

# Искусственный интеллект

*"Человек благоразумный подстраивает себя под окружающий мир, тогда как безрассудный человек упорно подстраивает этот мир под самого себя. Так что весь прогресс опирается на людей безрассудных".*

*Бернард Шоу.*

# Вопросы

1. Искусственный интеллект: характеристика понятия. Этапы развития.
2. Основные направления исследований в области ИИ: состояние и тенденции.
3. Экспертные системы: структура и виды.

# Искусственный интеллект

Термин введен в 1956 г. Дж. Маккарти на Дартмусской конференции.

1950 г. Алан Тьюринг «Computing machinery and intelligence». Это исторически первая дата возникновения исследований по искусственному интеллекту (хотя сам термин в работе не использовался).

# Искусственный интеллект

Имеет два значения:

- Теория создания аппаратных и программных средств, способных осуществлять интеллектуальную деятельность.
- Аппаратные и программные средства, способные осуществлять интеллектуальную деятельность, а также выполняемая ими деятельность.

Трудность: не существует однозначного определения и понимания интеллекта естественного.

**ИИ** – набор программных и аппаратных средств, использование которых должно было бы привести к тем же результатам, к которым при решении данного класса задач приводит интеллектуальная деятельность человека.

# Когнитивная наука: теория и практика



# От интеллекта естественного к искусственному



*«Тезис Лавлейс»: машина  
никогда не сможет сделать  
то, чего не может сделать  
человек.*

*В действительности человек  
умеет делать гораздо больше,  
чем знает, как делать.*



# Парадокс «китайской комнаты»

## КИТАЙСКАЯ КОМНАТА

Аксиома 1. Компьютерные программы — это формальные (синтаксические) объекты.

Аксиома 2. Человеческий разум обладает смысловым содержанием (семантикой).

Аксиома 3. Синтаксис сам по себе не является сущностью семантики и его не достаточно для семантики.

Заключение 1. Программы не являются сущностью разума и их наличия не достаточно для существования разума.

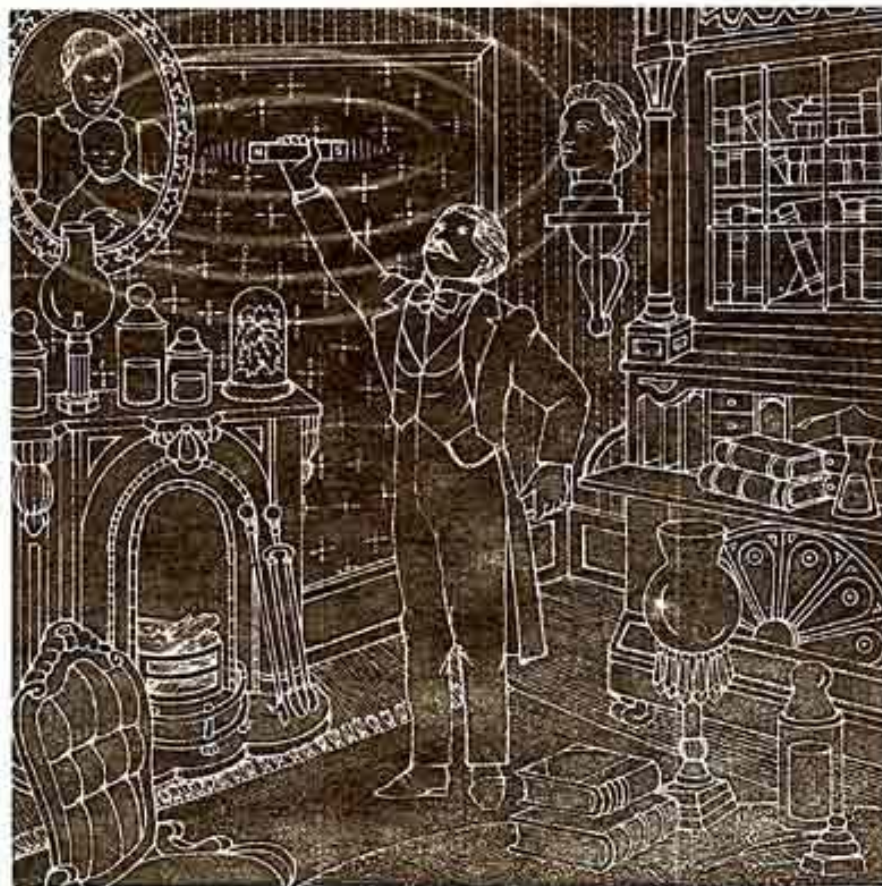
## СВЕТЯЩАЯСЯ КОМНАТА

Аксиома 1. Электричество и магнетизм — это физические силы.

