

# Интеллектуальные системы

Лекция 1



# История искусственного интеллекта



# Основные определения

- ▣ **Определение 1.** Система с интеллектуальной поддержкой – это система, способная самостоятельно принимать решения.
- ▣ **Определение 2.** Интеллектуальная система – это информационно-вычислительная система с интеллектуальной поддержкой при решении задач без участия оператора.
- ▣ **Определение 3.** Интеллектуализированная система – это информационно-вычислительная система с интеллектуальной поддержкой при решении задач с участием оператора.

# Философские основания ИИ

Может ли машина мыслить?

- Гипотеза слабого искусственного интеллекта
- Гипотеза сильного искусственного интеллекта

# История ИИ

- Появление предпосылок ИИ (до 1956)
- Рождение ИИ (1956)
- Период раннего энтузиазма (1952-1969)
- Зима искусственного интеллекта (1966-1973)
- Разработка систем, основанных на знаниях экспертов (1969-1979)
- Второй бум искусственного интеллекта(1980-1987)
- Возвращение к нейронным сетям (1988-1995)
- Появление интеллектуальных агентов (1995)

# Появление предпосылок ИИ (1943-1955)

## ▣ Изобретение ЭВМ:

- 1940 – Алан Тьюринг (Англия)
- 1941 – Конрад Цузе (Германия)
- 1940-1942 - Джон Атанасов, Эдвард Бэри (США)

## ▣ Тест Тьюринга (1950)

## ▣ Изобретение электроэнцефалографа (Ганс Бергер 1929 г.)

## ▣ Первая модель искусственных нейронов (Уоррен Маккалок, Уолтер Питтс 1943 г.)

# Рождение ИИ (1956)

## Дартмутский семинар – лето 1956 года

1. Джон Маккарти, *Дартмутский колледж*
2. Марвин Минский, *Гарвардский университет*
3. Клод Шеннон, *Bell Laboratories*
4. Натаниэль Рочестер, *IBM*
5. Артур Самюэль, *IBM*
6. Аллен Ньюэлл, *Университет Карнеги*
7. Герберт Саймон, *Университет Карнеги*
8. Тренчард Мур, *Принстонский университет*
9. Рей Соломонов, *Массачусетский технологический институт*
10. Оливер Селфридж, *Массачусетский технологический институт*

# Рождение ИИ (1956)

В **1954** г. в МГУ начал свою работу семинар "Автоматы и мышление" под руководством академика Ляпунова А. А. В этом семинаре принимали участие физиологи, лингвисты, психологи, математики.



# Период раннего энтузиазма (1956-1969)

## Гербертом Саймоном

- 1956 г. – компьютерная программа **Logical Theorist**, которая смогла автоматически доказать 38 законов из книги Рассела и Уайтхеда «Принципы математики»
- 1956 г. – компьютерная программа **General Problem Solver (GPS)**. Она могла не только вести доказательства, но и играть в шахматы и ханойскую башню.
- 1958 г. - «в течение 10 лет цифровой компьютер будет мировым чемпионом по шахматам».
- 1965 г. - «Машины смогут через 20 лет делать любую работу, которую делает человек»

## Марвин Минский

- 1967 г. - «Через поколение проблема создания искусственного интеллекта будет в основной своей части разрешена».
- 1969 г. - «От 3 до 8 лет надо, чтобы у нас была машина, у которой уровень интеллекта будет соответствовать среднему человеческому».

# Столкновение с реальностью (1966-1973)

- Программы не содержали (или почти не содержали) знания о предметной области; их временные успехи достигались за счет простых синтаксических манипуляций.
- Не учитывалась вычислительная сложность решаемых задач (Отчёт Лайтхилла – 1973 г.)
- Фундаментальные ограничения базовых структур (М. Минский – перцептрон может обучиться лишь тому, что он позволяет описать)

- ▣ **Разработка систем, основанных на знаниях экспертов (1969-1979)**
- ▣ **Второй бум искусственного интеллекта(1980-1987)**

**Экспертные системы** – это направление исследований в области искусственного интеллекта по созданию вычислительных систем, умеющих принимать решения, схожие с решениями экспертов в заданной предметной области.

