

**Реинжиниринг мозга,
информация и квантовая информация.
Почему не работает
по-до теорема Пенроуза об ИИ**

А.Д. Панов
НИИЯФ МГУ

К чему приводит длительная эволюция «цивилизации»
после технологического взрыва:

- Живой разум?
- Искусственный интеллект?
- Симбиоз?

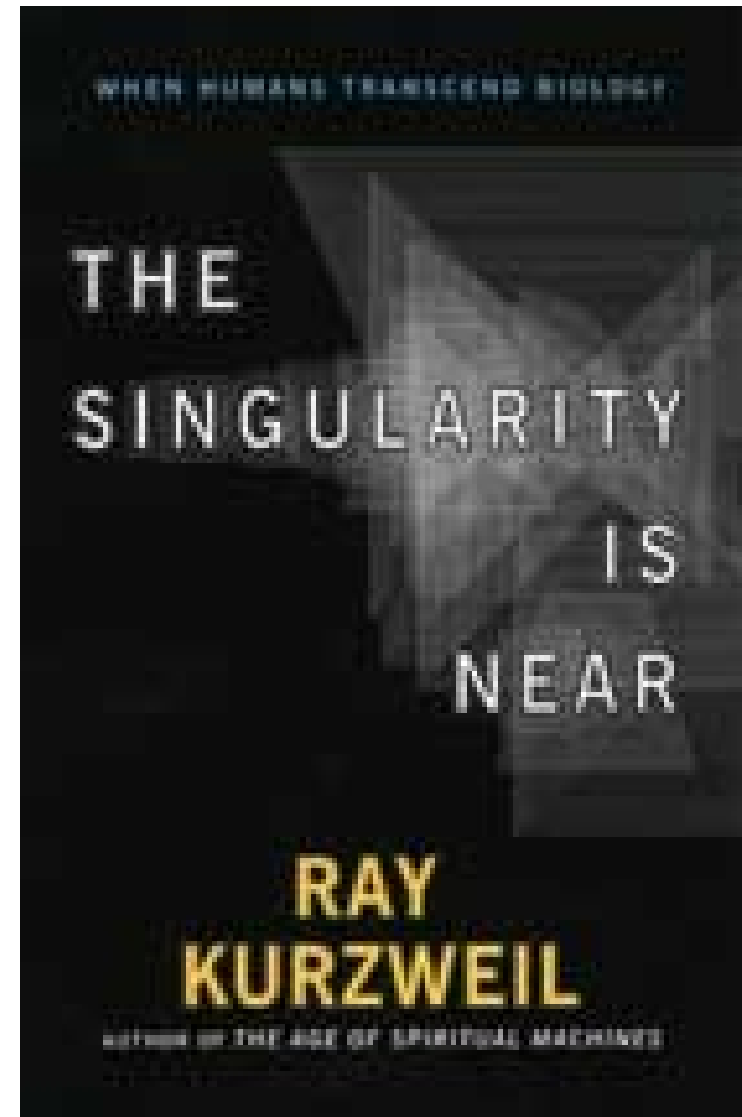
Технологическая сингулярность

Irving John Good, 1965 -
intelligence explosion

Vernor Vinge, 1988 -
technological singularity,
2005-2030

Hans Moravec, 1988 -
technological singularity,
2030-2040

Ray Kurzweil, 1990th -
technological singularity,
2045



2005



Виталий Львович Дунин-Барковский

Реинжиниринг мозга

Обратная инженерия мозга

Создание нейроморфного
искусственного интеллекта

Лаборатория обратной инженерии мозга имени Дэвида Марра (2012)

Предполагается, что к концу 2015 года в результате работ по проекту будет создано полное (детальное) описание механизмов работы мозга человека. Это описание можно будет использовать для создания (к 2018-2020 гг.) полномасштабного функционального макета человеческого мозга на базе технологических информационных элементов и устройств.

<http://rebrain.2045.ru/>

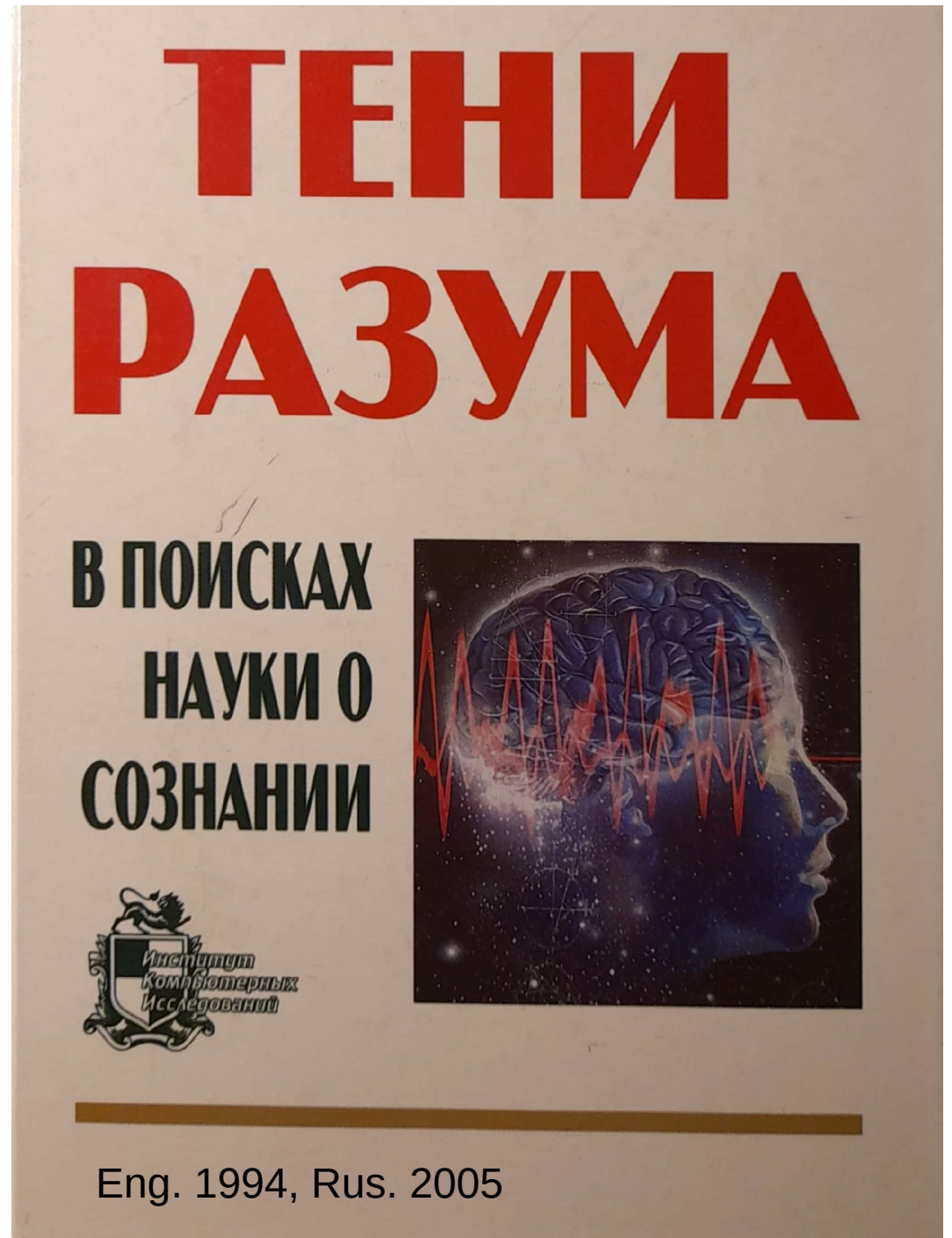
<http://aihandbook.intsys.org.ru/index.php/activity/ai-projects/projects-blog/215-proj-31>



Роджер Пенроуз

Но-го теорема об ИИ

J. Lukas.
Minds, Machines and Gödel.
Philosophy, 1961, V. 36, P.112-127





П Р И Л О Ж Е Н И Е К Ж У Р Н А Л У
«И Н Ф О Р М А Ц И О Н Н Ы Е Т Е Х Н О Л О Г И И» № 5/2014

А. Д. Панов

**ТЕХНОЛОГИЧЕСКАЯ
СИНГУЛЯРНОСТЬ,
ТЕОРЕМА ПЕНРОУЗА
ОБ ИСКУССТВЕННОМ
ИНТЕЛЛЕКТЕ И КВАНТОВАЯ
ПРИРОДА СОЗНАНИЯ**



**Почему не работает по-до теорема
Пенроуза об ИИ**

Сильный искусственный интеллект (AGI) (в контексте технологической сингулярности)

- Wikipedia:
Artificial general intelligence (AGI) is the hypothetical ability of an intelligent agent to understand or learn any intellectual task that a human being can.
- AGI не уступает человеческому интеллекту во всех его выражениях и в максимальных проявлениях.
- Иначе говоря, AGI не уступает интеллекту **всего человечества**, включая наиболее сильных его представителей и во все времена.

1-я теорема Гёделя о неполноте

Для любой аксиоматической системы, которая

- 1) Непротиворечива
- 2) Содержит в себе формальную арифметику (достаточно сильна)

можно сформулировать осмысленное утверждение, которое нельзя ни доказать, ни опровергнуть средствами этой системы.

Доказательство конструктивно - это утверждение строится в явном виде (и является истинным по построению)

Используются два основных метода:

- Гёделева нумерация и
- Диагональный метод Кантора

Теорема Гёделя-Тьюринга

Для любого конечного автомата, который

- 1) Реализует обоснованные процедуры
- 2) Достаточно силен, чтобы реализовывать алгоритмы, анализирующие другие алгоритмы на предмет их остановки

можно сформулировать осмысленное утверждение, истинность которого не может быть вычислена этим автоматом.

