

Лекция №2

Модели сознания

(проблема сознания и ее философские решения)

- Классические модели сознания
- Современные концепции сознания
- Разумен ли искусственный интеллект?



Сознание

- совокупность представлений о мире и человеке в нем
- Это состояние, в котором человеку одновременно доступен и мир, и он сам. Мгновенно связывает, соотносит то, что человек увидел, услышал, и то, что он почувствовал, подумал, пережил. Продукт мышления
- На обыденном уровне сознание связано с понятиями душа, психика, мозг
- Разброс представлений о нем широк: от его отрицания до панпсихизма («Все есть сознание»)
- *«Наше сознание – основа нашего существования. Нет ничего более доступного нам напрямую, и все же оно до сих пор остается величайшей загадкой во Вселенной»*

Д. Чалмерс

Модели сознания в истории философии

Сознание является для человека не только «окном в мир», но и «окном в себя». Проблема сознания рассматривалась в различных контекстах: души, разума, психики

Платон (429/427-347 до н.э.): в отличие от смертного тела, душа вечна, нематериальна; после смерти своего носителя воплощается в новом теле. Она есть часть мировой души.

Объективный идеализм

- В промежутке между перевоплощениями души оказываются в мире эйдосов, вечных и совершенных идей. Наш мир является своего рода производной от мира эйдосов. Знание как припоминание (анамнезис).
- Душа состоит из трех частей: одна разумная (помещается в голове - это душа философа) и 2 части неразумные (благородная воля, живущая в груди - это душа война; неблагородная, расположенная в желудке - душа ремесленника/землепашца).

Модели сознания в истории философии

Р. Декарт (1596-1650): *«Я мыслю, следовательно, существую»*

- Переход от исследования души к исследованию сознания. Рационализм.
- Две субстанции – мыслящая (духовная) и протяженная (телесная), не зависят друг от друга. Рождение психофизической проблемы.
- Душа и тело соединяются в теле. Душа воздействует на тело через эпифиз (шишковидную железу). Тело воздействовать на душу не может. Мыслящее «Я» не нуждается для своего существования ни в каком месте и не зависит ни от какой материальной вещи, оно зависит только от мышления.
- Идеи ума возникают из трех источников: образованные самим человеком, приобретенные и врожденные (напр., идеи о субстанции, о боге, о фигуре, о прямой линии, целое больше части и т.д.). Источником врожденных идей является Бог.

Модели сознания в истории философии

Дж. Локк (1632-1704): *«В сознании нет ничего, чего бы не было в ощущениях».*

- Эмпиризм/сенсуализм (ощущение – единственный достоверный источник знания)
- Основа сознания - опыт, состоящий из единичных восприятий. Разум есть «*tabula rasa*» (чистая доска), на которую опыт наносит свои письмена. Сознание – это сумма идей.
- Знания и идеалы не даны нам в готовом виде, но являются результатом воспитания.
- Выступал против Декарта, который обосновывал свою теорию познания наличием у человека врожденных идей: если бы идеи были врожденными они были бы известны и взрослому, и ребенку, и нормальному человеку, и глупцу; все люди придерживались бы одних и тех же моральных и политических убеждений, имели бы одно представление о Боге, а этого нигде не наблюдается

Модели сознания в истории философии

- Ж. Ламетри** (1709-1751): отвергает дуализм Декарта и приписывает материи способность к ощущению и мышлению. Души как особой субстанции не существует; существует только единая материальная субстанция. Мышление обусловлено самой организацией мозговой ткани
- Человек – «перпендикулярно ползающая машина»: его ум имеет механическую материалистическую природу. Мысль – функция мозга
 - *«Душа – это лишенный содержания термин, за которым не кроется никакой идеи и которым здравый ум может пользоваться лишь для обозначения той части нашего организма, которая мыслит»*