# Преимущества экспертных систем перед человеком-экспертом

## 1. Постоянство

Человеческая компетенция ослабевает со временем. Перерыв в деятельности человека-эксперта может серьёзно отразиться на его профессиональных качествах.

## 2. Лёгкость передачи

Передача знаний от одного человека другому – долгий и дорогой процесс. Передача искусственной информации – это простой процесс копирования программы или файла данных.

## 3. Устойчивость и воспроизводимость результатов

Экспертные системы устойчивы к «помехам». Человек же легко поддается влиянию внешних факторов, которые непосредственно не связаны с решаемой задачей. Эксперт-человек может принимать в тождественных ситуациях разные решения из-за эмоциональных факторов. Результаты экспертной системы – стабильны.

## 4. Стоимость

Эксперты, особенно высококвалифицированные обходятся очень дорого. Экспертные системы, наоборот, сравнительно недороги. Их разработка дорога, но они дешёвы в эксплуатации.

# Недостатки экспертных систем

- Передача экспертным системам «глубоких» знаний о предметной области является большой проблемой. Как правило, это является следствием сложности формализации эвристических знаний экспертов.
- Экспертные системы неспособны предоставить осмысленные объяснения своих рассуждений, как это делает человек. Как правило, экспертные системы всего лишь описывают последовательность шагов, предпринятых в процессе поиска решения.
- Отладка и тестирование любой компьютерной программы является достаточно трудоемким делом, но проверять экспертные системы особенно тяжело.
- Экспертные системы неспособны к самообучению. Для того, чтобы поддерживать экспертные системы в актуальном состоянии необходимо постоянное вмешательство в базу знаний инженеров по знаниям. Экспертные системы, лишённые поддержки со стороны разработчиков, быстро теряют свою востребованность.

# Возвращение к нейронным сетям (1988-1995)

- ▣ 1986 г. – алгоритм обратного распределения ошибки
- ▣ Нейронные сети Хопфилда и Хэмминга

# Появление интеллектуальных агентов (1995)

- В искусственном интеллекте, под термином **интеллектуальный агент** понимаются разумные сущности, наблюдающие за окружающей средой и действующие в ней, при этом их поведение рационально в том смысле, что они способны к пониманию и их действия всегда направлены на достижение какой-либо цели. Такой агент может быть как роботом, так и встроенной программной системой.

# Проблемная область ИИ

- Представление знаний
- Манипулирование знаниями
- Общение
- Восприятие
- Обучение
- Поведение



# Проблемная область ИИ

## ▣ Представление знаний

- *Формализация знаний*
- *Представление знаний в памяти ИИС*

## ▣ Манипулирование знаниями

- *Оперирование знаниями*
- *Пополнение знаний*
- *Классификация знаний*
- *Обобщение знаний*
- *Формирование новых абстрактных понятий*
- *Рассуждения, имитирующие особенности рассуждений человека*

## ▣ Общение

## ▣ Восприятие

## ▣ Обучение

## ▣ Поведение

# Интеллектуальные агенты



# Основные определения

*Агентом* является все, что может рассматриваться как воспринимающее свою *среду* с помощью *датчиков* и воздействующее на эту среду с помощью *исполнительных механизмов*.

*Последовательностью актов восприятия* агента называется полная история всего, что было когда-либо воспринято агентом.

*Функция агента* - это отображение любой конкретной последовательности актов восприятия на некоторое действие.

*Программа агента* – это конкретная реализация, действующая в рамках архитектуры агента.

# Концепция рациональности

*Рациональным агентом* является такой агент, который выполняет *правильные* действия.

- ▣ *Разработка показателей производительности*
- ▣ *Отказ от всезнания*
- ▣ *Обучаемость*
- ▣ *Автономность*

# Концепция рациональности

**Показатели производительности** – критерии оценки успешного поведения агента.

*Необходимо разрабатывать показатели производительности в соответствии с тем, чего действительно необходимо добиться в данной среде, а не в соответствии с тем, как, по мнению проектировщика, должен вести себя агент.*

# Концепция рациональности

## Отказ от всезнания

*Рациональность* — это максимизация ожидаемой производительности.

*Совершенство* — это максимизация фактической производительности.

