясняющих результаты вычислений в терминах понятий, например, в форме визуализации протекающих процессов

Анимация, «больших данных» полученных в результате прямого цифрового моделирования пример **применения суррогатной модели** для интерпретации результатов расчета

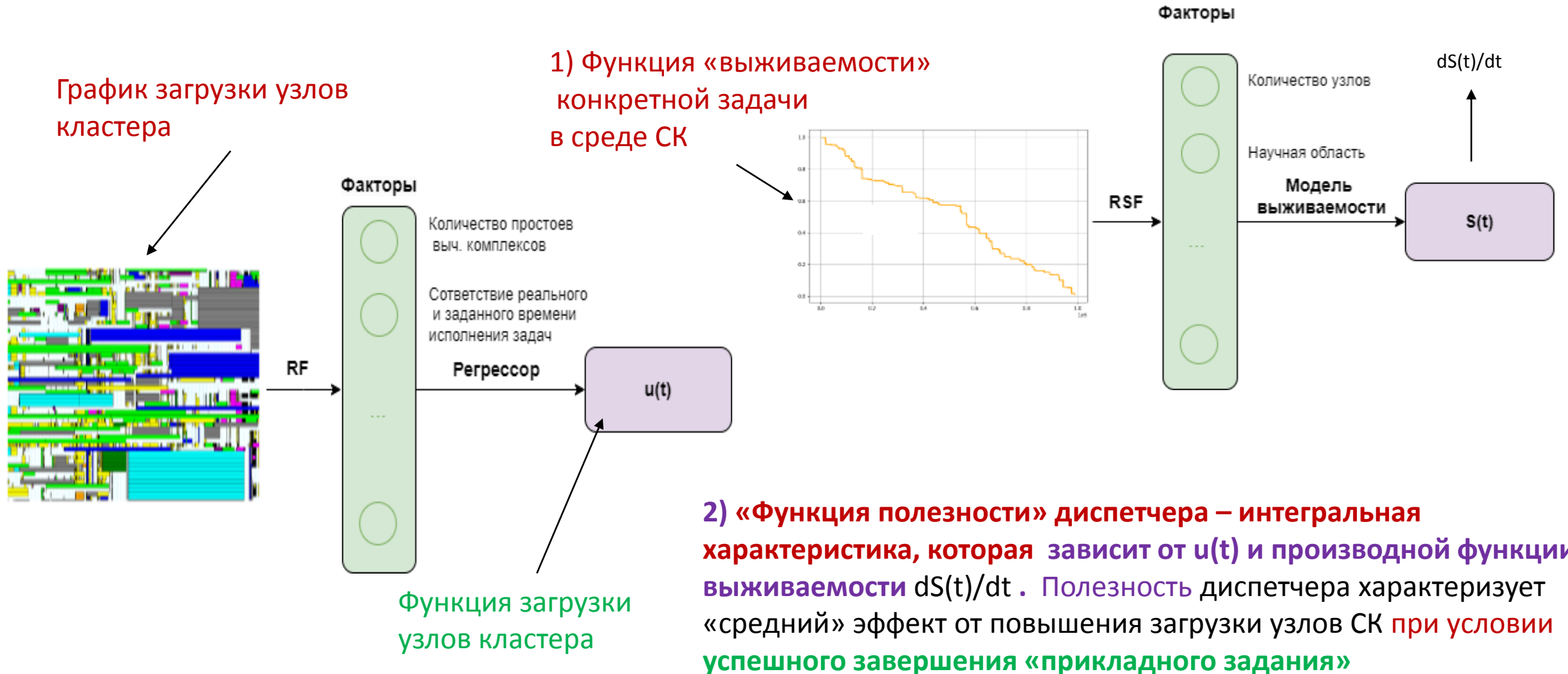


Мультимодальные concept-bottleneck model в обучении СК

– поиск путей решения «алгоритмически неразрешимых задач»:



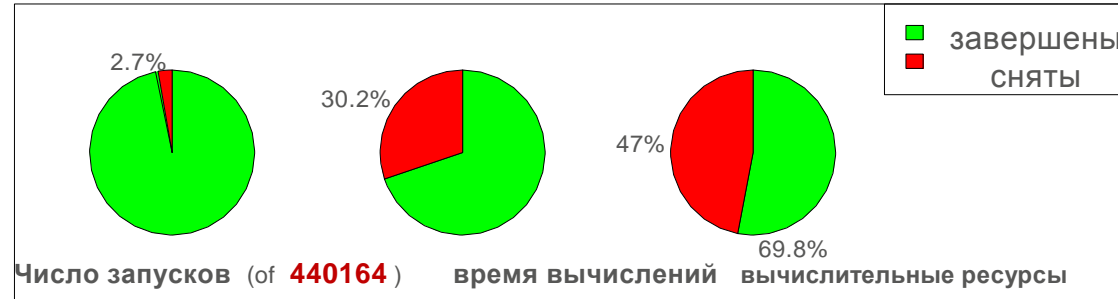
Процесс «объяснения» результатов решения СК



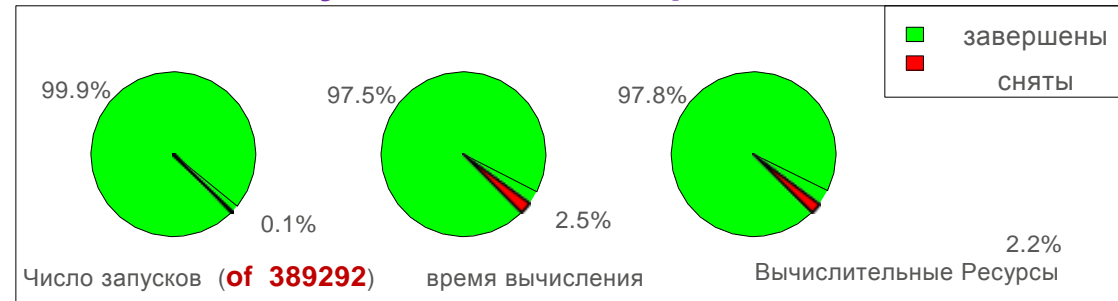


МАШИННОЕ ОБУЧЕНИЕ СК – ВОЗМОЖНОСТЬ ПОВЫШЕНИЯ «ПРИКЛАДНОЙ» ПРОИЗВОДИТЕЛЬНОСТЬ

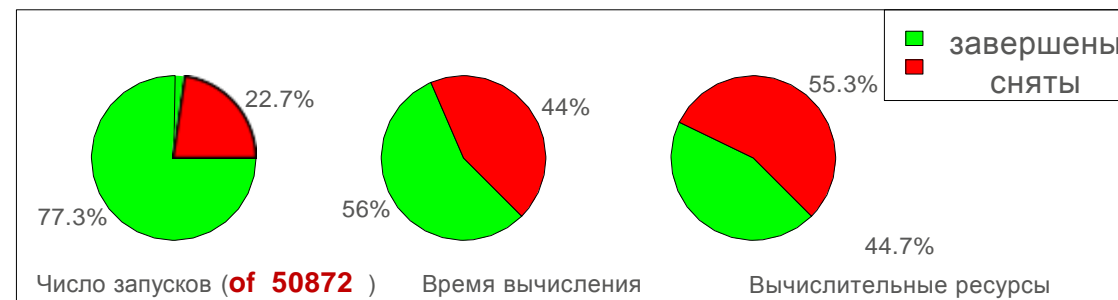
Общая характеристика эффективности для всех видов запусков заданий



Класс эквивалентности 1: Автоматический запуск известных заранее заданий



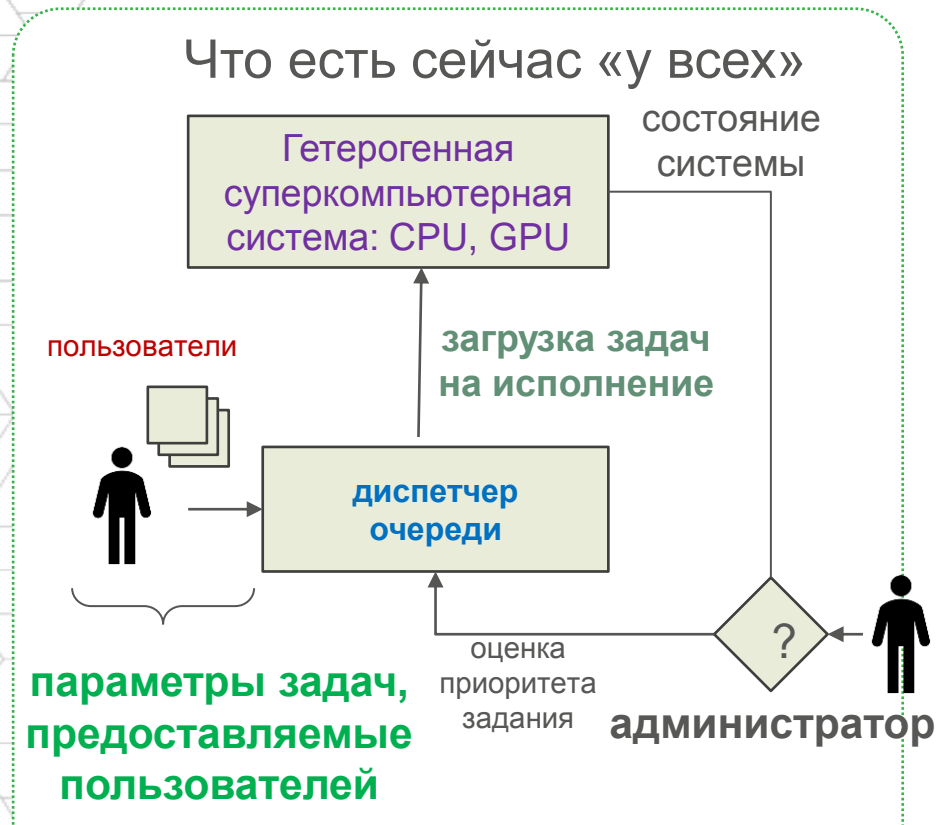
Класс эквивалентности 2: Запуск прикладных задания «вручном» режиме



Если параметры заданий пользователей, точно известны, («энтропия» знаний о СК/задача =0): то **настройки процессов вычислений можно автоматизировать (есть алгоритм решения)**