Модели сознания в истории философии

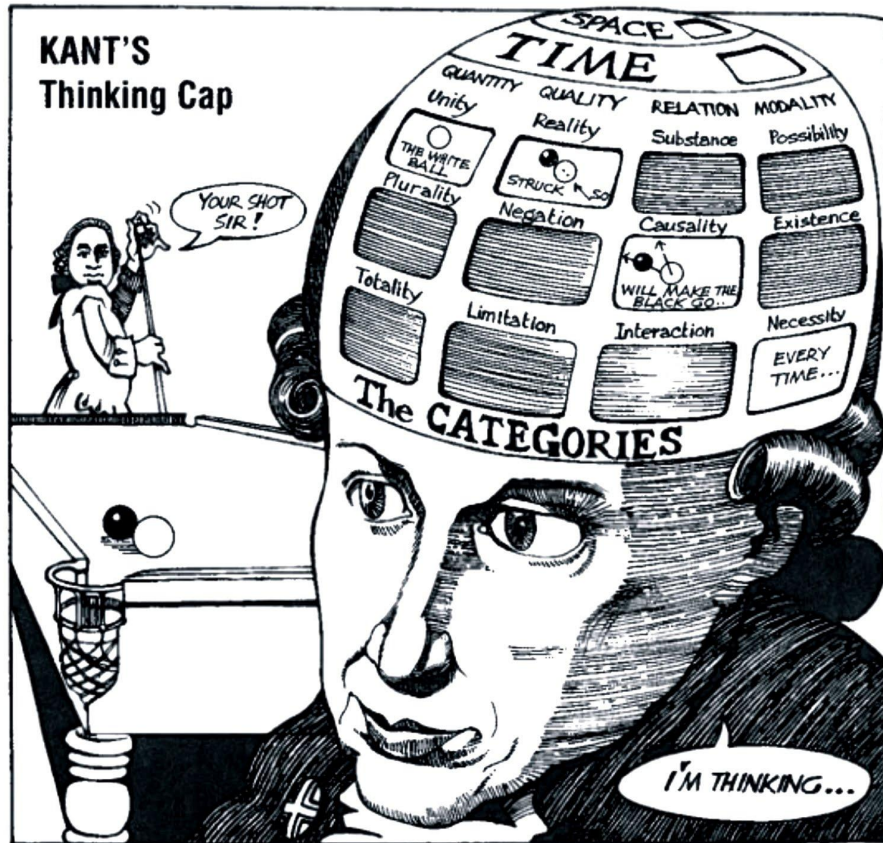
И.Кант (1724-1804). Трансцендентальный идеализм (с лат. *transcendens* — выходящий за пределы).

- Кант интерпретирует сознание как познание (в этом близок когнитивизму)
- Сознание и познание Кант рассматривает как трансцендентальные условия существования мира явлений (мира феноменов) – того мира, который мы воспринимаем, мыслим и познаем
- Кроме того, Кант допускает еще и существование некоторого мира трансцендентных вещей в себе (мир ноуменов) . Об этом мире мы не можем ничего знать, он вне нашего сознания. Нам доступны только явления, выступающие результатом переработки неизвестных вещей в себе нашей познавательной способностью, т.е. нашим сознанием.

- Кант выделяет три трансцендентальных способности сознания: чувственное восприятие, рассудок и разум. Именно благодаря им в сознании конституируется известный нам мир явлений. Предполагает наличие априорных (доопытных) форм познания
- До Канта: наши представления соотнобразуются с предметами. Центр познания – объект
- С Канта: предметы соотнобразуются с нашими представлениями. Центр познания – субъект, именно он определяет объект
- Процесс познания – не пассивное отражение, в нем присутствуют характеристики субъекта

Процесс познания по Канту

- вещь, воздействуя **на чувственность** человека, порождает **многообразие ощущений**;
- они упорядочиваются с помощью *априорных форм чувственности* (пространства и времени) **в предмет восприятия**;
- восприятие субъективно, чтобы оно стало общезначимым, надо подключить **рассудок** (мышление, оперирующее понятиями);
- рассудок **конструирует предмет** в соответствии с *априорными формами мышления* – категориями. Но рассудку недостает цели;
- руководит рассудком, ставит ему цели **разум**, выполняющий регулятивные функции в процессе познания.



- До Канта: наши представления соотносятся с предметами. Центр познания – объект
- С Канта: предметы соотносятся с нашими представлениями. Центр познания – субъект, именно он определяет объект
- Процесс познания – не пассивное отражение, в нем присутствуют характеристики

Модели сознания в истории философии

- **Диалектический материализм** – философское направление, выведенное из материалистических идей К. Маркса и Ф. Энгельса, развитое Лениным и другими философами-марксистами
- Материя – это объективная реальность, существующая вне и независимо от сознания человека и человечества, она первична относительно сознания.
- Сознание – неотъемлемое свойство высокоорганизованной материи. Носитель сознания – мозг; сознание есть функция мозга, заключающаяся в отражении мира



