

МОСКОВСКИЙ ТЕХНОЛОГИЧЕСКИЙ УНИВЕРСИТЕТ

**Кафедра философии, социологии и политологии
Дисциплина «Философия»**



**«Китайская комната»
Дж. Сёрля**

**Выполнил: Харитонов И.А.,
студент группы ККСО-01-14
Института кибернетики
МИРЭА**

Содержание

- 1. Краткая биография Джона Сёрла**
- 2. Что такое мысленные эксперименты?**
- 3. Цель эксперимента и его предпосылки**
- 4. Описание и аргументация «Китайской комнаты»**
- 5. Дискуссии в академических кругах**
- 6. Заключение**
- 7. Список литературы**

Мысленные эксперименты

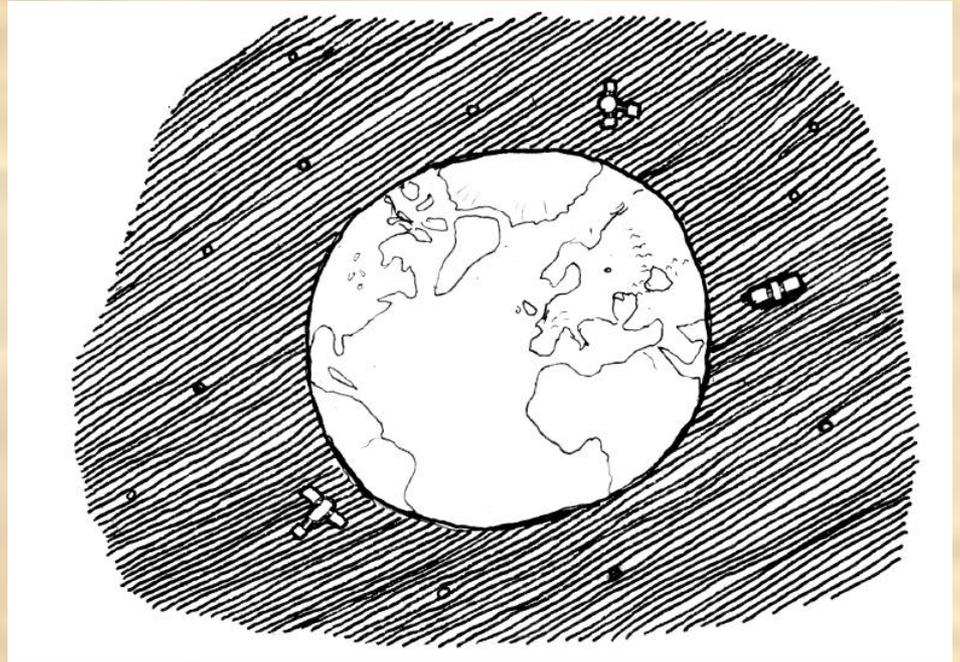
Понятие «Мысленный эксперимент» ввел австрийский физик, механик и философ –