Доказательство конструктивно - это утверждение строится в явном виде если известна структура автомата (и является истинным по построению)

Используются два основных метода:

- Гёделева нумерация и
- Диагональный метод Кантора

Но-го теорема Пенроуза об ИИ (полное доказательство)

1. Предположим, что некоторый компьютер, имеющий архитектуру конечного автомата и реализующий обоснованные процедуры, обладает всеми интеллектуальными способностями всего человечества (представляет собой AGI в сильном смысле).
 2. Тогда, любой математик, **используя свои математические способности**, на основе теоремы Гёделя-Тьюринга может построить истинное утверждение, истинность которого не может быть проверена **этим** компьютером, но которая ясна для математика (по построению). Построение всегда возможно, так как доказательство теоремы Гёделя-Тьюринга имеет конструктивный характер.
 3. Следовательно, предполагая, что компьютер обладает всеми способностями людей, мы немедленно указываем способность человека, которой этот компьютер не обладает.
 4. Это есть противоречие, и оно доказывает, что такой компьютер (AGI в сильном смысле) не может существовать.
- **Сильный ИИ невозможен для компьютеров на основе архитектуры конечного автомата.**
 - **Следствие: Мозг человека реализует невычислимую активность.**

Возражения (Пенроуз рассмотрел 20 возражений)

- ◆ Какова природа теоремы Пенроуза (все человечество не является математическим понятием...)? Это вообще не теорема?
 - Правильная теорема математической логики.
- ◆ Теорема Пенроуза адресует не реальные вычислительные машины, но идеальные конечные автоматы - машины Тьюринга.
 - Методическая ошибка, природа познаваема только на основе моделей
- ◆ Теорема Пенроуза основана на методе доказательства от противного, который не признается конструктивной математикой.
 - Откажитесь от математического анализа...
- ◆ И так далее...

**Возражения против теоремы не проходят
потому что теорема ПРАВИЛЬНАЯ**

2-я теорема Гёделя о неполноте

Если аксиоматическая система, которая содержит формальную арифметику, непротиворечива, то ее непротиворечивость не может быть доказана средствами этой системы.

И не может быть доказана вообще.

Уверенность в непротиворечивости математики основана только на опыте

2-я теорема Гёделя-Тьюринга

Если конечный автомат «достаточно силен» и при этом он реализует обоснованные процедуры, то его обоснованность не может быть установлена средствами этого автомата.

И не может быть установлена вообще.

Недостаточно сильные автоматы не представляют интереса для AGI!

Для любого конечного автомата, который может претендовать на функциональность AGI, обоснованность его процедур не может быть установлена, следовательно к нему по-го теорема Пенроуза об ИИ не приложима.

Замечание:

Трудно сомневаться в том, что процедуры человеческого мозга «являются необоснованными»: человеческое мышление противоречиво.

Наука не дает абсолютных доказательств чего бы то ни было...

*Перед Вами — единственная в своем роде книга.
В ней содержится полная хроника споров менталистов
и механицистов вокруг вопроса: кого же вторая теорема
Гёделя считает умнее — человека или машину?
И на чьей стороне оказался сам Гёдель?*

В. В. Целищев (2021)

АЛГОРИТМИЗАЦИЯ МЫШЛЕНИЯ

Гёделевский аргумент

- Математика и мышление
- Человек, машина и гёделево предложение
- Трансфинитный аргумент и принципы рефлексии
- Компьютер, истина и доказуемость
- Алгоритмическая версия ментализма
- Непротиворечивость и полнота
- Истинность гёделева предложения
- Человеческая математика
- «Сознание» машины



НАУКИ об
ИСКУССТВЕННОМ

№37



Уже у Дж. Лукаса (1961) был аргумент от 2-й теоремы Гёделя о неполноте, но вывода о том, что 1-я теорема неприменима к реально интересным компьютерам нет.

Его нет и у Целищева.

Вместо этого есть (запутанный) анализ по поводу того, можно ли считать наше мышление непротиворечивым

Реинжиниринг мозга, информация и квантовая информация

Два основных пути в создании ПО для ИИ:

- 1. Синтетическое направление**
- 2. Обратная инженерия мозга**

Основная проблема обратной инженерии мозга:

- Задача не понята, так как используются чрезвычайно упрощенные представления о работе мозга
- Предпринимаемые в этом направлении попытки неадекватны в смысле создания AGI

Рэй Курцвейл: Когда технологическая сингулярность?

Как только мощность коммерческих компьютеров, выраженная в операциях в секунду, превзойдет совокупную вычислительную мощность мозга всего человечества, AGI будет создан, и технологическая сингулярность может стать реальностью.

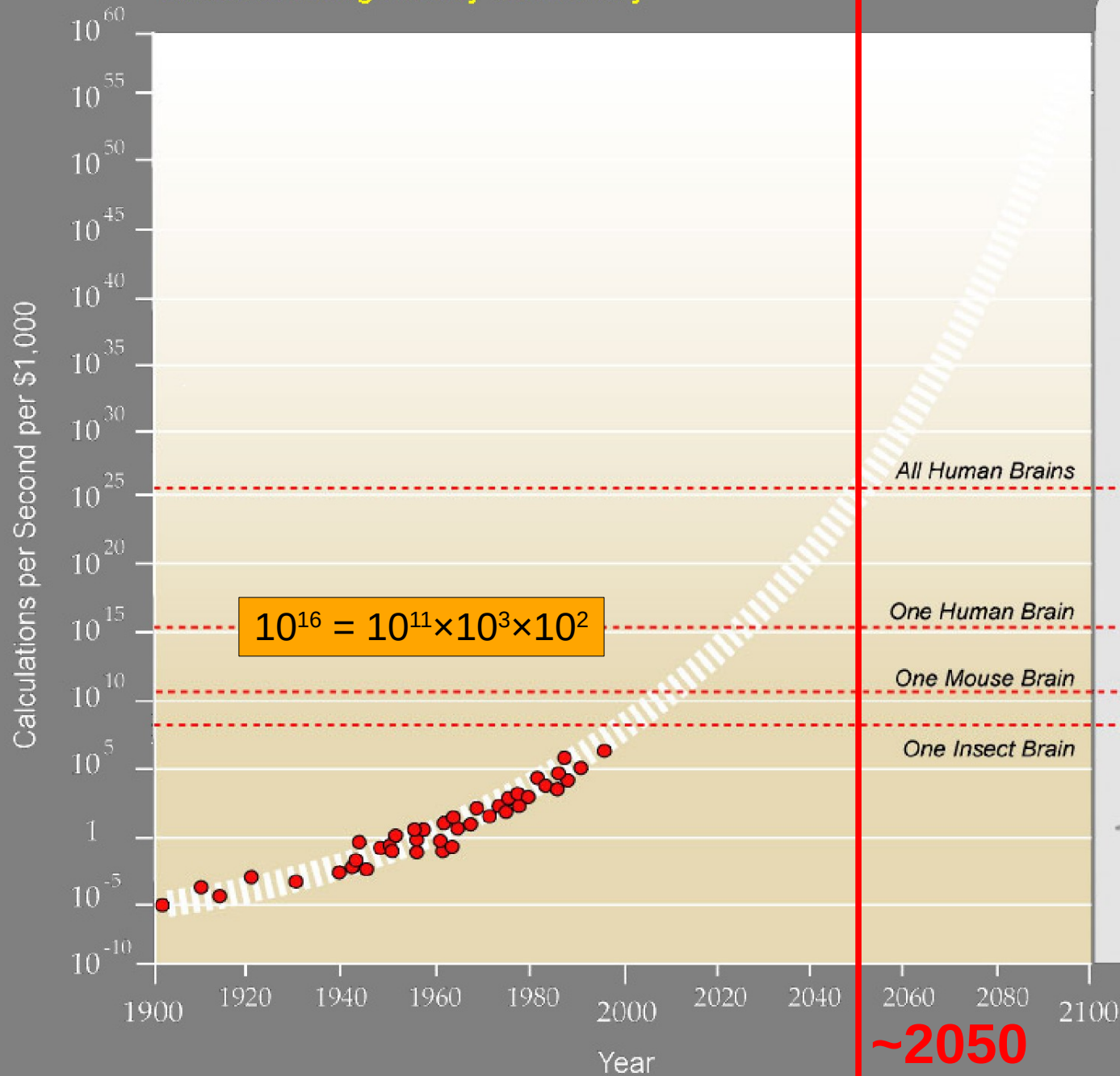
Вообще-то надо бы еще знать,
как запрограммировать AGI...

Вопросы:

1. Как растет скорость компьютеров?
2. *Какова вычислительная мощность мозга?*

Exponential Growth of Computing

Twentieth through twenty first century



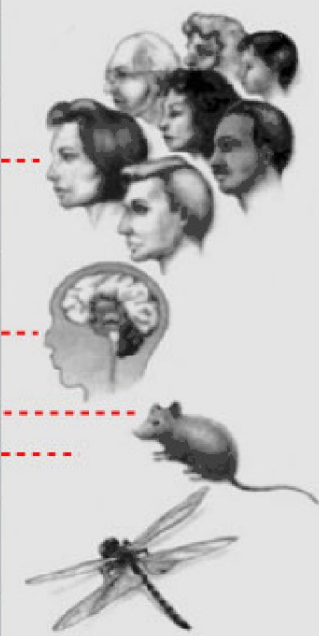
Logarithmic Plot

All Human Brains

One Human Brain

One Mouse Brain

One Insect Brain

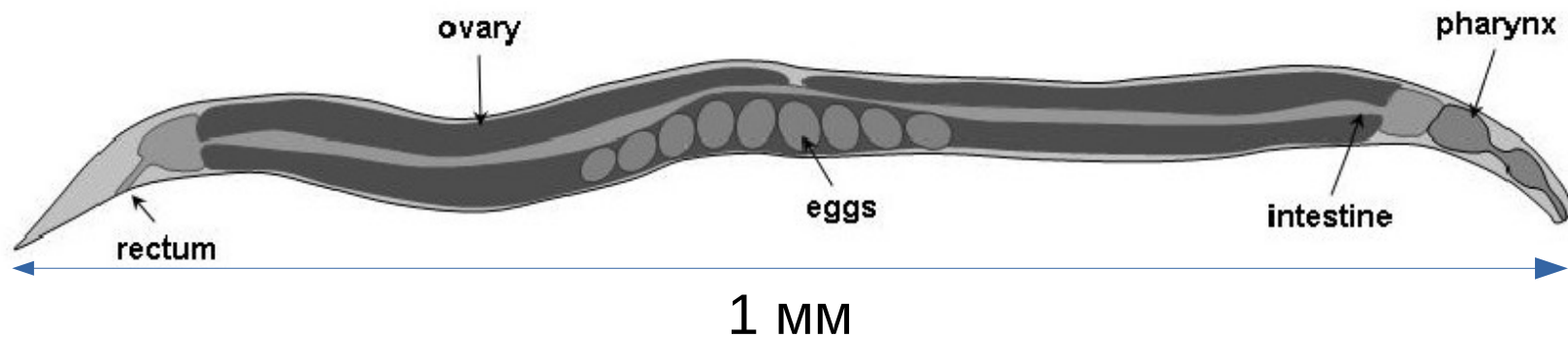


Обратная инженерия мозга:

Достаточные компьютерные мощности не гарантируют решения задачи.