*От агента нельзя требовать, чтобы он выполнял действия, которые оказываются наилучшими после их совершения.*

# Концепция рациональности

## **Обучаемость**

Начальная конфигурация агента может отражать некоторые предварительные знания о среде, но по мере приобретения агентом опыта эти знания могут модифицироваться и пополняться.

## **Автономность**

Рациональный агент должен обучаться всему, что может освоить, для компенсации неполных или неправильных априорных знаний.

После достаточного опыта существования в своей среде поведение рационального агента может по сути стать независимым от его априорных знаний.

# Определение проблемной среды

**Проблемная среда** объединяет в себе агента, показатели производительности, среду, исполнительные механизмы и датчики агента.

Тип агента	Показатели производительности	Среда	Исполнительные механизмы	Датчики
Водитель такси	Безопасная, быстрая, комфортная езда в рамках правил дорожного движения, максимизация прибыли	Дороги, другие транспортные средства, пешеходы, клиенты	Рулевое управление, акселератор, тормоз, световые сигналы, клаксон, дисплей	Видеокамеры, ультразвуковой дальномер, спидометр, глобальная система навигации и определения положения, одометр, акселерометр, датчики двигателя, клавиатура



# Свойства проблемной среды

## Полностью наблюдаемая или частично наблюдаемая среда

Проблемная среда является полностью наблюдаемой, если датчики выявляют все данные, которые являются релевантными для выбора агентом действия; релевантность, в свою очередь, зависит от показателей производительности.

Среда может оказаться частично наблюдаемой из-за создающих шум и неточных датчиков или из-за того, что отдельные характеристики ее состояния просто отсутствуют в информации, полученной от датчиков.

# Свойства проблемной среды

## Детерминированная или стохастическая среда

Если следующее состояние среды полностью определяется текущим состоянием и действием, выполненным агентом, то такая среда называется детерминированной; в противном случае она является стохастической.

Необходимо классифицировать среду как детерминированную или стохастическую с точки зрения агента, а не разработчика.

# Свойства проблемной среды

## Эпизодическая или последовательная среда

В эпизодической проблемной среде опыт агента состоит из неразрывных эпизодов. Каждый эпизод включает в себя восприятие среды агентом, а затем выполнение одного действия. При этом следующий эпизод не зависит от действий, предпринятых в предыдущих эпизодах.

В последовательных вариантах среды текущее решение может повлиять на все будущие решения.

# Свойства проблемной среды

## Статическая или динамическая среда

Если среда может измениться в ходе того, как агент выбирает очередное действие, то такая среда называется динамической для данного агента; в противном случае она является статической.

Если с течением времени сама среда не изменяется, а изменяются показатели производительности агента, то такая среда называется *полудинамической*.

# Свойства проблемной среды

## Дискретная или непрерывная среда

Различие между дискретными и непрерывными вариантами среды может относиться к состоянию среды, способу учета времени, а также восприятиям и действиям агента.

# Свойства проблемной среды

## Одноагентная или мультиагентная среда

На каком основании некоторая сущность должна рассматриваться как агент?

Объект В рассматривается как агента, если его поведение описывается как максимизирующее его личные показатели производительности, значения которых зависят от поведения агента А.

*Конкурентная* мультиагентная среда — сущность В пытается максимизировать свои показатели производительности за минимизации показателей производительности агента А.

*Кооперативная* мультиагентная среда — максимизируется показатели производительности всех агентов.

# Структура агентов

Агент = Архитектура + Программа

Основные виды программ агентов:

- простые рефлексные агенты;
- рефлексные агенты, основанные на модели;
- агенты, действующие на основе цели;
- агенты, действующие на основе полезности.

# Простые рефлексные агенты

Простой рефлексный агент выбирает действия на основе текущего акта восприятия, игнорируя всю остальную историю актов восприятия.





