Artificial Intelligence

CS3AI18 / CSMAI19

Lecture - 1/10: Introduction (Nature and Goals of AI)

DR VARUN OJHA

Department of Computer Science

University of Reading

Learning Objectives

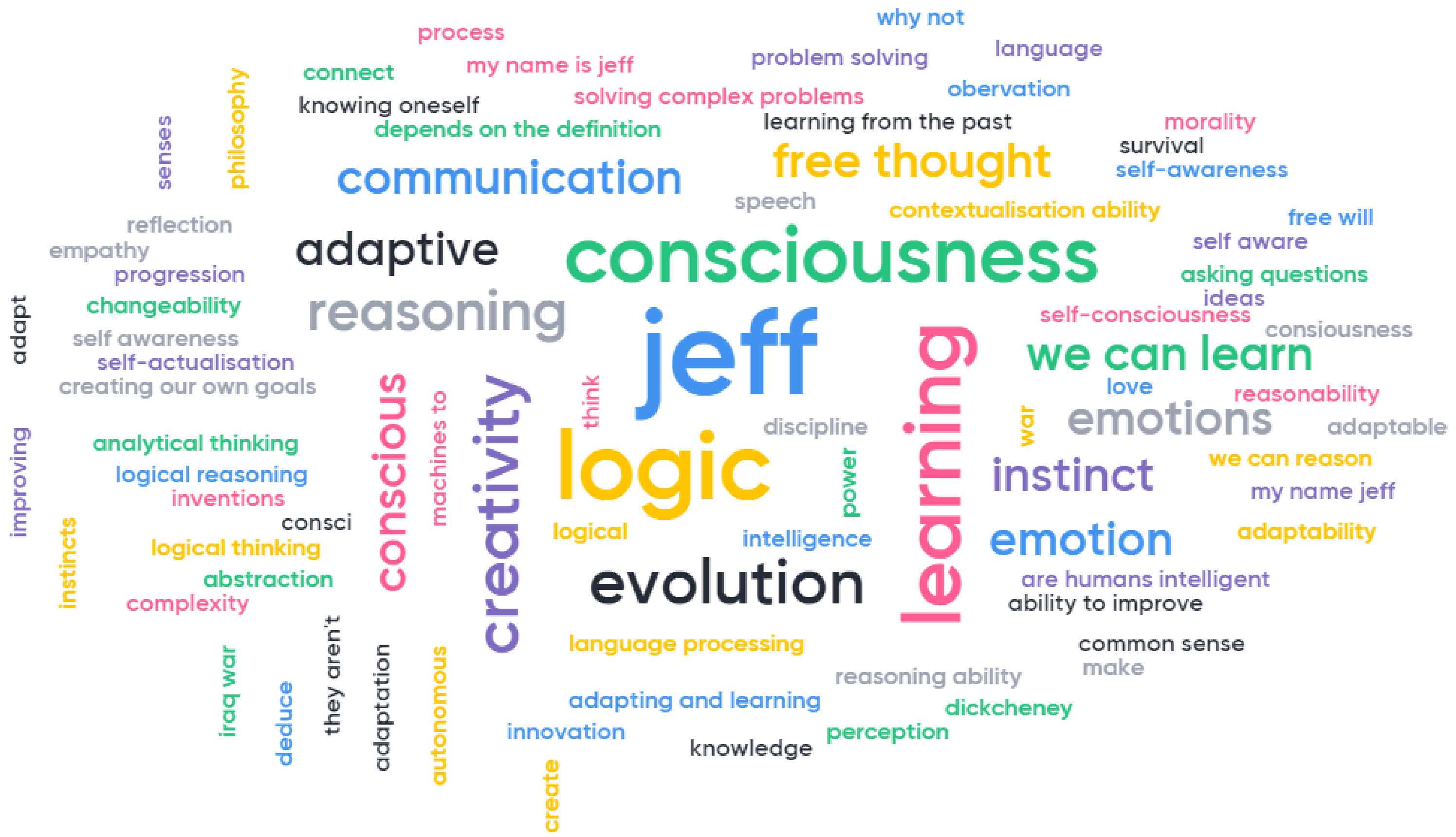
- On completion of this week, you will be able to
  - Understating definition of Artificial Intelligence
  - Aware of the Turing test
  - Know the history and applications of artificial intelligent
  - Learn about intelligent agents

Content of this Lecture

- Part – I : Definitions of AI
- Part – II : Turing Test
- Part – III: History and Applications of AI
- Part – IV: Intelligent Agent

Human Intelligence?

Why humans are intelligent?



Part 1
Definitions

What is Artificial Intelligence?

Exciting!!!...but not really useful

The exciting new effort to make computers think … machines with minds, in the full literal sense.

Haugeland, 1985

The study of mental faculties through the use of computational models.

Charniak and McDermott, 1985

What is Artificial Intelligence?

A field of study that seeks to explain and emulate intelligent behaviour in terms of computational processes.

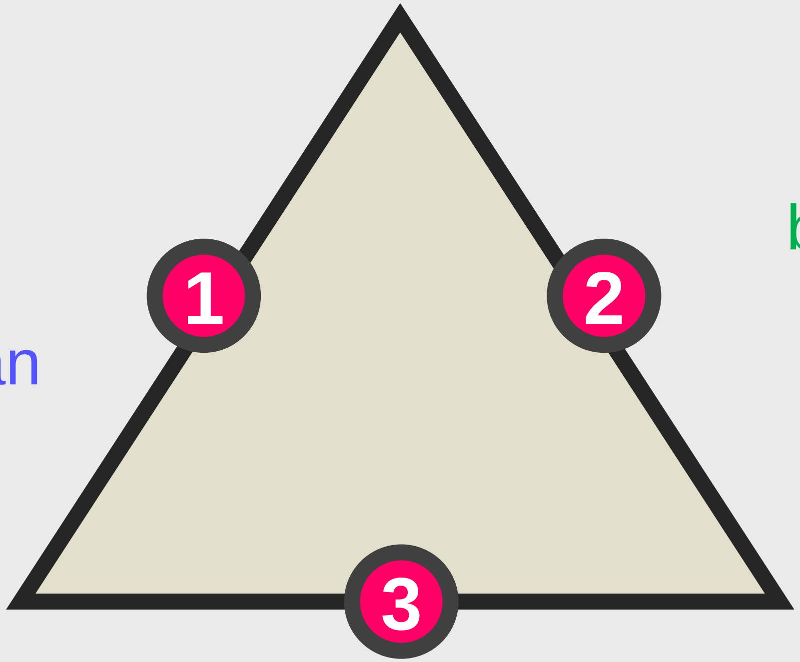
Schalkoff, 1990

The study of how to make computers do things at which, at the moment, people are better.

Rich and Knight, 1991

Dimensions of AI Definitions

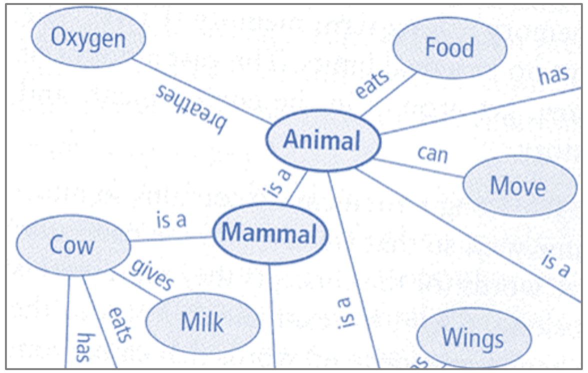
Building intelligent artefacts
vs.
understanding human behaviour.



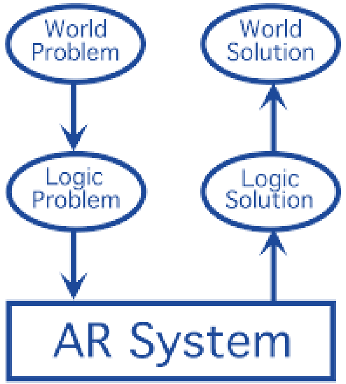
Should the system
behave like a human
Or
behave *intelligently*?

Does it matter how I built it
as long as it does the job well?

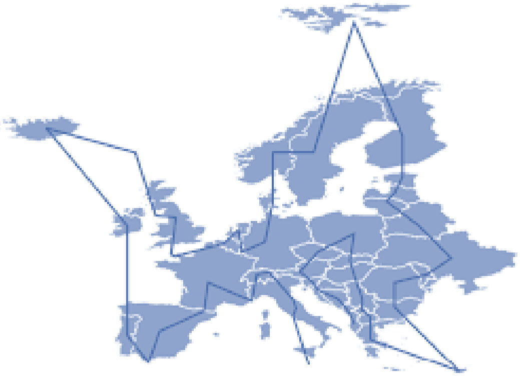
What Does AI Really Do?