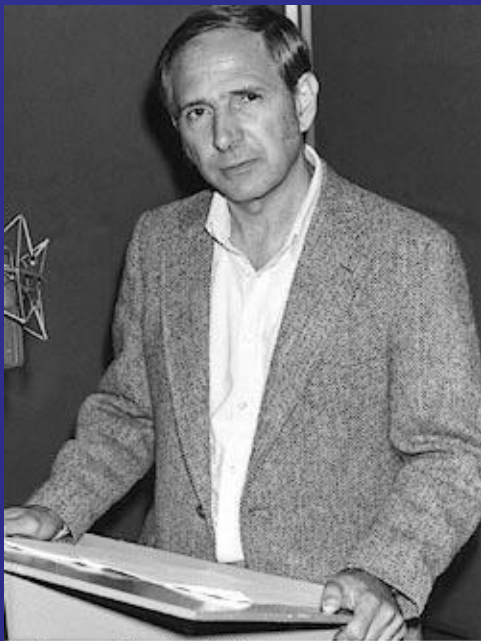
Аксиома 2. Существенное свойство света — это свечение.

Аксиома 3. Силы сами по себе не являются сущностью эффекта свечения и не достаточны для его наличия.

Заключение 1. Электричество и магнетизм не являются сущностью света и не достаточны для его наличия.



# Джон Сирл (Беркли, 1984): парадокс «китайской комнаты»



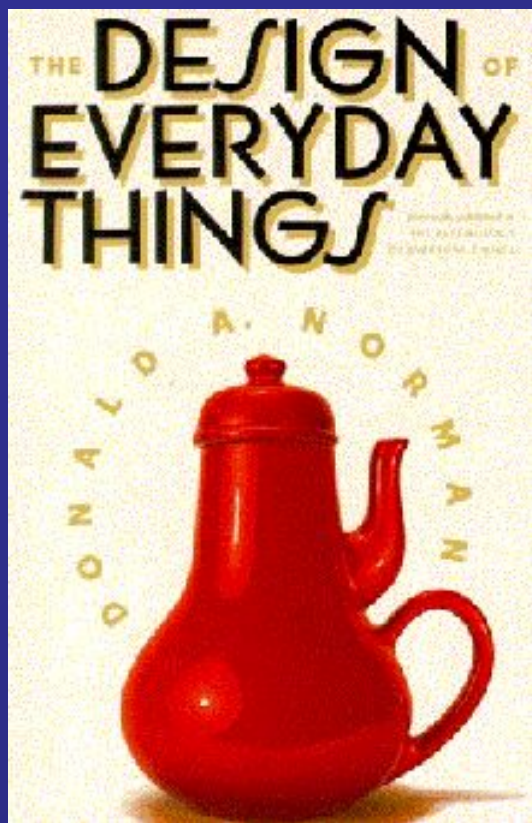
Searle recording the Reith Lectures for BBC in London, summer of 1984, where he first presented the "Chinese Room" argument. © BBC photo



«Иллюзия понимания» снаружи  
при полном непонимании внутри.