Полное знание коннектома не решает задачи моделирования.

Caenorhabditis elegans (C. Elegans) - нематода



- Нервная система взрослой особи — гермафродита состоит из **302 нейронов**
Проблема мощности компьютера полностью отсутствует!
- Сложный репертуар поведений:
навигация, поиск пищи, спаривание, обучение, социальное поведение, сон.
- Нейронная система полностью картирована, каждый нейрон и синапс имеют свое имя.
- С середины 1990-х делаются попытки симулировать нервную систему *C. Elegans* на компьютере.
- Результат очень ограничен. Только симуляция движений тела (есть статьи).

Research



Cite this article: Palyanov A, Khayrulin S, Larson SD. 2018 Three-dimensional simulation of the *Caenorhabditis elegans* body and muscle cells in liquid and gel environments for behavioural analysis. *Phil. Trans. R. Soc. B* **373**: 20170376.
<http://dx.doi.org/10.1098/rstb.2017.0376>

Accepted: 11 July 2018

One contribution of 15 to a discussion meeting issue 'Connectome to behaviour: modelling *C. elegans* at cellular resolution'

Subject Areas:

biomechanics, computational biology, biophysics

Keywords:

Caenorhabditis elegans, simulation, swimming, crawling, Sibernetica, OpenWorm


Three-dimensional simulation of the *Caenorhabditis elegans* body and muscle cells in liquid and gel environments for behavioural analysis

Andrey Palyanov^{1,2,3}, Sergey Khayrulin^{1,2,3} and Stephen D. Larson³

¹Laboratory of Complex Systems Simulation, A.P. Ershov Institute of Informatics Systems, Acad. Lavrentiev ave. 6, 630090 Novosibirsk, Russia

²Laboratory of Structural Bioinformatics and Molecular Modeling, Novosibirsk State University, Pirogova str. 2, 630090 Novosibirsk, Russia

³OpenWorm Foundation, % Software Freedom Law Center, 1995 Broadway, 17th Fl., New York, NY 10023, USA

 AP, 0000-0003-1108-1486; SK, 0000-0001-7636-7835; SDL, 0000-0001-5397-6208

To better understand how a nervous system controls the movements of an organism, we have created a three-dimensional computational biomechanical model of the *Caenorhabditis elegans* body based on real anatomical structure. The body model is created with a particle system-based simulation engine known as Sibernetica, which implements the smoothed particle-hydrodynamics algorithm. The model includes an elastic body-wall cuticle subject to hydrostatic pressure. This cuticle is then driven by body-wall muscle cells that contract and relax, whose positions and shape are mapped from *C. elegans* anatomy, and determined from light microscopy and electron micrograph data. We show that by using different muscle activation patterns, this model is capable of producing *C. elegans*-like behaviours, including crawling and swimming locomotion in environments with different viscosities, while fitting multiple additional known biomechanical properties of the animal.

This article is part of a discussion meeting issue 'Connectome to behaviour: modelling *C. elegans* at cellular resolution'.



Contribute today to OpenWorm. Thank you!

OpenWorm

Building the first digital life form. Open source.

Get Started

Explore the Worm

[Blog](#) [Projects](#) For **real** worm-geeks

Star 1,544

Like 1.5K

Follow @openworm

3,021 followers

Tweet

Enter the worm.



OpenWorm is an open source project dedicated to creating the first virtual organism in a computer.

- Характерная проблема: мало симулировать нервную систему. Надо симулировать все тело и среду обитания.
- Что вообще надо моделировать для «симуляции мозга»?

А. Г. СВЕРДЛИК

КАК ЭМОЦИИ
ВЛИЯЮТ НА
АБСТРАКТНОЕ МЫШЛЕНИЕ
И ПОЧЕМУ
МАТЕМАТИКА
НЕВЕРОЯТНО
ТОЧНА



КАК УСТРОЕНА КОРА ГОЛОВНОГО МОЗГА,
ПОЧЕМУ ЕЁ ВОЗМОЖНОСТИ ОГРАНИЧЕНЫ
И КАК ЭМОЦИИ, ДОПОЛНЯЯ РАБОТУ КОРЫ,
ПОЗВОЛЯЮТ ЧЕЛОВЕКУ
СОВЕРШАТЬ НАУЧНЫЕ ОТКРЫТИЯ



«Внутренняя ось» - тело - принимает
существенное участие в мышлении

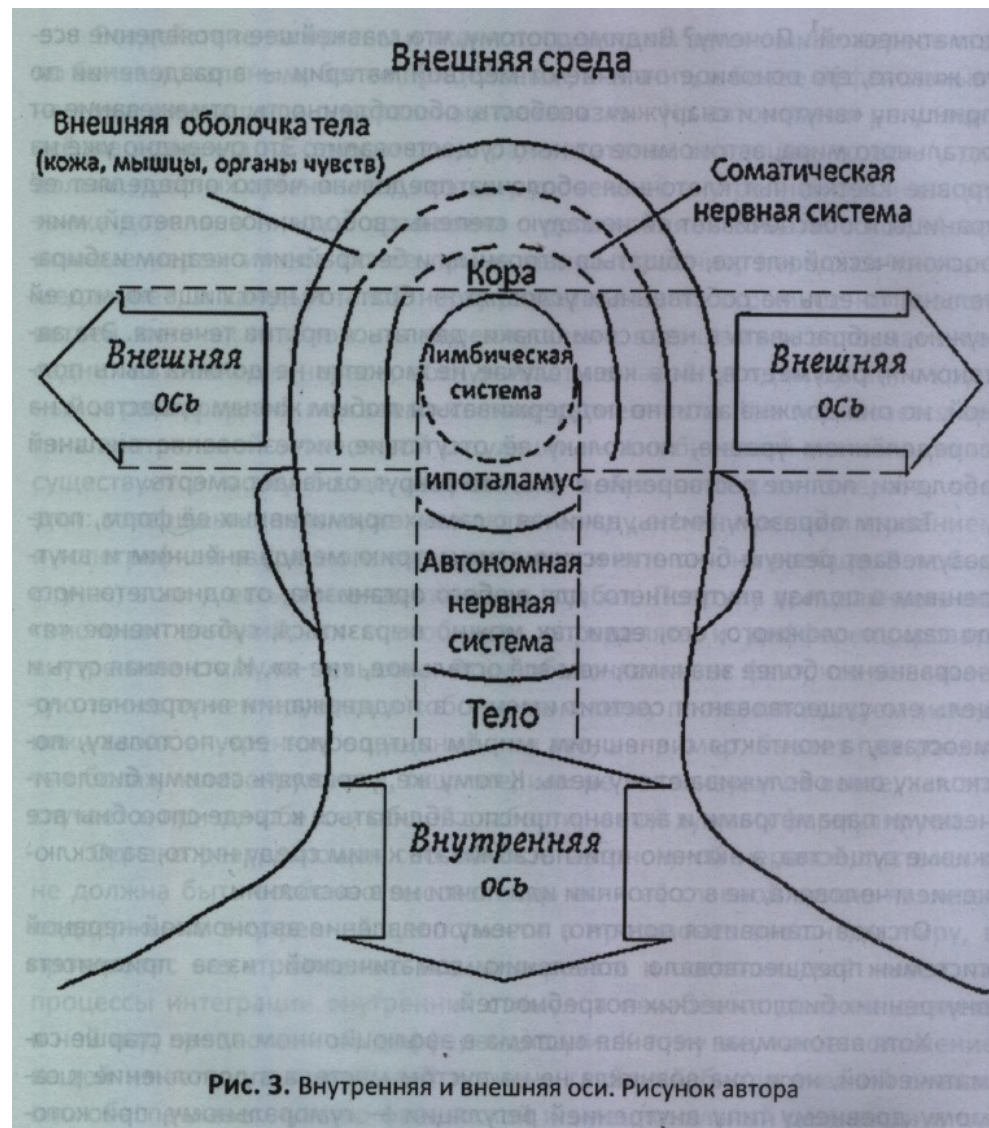
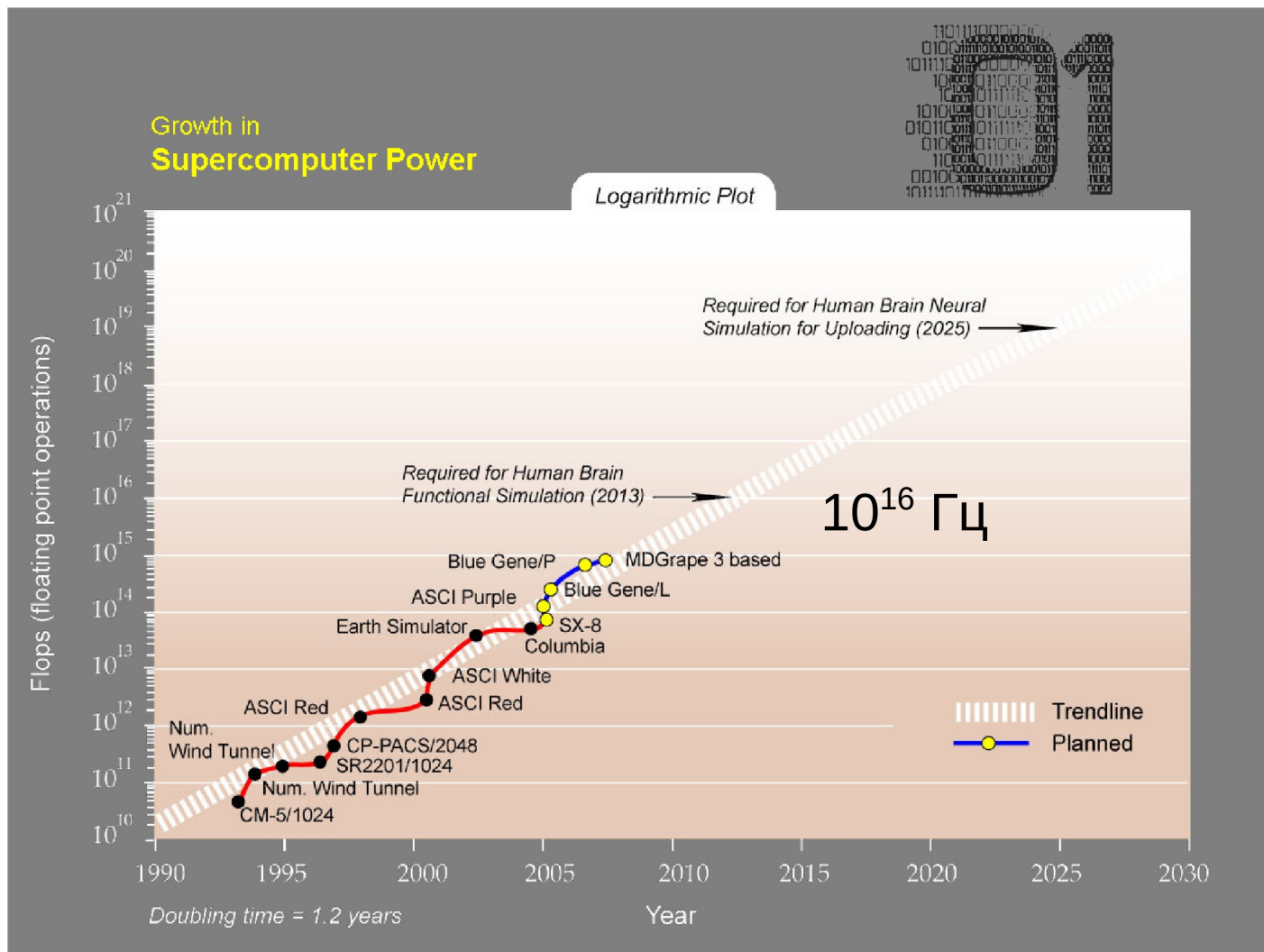