Knowledge Representation

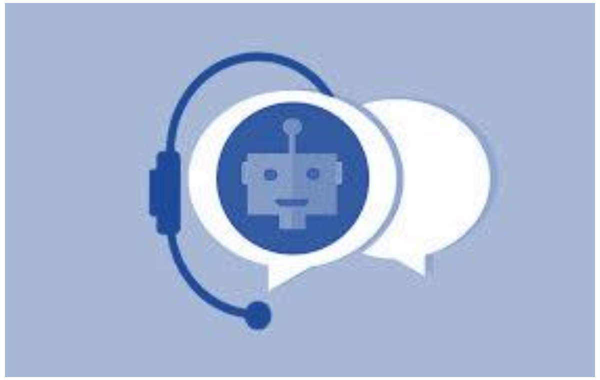


Automated reasoning

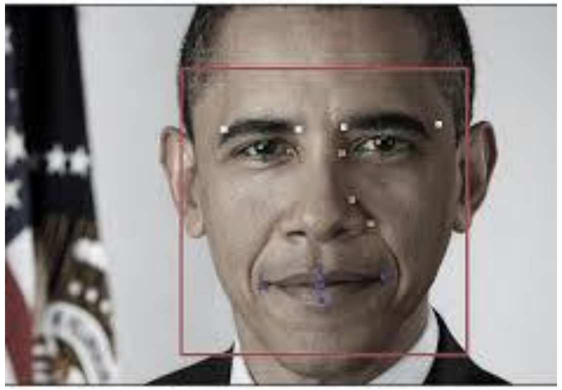


Planning

Machine Learning



Natural language understanding



Machine vision

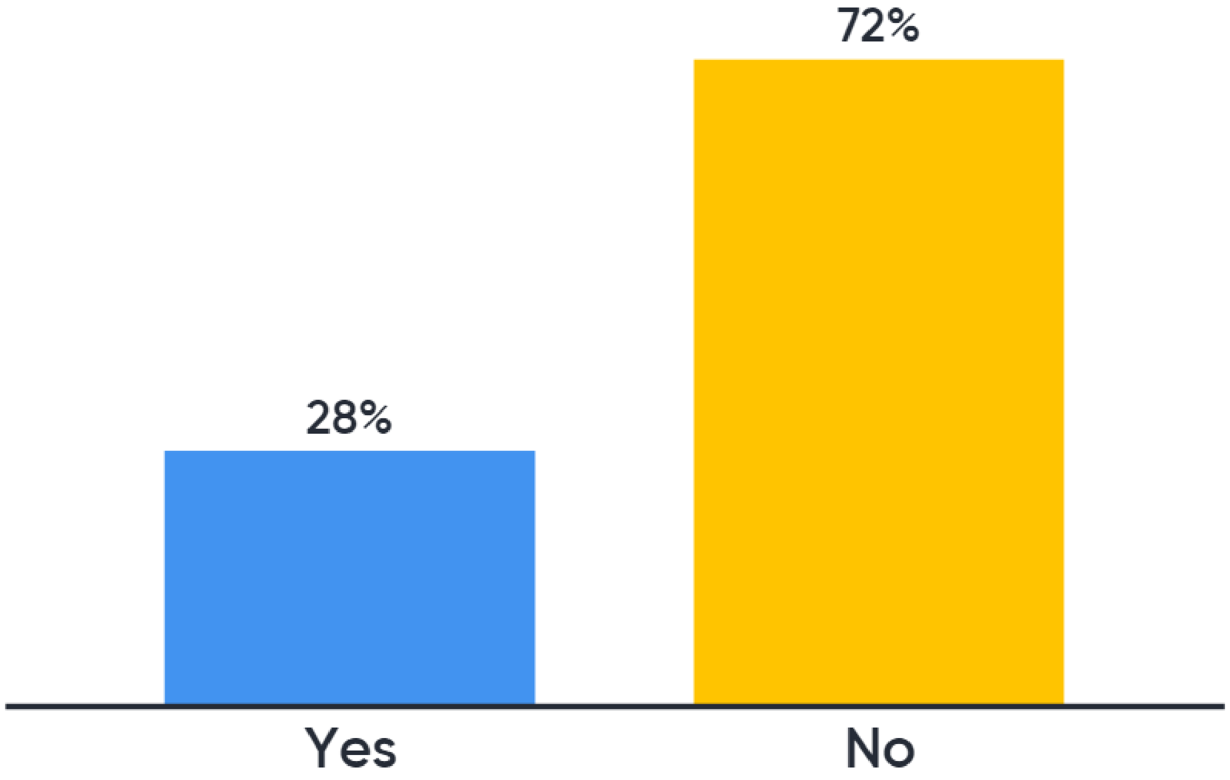
Robotics

Google

Web Search

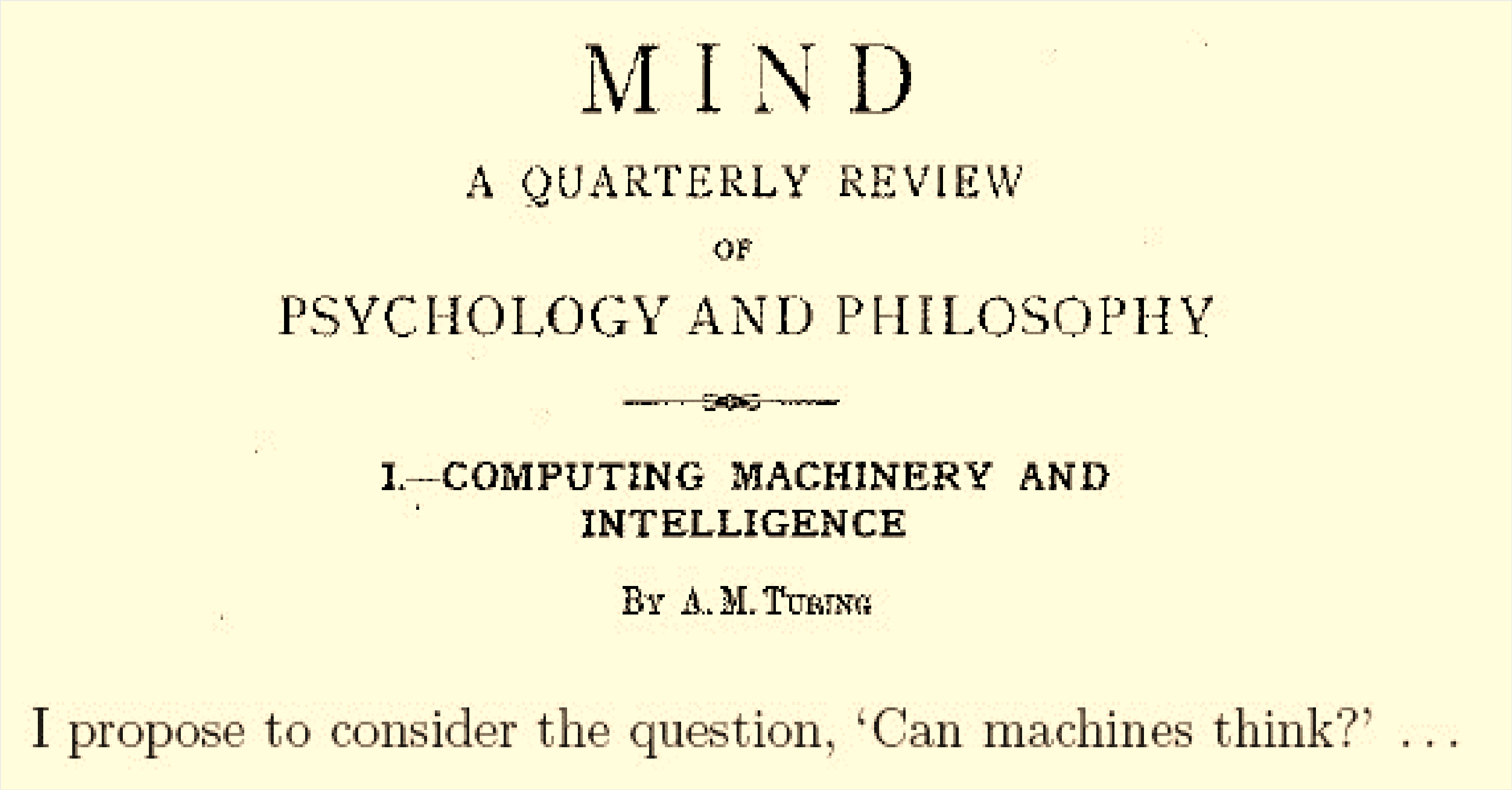
Is AI better than humans in solving problem?

Is AI is better than Humans in solving a problem?