Эрнест Мах

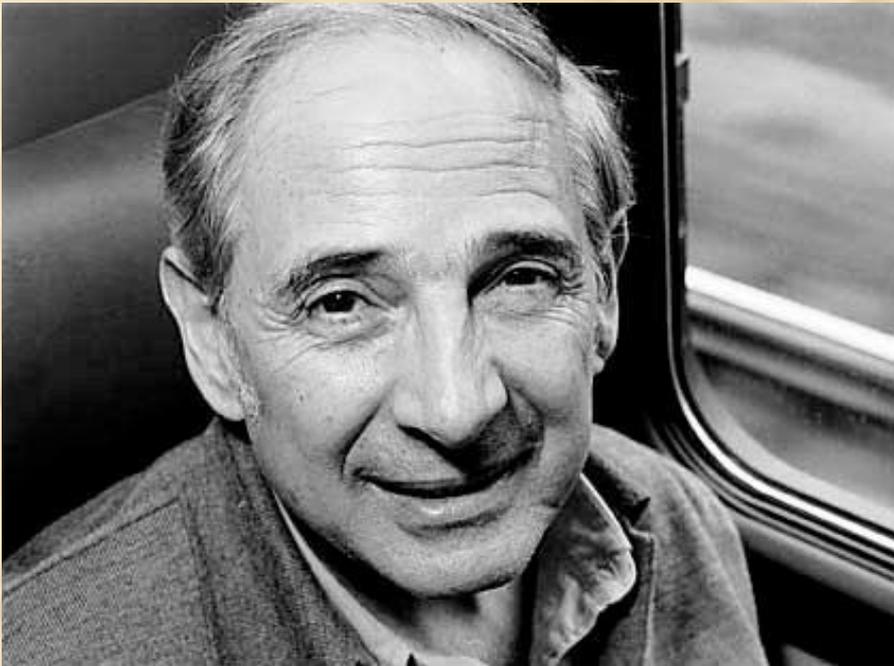


Мысленные эксперименты

Мысленный эксперимент — вид познавательной деятельности, в которой ключевая для той или иной научной теории ситуация разыгрывается не в реальном эксперименте, а в воображении.



Джон Сёрл



**Джон Роджерс
Сёрл**

Американский философ 20 века.

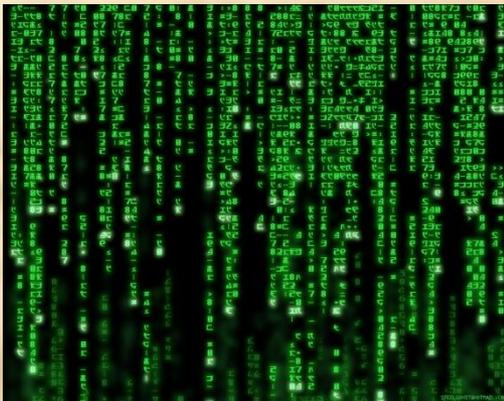
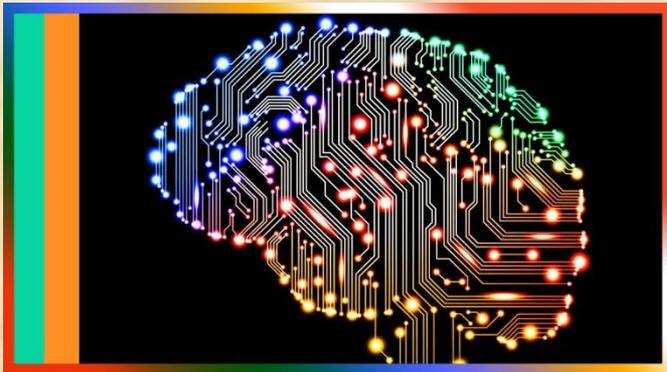
С 1980-х годов стал ведущим специалистом по философии искусственного интеллекта.

Профессор философии Калифорнийского университета, Беркли.

Автор эксперимента «Китайская комната»

Цель эксперимента

Опровержение предположения, что компьютеры или другие виды искусственного интеллекта могут думать и понимать.



Предпосылки

У эксперимента Джона Сёрла было множество предшественников, но стоит выделить три самых важных:

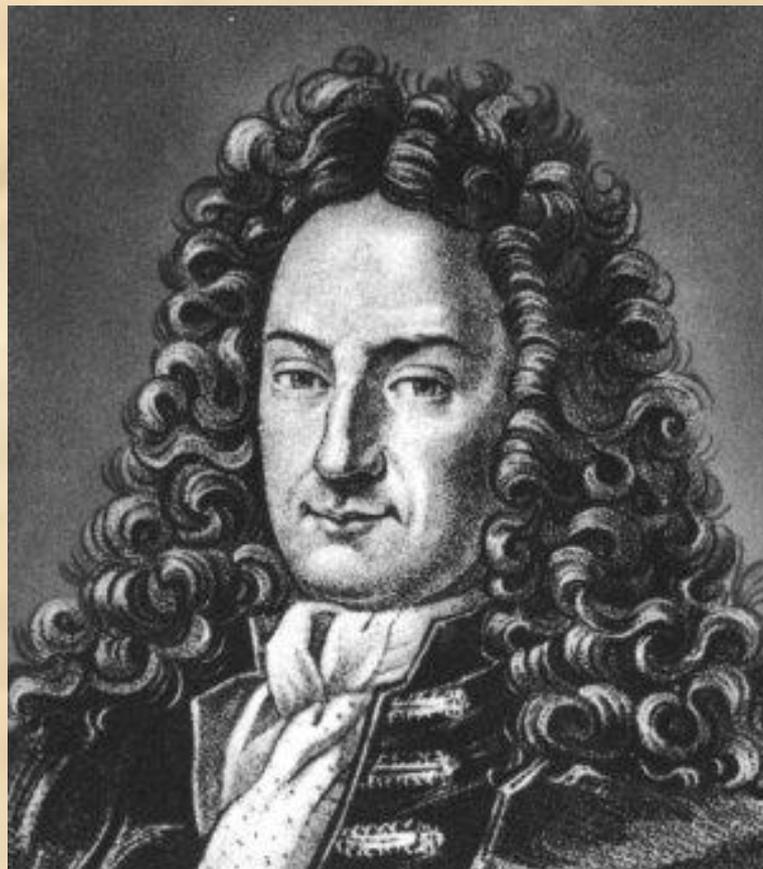
1. Мельница Лейбница

2. Бумажная машина Тьюринга

3. Китайская нация

Мельница Лейбница

В этом мысленном эксперименте Лейбниц предлагает представить машину размером с мельницу, симулирующую чувства, мысли и восприятие. И если зайти внутрь такой машины, то ни одна из составляющих механизмов не будет являться сознанием или объяснением восприятия.



Бумажная машина Тьюринга

Бумажная машина Тьюринга является машинным алгоритмом для игры в шахматы, где в роли машины-исполнителя выступал человек.

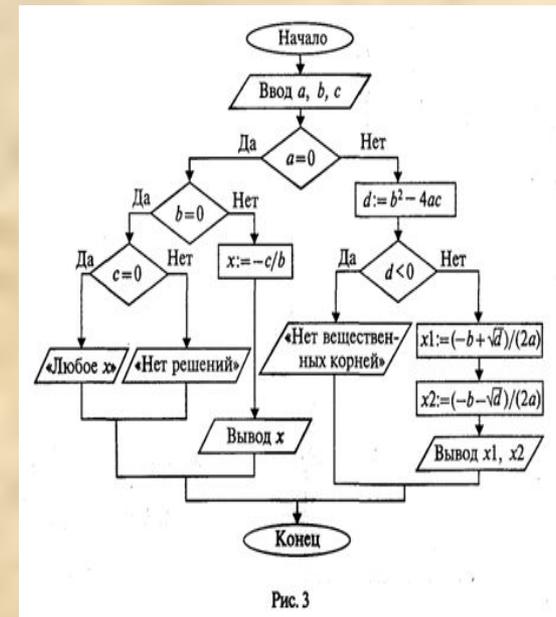
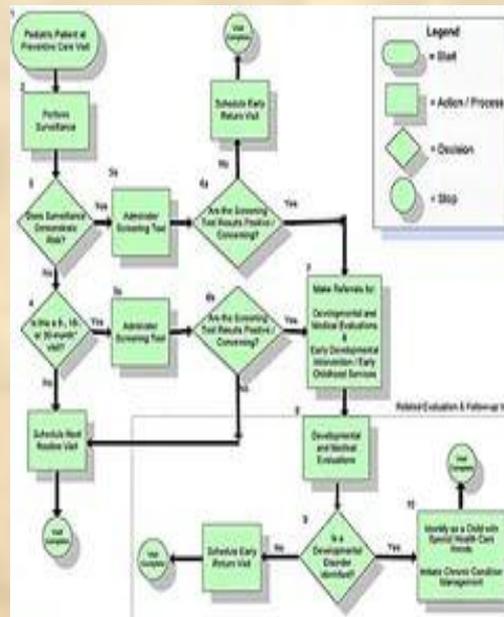
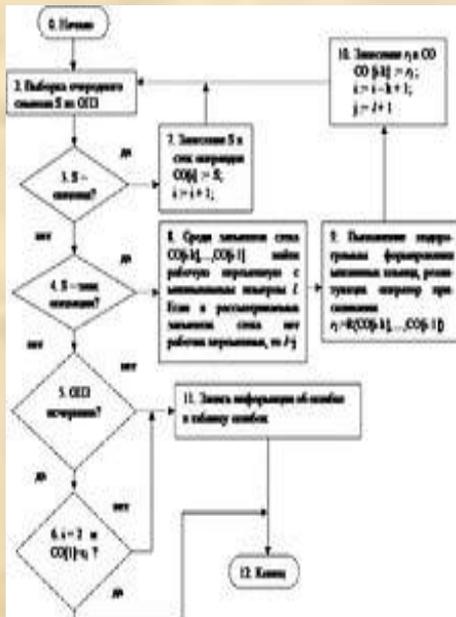


Рис. 3

Бумажная машина Тьюринга

Человеку требовалось более получаса, чтобы сделать один ход. Алгоритм был довольно условный, и сохранилась даже запись партии, где «бумажная машина» Тьюринга проиграла одному из его коллег. За отсутствием доступа к компьютеру, программа ни разу не проверялась в работе.