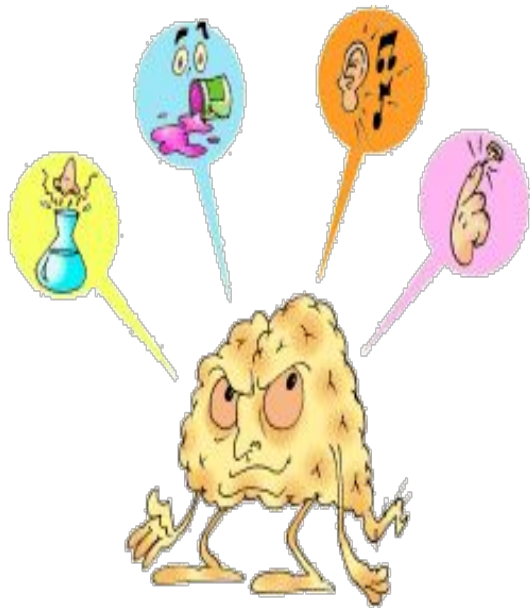
Каждый – эксперт в области своего сознания, но не знает какво быть другим...

- Каково быть летучей мышью?
- Каково быть братом Моцарта?
- Каково быть президентом Сибири?...

Характеристики сознания

- Субъектность и приватность
- Интенциональность (направленность на что-то)
- Идеальность
- Отсутствие пространственной протяженности
- Идеаторность
- Осведомленность (я знаю, что я в сознании)

Дискутируемая характеристика сознания: **КВАЛИТАТИВНОСТЬ**



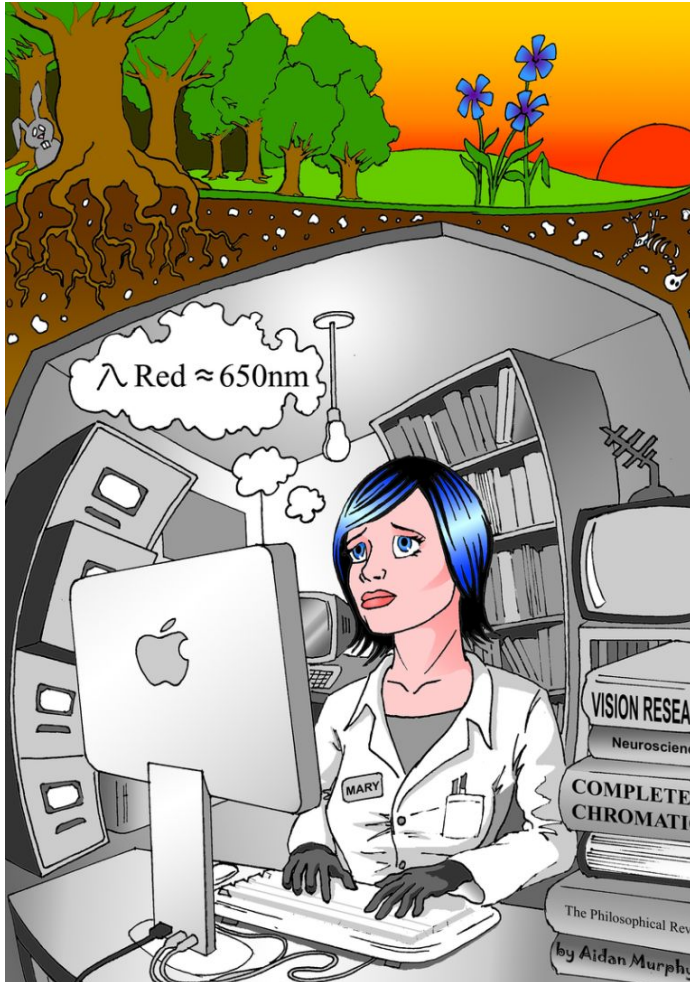
- **Квалиа** - термин, означающий субъективные (феноменальные) аспекты нашего восприятия мира
- Человека нельзя свести к материальным процессам. В остатке всегда будут непосредственные субъективные переживания: красота «красного», сладость «сладкого», боль «боли».
- Это ответ на вопрос «**Каково быть мной?**»

Квалиа

(с лат. «свойства, качества»)

- *Квалиа* – это данность наших чувств (воспринимаемые цвета, ощущаемые звуки, боль и т.д.). Это ментальное, субъективное, доступное только их носителю. Сознание – это совокупность квалиа
- Квалиа приватны
- Они невыводимы из физических свойств и несводимы к ним.
- Они доступны интроспекции (от лат. «смотреть внутрь»). Мы можем направить свой внутренний взор на эти свойства.
- Опыт квалиа называют квалитативным или феноменальным опытом.