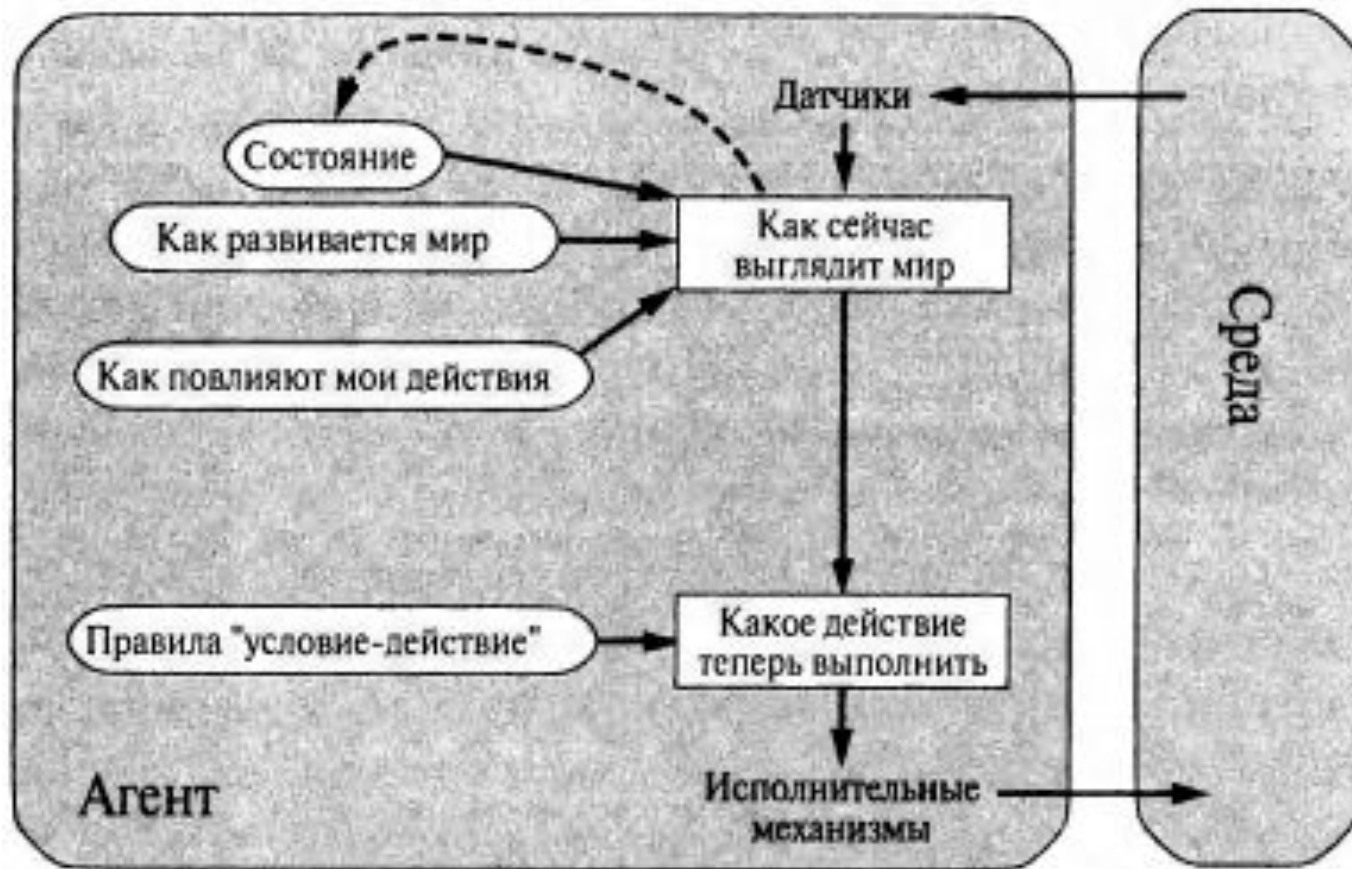
# Рефлексные агенты, основанные на модели

Агент должен отслеживать ту часть мира, которая воспринимается им в текущий момент, т.е. агент должен поддерживать *внутреннее состояние*.