Part 2
Alan Turing

Alan Turing - Father of AI



MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY

I.—COMPUTING MACHINERY AND INTELLIGENCE

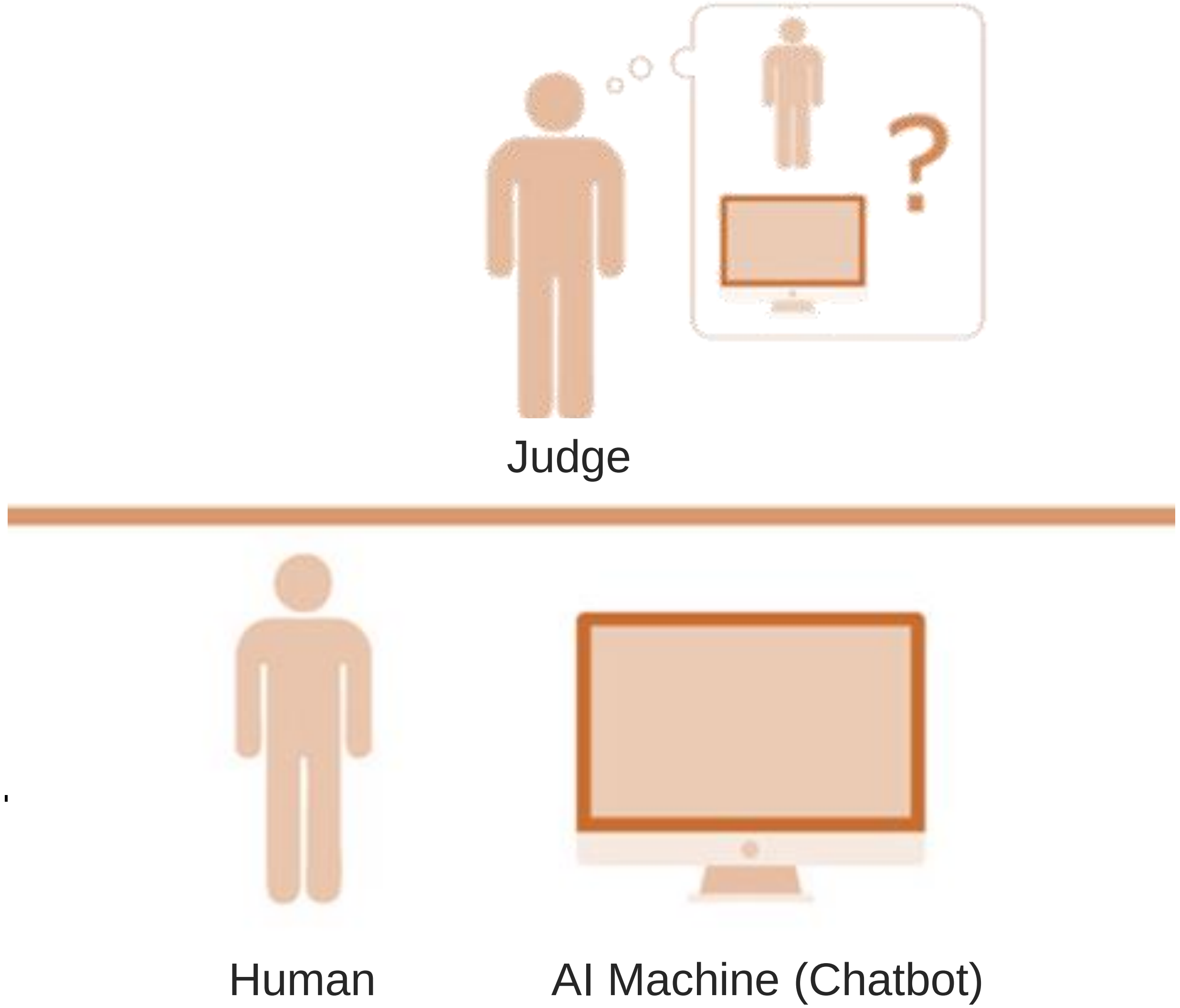
BY A. M. TURING

I propose to consider the question, 'Can machines think?' . . .

Turing, A.M. (1950), Computing machinery and intelligence, Mind, Vol.59, pp. 433-460

Turing Test

1. Judge (Human) communicates with a human and a machine over text-only channel.
2. Both human and machine try to act like a human.
3. Judge tries to tell which is which.



Is Turing Test the Right Goal?

“Aeronautical engineering texts do not define the goal of their field as making "machines that fly so exactly like pigeons that they can fool even other pigeons."

Russell and Norvig

Chinese Room Argument [Searle 1980]

If you see this shape,	then produce this shape,
"什麼"	"爲天"
followed by this shape,	followed by this shape,
"帶來"	"下式".
followed by this shape,	
"快樂"	

AI Definition Revisited

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Choose which defines AI best.

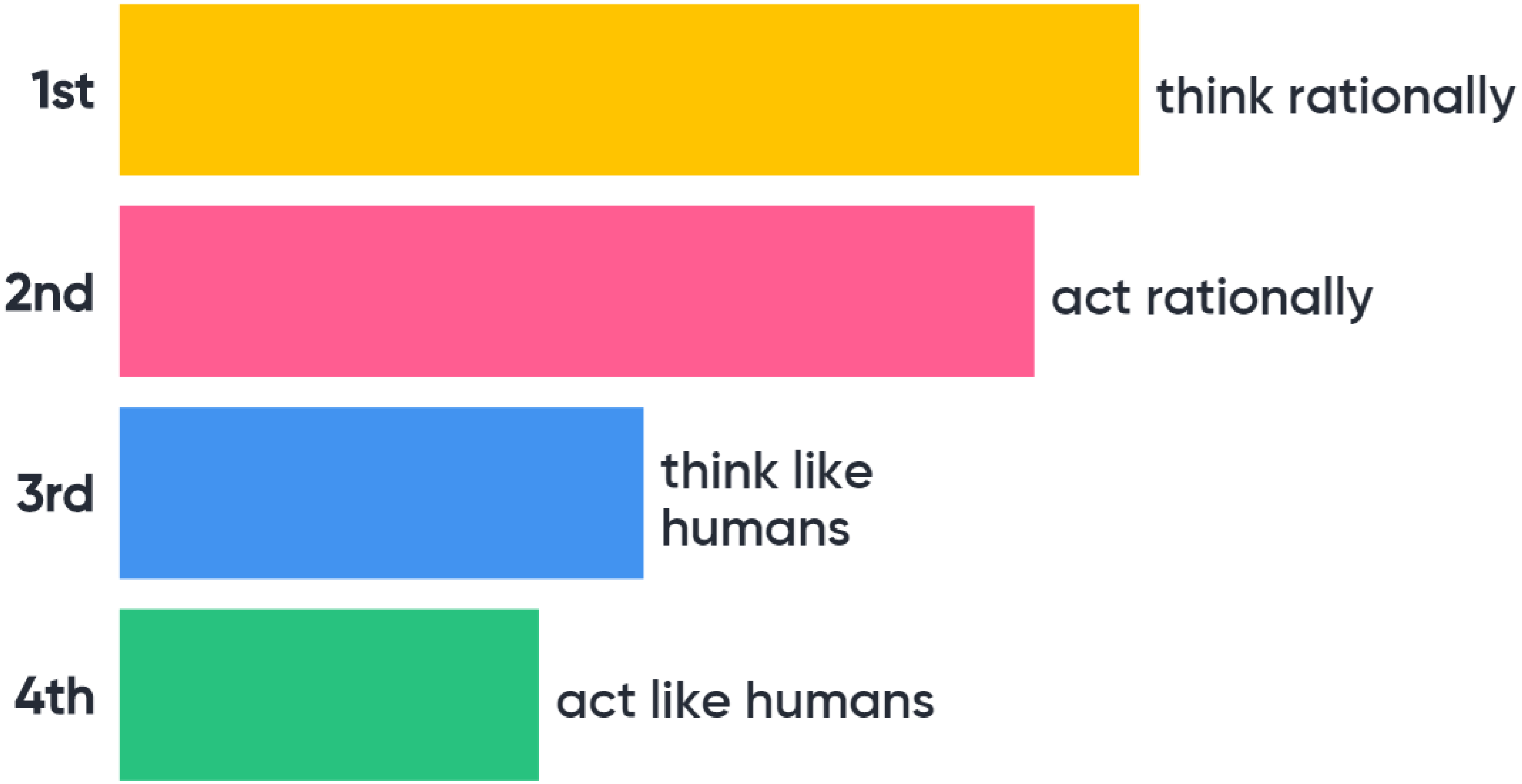
AI Definition Revisited

Systems that think like humans	Systems that think rationally
Systems that act like humans	**Systems that act rationally**

- Focus on action (act rationally).
- Avoids philosophical issues such as "is the system conscious."
- Distinction may not be that important
  - acting rationally / like a human presumably requires (some sort of) thinking rationally / like a human,
  - humans much more rational in complex domains

8:31 PM

Chose, which defines AI best.



Lessons from AI Research

What's Easy?