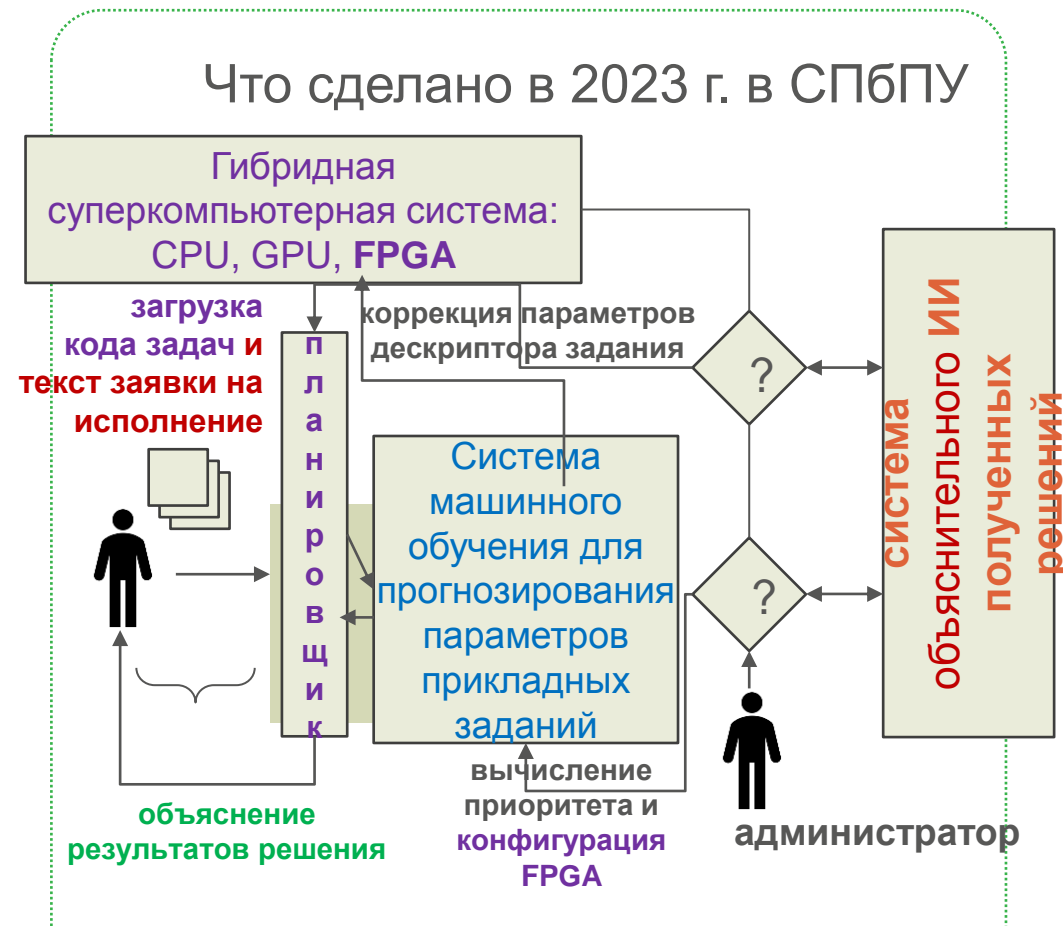


ВОПЛОЩЕНИЕ МАШИННОГО ОБУЧЕНИЯ В ГИБРИДНЫЙ СК



Проблемы:

- неточность оценок параметров заданий, которые формирует пользователи
- Неточные оценки порождают «ошибки» диспетчера в выделении ресурсов для пользовательских заданий



Как проблему решить : встроить систему машинного обучения диспетчера кластера, которая использует модель компетенций пользователя и функцию загрузки узлов кластера, чтобы прикладные задания завершились успешно



ОБУЧЕНИЕ ДИСПЕТЧЕРА СК В ТЕРМИНАХ «ВРЕМЕНИ ЖИЗНИ ЗАДАЧИ» НА ОСНОВЕ «ЦЕНЗУРИРОВАННЫХ ДАННЫХ»

Выживание (успешное завершение) задачи в ЦКП – это пара времен «отказа» и «цензурирования» (T, U) :

T – время отказа (задача за отведенное пользователем завершилась с ошибкой)

U – время цензурирования (задача превысила запрошенное пользователем время и была снята «снята» диспетчером)

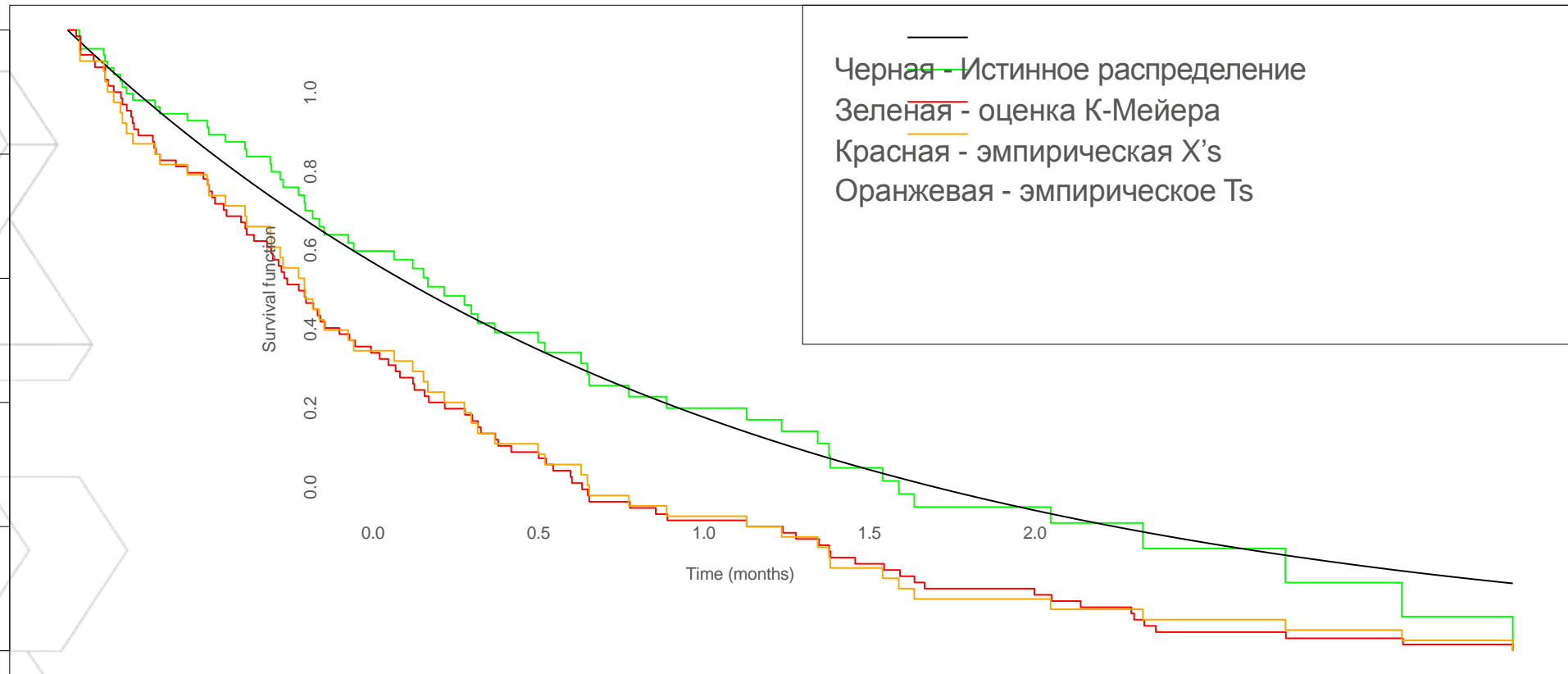
В статистике цензурированными называют данные, в которых значения признаков известны только на некоторых интервалах наблюдения. В нашем случае, используемое для «обучения» диспетчера ЦКП данные, представляет собой (X, δ) , где

$$X = T \wedge U; \delta = 1 \parallel \text{—признак цензурирования} \\ \{ T \leq U \}$$

Реально для наблюдения доступны **наименьшее из времен** T или U , но системе обучения известно, **какое из двух событий** (отказ или цензурирование) произошло раньше, что **позволяет ввести в модель ЦКП абстракцию «времен жизни прикладной задачи»**



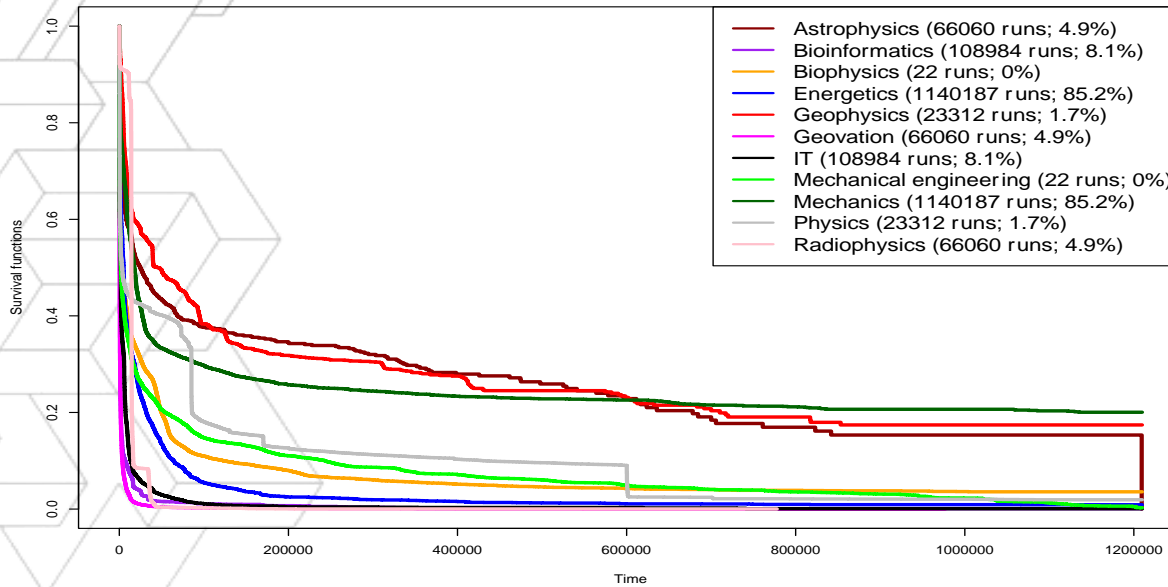
ИМЕРИЧЕСКИЕ РАСПРЕДЕЛЕНИЯ ДЛЯ ЦЕНЗУРИРОВАННЫХ $X = T \wedge U$ (СТАТИСТИКА КАПЛАНА–МЕЙЕРА) И ДАННЫХ НЕПОСРЕДСТВЕННО ВРЕМЕНИ «ОТКАЗА» (T)