Рис. 3. Внутренняя и внешняя оси. Рисунок автора

- Социальный аспект:
Дети-маугли, коллективность и прочее...
- **Мышление – свойство Вселенной, которое появляется на определенном этапе ее развития.**
- Единичный мозг есть только интерфейс, в котором процесс мышления проявляется наиболее отчетливо.
- Надо ли для моделирования мышления моделировать Вселенную в целом?

К «быстродействию» мозга



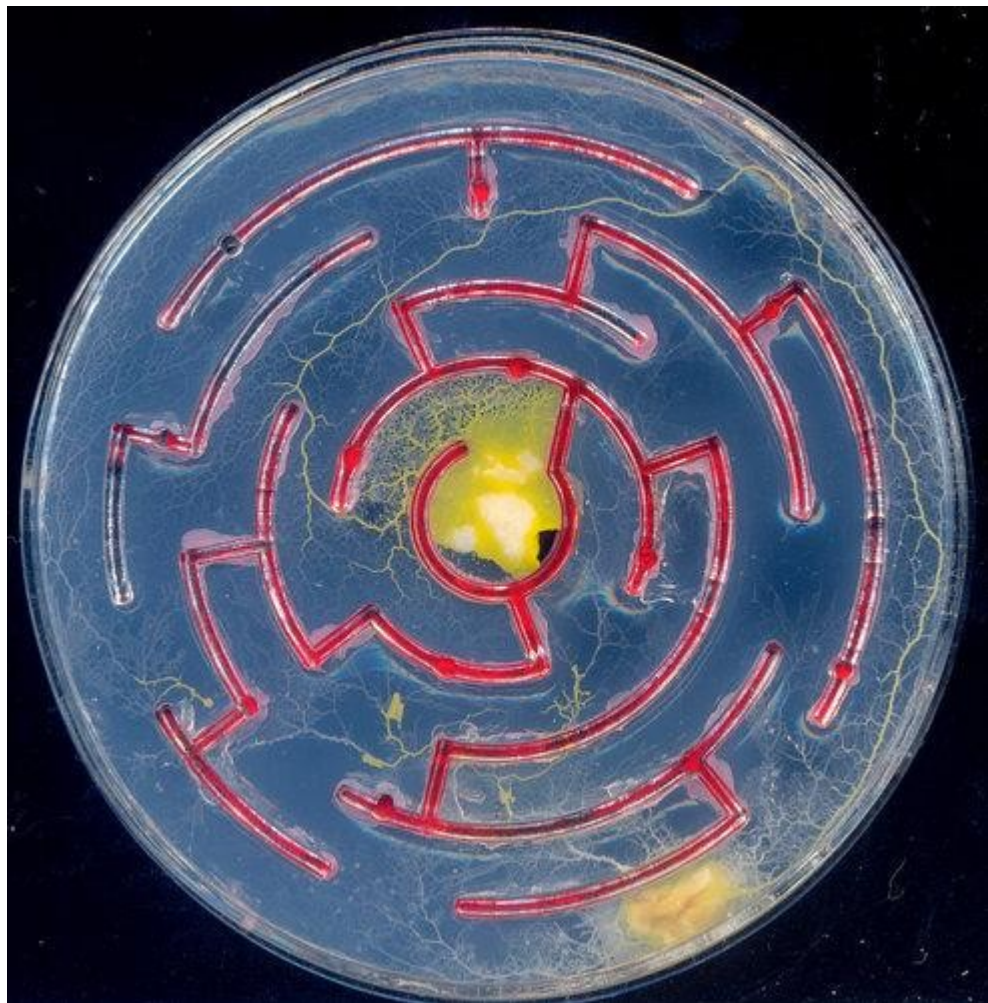
$$10^{11} \text{ нейронов} \times 10^3 \text{ синапсов} \times 10^2 \text{ Гц} = 10^{16} \text{ Гц}$$

Быстрые процессы - возбуждение, нейросетевая активность
 Нейронная парадигма, парадигма Кахаля.

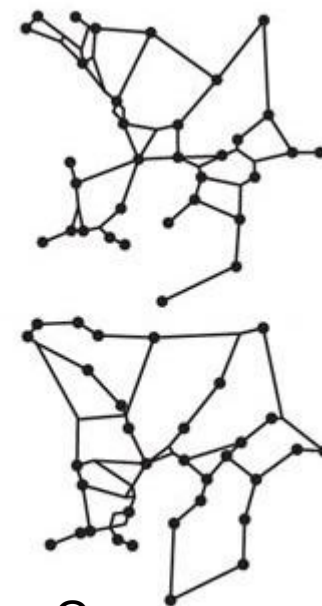
Медленные процессы - синапсы и шипики? (<http://elementy.ru/news/431207>)

**Является ли только
нейросетевая активность
«носителем» сознания?**

Слизевик: лабиринты, оптимизация, обучение.



Сеть слизевика



Сеть дорог

Влажный воздух заставляет миксомицетов двигаться быстрее, а сухой – наоборот, замедляет перемещение. Чередую поток влажного и сухого воздуха, ученые со временем зафиксировали интересную особенность: перед очередной подачей сухого воздуха слизевики снижали скорость.

Эволюционный консерватизм.

Ничто, созданное эволюцией, не пропадает зря.

Механизмы «мышления» одноклеточных должны работать и в нейронах, в том числе нейронах мозга

Доклады Академии наук СССР
1971. Том 198, № 4

УДК 62-506+612.815.1

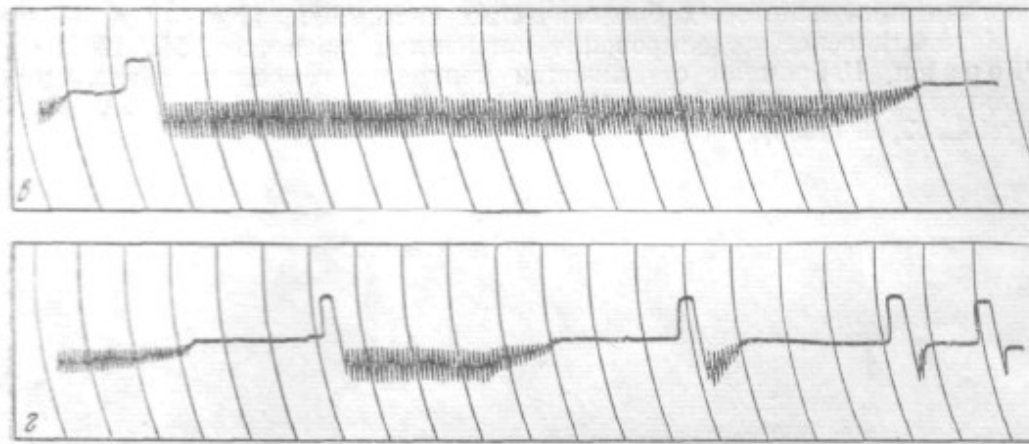
ФИЗИОЛОГИЯ

Б. Г. РЕЖАБЕК

О ПОВЕДЕНИИ МЕХАНОРЕЦЕПТОРНОГО НЕЙРОНА
В УСЛОВИЯХ ЗАМЫКАНИЯ ЕГО ЦЕПЬЮ ИСКУССТВЕННОЙ
ОБРАТНОЙ СВЯЗИ

(Представлено академиком Е. М. Крепсом 21 IV 1970)

Нейрорецептор растяжения речного рака — не имеет синаптических связей с другими нейронами, не входит в нейронную сеть.