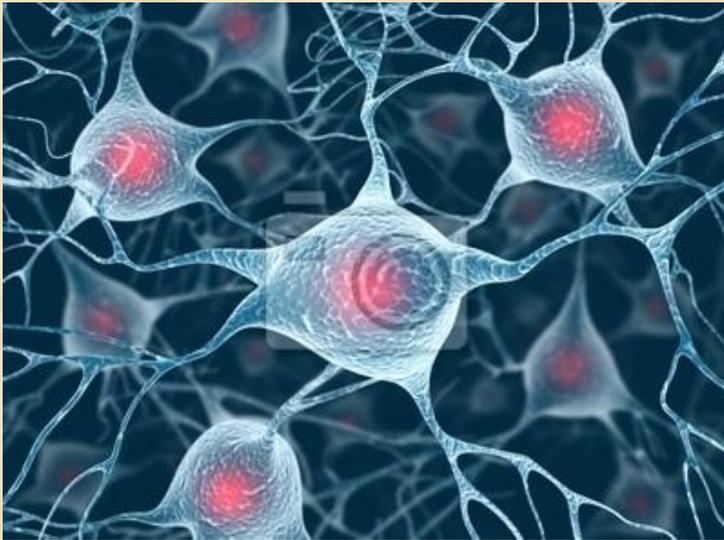
Китайская нация

Мысленный эксперимент, предложенный Недом Блоком в 1978 году для опровержения функционализма

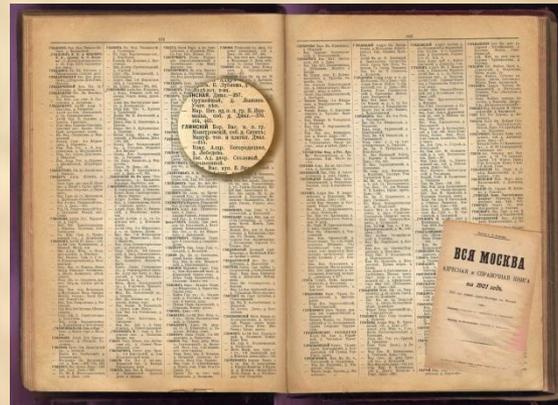
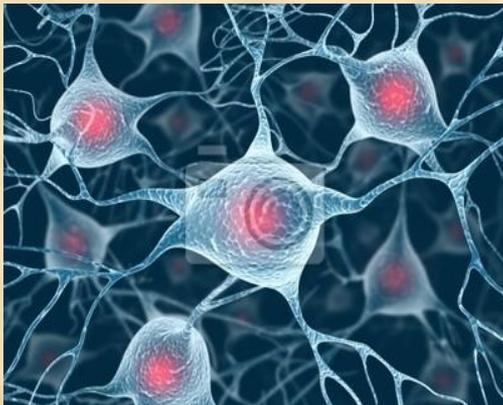


Китайская нация

Нед блок предлагал использовать множество китайцев, каждый из которых эмулировал бы нейрон искусственного мозга.



Китайская нация

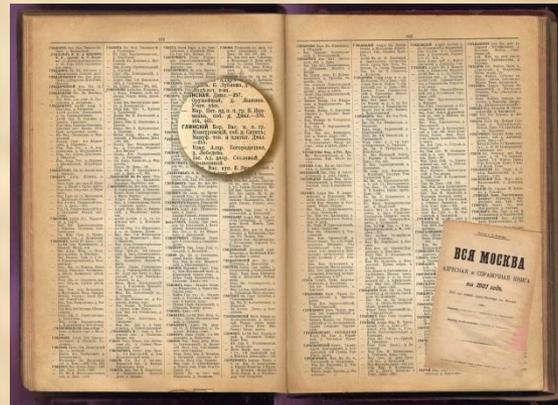
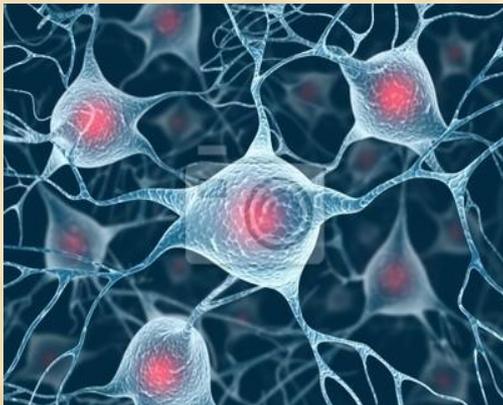


Каждый китаец в этой системе имеет телефон и список номеров других китайцев. В зависимости от того, когда зазвонил телефон конкретного китайца, он звонит какому-либо из своего списка.