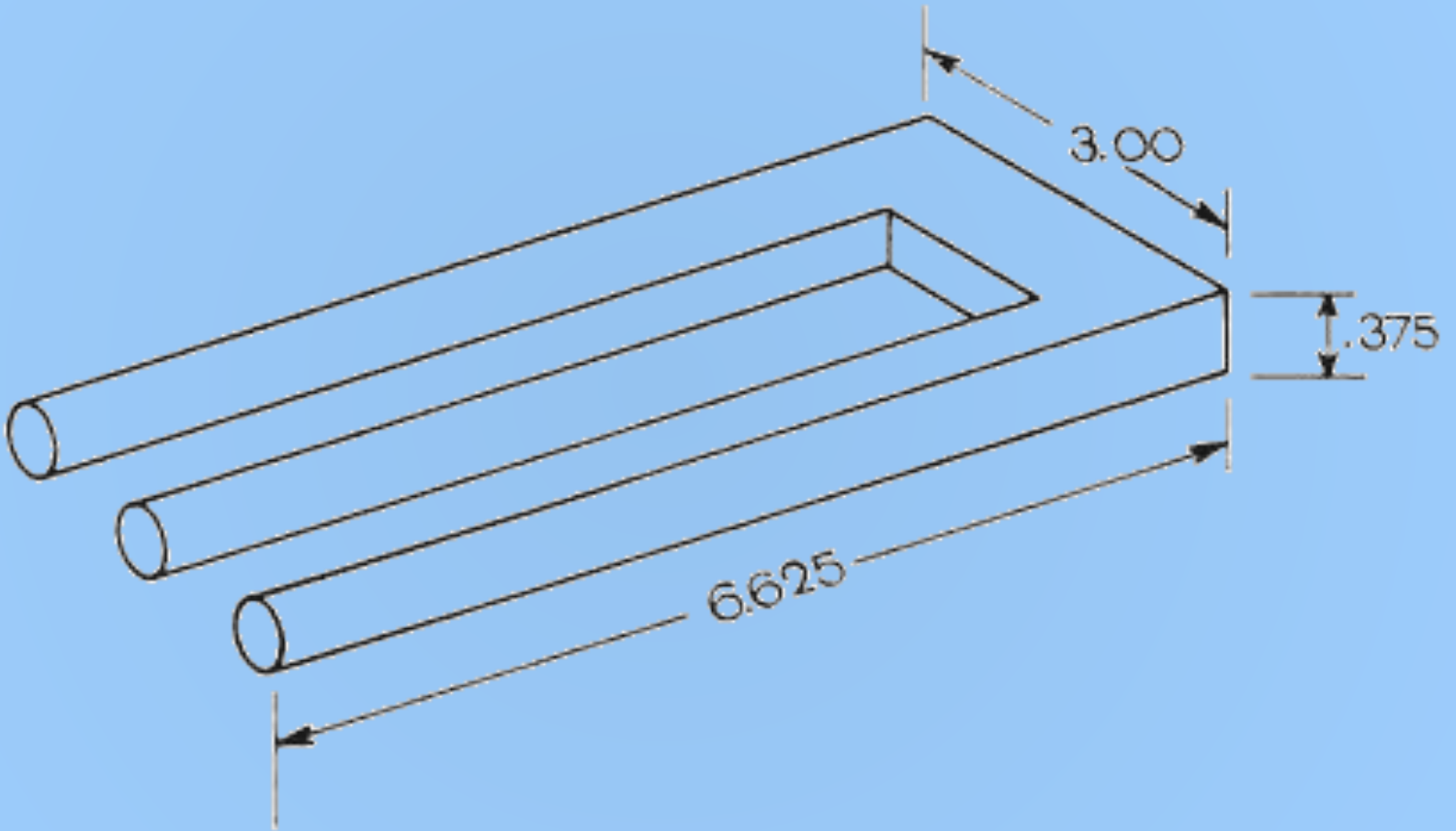


# Когнитивная наука и рукотворный мир: «психопатология обыденных вещей»

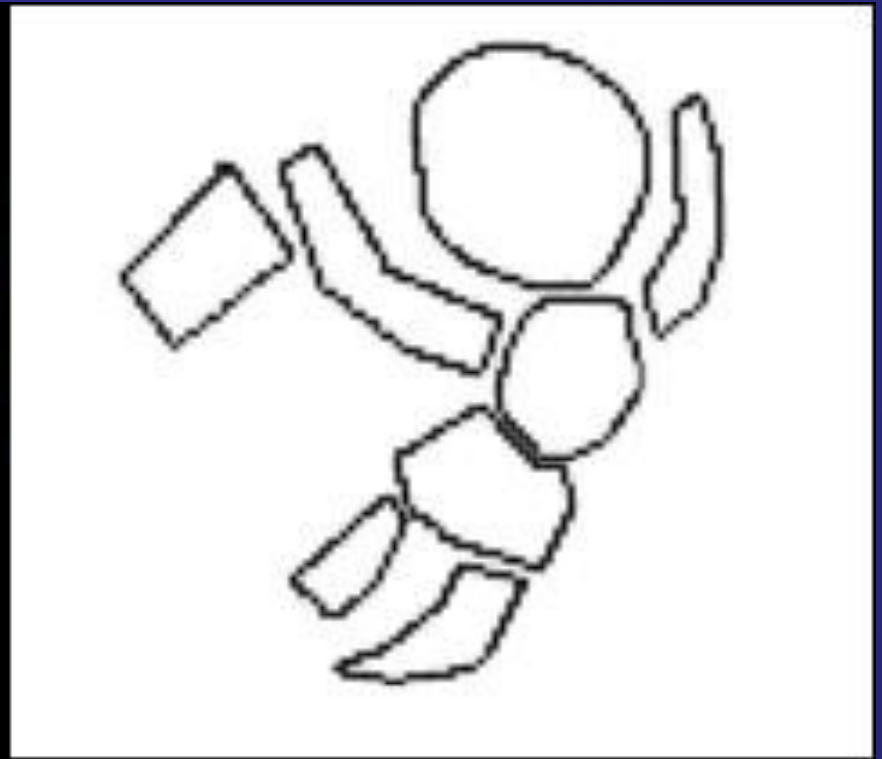


Предметы из «Каталога  
невозможных объектов»  
Жака Карельмана

# Невозможные объекты



# Тайны природы





Существуют гигантские  
базы знаний, мощные  
экспертные системы,  
содержащие тысячи  
правил решения задач

