Introduction to Artificial Intelligence

NCTU - Spring 2020

Before we actually start the course, let's look at some recent developments in "AI":

StyleGAN (nVidia, 2018/12)

Are these real people?

https://petapixel.com/2017/11/07/ai-creates-photo-realistic-faces-people-dont-exist/
https://www.thispersondoesnotexist.com/

StyleGAN applied to cats (nVidia, 2019/2)

Are these real cats?

https://www.livescience.com/64771-ai-generated-cats-from-hell.html

Artificial Intelligence?

First, consider these two subjects / products of AI research:

Robot Mouse vs. **Robothespian**

- Are they intelligent?
- Which is closer to your original idea of AI?
- What are their goals?
- What are their required skill sets? Which of these skills are considered intelligent?

Goals of Artificial Intelligence

- "Artificial" means "man-made". So, what kind of artificial intelligence do we want to achieve?
- Two dimensions:
  - To be **human-like**, or to be **rational**? (After all, human beings are not always rational.)
  - To focus on the process of **thinking**, or to focus on the **behaviors**?
- Therefore, four different possible goals of AI:
  - Thinking humanly
  - Acting humanly
  - Thinking rationally
  - Acting rationally

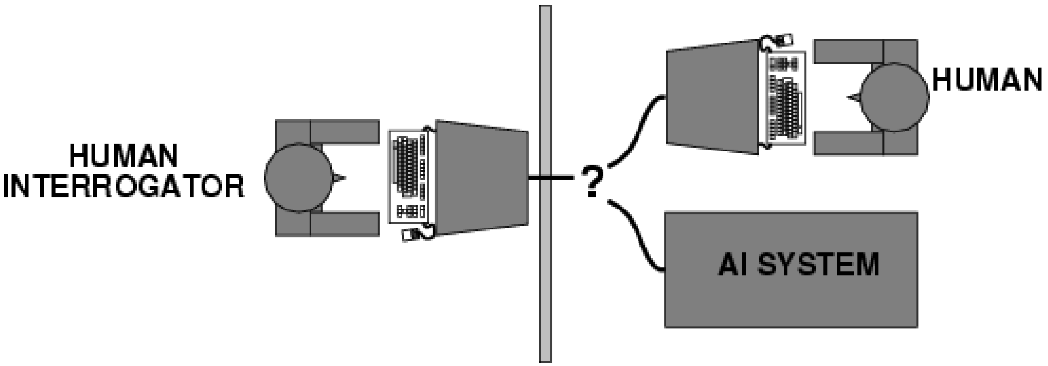
General vs. Narrow AI

- **General AI**:
  - Human-like intelligence: Thinking, reasoning, learning, creativity, etc., that are not limited to specific fields.
  - A popular topic of science fictions.
  - Generalized learning is the key.
  - Still out of reach.
- **Narrow AI**:
  - Designed for specific tasks.
  - Human experts control "what" and "how" to learn.
  - The focus of "practical AI".

Acting Humanly: Turing Test

Turing's (1950) "*Computing machinery and intelligence*":

- "Can machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game



- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in the following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning.

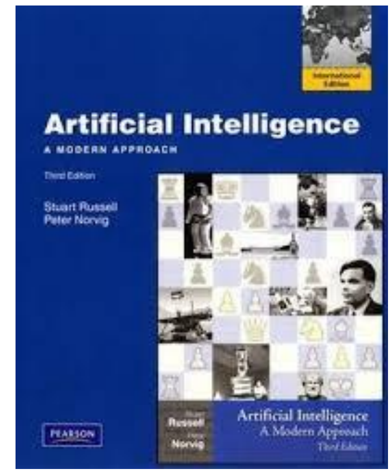
Q: How much do you think today's technology can pass the Turing Test?

A Brief History of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's “Computing Machinery and Intelligence”
- 1950s Early AI programs: checkers, Logic Theorist, etc.
- 1956 Dartmouth meeting: The term “Artificial Intelligence” adopted
- 1965 Robinson's complete algorithm for logical reasoning (a theorem prover)
- 1966-74 AI discovers computational complexity; neural network research almost disappears; the start of "AI winter"
- 1980-93 Expert systems industry booms and busts
- 1985 Neural networks return to popularity
- 1995 The start of "data science"
- 1997 Deep Blue
- 2009 The "Big Bang" of deep learning with GPU computing
- 2016+ AlphaGo, etc.