Китайская нация



При этом содержание разговора совершенно не важно: важен сам факт звонка.



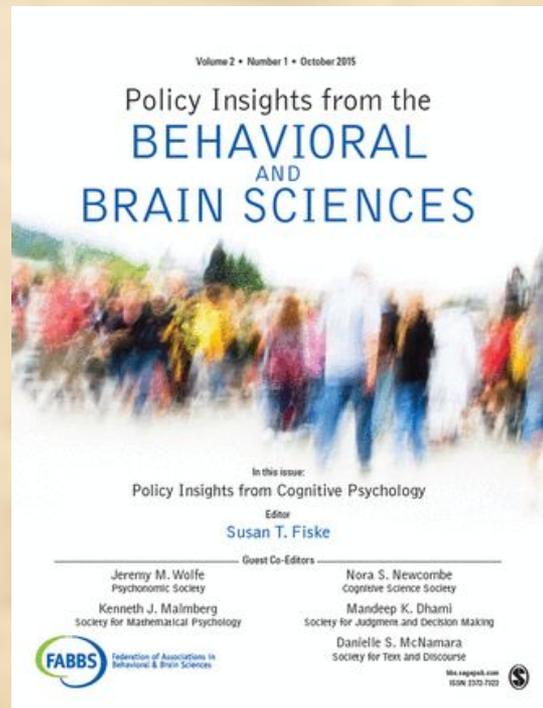
Китайская нация

Тогда, если в такой системе создать функциональное состояние, которое появляется при восприятии боли, вся «Китайская нация» окажется в состоянии боли, при этом ни один из китайцев не будет эту боль чувствовать.



Описание эксперимента

Эксперимент был опубликован в 1980 году в статье «Minds, Brains, and Programs» журнала «The Behavioral and Brain Sciences».



Описание эксперимента

Представим себе изолированную комнату, в которой сидит человек (испытуемый), не знающий ни одного китайского иероглифа, но имеет книгу с точными инструкциями по манипуляции иероглифами вида: «такие-то иероглифы поместите рядом с такими-то».