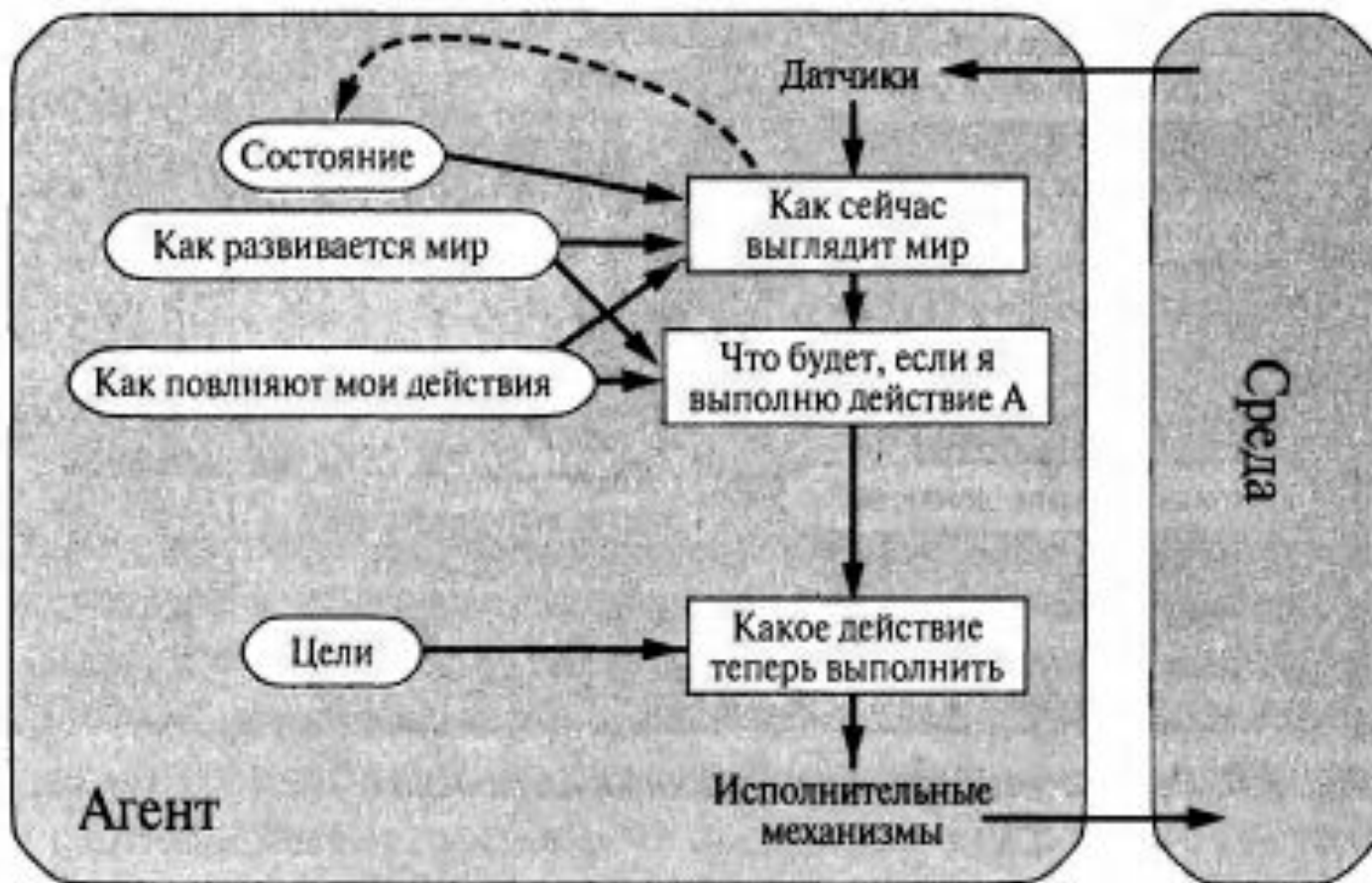
Агент должен хранить *модель мира*, т.е. знания о том, как:

- мир изменяется независимо от агента,
- влияют на мир собственные действия агента.