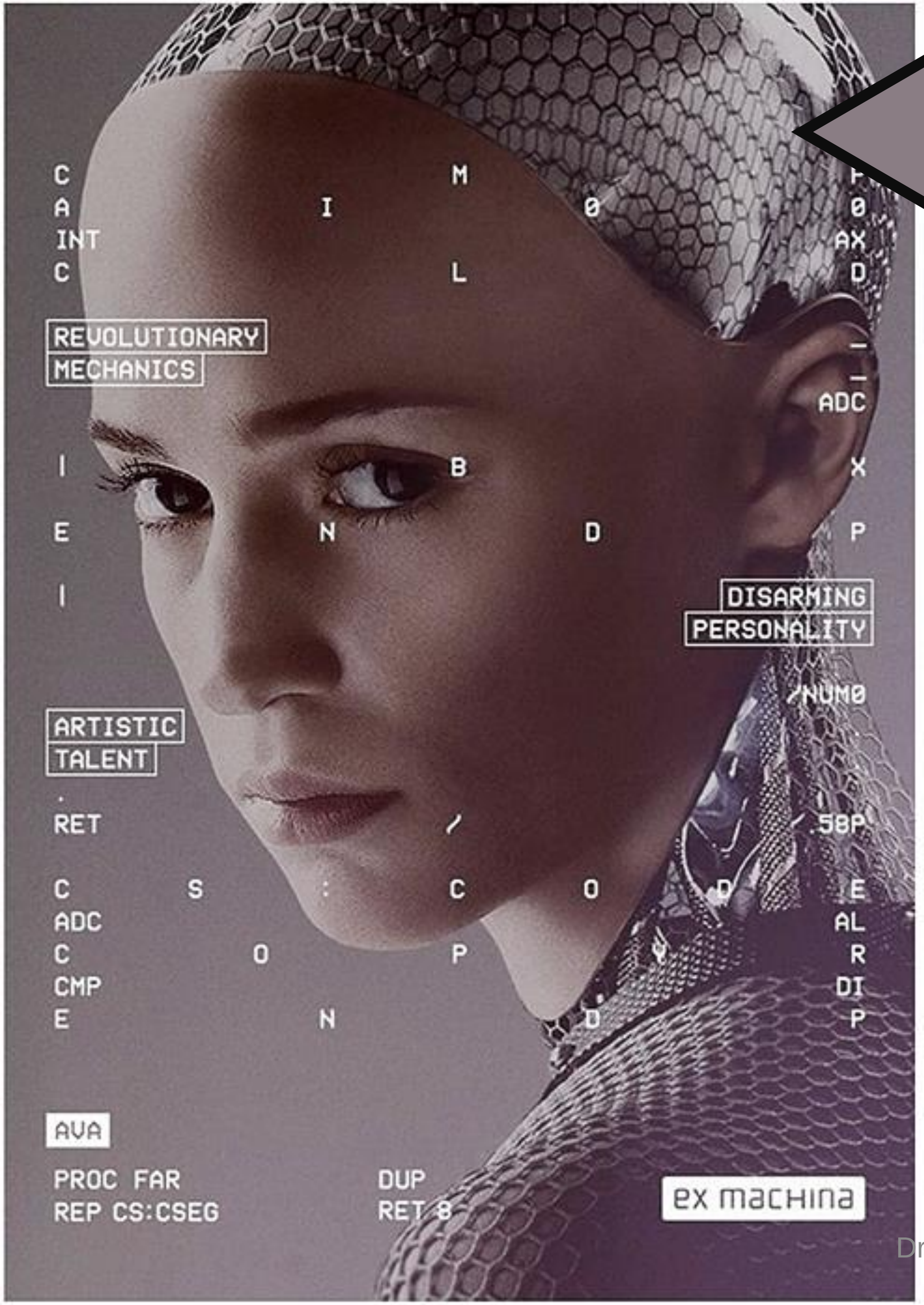
Clearly-defined tasks that we think require intelligence and education from humans tend to be doable for AI techniques

What's Hard?

Complex, messy, ambiguous tasks that are natural for humans (in some cases other animals) are much harder

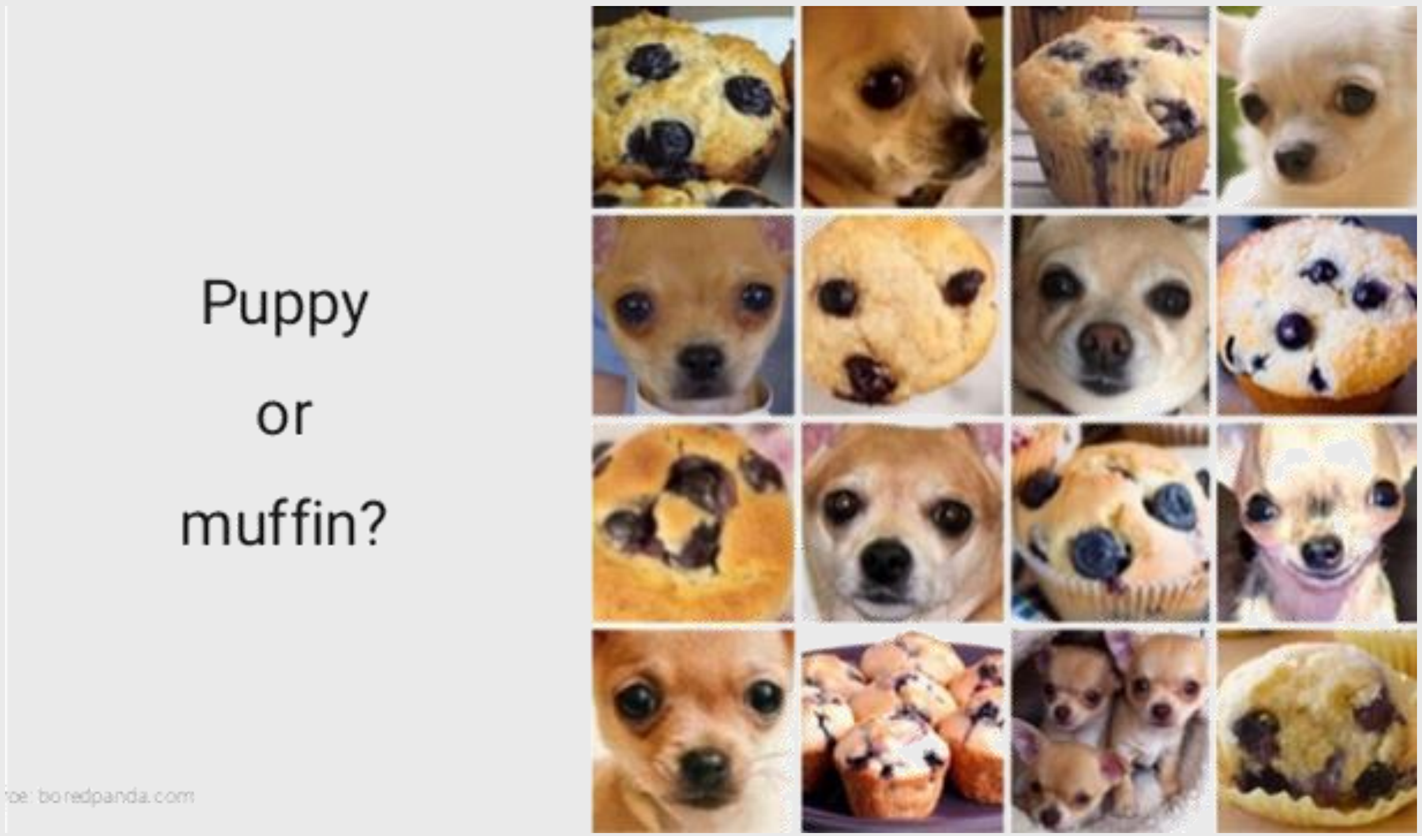
Types of AI

- **General-purpose AI** like the robots of science fiction is incredibly hard.
  - Human brain appears to have lots of special and general functions, integrated in some amazing way that we really do not understand at all (yet)
- **Special-purpose AI** is more doable (nontrivial?)
  - E.g., chess/poker playing programs, logistics planning, automated translation, voice recognition, web search, data mining, medical diagnosis, keeping a car on the road



The Goal

But busy in…



What Humans are Better At?

1. **Humans better at coming up with reasonably good solutions** in complex environments

2. **Humans better at adapting/self-evaluation/creativity** (“My usual strategy for chess is getting me into trouble against this person… Why? What else can I do?”)

Part 3

History & Applications

Early AI History

- 50s/60s: **Early successes!**
  - AI can draw logical conclusions, prove some theorems, create simple plans, initial work on neural networks (perceptron)
- Overhyping: researchers promised funding agencies spectacular progress, but started running into difficulties:
  - **Ambiguity**: highly funded translation programs (Russian to English) were good at syntactic manipulation but bad at disambiguation.
    - “The spirit is willing but the flesh is weak” becomes “The vodka is good but the meat is rotten”
  - **Scalability/complexity**: early examples were very small, programs could not scale to bigger instances.
  - Limitations of **representations** used.