Пределы шахматных  
возможностей  
компьютера,  
«обыгравшего» А.  
Карпова, не ограничены?



# История:



«Романтический период" ИИ.

Серьезные научные исследования.



Практические задачи.

1950-ые гг.

Г. Саймон, А. Ньюэлл, Шоу –  
ЛОГИК-ТЕОРЕТИК, ОБЩИЙ  
РЕШАТЕЛЬ ПРОБЛЕМ (GPS –  
General Problem Solver)

Алан Ньюэлл, Герберт Саймон  
«Решение задач человеком» (1972)



Представители компьютерного  
пессимизма:

Х. Дрейфус «Чего не могут  
вычислительны машины». 1978

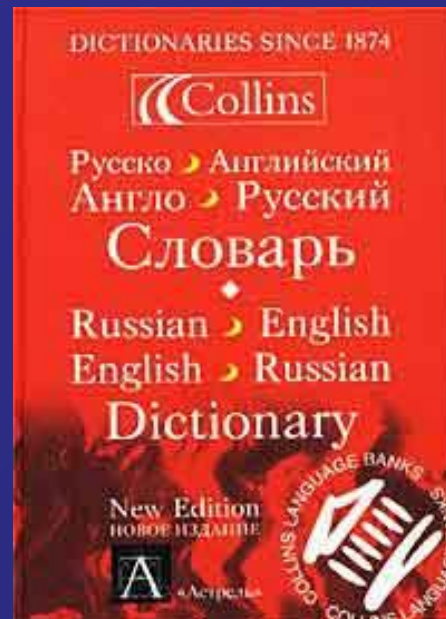
Х. Дрейфус, С. Дрейфус «Mind  
over mashine», 1986

# Обработка естественного языка: системы машинного перевода (1950-е)

Spirit is strong, but  
flesh is weak.

*Дух крепок, а  
плоть немощна.*

Out of sight,  
out of mind?



Vodka is good, but  
meat is rotten.

*Водка отменная, но  
мясо протухло.*

Invisible  
idiot!

Проверка качества перевода:  
обратный перевод!

**1960-ые гг.**

**Метод резолюций Робинсона.**

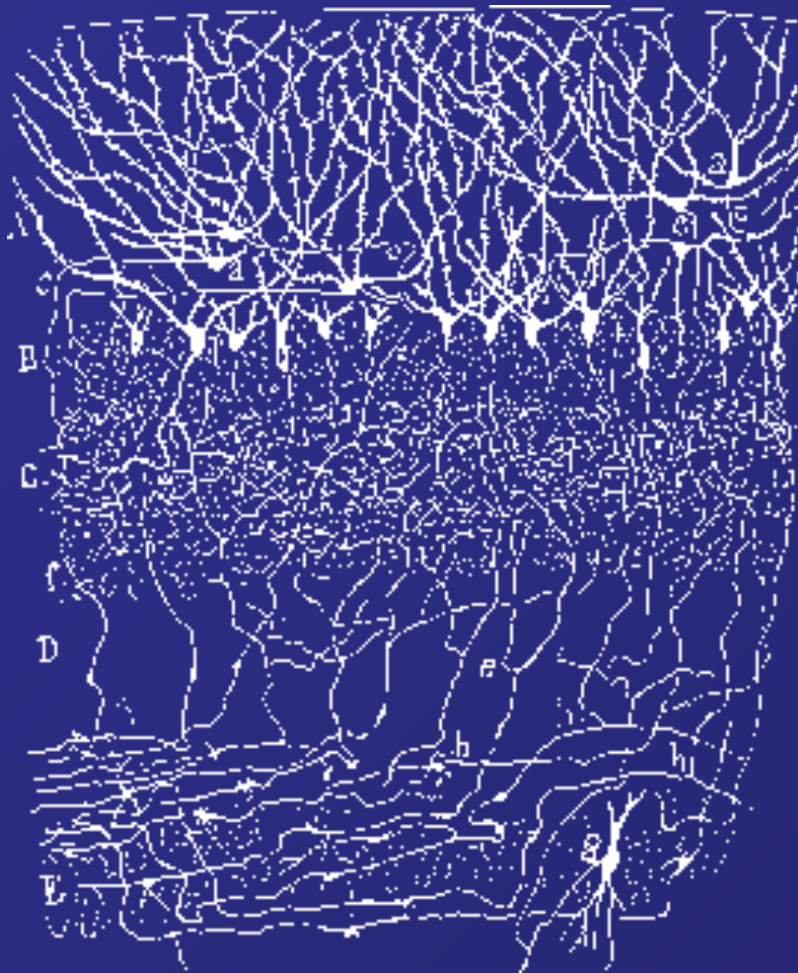
**Цель исследований: разработка программ,  
способных решать "человеческие задачи".**