Вывод: распределение на основе измеренного временем «отказов» T_s **недооценивают** реальную величину времени T , что приводит к «ошибкам» прогнозов



РАСПРЕДЕЛЕНИЯ ВРЕМЕН УСПЕШНОГО ЗАВЕРШЕНИЯ ЗАДАЧ В ЦКП

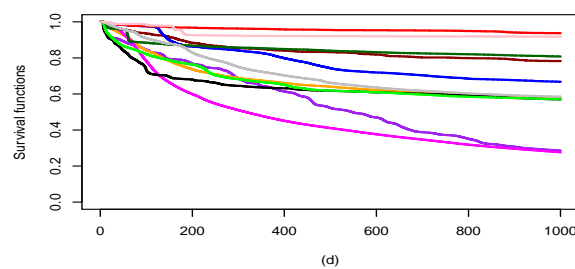
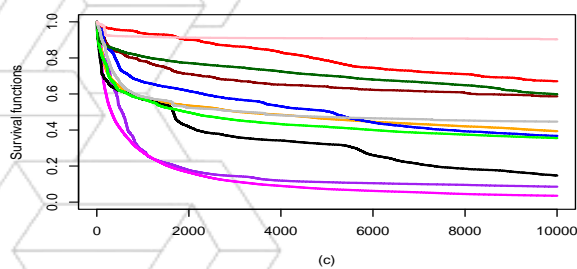
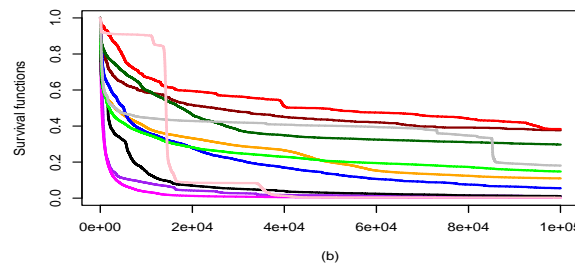
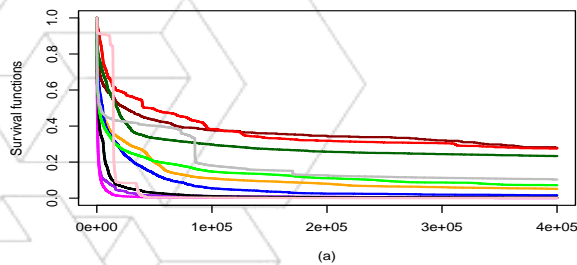


Показано распределение **времен успешного завершения задач** для различных групп пользователей ЦКП.

Значение по оси «у» равно «вероятности» (доли по отношению к всему множеству задач) не завершения ПРИКЛАДНЫХ задач рассматриваемого класса к моменту времени t (по оси x).

Если в некоторый конкретный момент времени t функция «отказа» имеет достаточно «большое значение», это означает, что большая часть прикладных задач этого класса:

- еще не завершилась или
- завершилась неудачно (цензурированы).

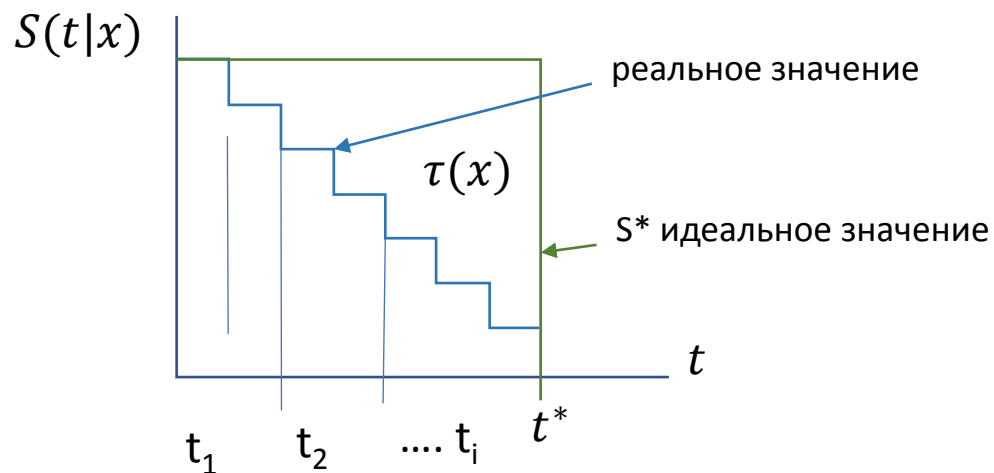


Вывод:

Если бы не было цензурирования и все задачи имели конечное время исполнения, то все функции отказа пришли бы в 0

«Абстракции», которые используются для обучения

Функции
1) «выживаемости»
задач



q «персональный» фактор
для конкретного пользователя
и задания

2. «Функция полезности» диспетчера СК характеризует «средний» эффект от влияния различных факторов на «успешное» завершения «прикладного задания»

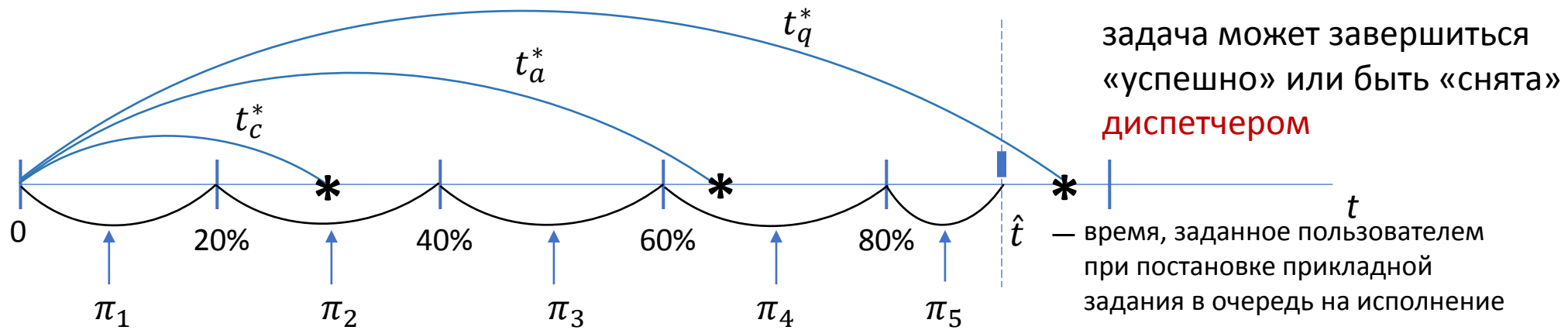
Оценка любого «фактора» $\{q\}$, влияющего на «реальную производительность» :

$$IU = \int_0^{t_{max}} q(t) \frac{\partial S(t)}{\partial \tau} dt = \int_0^{t_{max}} q(t) \rho(t) dt$$

в случае
диспетчера СК

$$IU = \sum_{i-\text{пользователь}}^N \sum_{j-\text{задача}}^M \int_0^{t_{max}(i,j)} u_{ij}(t) \rho_{ij}(t|x_{ij}) dt$$

Формализация задачи повышения «прикладной производительности»

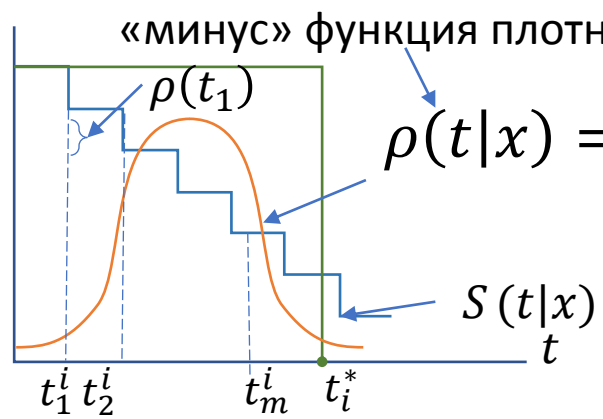


π_i - доля задач, которые попали в i -й интервал (оценка вероятности того, что)

$\gamma_1 \quad \gamma_2 \quad \dots \quad \gamma_m$

«минус» функция плотности

график функция выживаемости аналог гистограммы распределения: дискретные «приращения» графика в точках «событие»: $\rho(t_1) = 0.05 \quad \rho(t_2) = 0.08$



$$\rho(t|x) = \frac{\partial S(t|x)}{\partial t}$$

задача q превысила время \hat{t}_q произошло прерывание (деление на 0) заикливание задачи

Функция «полезности диспетчера»:

$$IU = \int_0^{t_{max}} u(t) \cdot \rho(t|x) dt$$

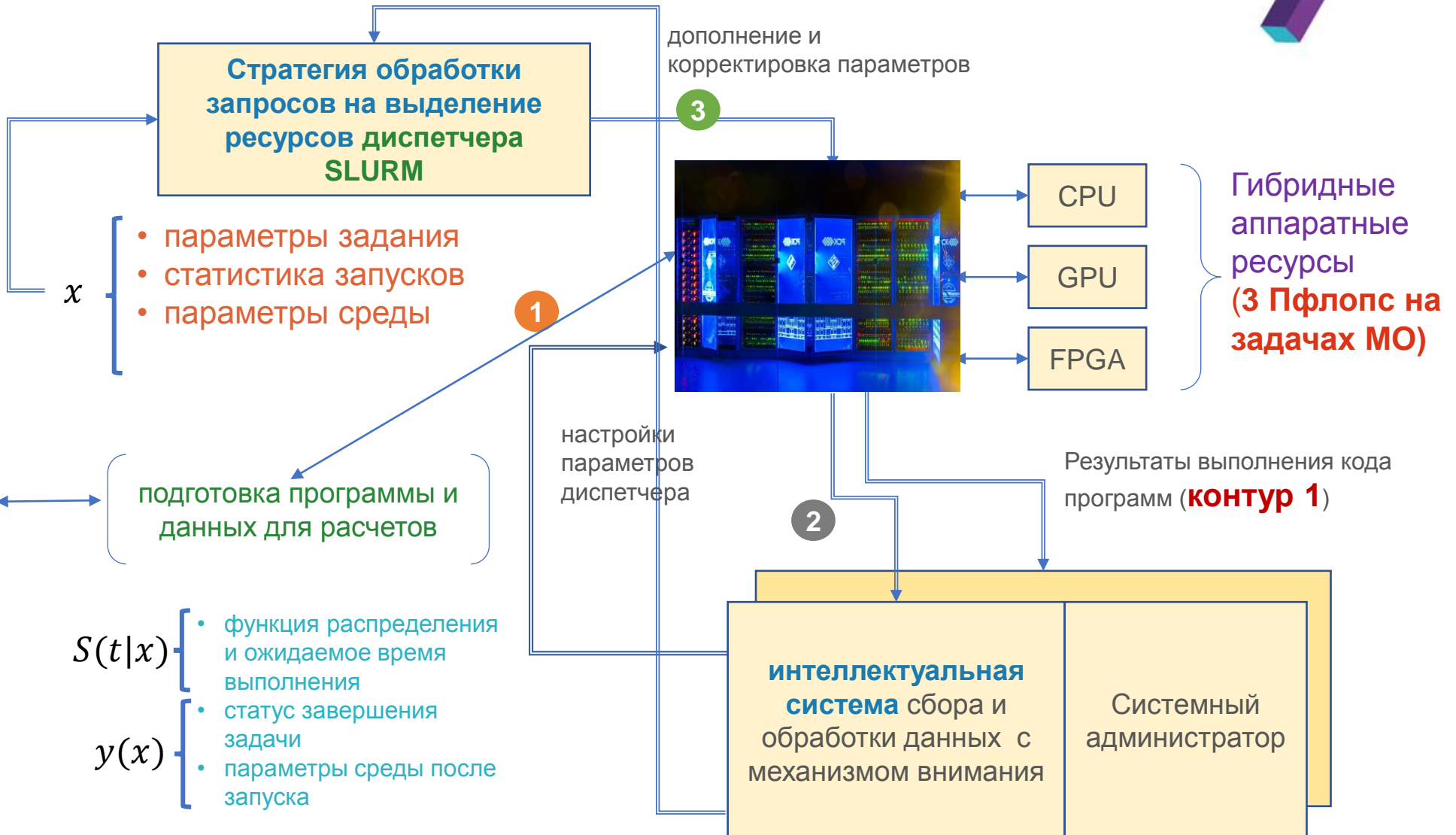
обозначения :