About This Course

- Instructor information:
  - Contact: EC709, ext. 56689, wangts@cs.nctu.edu
  - Office hour: 2DX or by appointment
- The main topics:
  - Searching (including some game playing)
  - Logical reasoning
  - Learning



- Textbook:
  - *Artificial Intelligence: A Modern Approach* (3rd Ed., 2003) by Russell & Norvig.
  - A comprehensive book for good foundations, but a little too old for current developments.
  - We will add some material that cover some more recent topics.

About This Course

- Exams and homework (50%):
  - Two written exams
  - 2~4 quizzes (for review purpose)
- Projects (50%):
  - We expect to have <u>three individual projects</u>. Each will involve programming and experiments with your programs.
  - A group project (up to 3 students per group) on a game agent. There will be a tournament.
  - You need to submit your code and a report for each project you do.

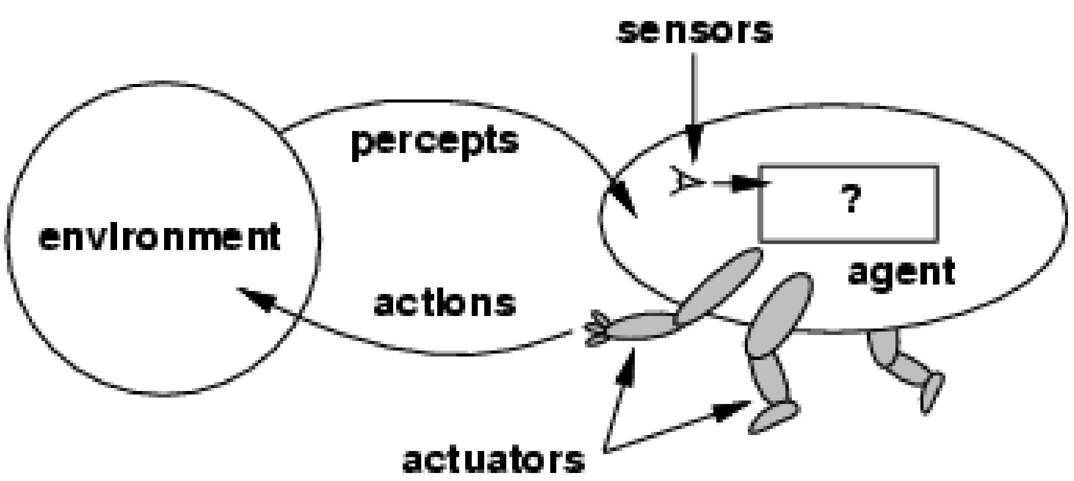
The Concept of "Agents"

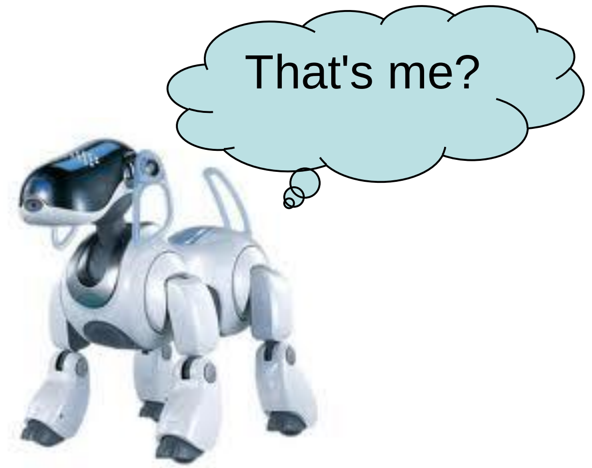
- An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators

- Human agents:
  - Sensors: eyes, ears, etc.
  - Actuators: hands, legs, mouth, etc.

- Robotic (nonhuman) agents:
  - Sensors: cameras, range finders, etc.
  - Actuators: motors, speakers, etc.

The Concept of "Agents"





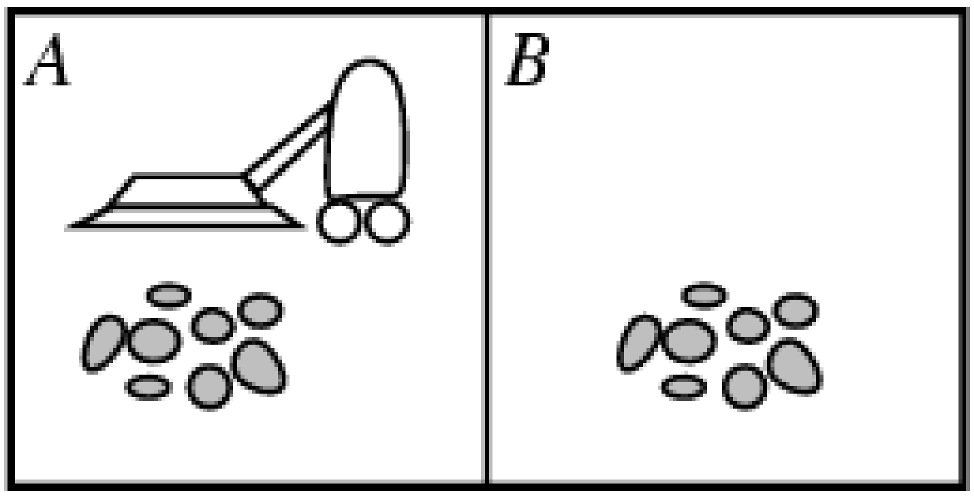
- Abstractly, an agent is a function from percept histories to actions:

$$f : P^* \rightarrow A$$

- This course is about designing rational agents: For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance.

Example: Vacuum-Cleaner World

- Percepts: location and contents, e.g., [*A*, *Dirty*]
- Actions: *Left*, *Right*, *Suck*, *NoOp*.



Example: A Vacuum-Cleaner Agent

Percept sequence	Action
$[A, Clean]$	$Right$
$[A, Dirty]$	$Suck$
$[B, Clean]$	$Left$
$[B, Dirty]$	$Suck$
$[A, Clean]$, $[A, Clean]$	$Right$
$[A, Clean]$, $[A, Dirty]$	$Suck$
$\vdots$	$\vdots$

function REFLEX-VACUUM-AGENT([*location*,*status*]) **returns** an action

  if $status = Dirty$ **then return** *Suck*
  else if $location = A$ **then return** *Right*
  else if $location = B$ **then return** *Left*

Rationality of an Agent

- Example (generalized vacuum-cleaner world): Fixed performance measure evaluates the environment sequence
  - one point per square cleaned up in time T?
  - one point per clean square per time step, minus one per move?
  - penalize for > k dirty squares?
- Given the available percept sequence, a rational agent chooses the action that optimizes the expected value of the performance measure.
- Rational ≠ omniscient (percepts may not supply all relevant information)
- Rational ≠ clairvoyant (action outcomes may not be as expected)
- Therefore, rational ≠ successful, and rational ≠ optimal

PEAS

To design a rational agent, we must specify the following.

Example: An automated taxi

- **P**erformance measure: safety, destination, profits, legality, comfort, …
- **E**nvironment: roads, traffic, pedestrians, weather, …
- **A**ctuators: steering, accelerator, brake, horn, speaker, display, …
- **S**ensors: video, accelerometers, gauges, engine sensors, GPS, …

PEAS Example

Additional examples: Game-playing agent

- Performance measure:
- Environment:
- Actuators:
- Sensors:

PEAS Example

Additional examples: Interactive route planning

- Performance measure:
- Environment:
- Actuators:
- Sensors:

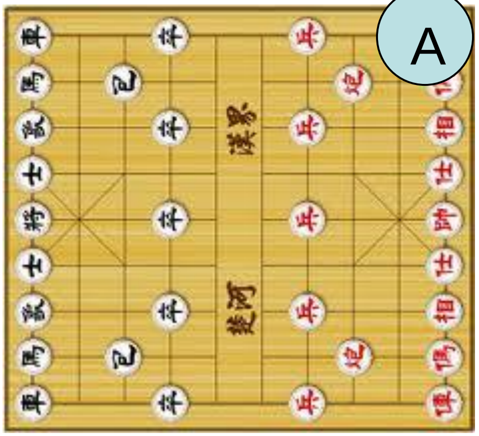
Environment Types

- Fully observable or partially observable:
- Single agent or multi-agent:
- Deterministic or stochastic: (Is the current state of the environment completely determined by its past states and the agents' actions?)
- Episodic or sequential: (Are future decisions affected by past ones?)
- Static or dynamic: (Does the environment change between actions by the agent?)
- Discrete or continuous:
- Known or unknown: (rules or "laws of physics" known to the agent)

Example Environment Types

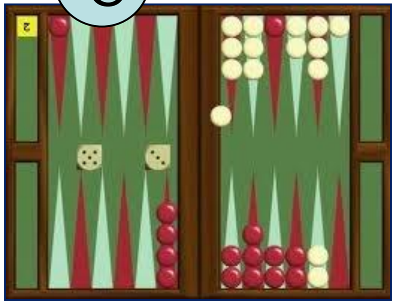
Try to characterize the environment types of these games:

A

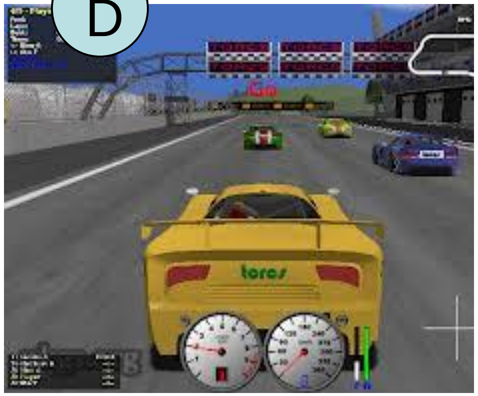


B

C



D



	A	B	C	D
Observable?				
Single-agent?				
Deterministic?				
Episodic?				
Static?				
Discrete?				
Known?				