*Познание – есть тип  
вычисления*

3. Пилишин – теоретическая концепция  
ИИ как машинного эквивалента  
человеческого мышления.  
«Вычисление и познание», 1984

# Л. Сутро, У. Килмер, Дж. Олбус – нейроподобные сети.



# Мозг человека: преимущества перед компьютером

- $10^{11}$  нейронов,  $10^{14}$ - $10^{15}$  связей между нейронами.
- Частота импульсации --  $10^2$  Гц (современные персональные компьютеры -- до  $10^9$  Гц).

***NB!*** Медлительность и ненадежность отдельных нейронов компенсируется их количеством.

- Параллельная переработка информации (в компьютерах -- преимущественно последовательная).
- «Переход количества в качество»: богатство поведения.
- Нельзя сказать, что мозг исходно «готов к использованию»: велика роль обучения.

# Нейронные сети: рождение идеи (1943)

Уоррен Маккаллох



Уолтер Питтс



# Нейросетевой подход: ОСНОВНЫЕ ПОЛОЖЕНИЯ

- Процессы познания -- результат взаимодействия большого числа простых перерабатывающих элементов, связанных друг с другом и организованных в слои («модули»).  
«Переработка информации» -- определенный ответ элемента на воздействия извне.
- Знания, управляющие процессом переработки, хранятся в форме *весовых коэффициентов связей* между элементами сети. Главное -- не элементы, а связи между ними.
- Обучение -- процесс изменения весовых коэффициентов связей между элементами сети (приспособления их к решению определенной задачи).



**Начало 1980 –ых гг.**

**Дж. Маккарти – идея немонотонной  
логики с учетом изменения  
ситуаций.**



1969 -- Марвин Мински,  
Сеймур Пейперт  
«Перцептроны»:  
приговор нейронным сетям?



Критика перцептронов: математическое обоснование их неэффективности в решении задач распознавания образов (в ходе поэлементного анализа связанных и несвязанных изображений теряется информация о связности, которую невозможно задать линейно).



М. Мински, Р. Шенк – фреймовые  
системы.

Фрейм – целостная структура,  
содержащая информацию об  
ОСНОВНЫХ СВОЙСТВАХ ПОНЯТИЯ.

# Концепция компьютерного зрения

Дэвид Марр «Зрительное восприятие: вычислительное исследование отображения и обработки зрительной информации у человека» (1982 г.).

Конструирование зрительной информации представляется различными функциями.

Получена система уравнений, описывающая зрительное восприятие.

Уравнения могут быть ранжированы и использованы при конструировании зрительных процессов компьютера.

# Этапы развития систем ИИ:

Первый этап – игры. Математические головоломки. 1956 – конец 1960-ых гг.

Второй этап – создание интегральных роботов. Конец 1960-ых – сер. 1970-ых гг.

Третий этап – сложный план поведения, общение на естественном языке. Создание человеко-машинных систем конец 1970-ых гг.

# Искусственный интеллект

- решение задач
  - экспертные системы и системы поддержки принятия решения: инженерия знаний
  - системы распознавания образов, движения и т.д.
  - системы обработки текста и машинного перевода
  - игровые программы (шахматы и т. д.): Гарри Каспаров против *Deep Blue*
  - обучаемые и обучающие системы
  - робототехника и автономные агенты...
- проблема взаимодействия с пользователем на естественном языке:
    - ~ справочные системы
    - ~ поисковые системы
    - ~ компьютерные игры
    - ~ компьютерные «психотерапевты»

# Экспертные системы и нейронные сети



Их использование сочетается с технологией традиционного программирования.

Преимущество: динамическая модификация приложений пользователем.

# Машинный интеллект и робототехника

Роботы первого поколения  
(промышленные роботы).