«Ученая Мери» - мысленный эксперимент Фрэнка Джексона, косвенно подтверждающий теорию квалиа



- Проведя всю жизнь в черно-белой комнате, зная в теории все о цветовом восприятии, но никогда не видя цвета, она однажды выходит из комнаты и видит красную розу. Узнала она что-то новое о цветовом восприятии или нет? Да, узнала, и это означает, что ее исчерпывающее знание нуждается в каких-то дополнительных фактах - а именно, в квалиа.



Цвет как квалиа: это релятивное свойство, отражающее целый комплекс отношений (опасное, съедобное, полезное, спелое)



○ Природа цвета субъективна?



Еще о квалиа: эксперимент «Инвертированный спектр»: два человека описывают то, что они видят. Но у одного человека спектр инвертирован по отношению к спектру другого: где один видит зеленое – другой видит красное и т.д. Но открыть инверсию спектров невозможно, так как их речевое поведение вполне адекватно согласуются. Одни и те же вещи они называют зелеными, красными и т.д. Эта разница в опыте не может быть обнаружена на основании данных, доступных от третьего лица. Она м.б. познана только от первого лица.

Гносеологическая постановка проблемы сознания

«легкая и трудная проблемы сознания»

(Д.Чалмерс, 1995)

Лёгкая проблема сознания – то, что можно измерить, то что решается традиционными научными методами (то, что относится к способностям или функциям мозга)

Трудная проблема сознания : почему и каким образом мозг порождает сознание?

Онтологическая постановка проблемы

- *Физикализм*: всё, что существует, имеет физическую природу или основано/производно от физических свойств. отождествляет сознание с мозгом.
- *Антифизикализм*: существующее не привязано обязательным образом к физическим объектам или свойствам

Проблема соотношения тела и сознания

(Mind-body problem)

- Центральная проблема современной философии сознания; понимается как соотношение ментального и физического, сознания и тела с последующей интерпретацией этого соотношения
- Задача – уточнение онтологического статуса ментального или сознания.

Подходы к решению проблемы (1):

I. Редукционизм (сведение сложного к простому) *Монистические подходы*

- *Идеалистический монизм*. Существует только ментальная субстанция (напр., Джордж Беркли).
- *Материалистический монизм*. Существует только физическая субстанция (*физикализм* — взгляд, согласно которому ВСЁ во Вселенной, имеет физическую природу). Сознание как материальный объект есть продукт человеческого мозга.

Критика физикализма

- *"Философский зомби"* (Д. Чалмерс): эксперимент о несводимости субъективных аспектов сознания к функциональным фактам. Сознание есть *нечто большее*, чем то, что рисует нам физикалистская картина мира. Философский зомби - наш двойник. Но: *он не испытывает те субъективные ощущения, которые испытываем мы*. Это значит, что ощущения имеют особую природу, особое, субъективное

Подходы к решению проблемы(2):

II. Нередуционизм

Дуалистические подходы - сознание обладает особой природой, принципиально несводимой к физической материальной реальности. Примеры:

- **Субстанциальный дуализм Рене Декарта**: мир состоит из двух родов субстанций - материальных и духовных. Сознание и материя независимы друг от друга
- **Дуализм свойств**: существует лишь одна субстанция или одна реальность (материальная); сознание же связано с набором несводимых к материальной реальности свойств, которые эмерджентно (с англ. «возникающий, неожиданно появляющийся») порождаются на основе материальных систем (мозга) (*Томас Нагель*)

Современные модели сознания (1)

- Биологический натурализм. Сознание - результат совершенствования мозга в процессе эволюции природы. Оно неотделимо от тела и может быть описано как часть нашей биологической природы.

Джон Сёрл (1932). Сознание является реальным субъективным опытом, вызванным физическими процессами в мозге. Мозг порождает разум. Наше сознание оказывает влияние на наше поведение.
Ментальные состояния — высокоуровневые свойства физических систем.

Критик идеи искусственного интеллекта (автор мысленного эксперимента «Китайская комната»).

Современные модели сознания (2)

- Функционализм. Сознание - это не особое свойство материи, а функциональные отношения логических состояний, которые могут быть осуществлены на разных носителях – в теле человека, в компьютерной программе и др.