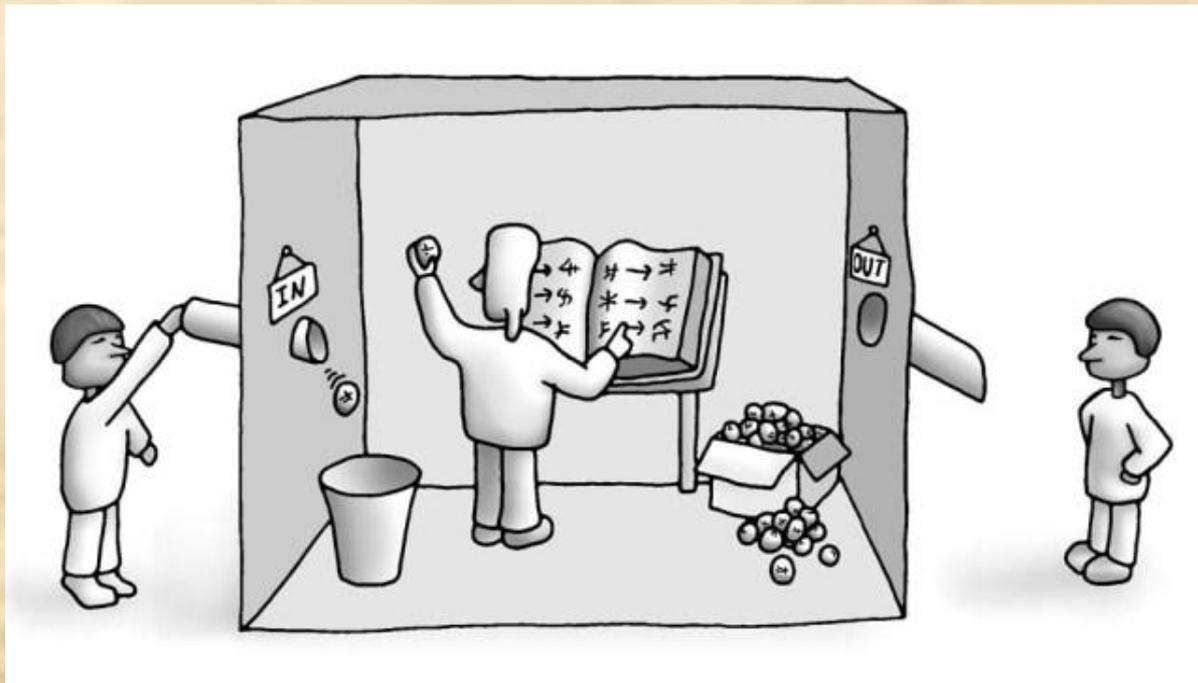
Таблица иероглифических ключей

一	丨	丶	ノ	乙	丨	二	上	人	儿
八	八	口	冂	冂	几	山	刀	力	夕
匕	匕	匕	十	卜	尸	厂	厶	又	口
日	土	士	又	夕	夕	夕	夕	夕	夕
寸	小	无	尸	夕	山	工	己	中	
干	艮	产	又	升	戈	弓	王	多	才
心	戈	尸	手	文	文	文	斗	斤	方
无	日	日	月	木	欠	止	夕	爻	母
比	毛	氏	气	水	火	爪	父	爻	月
月	牙	牛	犬	豕	玉	瓜	瓦	甘	生
用	田	疋	疋	火	白	皮	血	目	牙



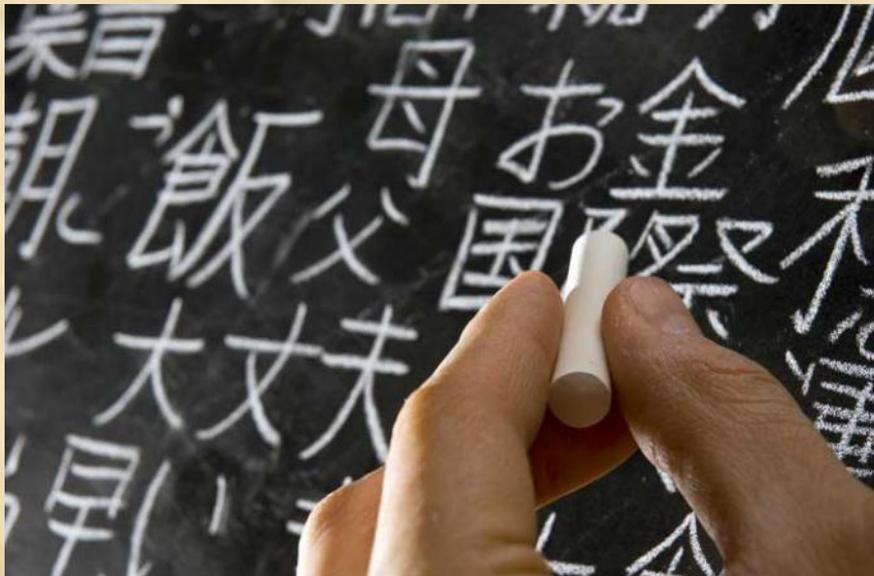
Описание эксперимента

Наблюдатель может задать любой вопрос и получить на него осмысленный ответ, но испытуемый при этом не имеет никаких знаний об иероглифах и не может научиться ими пользоваться, потому что не может узнать значение ни одного из иероглифов.



Описание эксперимента

Наблюдатель передает испытуемому через щель иероглифы с вопросами, а на выходе должен получить осознанный ответ. При этом после выполнения инструкций из книги с полученными иероглифами они преобразуются в ответ.