$u(t)$ – доля занятых узлов кластера; $S(t|x)$ – функция выживаемости

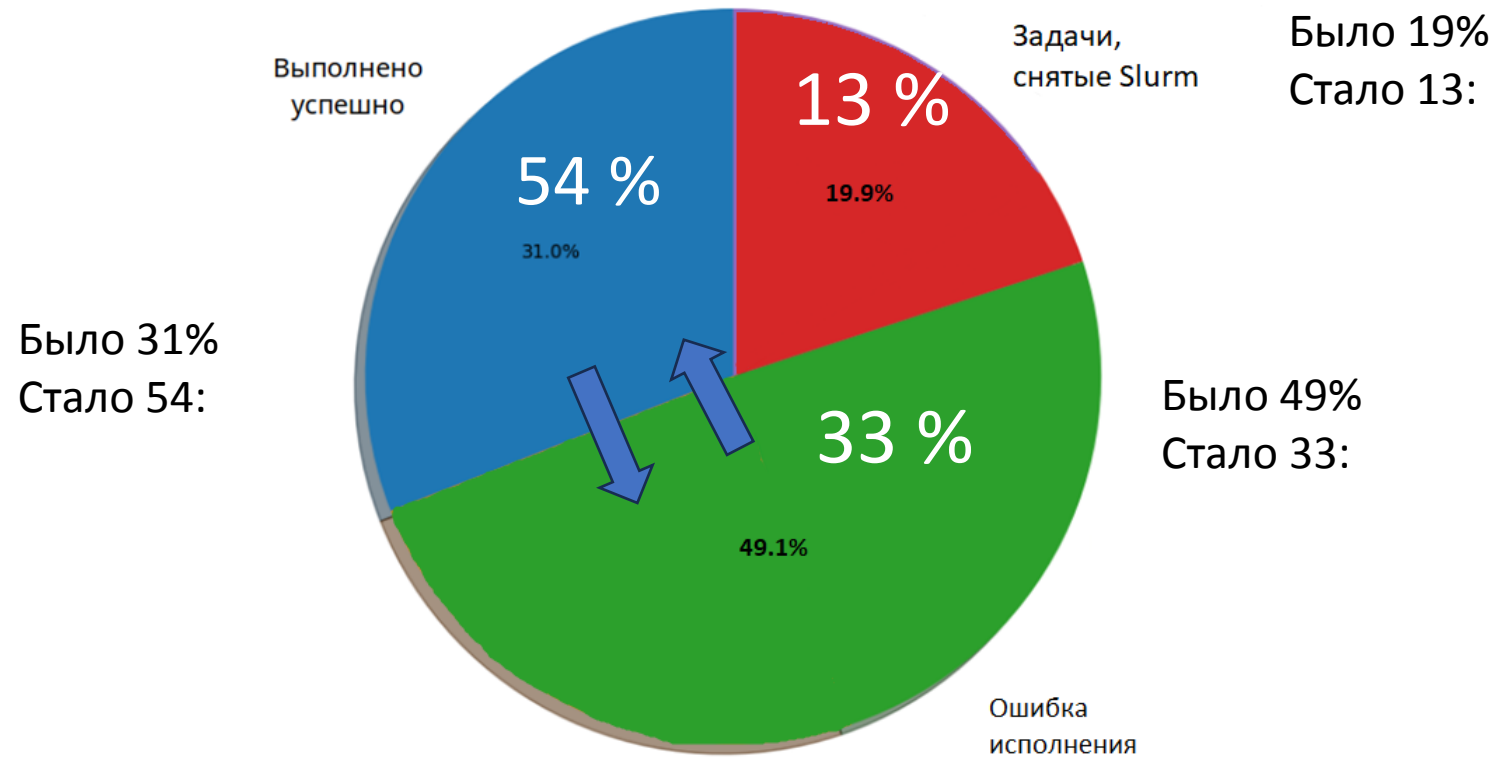
Структура «умного» суперкомпьютера, работающего в режиме ЦКП

Входные данные

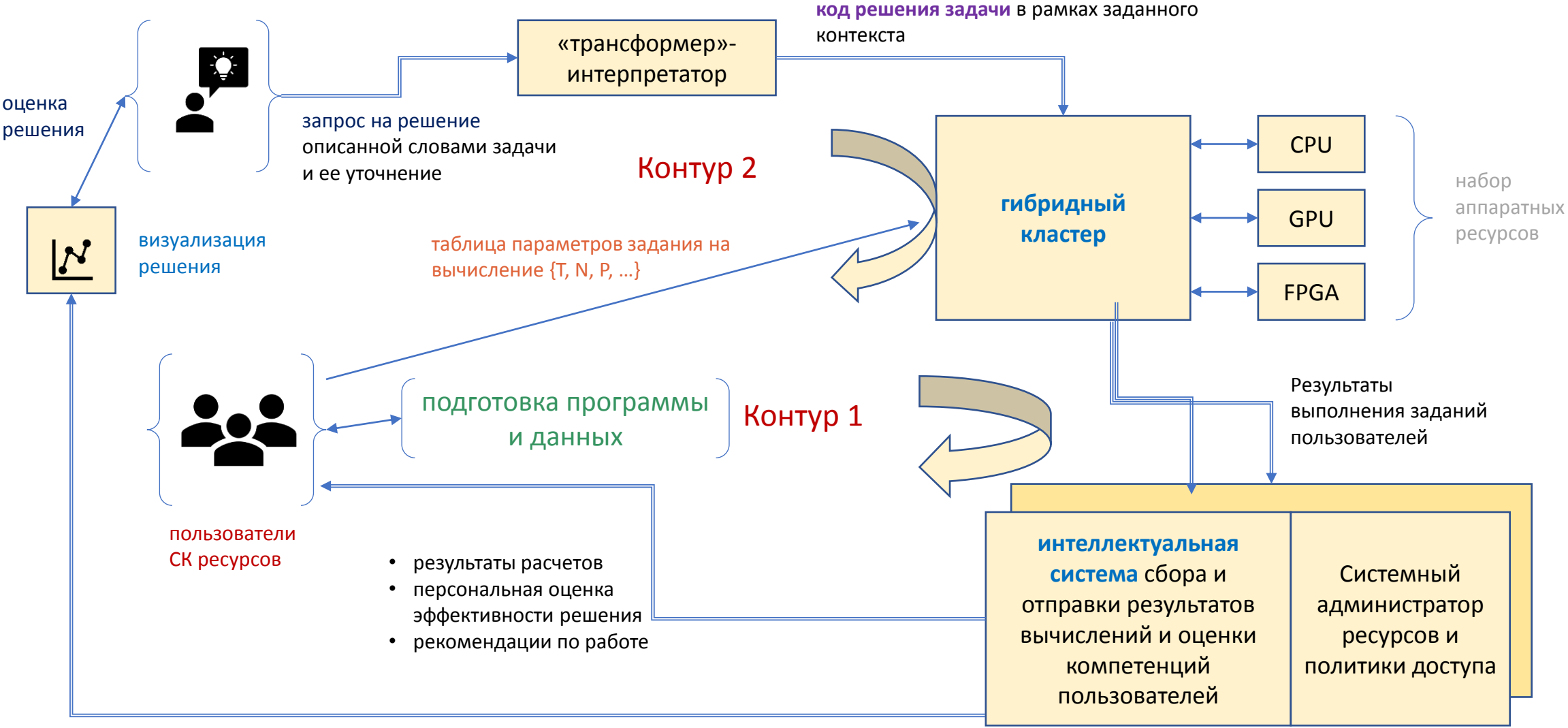
- [
- параметры задания,
 - модель пользователя
 - характеристики состояния выч. среды
-]