Дэниел Деннет (1942). Теория множественных проектов. Ни «Я», ни квалиа не существует. То, что мы воспринимаем как сознание — результат физиологических процессов, протекающих в мозге. В человеческом мозге нет единого центра, создающего сознание. Оно возникает в результате деятельности множества нейронов, образующих сложные функциональные сети. Эти сети и есть сознание. Так называемое «сознание» выполняет ту же функцию в мозге, что и компьютерная программа в процессоре. Наш мозг – это машина из миллиардов «роботов»-нейронов. Человеческое сознание не имеет принципиальных отличий от программ сложного робота.

Современные модели сознания (3)

- Дуализм свойств (натуралистический дуализм): сознание связано с набором нередуцируемых к материальной реальности свойств, которые эмерджентно порождаются на основе мозга

Дэвид Чалмерс (1966). Квалиа – фундаментальный элемент Вселенной, наряду со скоростью света, массой, гравитацией. Значит, теория сознания должна скорее основываться на фундаментальной физике, а не на биологии, и её главным элементом должны быть психофизические законы, описывающие связь квалиа с физическими качествами. При этом квалиа не являются частью материального мира. Поэтому созданная Чалмерсом теория сознания является дуалистической. Сознание имеется у всего во Вселенной (*панпсихизм*). Оно - нематериальное фундаментальное свойство Вселенной. ВСЁ имеет сознание (Д.Чалмерс).

Сознание – это...

- Всё то, что вы чувствуете и ощущаете..., пока не засыпаете, или не «теряете сознание» (Дж. Сёрл)
- 01
- Это совокупность элементов сенсорного опыта, субъективное состояние, сопровождаемое квалиа (в духе Д. Чалмерса)
- 02
- «Иллюзия» мозга, позволяющая каждому считать себя сознательным существом; иллюзия интенциональной установки поведения (по следам Д. Деннета)
- 03

Современные когнитивные исследования

- Ведущее направление в исследованиях сознания к. XX – нач. XXI вв. - *когнитивистика* или когнитивный подход (от лат. *cognitio* «познание»)
- Это междисциплинарный комплекс научных направлений, изучающих познание и высшие формы мыслительных процессов

Сильный и слабый искусственный интеллект (ИИ)

- ИИ – наука и технология создания интеллектуальных машин (в частности, компьютерных программ)
- Обладает ли сознанием машина, *полностью* воспроизводящая работу человеческого мозга?
- Развитие когнитивистики раскрывает возможности выявить, объяснить и описать процессы, происходящие в мозге человека, что позволит создать искусственный интеллект

- **Сильный ИИ** (понятие ввел Дж.Серл)– это программа, полностью воспроизводящая человеческий разум.
- **Слабый ИИ** – это инструмент изучения разума, способный моделировать некоторые ситуации, но не все процессы мозга.

Разумны ли роботы?