Early AI History

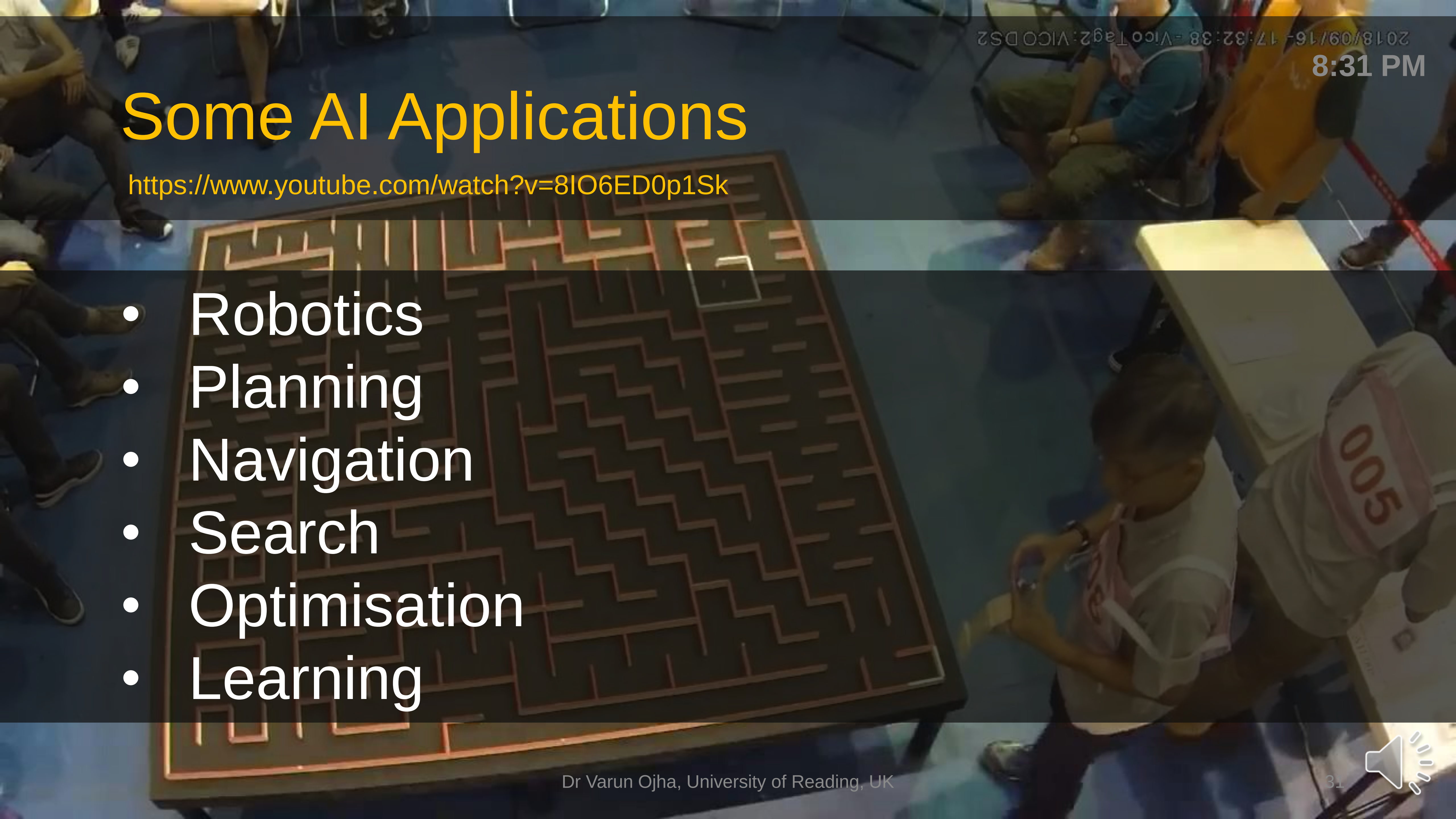
- 70s/80s:
  - Creation of **expert systems** (systems specialized for one particular task based on experts' knowledge), wide industry adoption.
- Overpromising: AI winter(s)
  - Research funding cutdown.
  - Bab reputation

Modern AI

- More rigorous, scientific, formal, and mathematical
- Fewer **grandiose** promises
- Divided into many **subareas** interested in particular aspects (e.g., speech, vision, language processing)
- More directly connected to “neighbouring” disciplines
  - statistics, economics, operations research, biology, psychology, neuroscience
  - Often leads to question “Is this really AI”?
- Some senior AI researchers are calling for re-integration of all these topics, return to more grandiose goals of AI

Example AI Applications

- **Search**
  - Solving a Rubik's cube
- **Constraint satisfaction/optimization problems**
  - Scheduling a given set of meetings (optimally)
- **Game playing**
  - Playing chess
- **Logic, knowledge representation**
  - Solving logic puzzles, proving theorems
- **Planning**
  - Finding a schedule that will allow you to graduate (reasoning backwards from the goal)
- **Probability, decision theory, reasoning under uncertainty**
  - Given some symptoms, what is the probability that a patient has a particular condition? How should we treat the patient?
- **Machine learning, reinforcement learning**
  - Recognizing handwritten digits



8:31 PM

Some AI Applications

https://www.youtube.com/watch?v=8IO6ED0p1Sk

- Robotics
- Planning
- Navigation
- Search
- Optimisation
- Learning

Video Source https://www.youtube.com/watch?v=nfIelJYOygY

Self-driving Car

https://www.youtube.com/watch?v=A8UCDfAheOQ

Kasparov vs. IBM Deep Blue

2/6

Recent Success of AI

8:31 PM

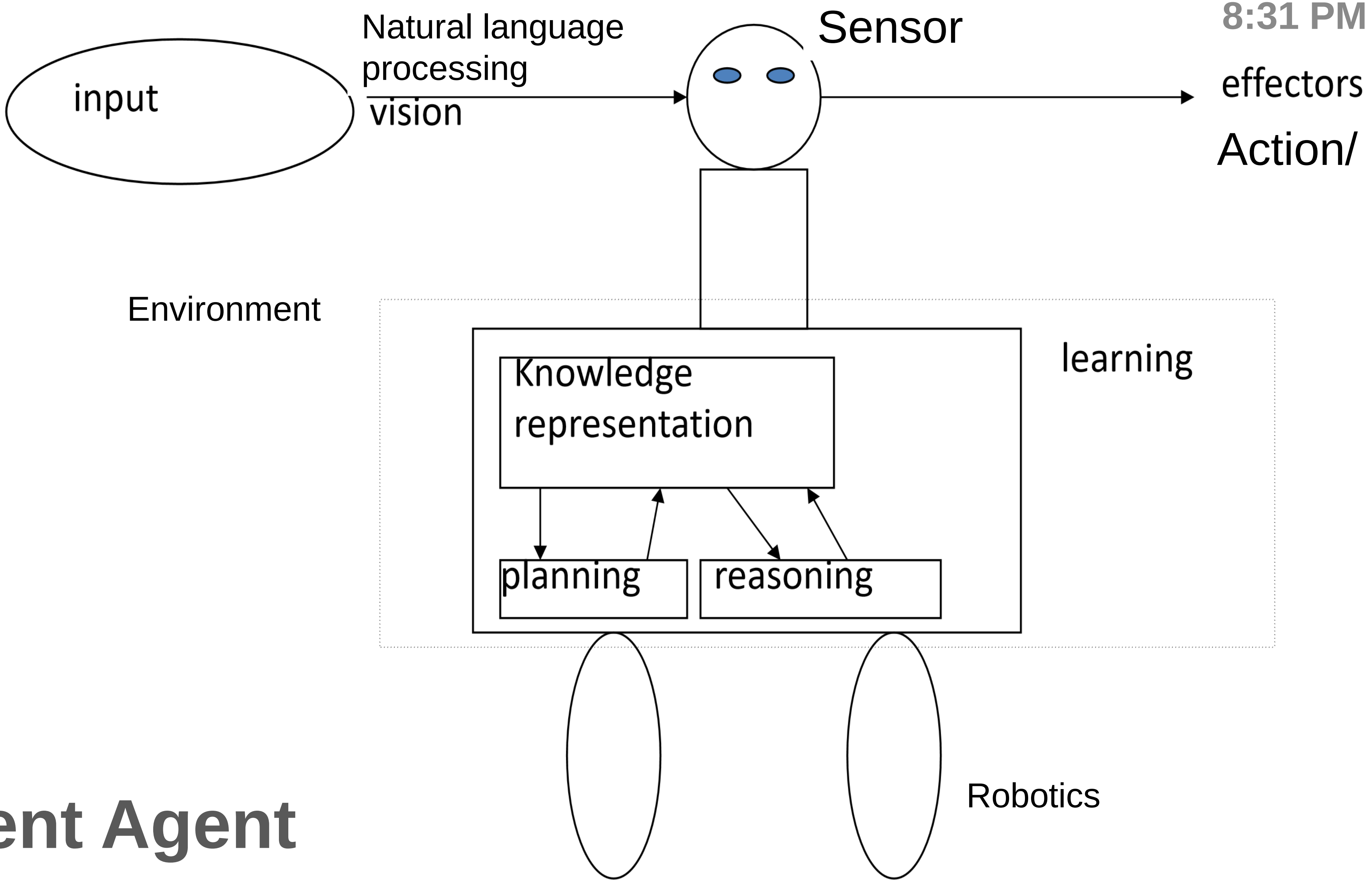
Part 4
Intelligent Agent

Agents

- An agent is anything that can be viewed as
  - perceiving its environment through sensors and
  - acting upon that environment through actuators.

- **Human agent:**
  - eyes, ears, and other organs for sensors;
  - hands, legs, mouth, and other body parts for actuators.

- **Robotic agent:**
  - cameras and infrared range finders for sensors;
  - various motors for actuators.