Роботы второго поколения  
(система глаз—рука).

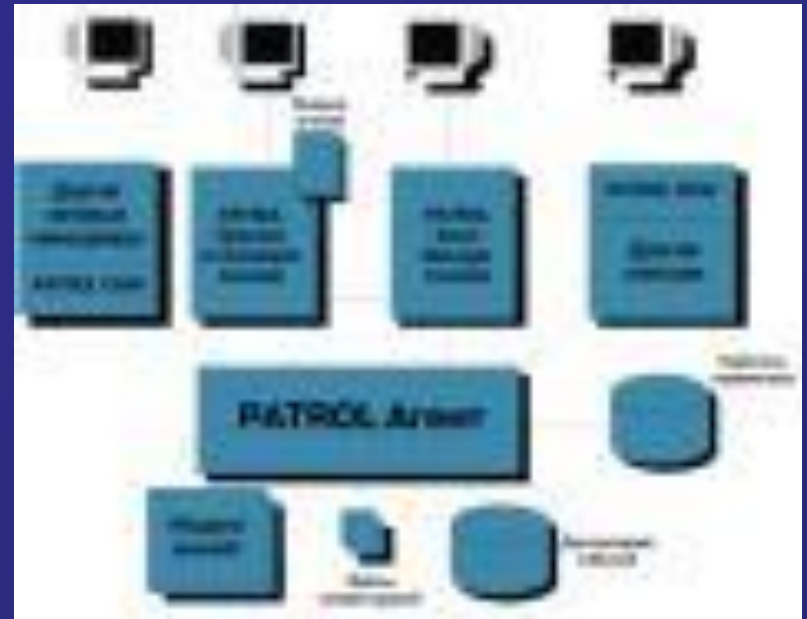
Роботы третьего поколения (с  
искусственным интеллектом;  
на стадии разработки)





# Автономные агенты

Генри Либерман  
(Лаборатория MIT) –  
автоматическое  
генерирование технической  
документации.



SAP разработала новую технологию интеллектуальных программных агентов, в задачу которых входит определение проблем в системах управления производственными цепочками.

# Генетическое программирование

Genetic programming –  
использование  
метафоры генной  
инженерии для описания  
различных алгоритмов.

Лидер – Стэнфордский  
университет, Джон Коз.



# Успех систем ИИ:

- Специализация
- Языки традиционного программирования
- Интегрированность
- Открытость и переносимость
- Архитектура клиент-сервер

# Направления исследований в области ИИ

Основа деления – две точки зрения на вопрос о том, как строить ИИ:

- важен результат;
- необходимо изучение механизмов человеческого мышления

# Искусственный интеллект

```
graph TD; A[Искусственный интеллект] --> B[Моделирование результатов интеллектуальной деятельности. Машинный интеллект]; A --> C[Моделирование биологических систем. Искусственный разум]; B --> D[Моделирование механизмов. Нейроподобные сети]; C --> E[Моделирование информационных процессов. Эвристическое программирование]; D --> F[Комплексный подход. Эвристическое моделирование]; E --> F; C --> F;
```