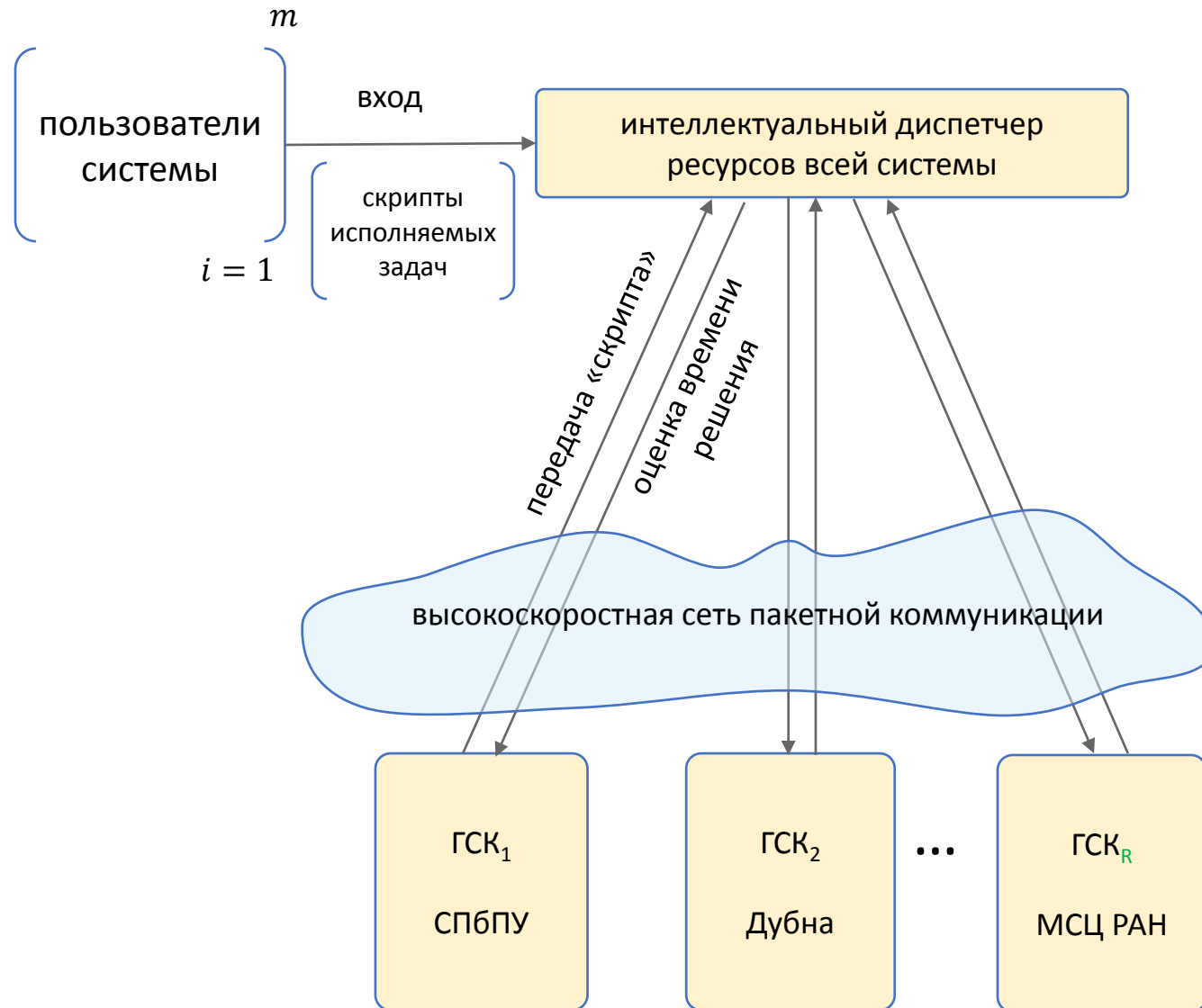
Результаты использования «обучаемых оценок» времени решения прикладной задачи



Развитие идеи: 2-х контурная интеллектуальная платформа



Архитектура сети из гибридных «интеллектуальных» СК



Гипотезы:

1. Каждый узел системы обрабатывает d_i классов прикладных задач
2. Каждый класс прикладных задач характеризуется апостериорной функцией «выживания» $S_i(y| \dots)$
3. На основе $S_i(y| \dots)$ строится модель прогноза времени решения новой задачи
4. Оценки времени решения от каждого узла ГСК системы передаются в диспетчер системы, который с использованием критерия «реальной производительности» и учетом «накладных расходов» на передачу данных выбирает узел системы, где данная задача будет исполняться

Оценка времени решения выполняется в каждом узле системы с использованием методов предсказания (регрессии) на основе использования функции «выживания», построенной для каждого конкретного класса прикладных задач, исполняемых в 1, 2, ..., R узлах системы

Выводы

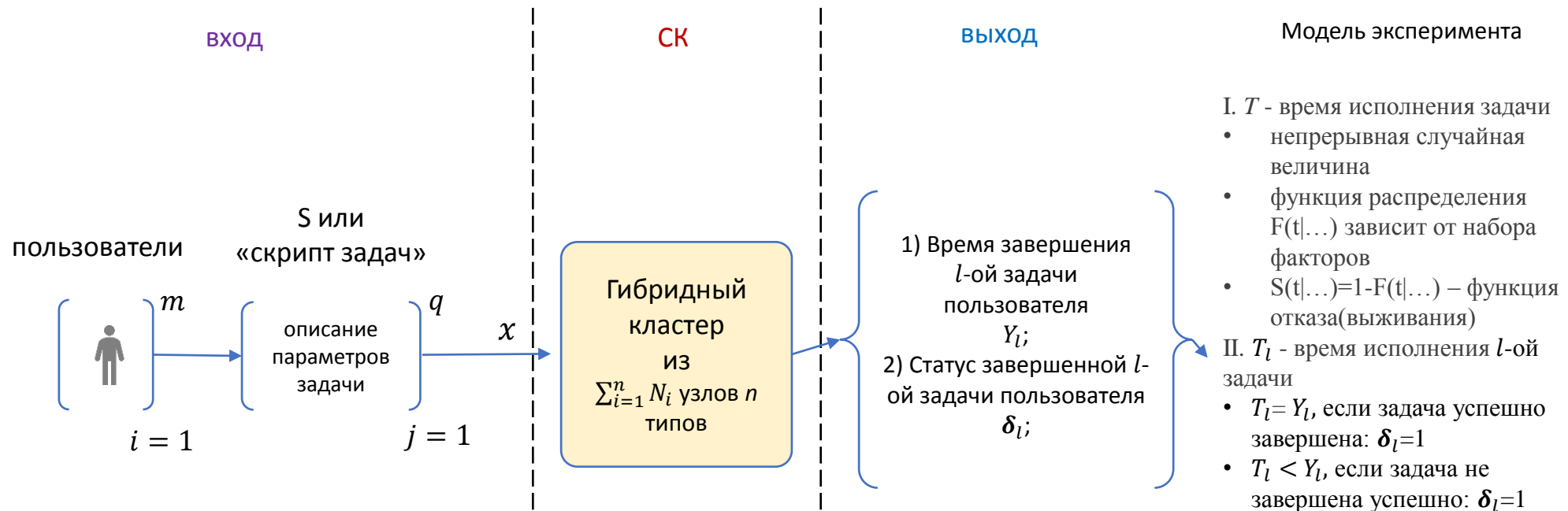
- При использовании моделей МО :
 - «прикладная» производительность гибридного СК ЦКП повышается на более **чем 10%**
- Новые абстракции, такие как как «выживаемость прикладных заданий» и «полезность диспетчера», **могут стать основой**
 - **новой «теории интеллектуальных СК».**
 - успешно использоваться для решения задачи «объяснительного интеллекта»:
- Перспективы
 - внедрить механизм самовнимания для **непараметрических «ВХОД-ВЫХОД»**
 - оптимизировать **параметрическую модель ЦКП** путём решения совместной задачи выживаемости и регрессии,

Спасибо за внимание!



Приложение 1

Модель «вход-выход» гибридного СК, используемая в задаче машинного обучения диспетчера



Выводы:

финитных методов для моделирования интеллекта с помощью методов машинного обучения недостаточно, а требуется трансфинитная индукция

Задачи:

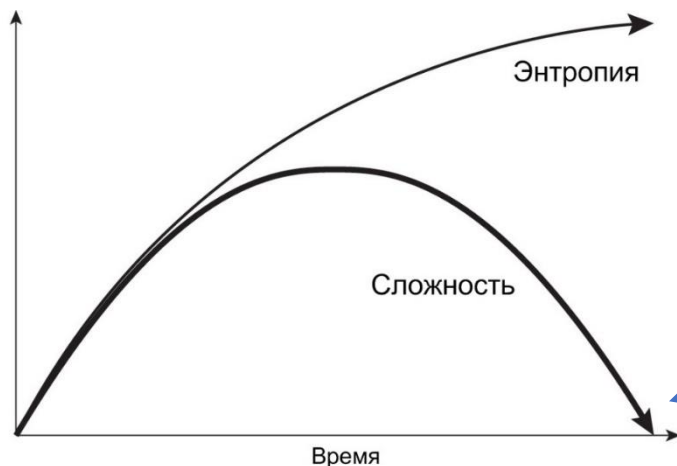
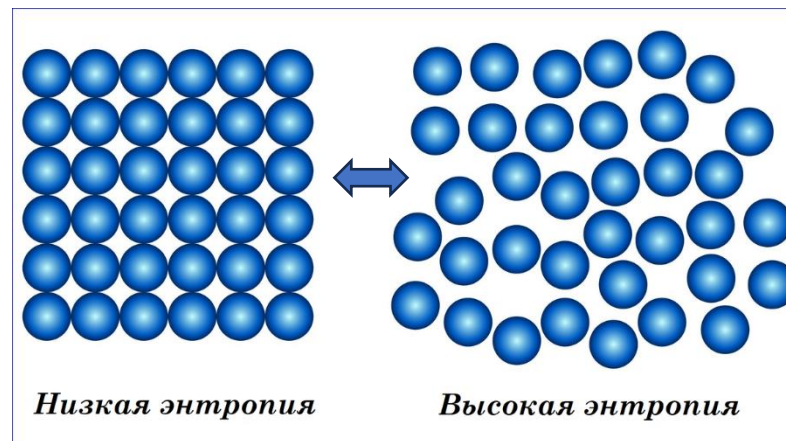
- 1) Оценивание распределения T для различных наборов факторов
- 2) Установление зависимости распределения T от набора факторов
- 3) Сравнение распределений T при различных наборах факторов
- 4) Изучение различий распределений T при различных наборах факторов

Приложение 2

Энтропия – поиск общего в термодинамике и теории передачи информации

Энтропия в термодинамике — мера рассеяния энергии.