An Intelligent Agent

Agent and Environment

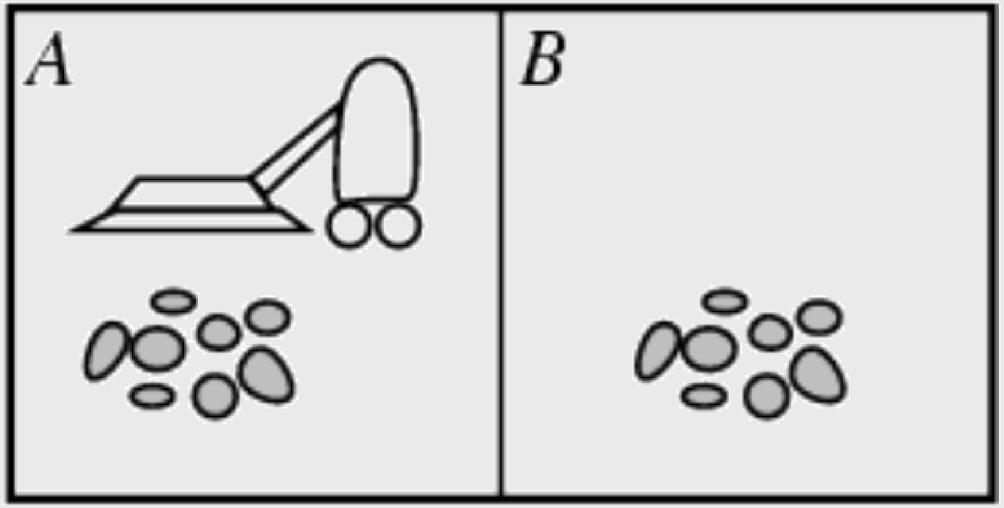
- The agent function f maps from `percept` histories $\mathcal{P}^*$ to `actions` $\mathcal{A}$:

$$f : \mathcal{P}^* \rightarrow \mathcal{A}$$

- The agent program runs on the physical architecture to produce f

- agent = architecture + program

Vacuum-cleaner world



- **Percepts:** location and state of the environment, e.g., [A,Dirty], [B,Clean]
- **Actions:** Left, Right, Suck, NoOp

Rational agents (1)

- For each possible percept sequence,
  - a rational agent should select an action that is expected to maximize its performance measure,
  - based on the evidence provided by the percept sequence and
  - whatever built-in knowledge the agent has.

- **Performance measure:** An objective criterion for success of an agent's behavior

- E.g., performance measure of a vacuum-cleaner agent could be
  - amount of dirt cleaned up,
  - amount of time taken,
  - amount of electricity consumed,
  - amount of noise generated, etc.

Rational agents (2)

- Rationality is distinct from omniscience (all-knowing with infinite knowledge)
- Agents can perform actions in order to modify future percepts so as to obtain useful information (information gathering, exploration)
- An agent is autonomous if its behavior is determined by its own percepts & experience (with ability to learn and adapt) without depending solely on build-in knowledge

Task Environment

- Before we design an intelligent agent, we must specify its “task environment”:
  - Performance measure
  - Environment
  - Actuators
  - Sensors

Task Environment (PEAS): Self-driving car

- **Performance measure:** Safe, fast, legal, comfortable trip, maximize profits
- **Environment:** Roads, other traffic, pedestrians, customers
- **Actuators:** Steering wheel, accelerator, brake, signal
- **Sensors:** Cameras, sonar, speedometer, GPS, engine sensors, keyboard

Task Environment (PEAS): Medical diagnosis system

- **Performance measure:** Healthy patient, minimize costs, lawsuits
- **Environment:** Patient, hospital, staff
- **Actuators:** Screen display (questions, tests, diagnoses, treatments, referrals)
- **Sensors:** Keyboard (entry of symptoms, findings, patient's answers)

Task Environment (PEAS): Part-picking robot

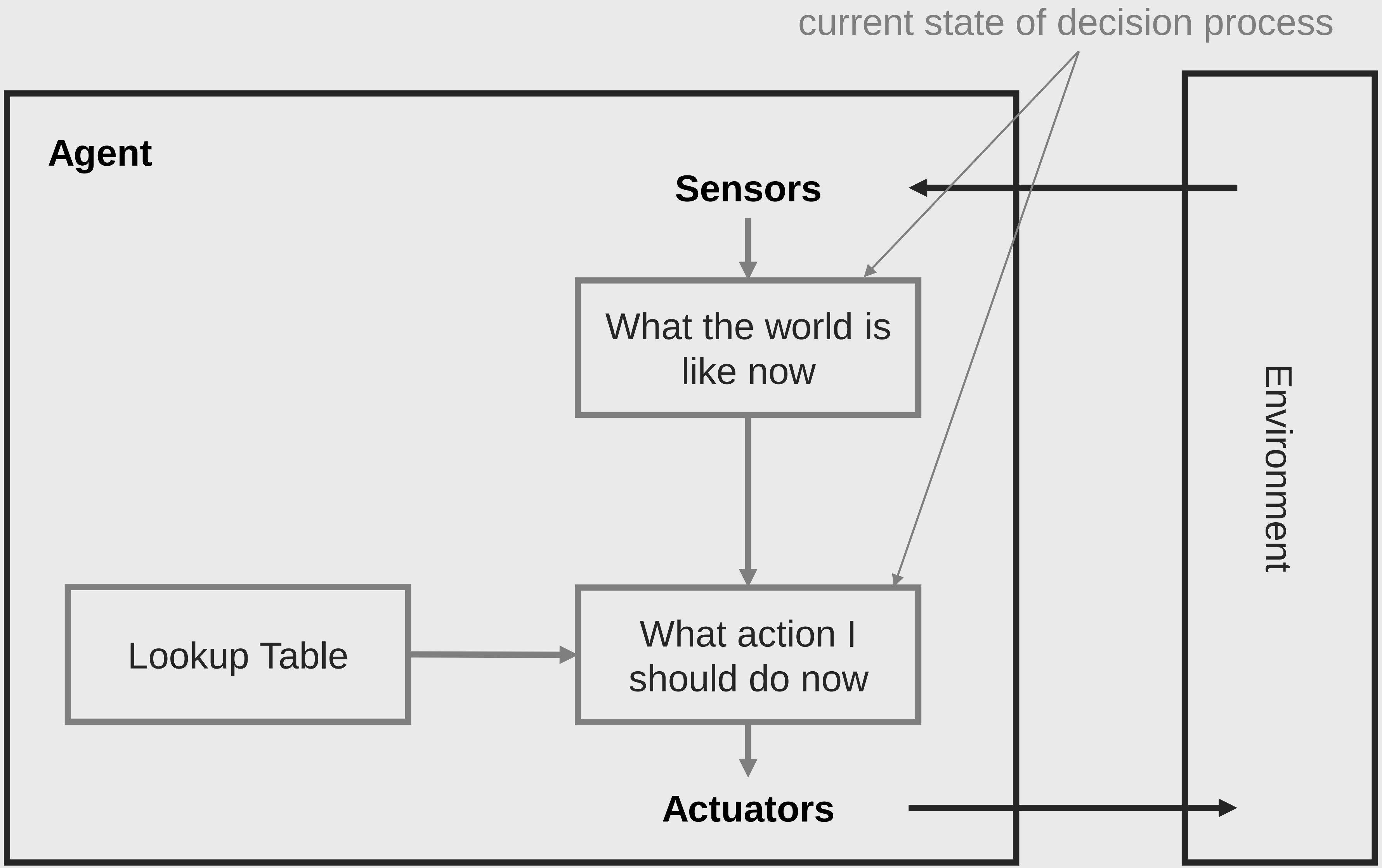
- **Performance measure:** Percentage of parts in correct bins
- **Environment:** Conveyor belt with parts, bins
- **Actuators:** Jointed arm and hand
- **Sensors:** Camera, joint angle sensors

Agents Type

- Five basic types in order of increasing generality:
  - Table Driven agents
  - Simple reflex agents
  - Model-based reflex agents
  - Goal-based agents
  - Utility-based agents