Обучение индивидуального нейрона.
Сохраняется несколько часов.

Медленные процессы нейронов очень похожи
на движение амеб

<https://www.youtube.com/watch?v=2TIK9oXc5Wo>

Где на субклеточном уровне может
помещаться механизм
«внутриклеточного сознания»?

Цитоскелет и микротрубочки

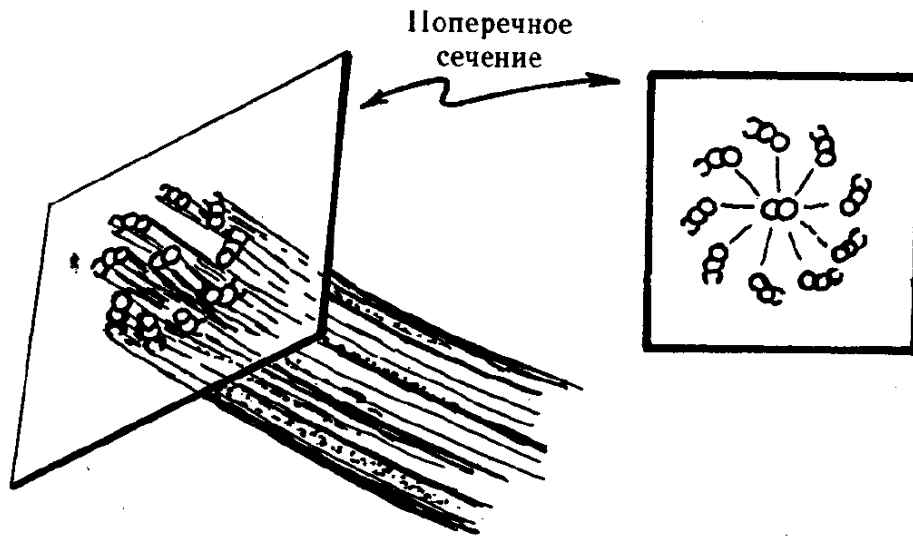


Рис. 7.3. Важной частью цитоскелета являются пучки крохотных трубочек (микротрубочек), организованных в структуры, напоминающие в поперечном сечении лопасти вентилятора. Такое строение имеют, например, реснички парамеции.

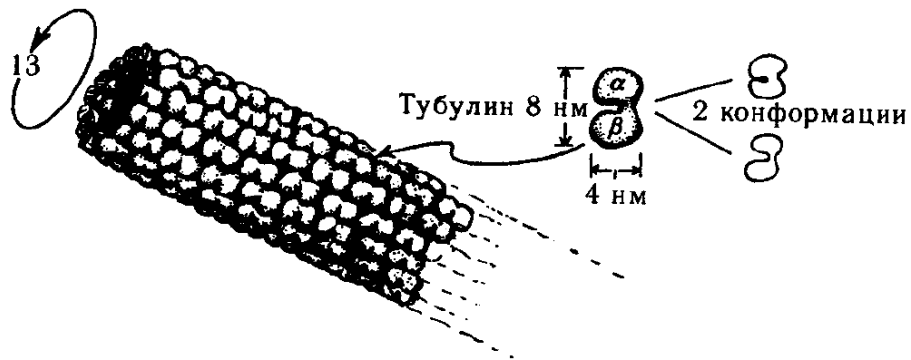


Рис. 7.4. Микротрубочка. Полая трубка, обычно состоящая из 13 рядов димеров тубулина. Каждая из молекул тубулина может существовать в двух (по крайней мере) конформациях.

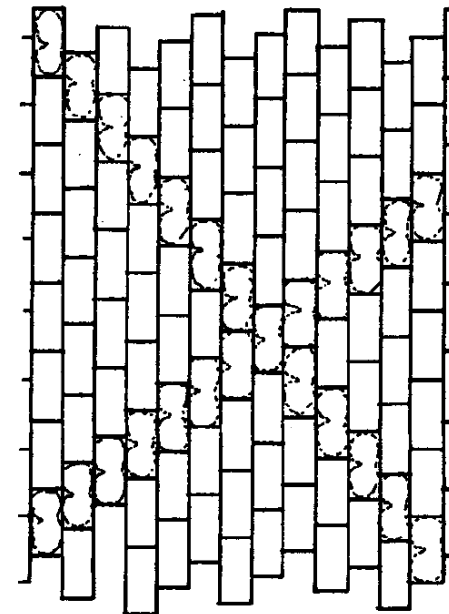
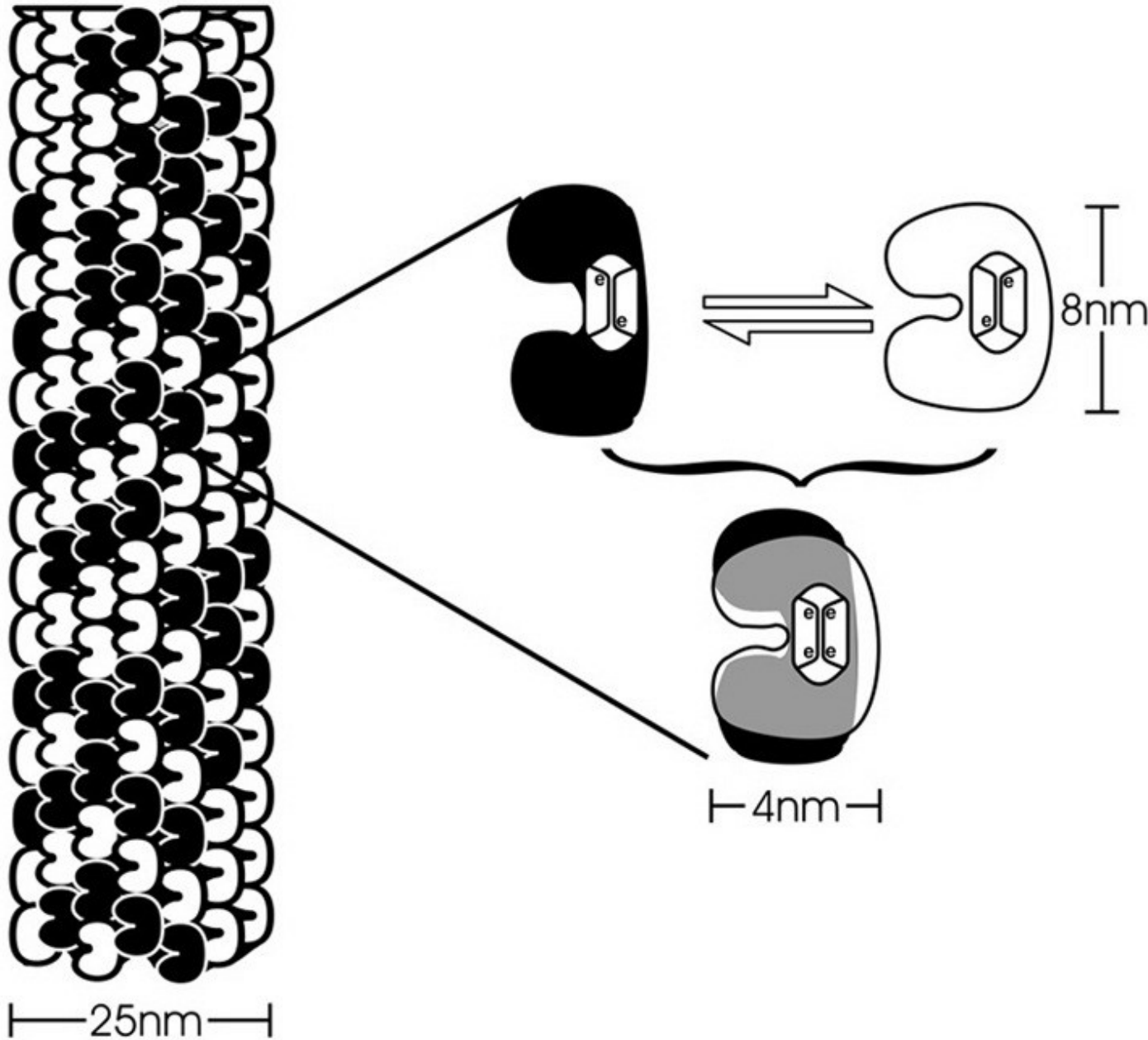


Рис. 7.8. Представим, что микротрубочка разрезана вдоль и затем развернута в полосу. Можно видеть, что молекулы тубулина располагаются вдоль наклонных линий, причем каждый новый виток смещен относительно предыдущего на 5 или 8 молекул (в зависимости от того, куда наклонена линия, вправо или влево).

Микротрубочка — клеточный автомат, состоящий из молекул тубулина — битов с двумя состояниями



α - тубулин
 β - тубулин
две
конформации,
0/1

Каково может быть быстроедействие «клеточного автомата» микротрубки?

10^7 Гц — основная частота микротубочек:

Pokorny, J., Hasek, J., Jelinek, F., Saroch, J. & Palan, B. (2001)
Electromagnetic activity of yeast cells in the M phase.
Electro Magnetobiol 20, 371–396.