**Моделирование результатов интеллектуальной деятельности. Машинный интеллект**

**Моделирование биологических систем. Искусственный разум**

**Моделирование механизмов. Нейроподобные сети**

**Моделирование информационных процессов. Эвристическое программирование**

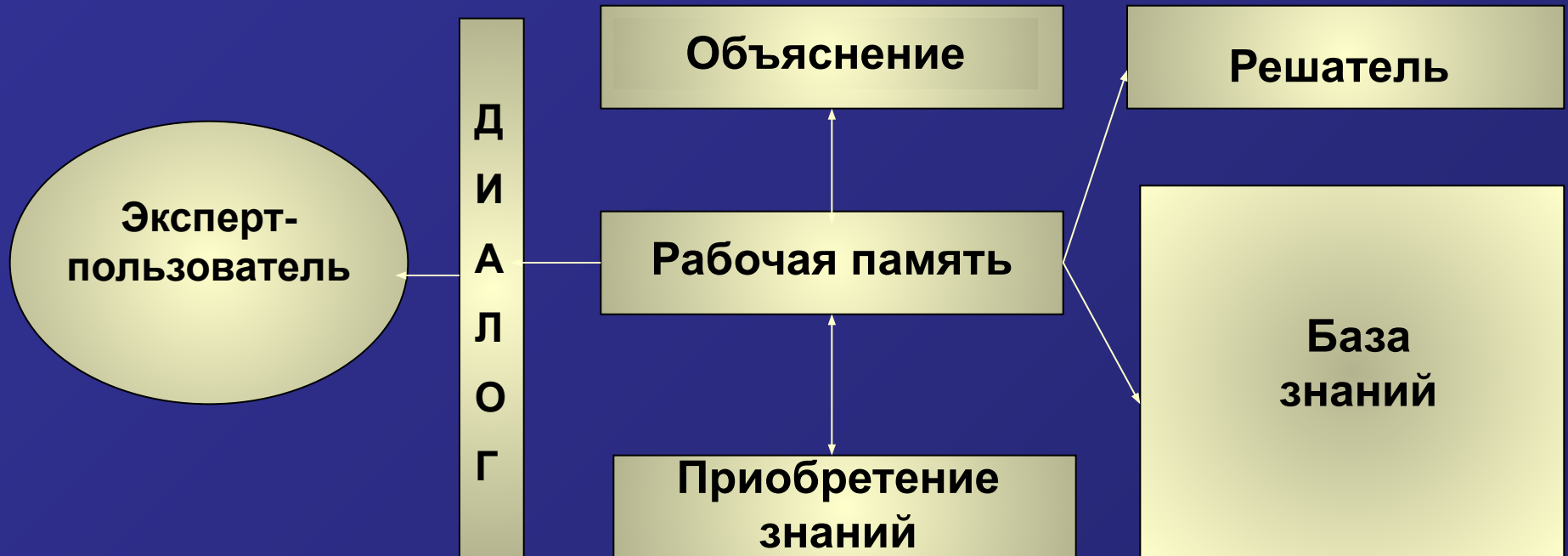
**Комплексный подход. Эвристическое моделирование**

# Экспертные системы

Относятся к системам ИИ общего назначения – системам, которые не только исполняют заданные процедуры, но на основе метапроцедур поиска генерируют и используют процедуры решения новых задач.

Задачи: исследование и разработка программ (устройств), использующих знания и процедуры вывода для решения задач, являющихся трудными для людей экспертов.

# Структура экспертной системы



## Сущность ЭС:

- В базе знаний записан перечень возможных проблемных ситуаций, в соответствие с каждой из которых поставлен какой-либо ответ, совет или непосредственное действие.
- Программная оболочка системы, получив информацию о проблемной ситуации, определяет ее соответствие одному или несколькими "клише" из записанных в базе знаний (или констатирует неизвестную ситуацию, если подходящего шаблона в базе нет).
- Информация "выведывается" у пользователя путем задавания ряда вопросов. Содержание последующих вопросов зависит от ответов на предыдущие.



# Объяснение в ЭС

**Цель:** обосновать, аргументировать ответ в максимально естественной форме.

## Что объяснять?

- как получено решение;
- как использована некоторая информация (факты, правила);
- почему не использована некоторая информация (факты, правила);
- что использовано в целом при решении задачи (факты, правила).

# Знания в ЭС

- Знания о предметной области
- Общие знания
- Знания как решать задачу
- Знания о том, как взаимодействовать с пользователем

Степень уверенности в достоверности факта или точности правила – **коэффициент уверенности.**