# Рефлексные агенты, основанные на модели



# Агенты, основанные на цели



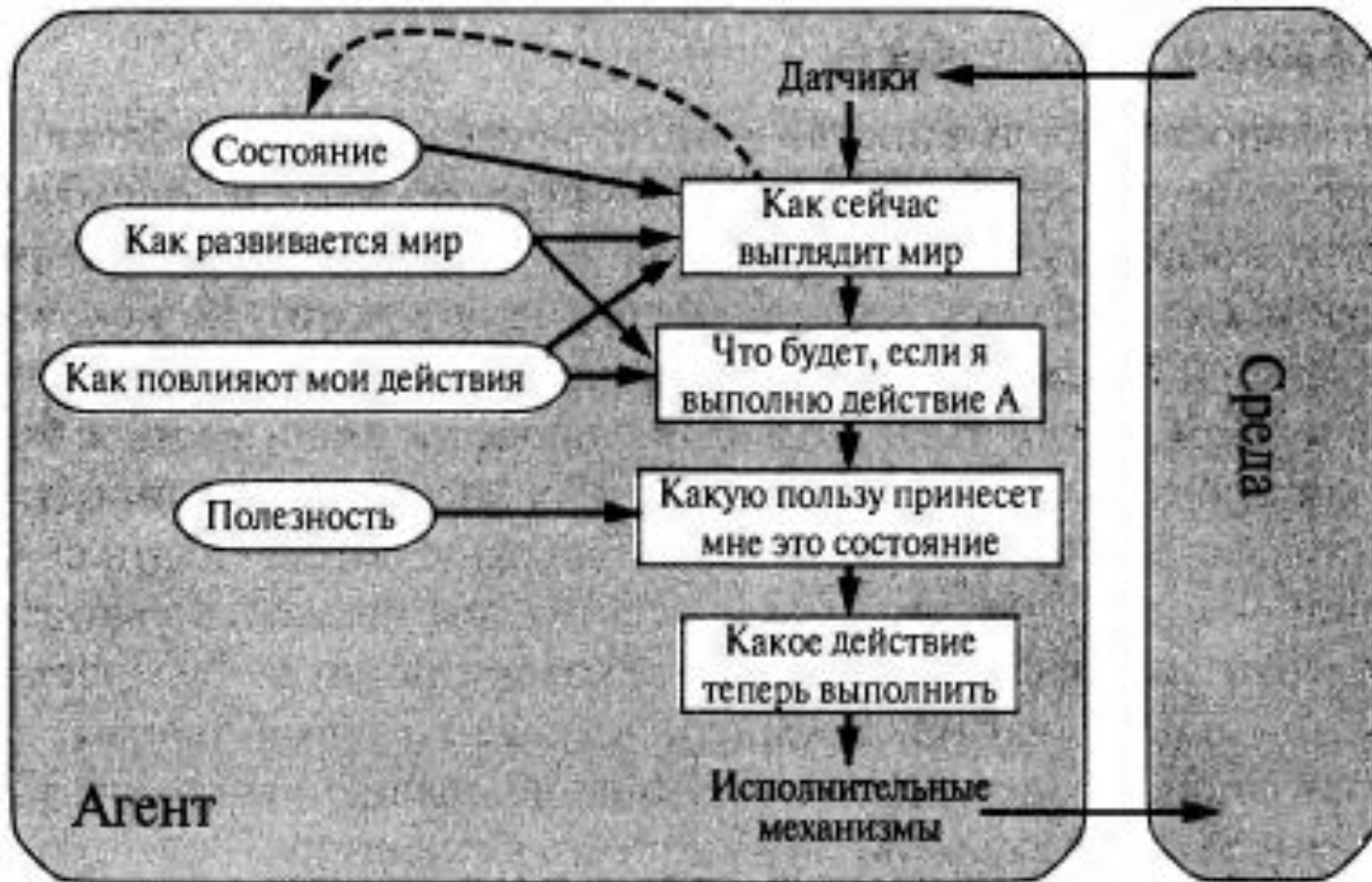
# Агенты, основанные на полезности

*Функция полезности* отображает состояние (или последовательность состояний) на вещественное число, которое обозначает соответствующую степень удовлетворенности агента.

Полная спецификация функции полезности обеспечивает возможность принимать рациональные решения в следующих случаях:

- если имеются конфликтующие цели,
- если имеется несколько целей, к которым может стремиться агент.

# Агенты, основанные на полезности



# Обучающиеся агенты

*Обучающий компонент* – отвечает за внесение усовершенствований.

*Производительный компонент* – обеспечивает выбор внешних действий.

*Критик* – определяет каким образом должен быть модифицирован производительный компонент для того, чтобы он успешнее действовал в будущем.

*Генератор проблем* – предлагает действия, которые должны привести к получению нового и информативного опыта.

# Обучающиеся агенты