Table Driven agents

Impractical



Simple reflex agents

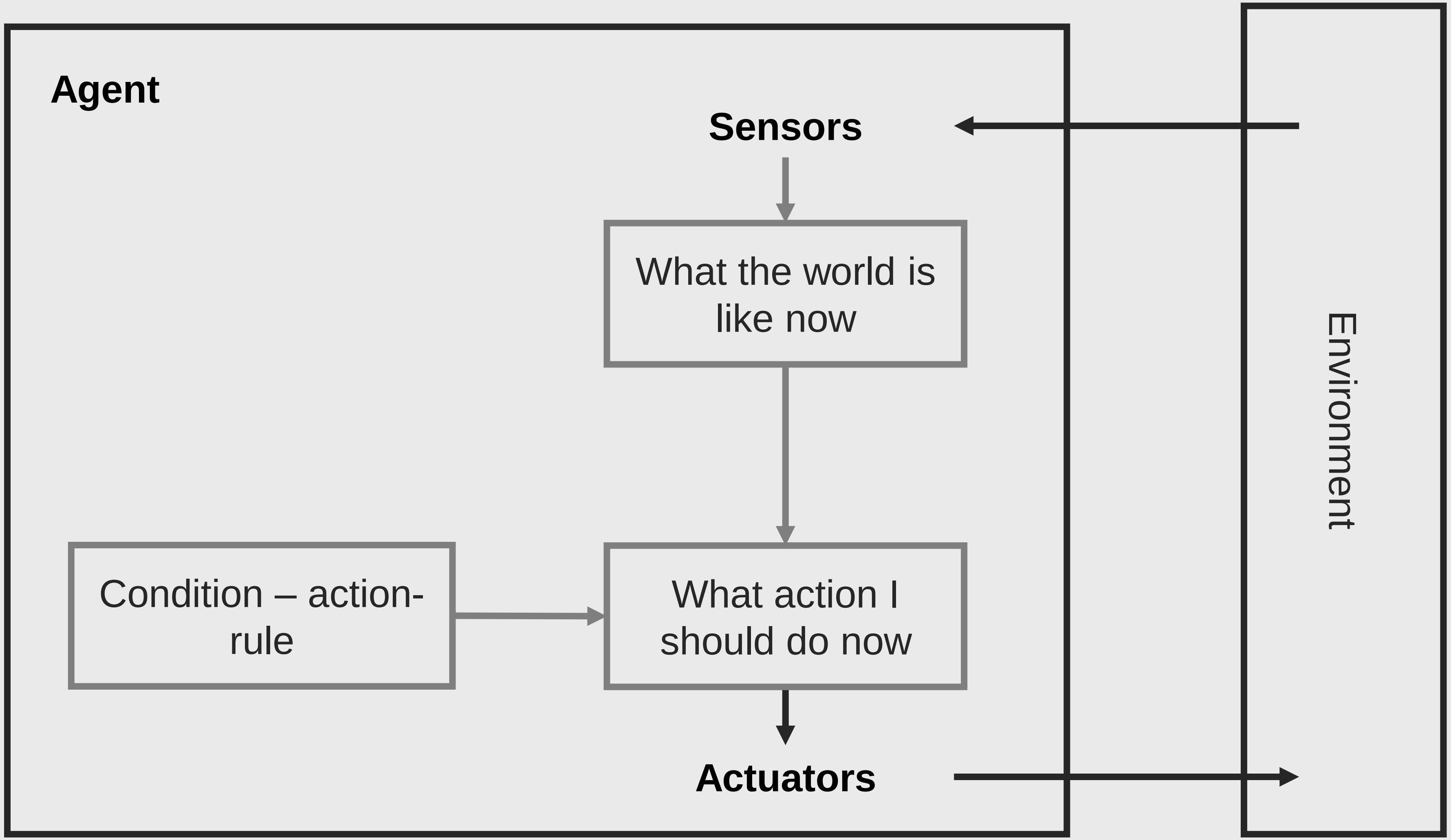
8:31 PM

Example: vacuum cleaner world

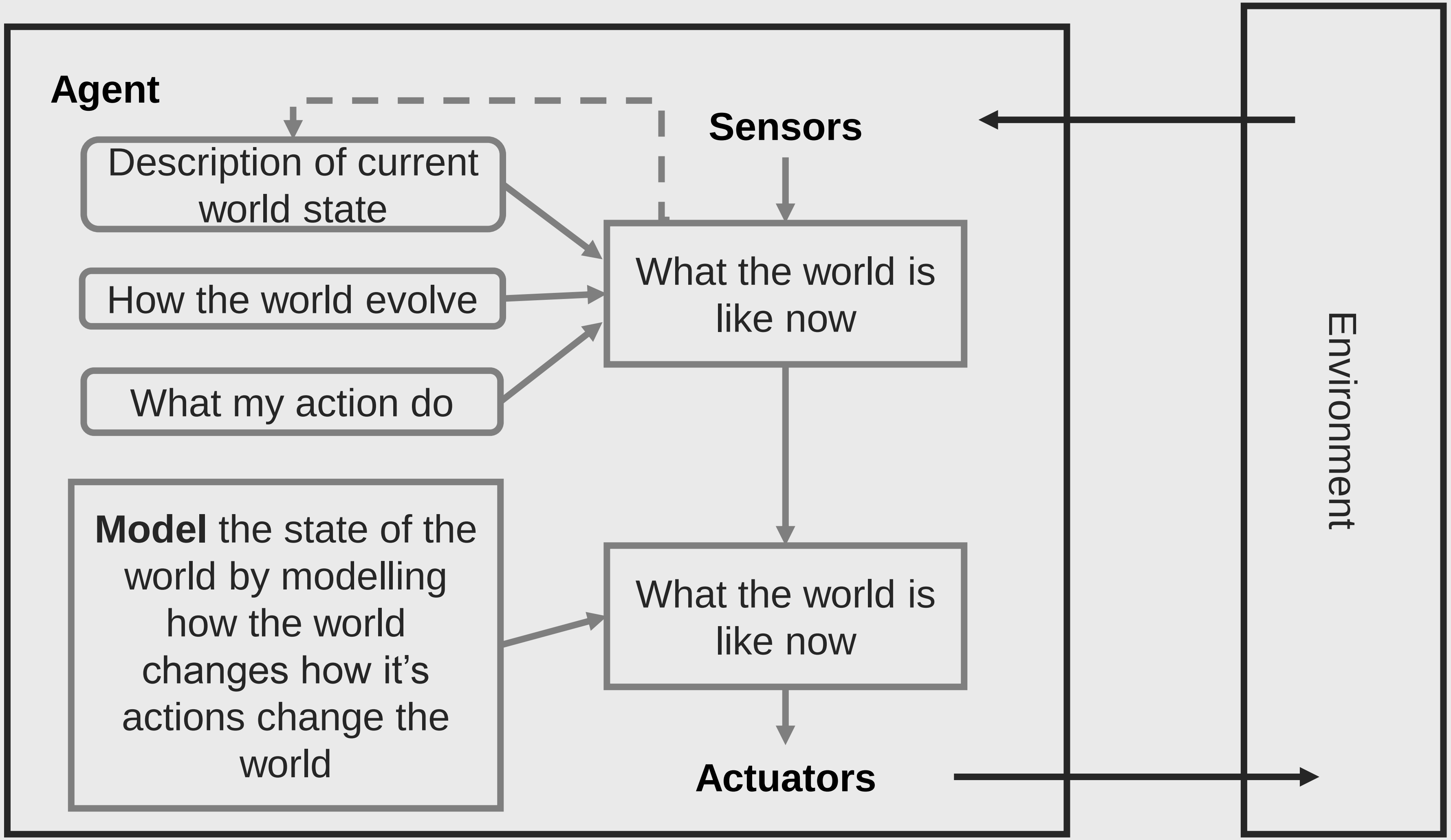
Fast but too simple

NO MEMORY

Fails if environment is partially observable



Model-based reflex agents



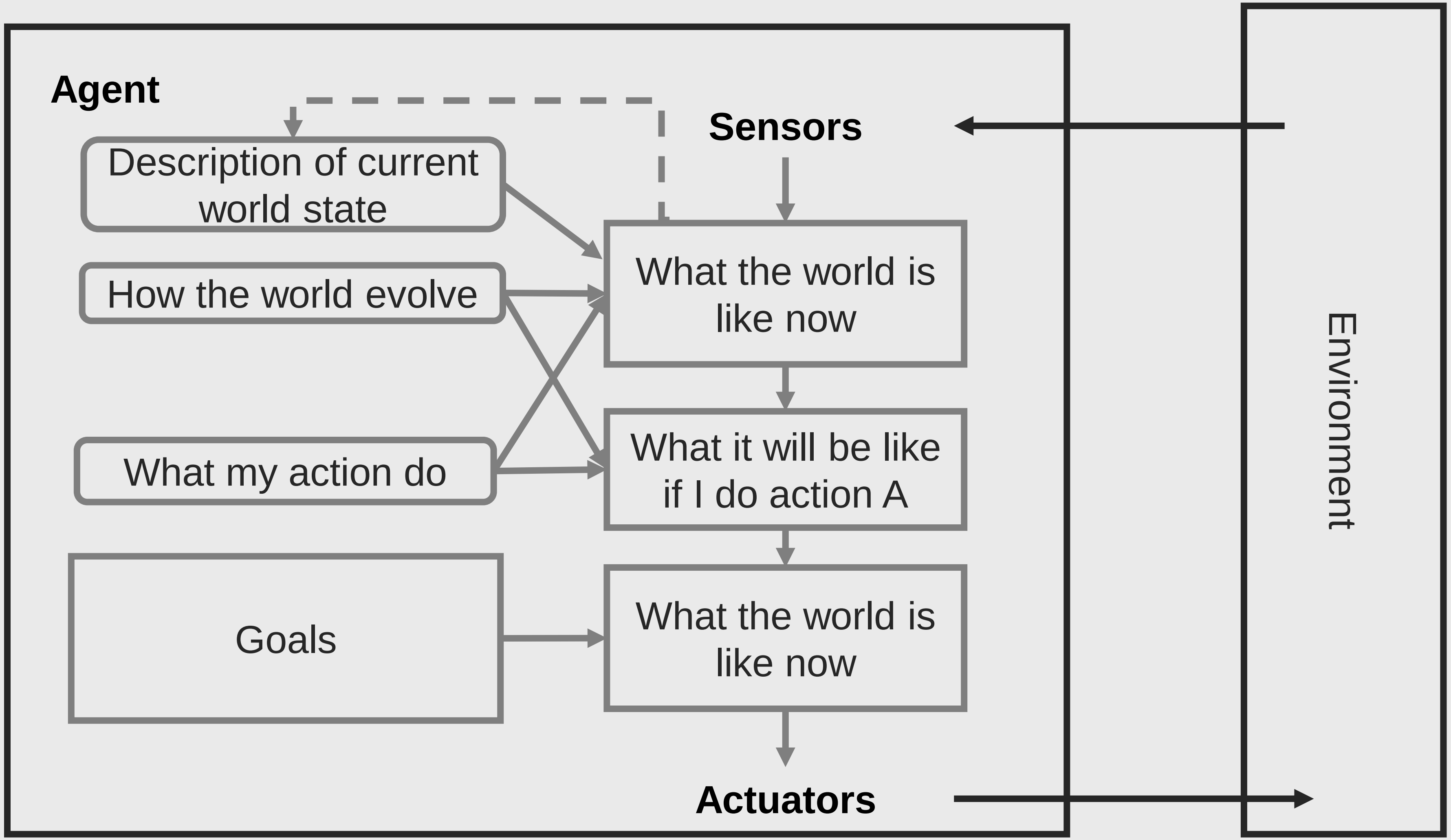
This can work even with partial information

It's is unclear what to do without a clear goal

Goal-based reflex agents

Goals provide reason to prefer one action over the other.