# Представление знаний в ЭС

**ЕСЛИ** (*условие*) – **ТО** (*действие*)

**ЕСЛИ** *пациент был по профессии изолирофщиком до 1988 г.,* **ТО** *пациент непосредственно работал с асбестом.*

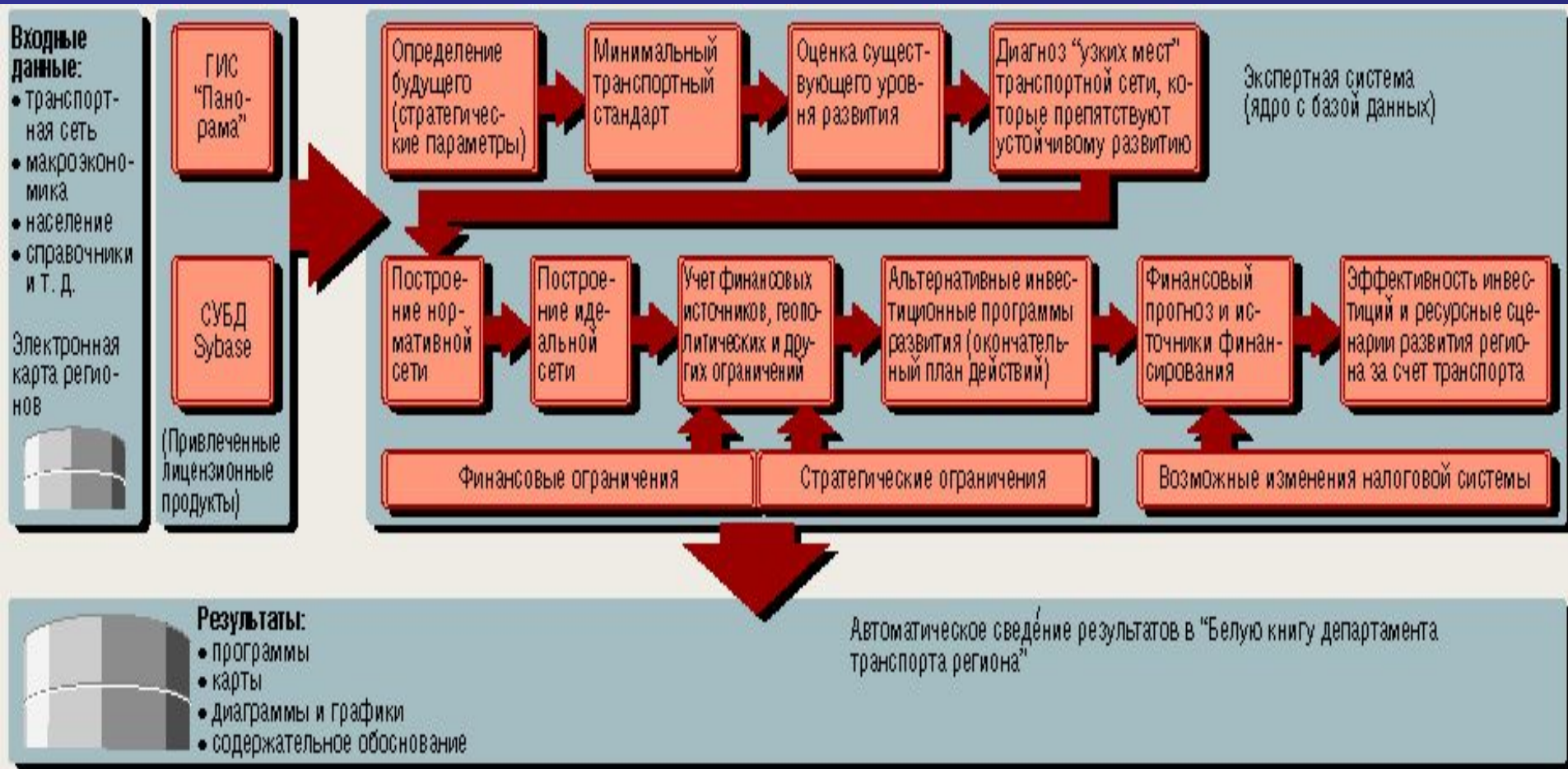
**ЕСЛИ** *пациент непосредственно работал с асбестом и при этом находился в закрытом помещении,* **ТО** *пациент получил большую дозу асбестовой пыли.*

# Разработка ЭС

- *Эксперт.*
- *Программист-специалист по разработке инструментальных средств.*
- *Инженер по знаниям.*
- *Пользователь.*

	<b>Предмет</b>	<b>Форма знания</b>	<b>Методы</b>	<b>Целевое назначение</b>
<b>Научное познание</b>	Природа, ее свойства, связи	Гипотеза, теория, закон, научный факт	Теоретическое исследование, наблюдение, эксперимент	Познание окружающего мира
<b>Техническое познание</b>	Артефакты	Гипотеза, теория, эмпирич. факты	Теорет. исследование, наблюдение, испытание	Познание свойств и связей артефактов
<b>Инженерное мышление</b>	Процесс создания артефактов	Правила конструирования, управления	Конструирование, системный анализ, расчет	Создание эффективных средств деятельности
<b>Обыденное познание</b>	Окружающий мир, быт	Здравый смысл, личный опыт	Обыденное наблюдение, освоение опыта	Достижение бытового благополучия

# Схема работы ЭС



# ЭС в исторических исследованиях

1. ГИДРОНИМИКОН (Ю.Е.Храмов) – анализ происхождения гидронимов Восточной славии. ПО – ЛОТТА.
2. Восстановление истории семей (Ж. Карвалью). ПО – PROLOG. Обладает открытой структурой.
3. АМСОР (Л.И.Бородкин) – решение задач аграрной типологии Евр. России конца XIX – нач. XX вв.



# ЭС в исторических исследованиях

4. SNARK ( -- решение проблем древней и средневековой истории.
5. Атрибуция литературных и исторических текстов (В.Кнопперт). ПО – KLS.HT.