Аргументация эксперимента

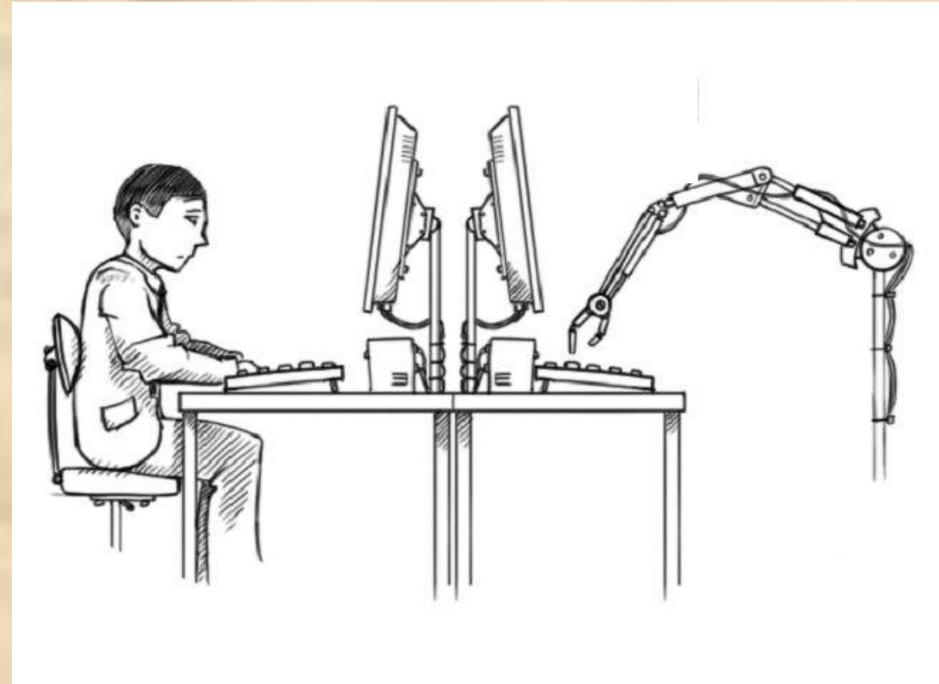
Таким образом Сёрл заключает, что эта система может пройти тест Тьюринга, но при этом никакого понимания языка не происходит, а значит тест Тьюринга не является адекватной проверкой мыслительных способностей.



Аргументация эксперимента

Формально аргументацию можно представить так:

- 1. Если гипотеза «сильного» ИИ верна, тогда существует такая программа для китайской письменности, при запуске которой в вычислительной системе эта система будет понимать китайскую письменность.**
- 2. Я могу исполнить программу для китайской письменности без понимания этой письменности.**
- 3. Следовательно, гипотеза «сильного» ИИ неверна.**