$$\Delta S = \frac{\Delta Q}{T}$$



Количество данных $K \ll N$ — число возможных комбинаций из этих данных

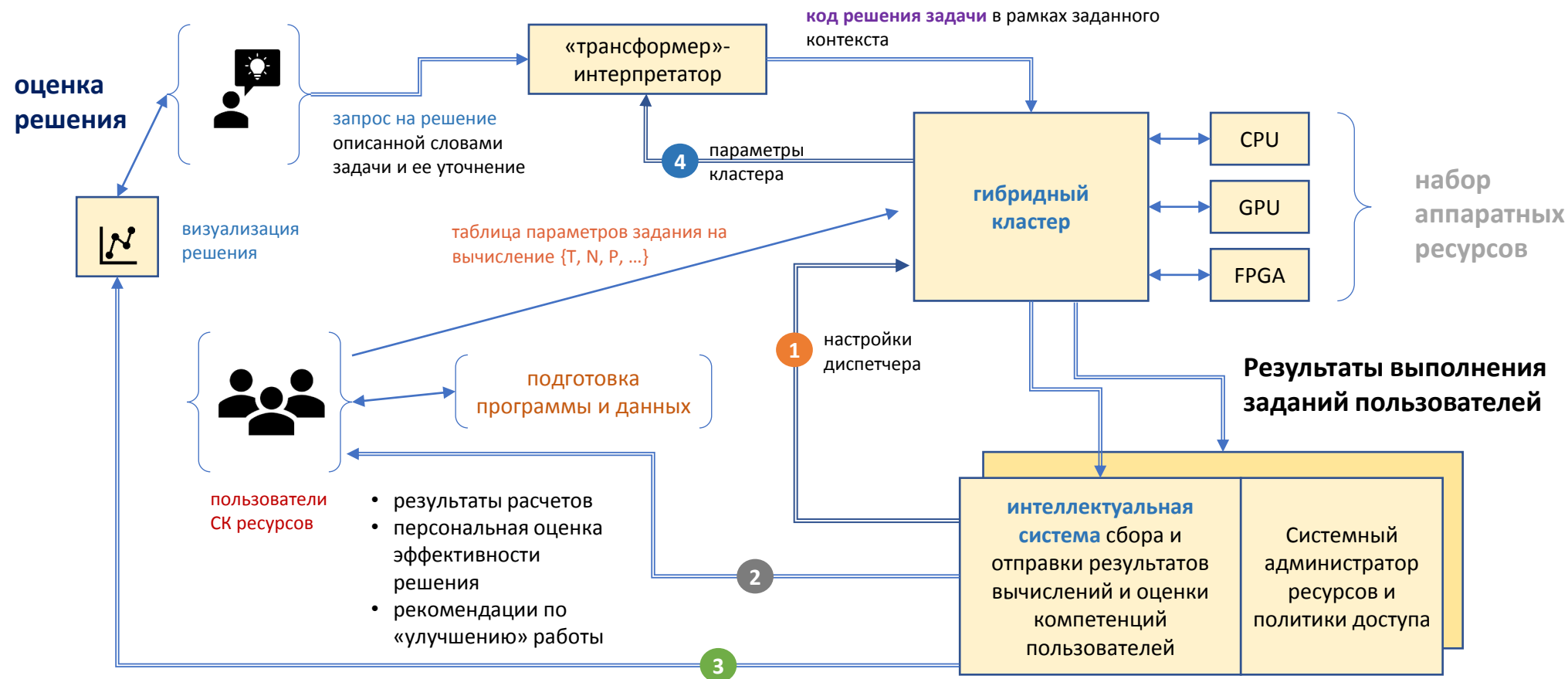
Энтропия в теории К. Шеннона мера «рассеивания» ... потери формы объекта

Вопрос: «сложность» это «форма»? Как сложность описания объекта растет по мере роста энтропии?!

изменение внутренней энергии системы всегда отрицательно (**энергия рассеивается**), а изменение энтропии всегда положительно (**неопределенность растет**).

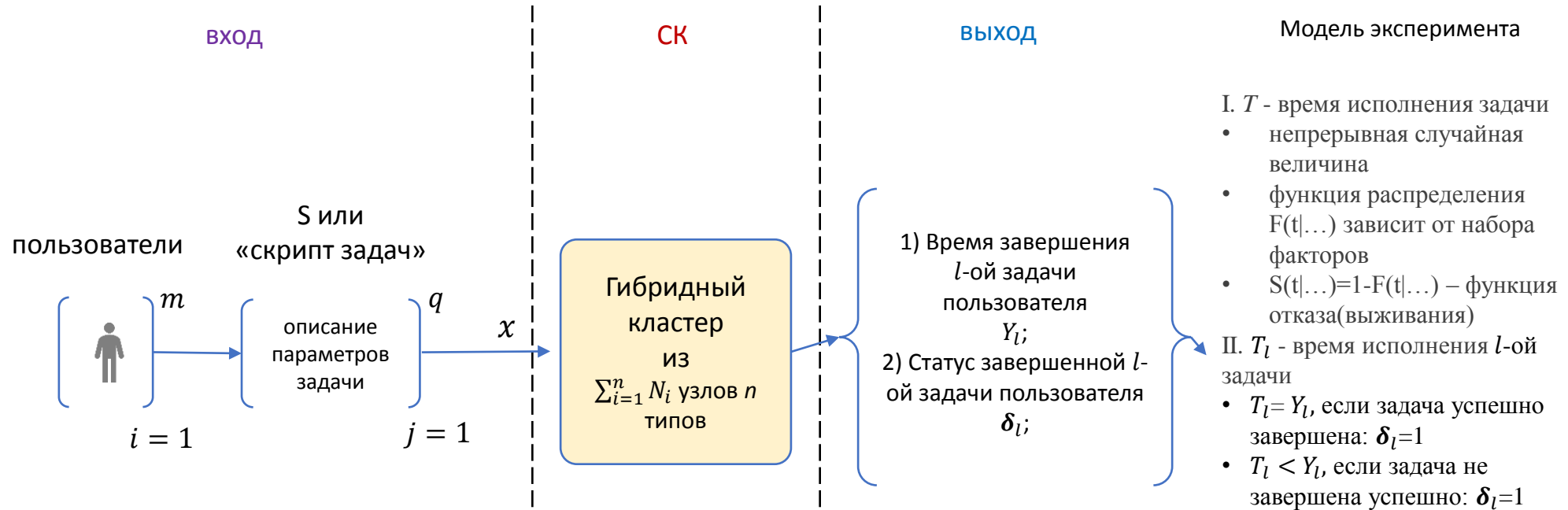
$$E = \sum -p_i \log_2(p_i) \Rightarrow \sum p_i \log_2 \frac{1}{p_i} \quad \text{ЛЭТЬ}$$

Что планируется сделать : Рекурсивная когнитивно-деятельностная модель процессов вычислений в ЦКП «Политехнический»



Приложение 2.2

Модель «вход-выход» гибридного СК, используемая в задаче машинного обучения диспетчера



Выводы:

финитных методов для моделирования интеллекта с помощью методов машинного обучения недостаточно, а требуется трансфинитная индукция

Задачи:

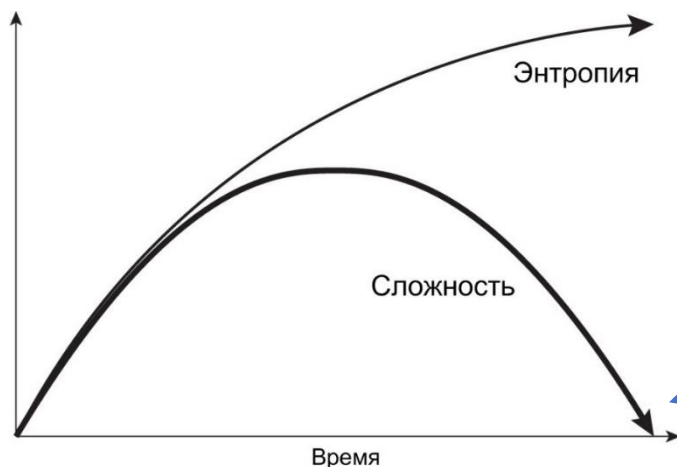
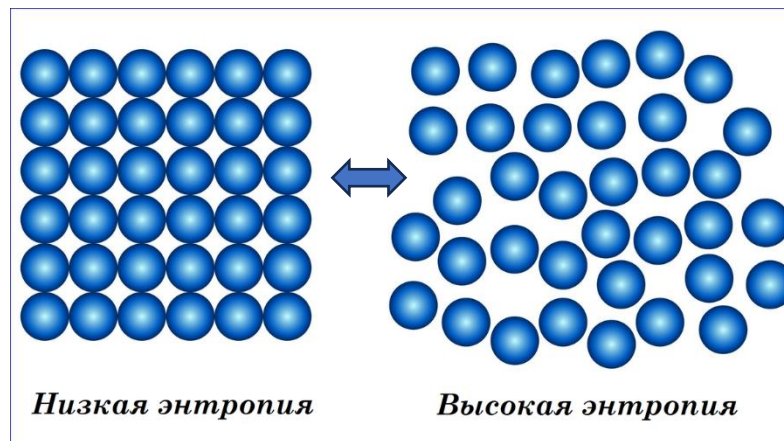
- 1) Оценивание распределения T для различных наборов факторов
- 2) Установление зависимости распределения T от набора факторов
- 3) Сравнение распределений T при различных наборах факторов
- 4) Изучение различий распределений T при различных наборах факторов

Приложение 2.3

Энтропия – поиск общего в термодинамике и теории передачи информации

Энтропия в термодинамике — мера рассеяния энергии.

$$\Delta S = \frac{\Delta Q}{T}$$



Вопрос: «сложность» это «форма»? Как сложность описания объекта растет по мере роста энтропии ?!

изменение внутренней энергии системы всегда отрицательно (**энергия рассеивается**), а изменение энтропии всегда положительно (**неопределенность растет**).