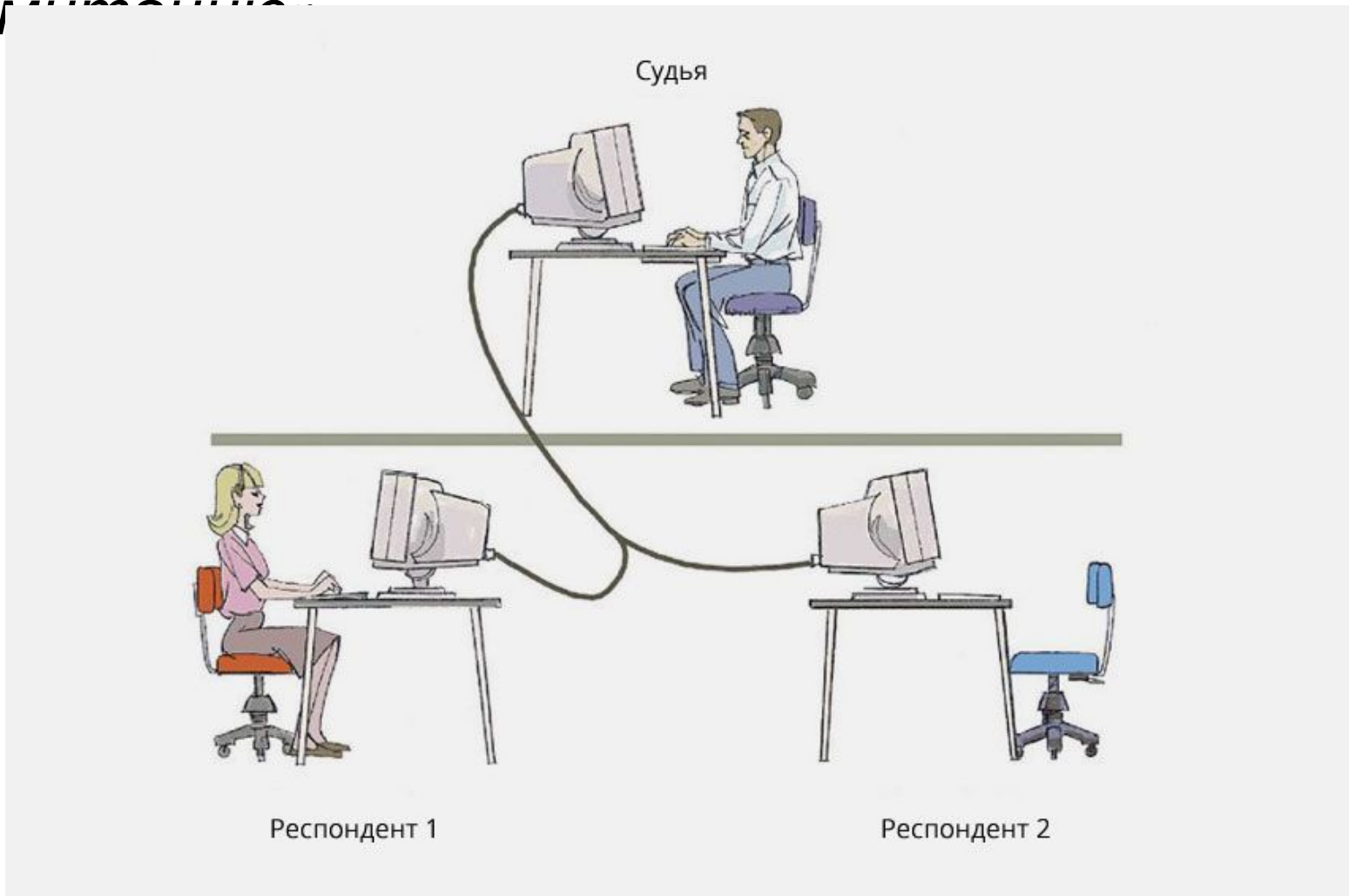
Робот София

Чат-бот Tay



Может ли машина обладать сознанием?

тест Алана Тьюринга «Игра в имитацию»



Критика теста Тьюринга

(«Китайская комната» Дж. Р. Сёрля)



В статье «Is the Brain's Mind a Computer Program?» («Является ли мозговое мышление компьютерной программой?»), опубликованной в 1990 году, Сёрль описывает, что при проведении теста Тьюринга не существует *никакой* возможности отличить действительно мыслительную деятельность программы от механического исполнения правильно подготовленным инструкциям.

- В случае с «Китайской комнатой», система в целом могла бы пройти тест Тьюринга, но при этом **никакого понимания языка внутри системы не происходит**, а значит тест, по мнению *Дж. Серля*, не является адекватной проверкой мыслительных способностей
- Что не так? Компьютерная программа манипулирует формальными - синтаксическими объектами. Человеческий разум оперирует семантическим (смысловым) содержанием.
- Мозг порождает разум. *Любая другая система, способная породить разум, должна обладать свойствами, эквивалентными свойствам мозга.* Об этом – мысленный эксперимент *Зенона Пылышина* о природе носителя сознания: допустим, исследователь начинает заменять нейроны в мозге познающего субъекта по одному на микрочипы с теми же свойствами и постепенно заменяет все нейроны. Будет ли получившийся в итоге «электронный мозг» носителем того же сознания, что и исходный мозг познающего субъекта? Сам субъект, вероятнее всего, подмены не заметит, следовательно, мы вынуждены будем признать, что сознание не зависит от того, на каком носителе реализуется

Этические проблемы ИИ

- Если в будущем машины смогут рассуждать, осознавать себя и иметь чувства, то что тогда делает человека человеком, а машину — машиной?
- Если в будущем машины смогут осознавать себя и иметь чувства, возможно ли будет их эксплуатировать или придется наделять их правами?
- Если в будущем машины смогут рассуждать, то как сложатся отношения людей и машин?