Pokorny, J. (2004)
Excitation of vibration in microtubules in living cells.
Bioelectrochem. 63: 321-326.

На один нейрон: 10^8 тубулинов $\times 10^7$ Гц = 10^{15} Гц
(это, конечно, максимум скорости)

На мозг: $10^{11} \times 10^{15} = 10^{26}$ Гц

Аналогия МОЗГ-классический компьютер

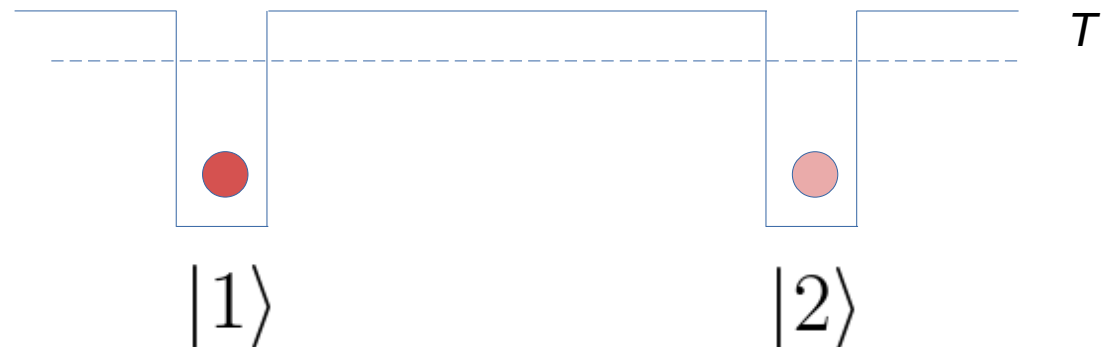
Аналогия мозг — это классический компьютер.

Является ли мозг классическим вычислительным устройством?

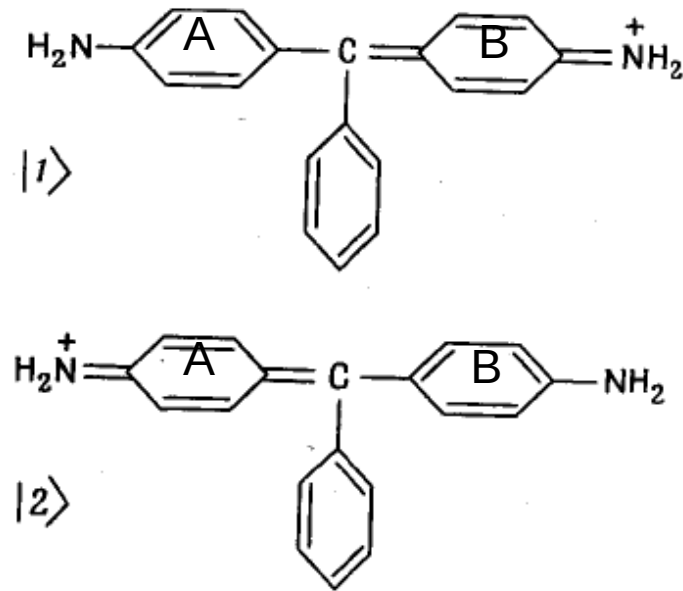
- Может ли мозг быть квантовым компьютером?
- Быстрые процессы нейронной сети —
(скорее всего) классические процессы.
- Могут ли на более глубоких уровнях работать квантовые процессы вычислений?
- «Мозг слишком мокрое и тёплое место» -
наивный контр-аргумент от спинтроники
(Рэй Курцвейл и многие другие).
- Атом гелия. Температура ионизации 25000 градусов.
Всё зависит от *энергетической щели*.

Кубит на квантовой точке

$$|\Psi\rangle = a_1|1\rangle + a_2|2\rangle$$



Квантовые суперпозиции в макромолекулах



Фуксин, краситель.
 При комнатной температуре —
 кристалл или раствор.

$$|I\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle); \quad E$$

$$|II\rangle = \frac{1}{\sqrt{2}}(|1\rangle - |2\rangle); \quad E + \Delta E$$

$$|1\rangle = |A3\rangle|B2\rangle$$

$$|2\rangle = |A2\rangle|B3\rangle$$

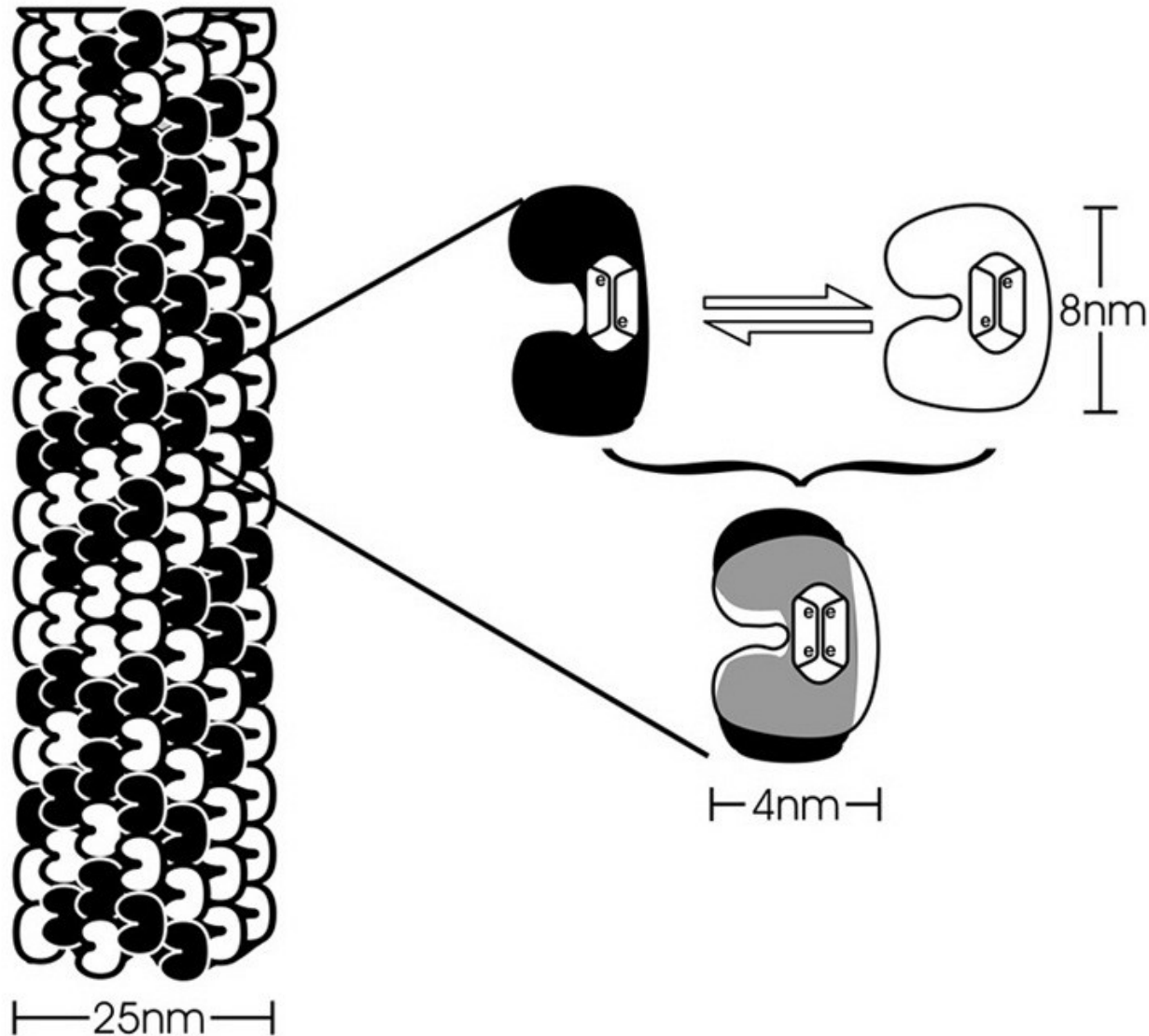
$$|I\rangle = \frac{1}{\sqrt{2}}(|A3\rangle|B2\rangle + |A2\rangle|B3\rangle)$$

$$|II\rangle = \frac{1}{\sqrt{2}}(|A3\rangle|B2\rangle - |A2\rangle|B3\rangle)$$

Скореллированные пары
 Эйнштейна-Подольского-Розена

Все зависит от энергетической щели, отделяющей состояния от окружения!

Микротрубочки: Может ли тубулин быть кубитом?



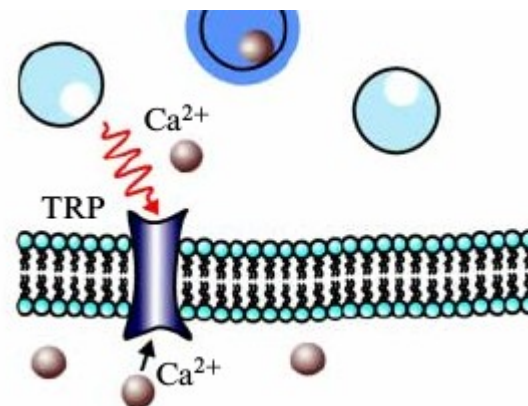
Кристаллическая вода
изолирует микротрубочку
от внешнего влияния.

Микротрубочка -
квантовый клеточный
автомат?