Количество данных $K \ll N$ — число возможных комбинаций из этих данных

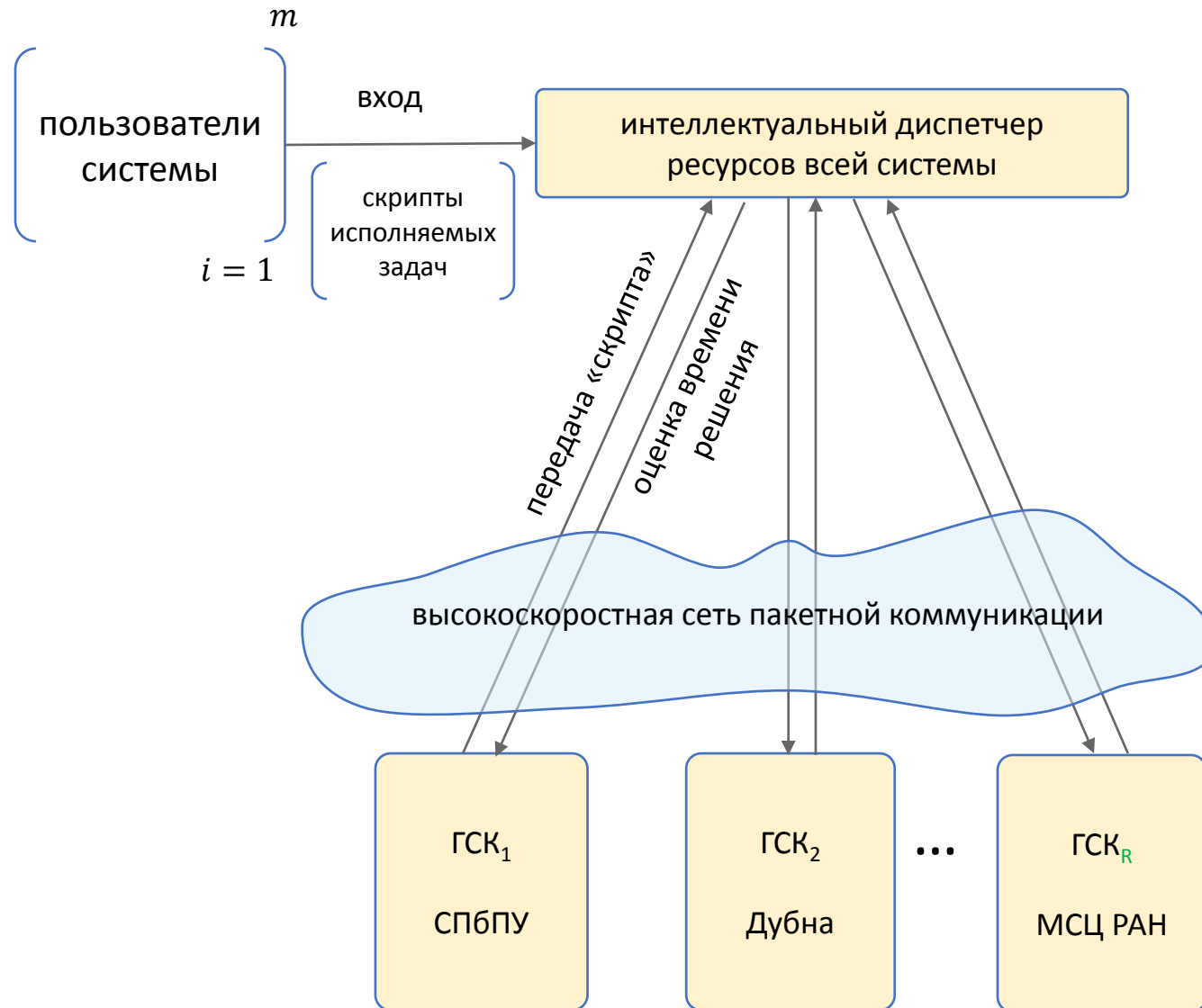
Энтропия в теории К. Шеннона мера «рассеивания» ... потери формы объекта

$$E = \sum -p_i \log_2(p_i) \Rightarrow \sum p_i \log_2 \frac{1}{p_i}$$

ЛЭТЬ
ΣМ

Приложение 3

Архитектура системы (сети из r гибридных СК)

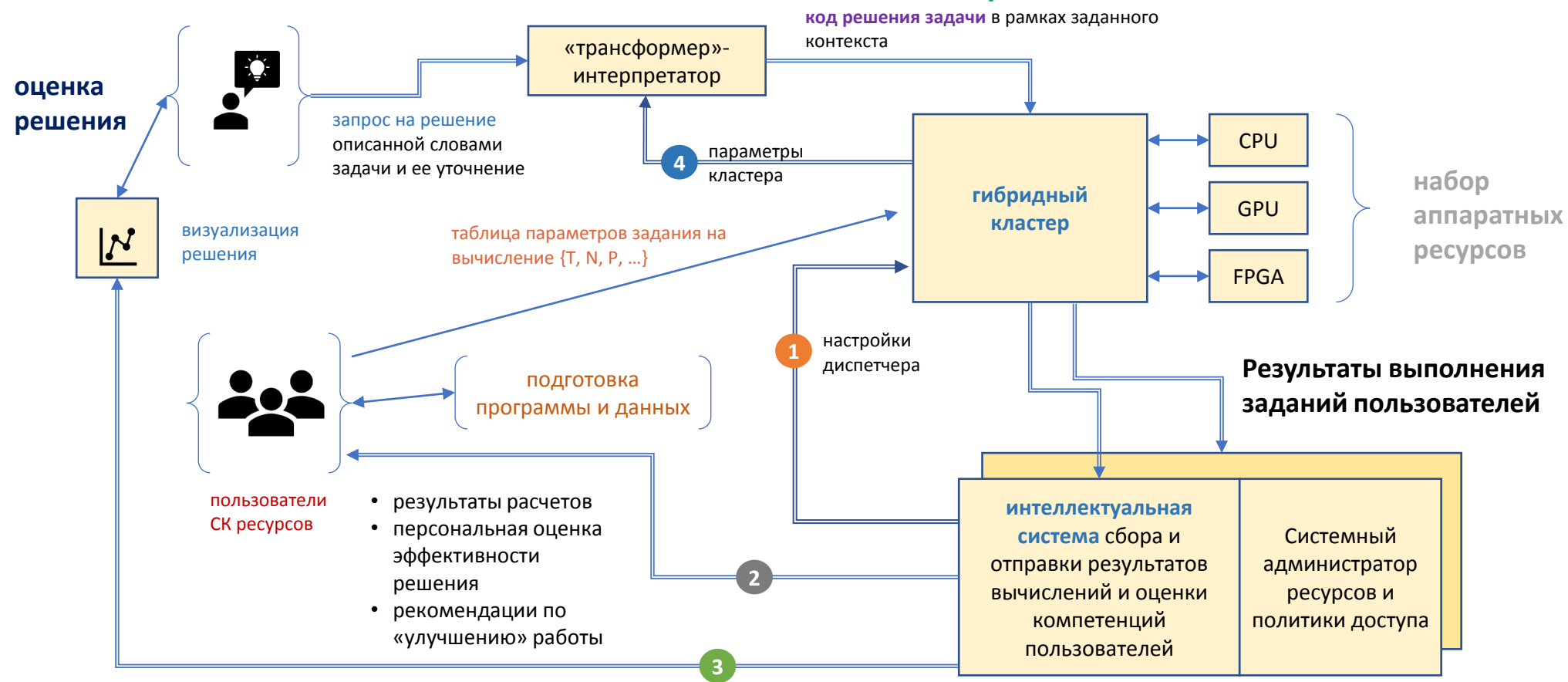


Гипотезы:

1. Каждый узел системы обрабатывает d_i классов прикладных задач
2. Каждый класс прикладных задач характеризуется апостериорной функцией «выживания» $S_i(y| \dots)$
3. На основе $S_i(y| \dots)$ строится модель прогноза времени решения новой задачи
4. Оценки времени решения от каждого узла ГСК системы передаются в диспетчер системы, который с использованием критерия «реальной производительности» и учетом «накладных расходов» на передачу данных выбирает узел системы, где данная задача будет выполняться

Оценка времени решения выполняется в каждом узле системы с использованием методов предсказания (регрессии) на основе использования функции «выживания», построенной для каждого конкретного класса прикладных задач, исполняемых в 1, 2, ..., R узлах системы

Приложение 4 самообучающаяся когнитивно-деятельностная модель процессов вычислений в ЦКП «Политехнический»



Приложение 5

Информационное воздействие

Далее будем называть информационным такое воздействие на систему, которое :

- **изменяет количество равновероятных состояний** в системе,

но при этом

- не изменяет количество частиц системы;
- не изменяет общую энергию системы;
- не изменяет температуру системы.

Обозначим первоначальную энтропию системы как S_0 , а энтропию после информационного воздействия S_1 .

Таким образом, $\Delta I = S_0 - S_1$

Формализация «информационного воздействия»

В случае отсутствия информационных воздействий на систему и ее изоляции от обмена энергией справедлива формула:

$$E = \Delta G + T\Delta S$$

Эта формула выражает известный из термодинамики закон не убывания энтропии. Энтропийное слагаемое в этой формуле при отсутствии внешних воздействий всегда неотрицательно, а слагаемое свободной энергии – всегда не положительно. Считая зависимость информации от свободной энергии аналогичной той, что представлена в формуле Больцмана, получим:

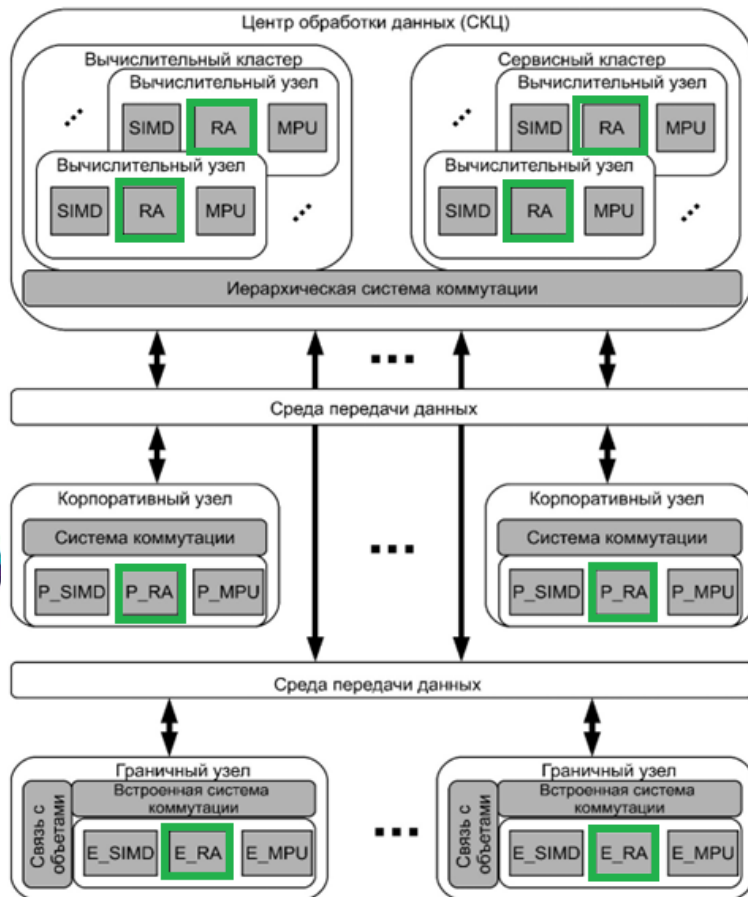
$$E = T(\Delta I + \Delta S)$$

Итого. Из второго закона термодинамики можно вывести закон не возрастания информации в объекте (системе) без внешнего информационного воздействия.