We need to predict the future: we need to plan & search

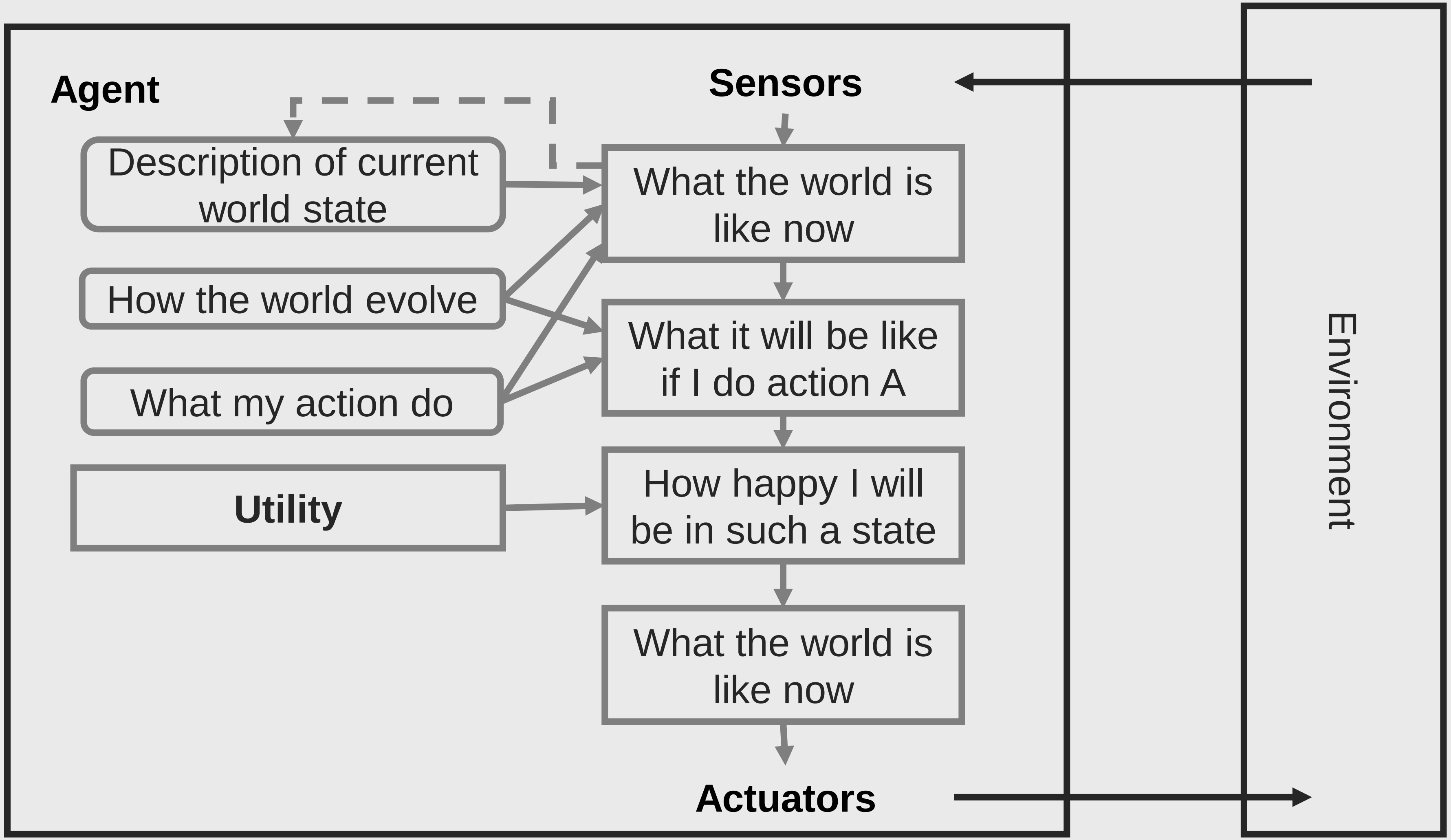


Utility-based reflex agents

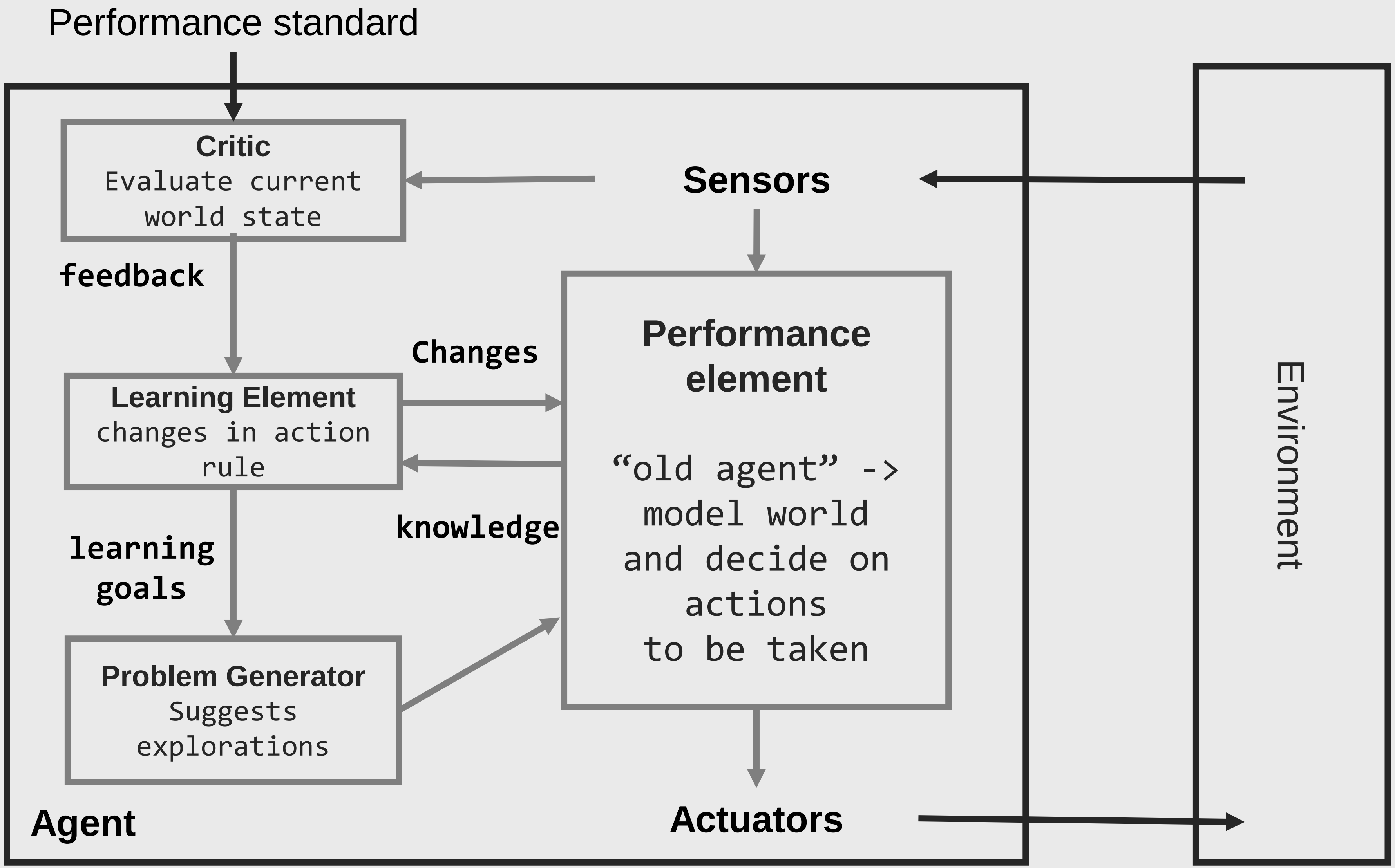
8:31 PM

Some solutions to goal states are better than others.

Which one is best is given by a utility function. Which combination of goals is preferred?



Learning agents



How does an agent improve over time? By monitoring it's performance and suggesting better modelling, new action rules, etc.

Gödel's Incompleteness Theorems

- **"Any consistent formal system F within which a certain amount of elementary arithmetic can be carried out is incomplete**; i.e., there are statements of the language of F which can neither be proved nor disproved in F. "

- Gödel's theorem assures us that humans will always be superior to machines.

- A robot/agent can never be aware of itself (be self-conscious).