**Может быть да,
Может быть нет.**

Другие места в клетке, где может иметь место квантово-информационная активность

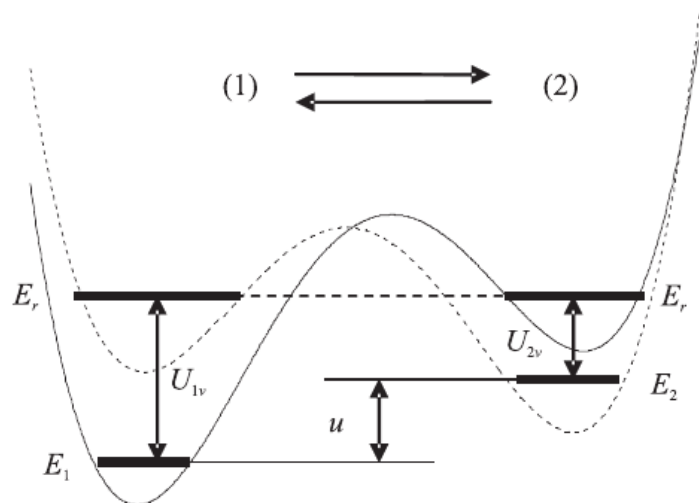
Мембранные ионные каналы.



А.М. Жёлтиков. УФН, Т.188(2018) С. 1119

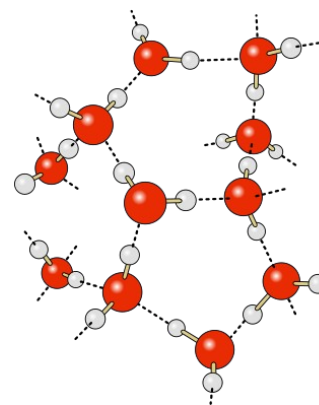
Двухъямный потенциал водородной связи ДНК, РНК, вода (!)

$U(x)$



П. М. Красильников
КОМПЬЮТЕРНЫЕ ИССЛЕДОВАНИЯ И
МОДЕЛИРОВАНИЕ Т. 1 (2009) № 3 С. 297

Вода не «простая»,
а «связанная жидкость»



Википедия



1972, рус.1978

Хьюберт Дрейфус.

Почему мозг не компьютер?

Машины обрабатывают информацию,
а человек работает со смыслами.

Вовсе не очевидно, что человеческие
смыслы могут быть закодированы
информацией.

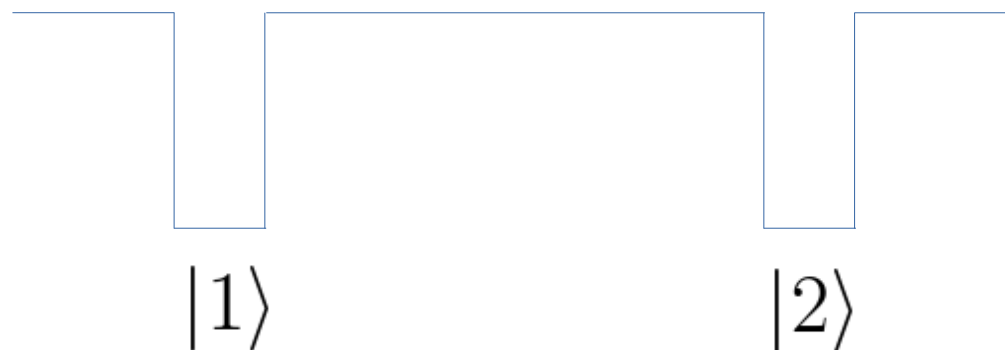
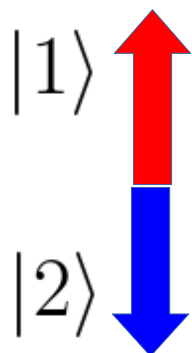
Мысль — не вычисление

Квантовая информация

Одна и та же информация может быть записана на разные носители.
Информация — это *абстракция* от распределения физических неоднородностей, кодирующих эту информацию.

Квантовые состояния одной и той же структуры могут быть реализованы на разных носителях.

$$|\Psi\rangle = a_1|1\rangle + a_2|2\rangle$$



Абстракция от структуры квантового состояния, независимо от природы системы - это **квантовая информация**.

Квантовая информация \neq Информация

Телепортация квантовых состояний (квантовой информации)

Копирование квантовой информации запрещено:

$$\langle A_X | \langle B_0 \rangle \rightarrow \langle A_X | \langle B_X \rangle$$

Телепортация квантовой информации разрешена:

$$\langle A_X | \langle B_0 \rangle \rightarrow \langle A? \rangle \langle B_X \rangle$$

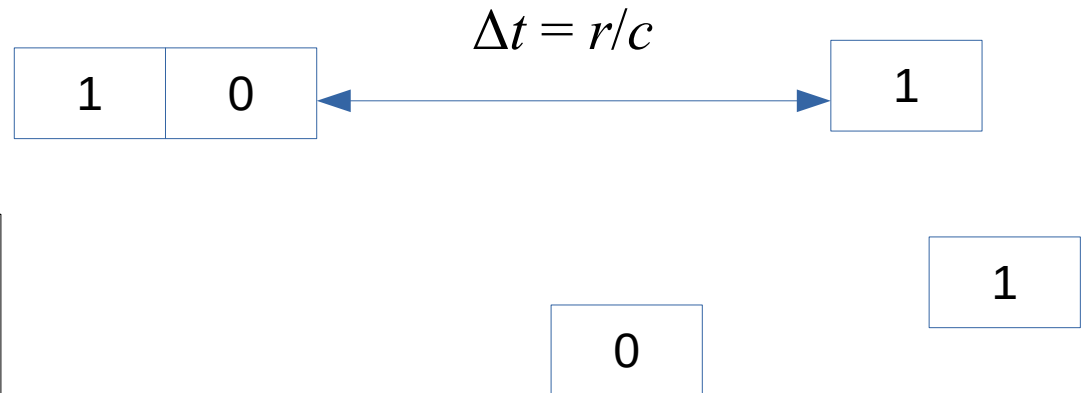
Квантовая информация не обладает основным свойством обычной информации – ее нельзя копировать, поэтому квантовая информация – НЕ информация

Релятивистское информационное поле

Информация – цепочка бит



Любой бит можно произвольно сопоставить с любым другим битом

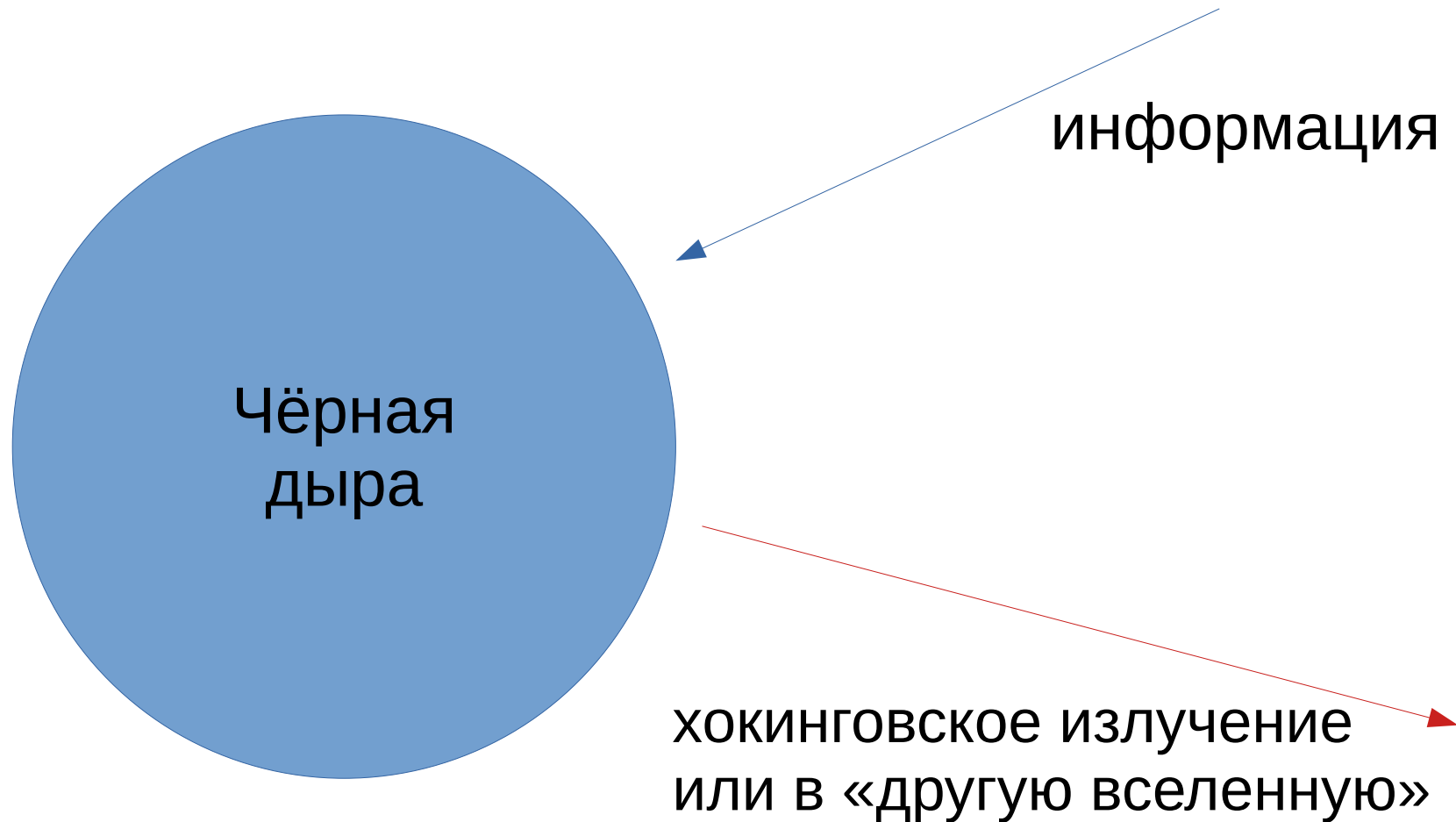


Релятивистское информационное поле – цепочка бит + локализация каждого бита в пространстве-времени