Приложение 7 It from bit: из информации обратно в физику

- Воспринимаемые наблюдателем образы организуются в многообразия, которые факторизуют реальность с точностью до а) алгебраических систем и б) отношений порядка, т.е. до «чисел и «понятий». Физика сводится к топологии !?
- При этом голографический принцип физических наук утверждает, что Вселенная де-факто двумерна:
 - Как голограмма на стеклянной пленке – **интерференционном поле**, воспроизводит трехмерный объект, так и
 - воспринимаемая наблюдателем трехмерная Вселенная может быть эквивалентна **многообразию** квантовых полей, «нарисованных» на удаленной двумерной поверхности.
 - **Задача физики – понять и объяснить математикам, что же там на самом деле нарисовано (закодировано)?**

Приложение 8_1 не пора ли «оглянуться назад»



- Современные супер-компьютеры на физическом уровне «умеют» делать всего **одну БАЗОВУЮ операцию** – сдвиг двоичного регистра процессора, но ... делают эту операцию **очень быстро!**
- Таким экстенсивным способом решить проблему «сложности» не удастся ...из за соотношения Ландауэра
- Может быть стоит увеличить «разнообразие» БАЗОВЫХ МАШИННЫХ КОМАНД, обучая этому один из 3-х уровней «процессора»
(технология была известна как микро-программирование_)

Понятие «малых» и «больших» данных

Данные «малые» :

структура простая, которая может быть закодирована» конечной» математической формулой

«Смысл» данных может кодироваться

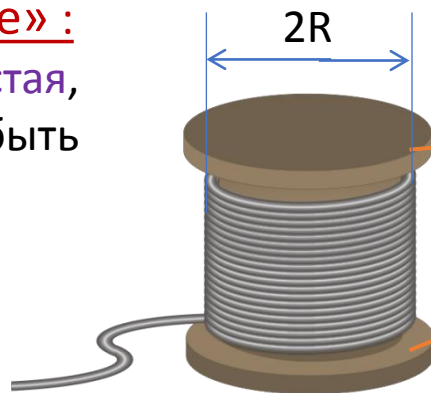
ЧИСЛОМ

Данные «большие»:

структура сложная и не имеет конечного формального (математического) описания

«Смысл» таких данных может кодироваться

словами (красиво, безобразно...)



Когнитивные функции -

понимание смысла символов и цифровых

мер

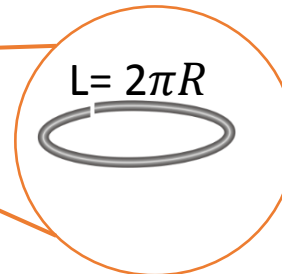


Оценка меры сложности



Аналитические функции -

Вычисление по конечным формулам



«распутывание» как процесс воплощения «конечного» опыта

??? алгоритма нет

но есть «опыт» решения аналогичных задач

«вычисления» как процесс «каузального понимания» обрабатываемых данных на основе количественной и семантической мер

объединение методов индуктивного обучения и дедуктивного научения

«...я оглянулся посмотреть не оглянулась ли она, чтобы посмотреть не оглянулся ли я ...» Песня «Видение»

М. Леонидов (1996)

понять перспективы «цифровых» технологий и возможности машинного обучения в решения как **прямых задач** цифрового моделирования , так и **обратных задач** структурно-параметрической идентификации используемых моделей. Идеи «научения» наследуют концепции, восходящие к истокам свременных научных знаний, а именно :

- **Пифагору** с концепцией иллюзорности материального, но реальности **чисел**
- **Платону** с концепцией реальности **идей**, способных к обретению **материальной формы**

Вывод :

- обратимый процесс воплощения результатов вычислений (цифровых «теней» реальности) в свойства моделируемых объектов можно рассматривать как механизм **научения**, описываемый в терминах способности системы к «самосозданию»
- результаты «научения» необходимо «регуляризовать», придав им ясный смысл

Приложение 10 от «машины Тьюринга» к «машине Геделя»: от концепции «ИИ» к концепции «экзо-интеллекта»:

Как эволюция живых организмов, которую связывают с естественным отбором в процессе изменения природных условий, так и процесс развития программно-управляемых вычислительных систем на пути обретения ими функций «самообучения». можно связать необходимым условием их **«выживания»** в результате радикального увеличения **масштабов цифровой эко-системы их существования**.

Суть фенотипического подхода: в современных условиях «закон Мура», феноменологический «двигатель «машины Тьюринга». больше не работает ... однако современная «цифровая» платформа способна воплотить в себе процессы научения. **СТАНОВЯСЬ «машинной Геделя»**, способна доказать необходимость своего «дообучения» и изменения.

Концепция **экзо-интеллекта** базируется на том, что т. н. процесс **реификации** связывается с принципом **множественности моделей** (В. В. Налимов), **сопоставляемых результатам решения «обратных задач» цифрового моделирования**.

Экзо-интеллектуальные системы способны из множества возможных решений **выбирать 1) те которые являются наиболее вероятными «с учетом текущего контекста» (технологии глубокого обучения), так и 2) те, которые:**

- 1) «повышают» **независимость** самих процессов обучения от свойств среды обитания,
- 2) формируют границу того, что принадлежит к обучаемой системе, а что нет.

Что «находится» между существованием и бытием – комбинаторика мыслимых сущностей

Парменид

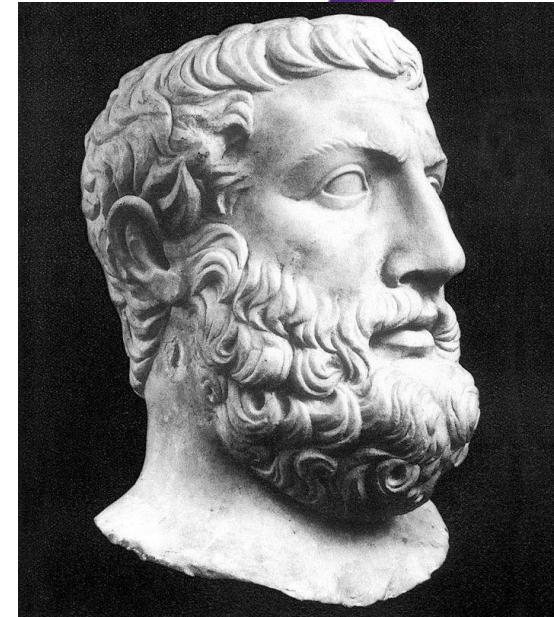
Цитаты:

«[мыслить](#) и быть — одно и то же»

«мыслимое есть Бытие, ибо нельзя мыслить ни о чём»

бытие и мышление неразделимы. Если мы можем думать о чем-то, это означает, что это «что-то» существует. В этом смысле, бытие есть существование мыслимых сущностей.

«Книга вымышленных существ», написанная [Хорхе Луисом Борхесом](#) в [1954](#) году может считаться образцом [антологий](#) реально несуществующего, но мыслимого, как экспериментов философов



Парменид Элейский.

515 год -450 до н.э.,
Греция

Цитаты: Есть — бытие,
а ничто — не есть

Приложение 11

Постановка задачи машинного обучения

Имеется множество векторов $\{\vec{x}\}$ из многомерного числового пространства, координаты которого будем называть признаками. Векторы независимо получены выборкой из некоторого (неизвестного) вероятностного распределения, обладающего плотностью $p(\vec{x})$.

Для части векторов известно значение (метка) класса из оговоренного конечного набора возможных классов. Требуется установить, к каким классам относятся оставшиеся векторы, а также любые другие векторы, порожденные процессом с такой же функцией распределения.

- Из векторов-примеров с известными метками можно образовать обучающую выборку, для которой с использованием алгоритмов обучения машин (например, искусственных нейронных сетей или деревьев правил), строится функция-классификатор $h(\vec{x})$. Выходом классификатора является метка класса для каждого входного вектора \vec{x} .
- С практической точки зрения, обучаемые классификаторы различаются по таким показателям, как точность, производительность, устойчивость к погрешностям в исходных векторах, а также по затратам на их обучение и тестирование.

Приложение 12

Дилемма - поиска ответа на вопрос

что проще — построить и обучить один глобальный классификатор, либо сегментировать проблему на серию задач меньшего масштаба и обучить, а затем объединить, большое количество моделей, желательно, меньшей сложности. Последнее может оказаться предпочтительным, если при объединении моделей достигается требуемая точности прогнозирования.

- Вопрос: Откуда может появиться **резерв точности множества моделей**, каждая из которых настроена на свою подзадачу, и, вообще говоря, не обязана быть высокоточной на всем исходном множестве данных?
- Ответ: это суть центральной предельной теоремы теории вероятности (ЦПТ).
- В рассматриваемом контексте результат состоит в том, что последовательности частичных средних, вычисленных по наборам из n независимых случайных величин, даже имеющих большую дисперсию σ , стремятся к нормальному распределению с дисперсией, в корень из n раз меньшей.

Приложение 13 современная концепция машинного обучения - «ИСТИНА В ПОВТОРЕНИИ ФАКТОВ»

- В современной науке принято называть «истинной» лишь ту информацию, которая подкреплена **формальными доказательствами**. Если эксперимент не подтверждается теорией ... то эксперимент проведен неверно. Однако, согласно теоремам Геделя, не всякая истина доказуема, а сами формальные доказательства трудны для понимания и требуют больших затрат времени.
- **имеется и другой когнитивно-эмпирический критерий «истины» или «правдивости»**, основанный на повторяемости воспринимаемых мозгом эффектов и доверию к повторяющимся утверждениям, которые обрабатываются мозгом быстрее и... вызывают у него большее доверие.
- Следование критерию «повторяемости фактов» логично с точки зрения «здравого смысла» или теории «условных рефлексов», но может приводить к глубоким когнитивным искажениям, известным как эффект «иллюзорности правды». Этот эффект под названием «дипфейк» в настоящее время легко реализуют современные системы ИИ
- Итак, концепция «машинного обучения» не отличается «идейно глубиной» и оригинальностью, но хорошо интерпретируема как следствие из правила: «повторение – мать учения». Машинное обучение на базе этой концепции приводит к эффекту «понимания» на основе корреляций, но лишь того, что многократно повторяется размеченной обучающей выборке (deep learning) или явно «подкрепляются» в потоке поступающих в систему данных (reinforcement learning).

What we have before us are some breathtaking opportunities disguised as insoluble problems

(перед нами открываются захватывающие дух возможности, замаскированные под неразрешимые проблемы)

John Gardner, 1965

- Итак, чтобы создать универсальные системы ИИ или так называемые AGI, необходимо:
 - **научить** системы ИИ справляться с формально **неразрешимыми проблемами** решения «обратных математических задач», формально имеющих множество «правильных» решений
 - **развивать** методы машинного обучения в направлении от использования статистических корреляций к использованию механизмов обобщения на основе объяснения фактов с учетом контекста принятия решений/
 - **Итого:** использовать методы обучения на основе формальных концепций или оперативно валидированных дедуктивных правил