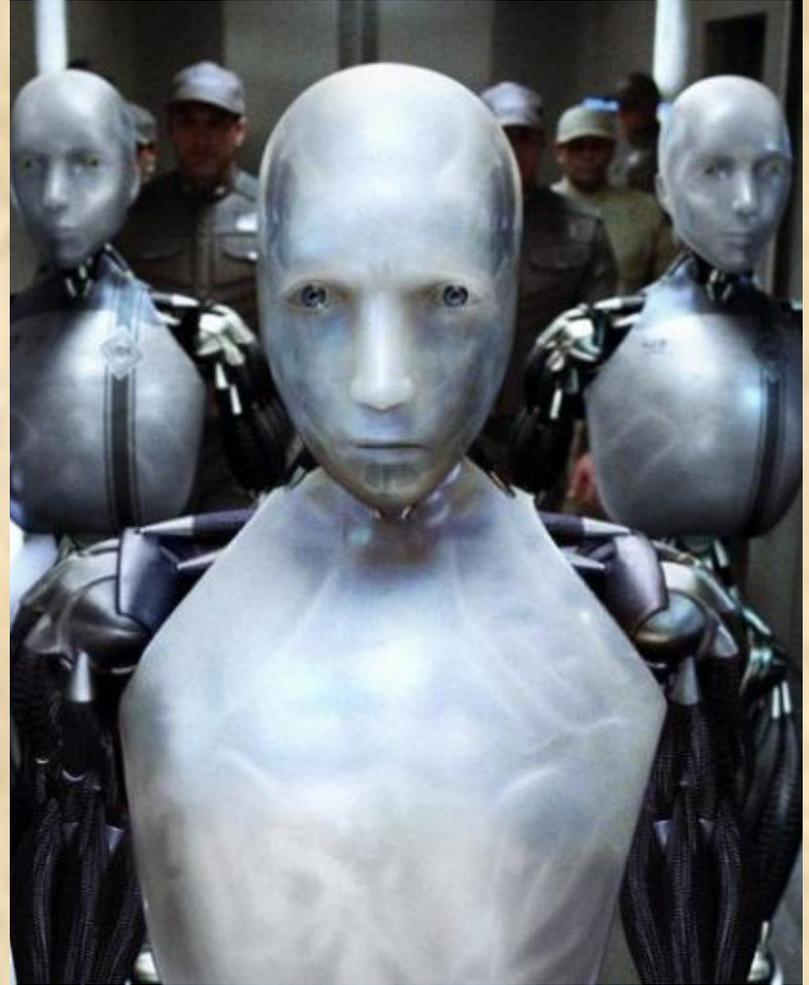
Дискуссии

Некоторые критики считают, что тест Тьюринга выдержал не человек в комнате, а система, состоящая из комнаты, книги правил и человека. Система из книги правил, человека и комнаты, по их мнению, является разумной и понимает китайский язык. В качестве контраргумента Сёрл предлагает заставить запомнить человека формальные правила ответов для системы. По мнению философа, система, которая теперь будет состоять только из человека, всё равно не сможет понять китайский язык.



Дискуссии

Часть критиков предлагала создать робота с самообучающейся программой, а также оснастить его сенсорами для восприятия мира. Такой робот, по их мнению, может взаимодействовать с миром, познавать его и быть способен обучиться подобно ребёнку. Сёрл же предлагал поместить человека внутрь робота, который, выполняя бы движение, следя информации с сенсоров, и в то же время он бы не имел никакого понимания происходящего, а лишь выполнял инструкции.



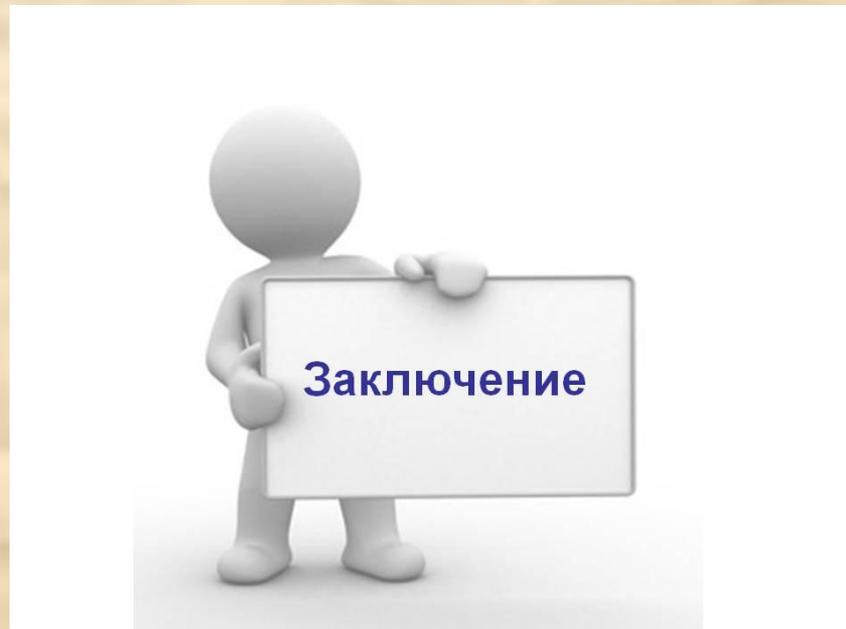
Дискуссии

Некоторые критики предлагают воссоздать нейронную сеть человека, владеющего китайской письменностью. Такая нейронная сеть должна понимать иероглифы. На это Сёрл предлагал модифицированную «китайскую комнату». В такой комнате вместо нейронов можно представить водопроводные трубы, а человек внутри комнаты переключает эти трубы согласно инструкции в зависимости от входных сигналов, но ни трубы, ни человек не понимают иероглифы.



Заключение

Несмотря на длительные дискуссии исследователей из различных областей, спустя более чем 30 лет консенсус относительно аргумента Сёрла до сих пор не достигнут. Одни называют этот мысленный эксперимент самым известным контрпримером в истории, нанёсшим удар по функционализму, от которого тот так и не смог оправиться. Его также называют классическим аргументом для когнитивной науки. Другие (например, Д. Деннет) считают этот аргумент явно ошибочным и вводящим в заблуждение.



Литература

1. Джон Сёрль, статья «Разум, мозг и программы»
2. Джон Сёрль, книга «Разум, мозг, наука»
3. Пол М. Черчленд и Патриция Смит Черчленд («Может ли машина мыслить?», «В мире науки»)
4. Журнал «Популярная механика»
и сайт www.popmech.ru



**СПАСИБО
ЗА ВНИМАНИЕ!**