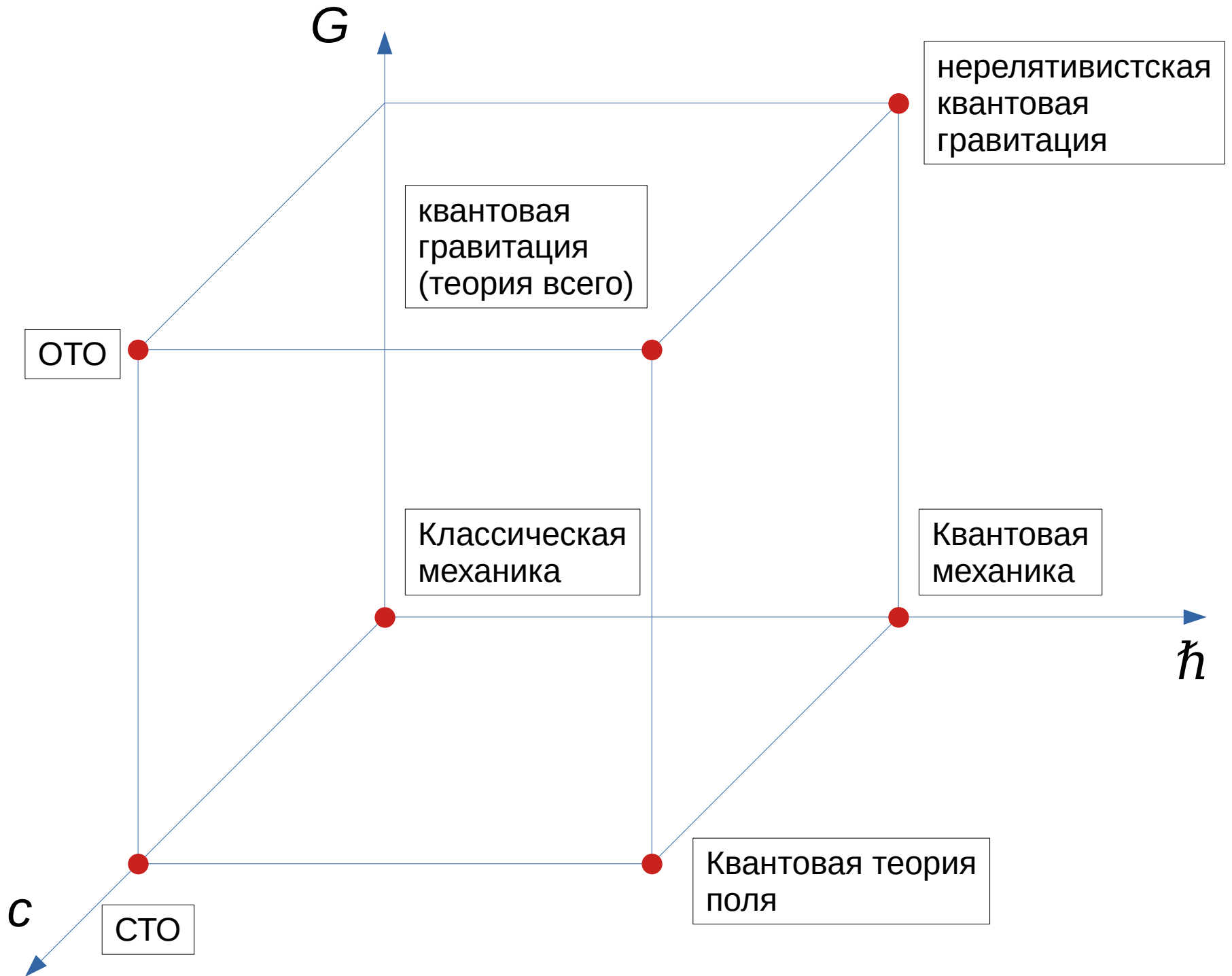
Любой бит можно сопоставить с другим битом за принципиально конечный промежуток времени.

Машина Тьюринга не может быть моделью для работы с информационным полем.

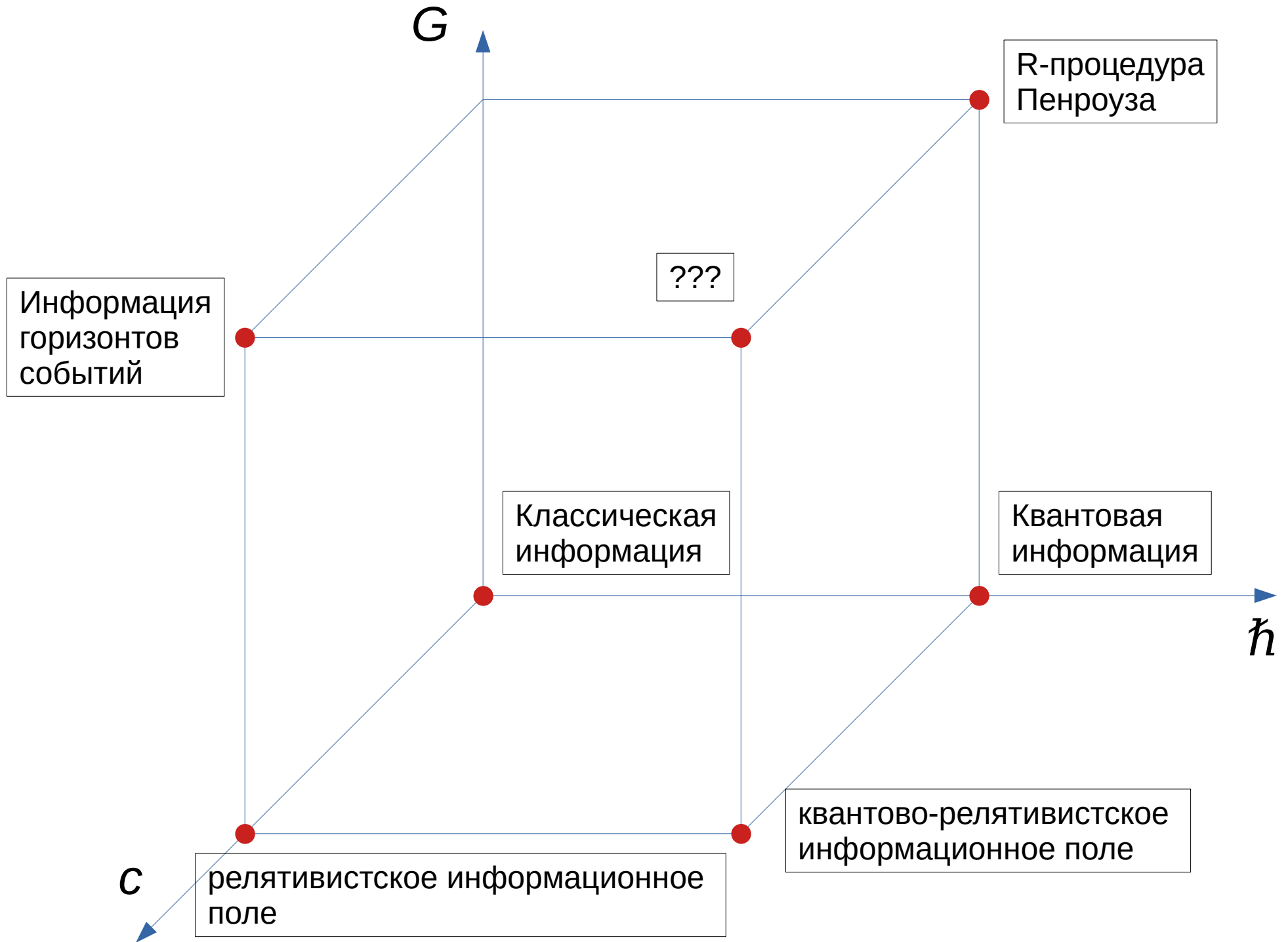
**Общерелятивистская информация –
под горизонтами событий
(или на горизонтах событий)**



Куб физических теорий



Куб физической информации



Тонкие вопросы, связанные с разными видами физической информации

Квантовая механика доступна только через классический интерфейс классических приборов.

Это не означает, что квантовая механика сводится к поведению некоторых классических приборов.

За поведением классических приборов стоит другая, квантовая онтология



Квантовая информация, релятивистское информационное поле и т. д. допускают модели в обычной информации

Это не означает, что они сводятся к обычной информации

За такими информационными моделями стоит онтология неклассической информации



1972, рус.1978

Хьюберт Дрейфус.

Почему мозг не компьютер?

Машины обрабатывают информацию,
а человек работает со смыслами.

Вовсе не очевидно, что человеческие
смыслы могут быть закодированы
информацией (стр. 114-115)

Мысль — не вычисление

Если в формировании и работе со смыслами задействованы квантовые внутринейронные моды, то в строгом смысле смыслы действительно не кодируются (классической) информацией — они кодируются квантовой информацией, которая не является информацией в обычном понимании

Парад аналогий

17-18-й в. Классическая механика. Механические часы, Солнечная система	Мозг - механизм
19-начало 20 в. Электродинамика, ЭЭГ	Мозг — электрическая цепь. Персональный магнетизм
Со второй половины 20-го века Компьютеры, информатика	Мозг – устройство обработки информации, компьютерная аналогия

- Компьютерная аналогия может быть не более, чем модой или нерerefлексированным стереотипом.
- Мозг может оказаться НЕ устройством обработки информации или даже квантовой информации – не компьютер, не квантовый компьютер, не квантово-гравитационный компьютер (Пенроуз)
- Тогда реинжиниринг мозга в компьютерных моделях не только трудно решаемая задача, но ложное направление мысли
- Компьютерная модель явления не тождественна самому явлению. Создание компьютерной модели может оказаться принципиально невозможным.

- Программа реинжиниринга мозга
бесперспективна с точки зрения создания AGI
- Программа реинжиниринга мозга полезна
для изучения работы мозга

Слон — это лошадь