Agent Types

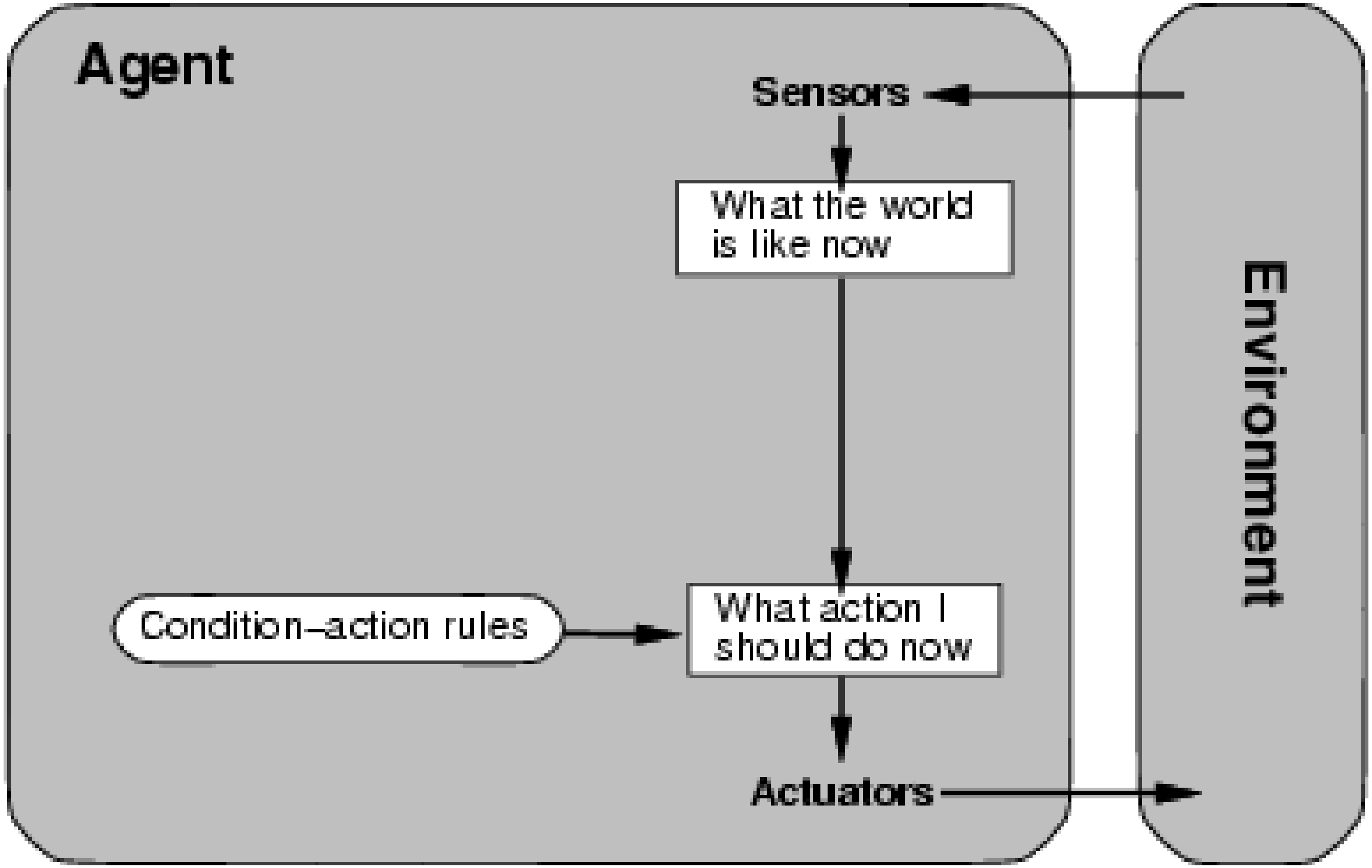
Four basic types, in order of increasing generality:

- Simple reflex agents
- Model-based reflex agents
- Goal-based agents
- Utility-based agents

All these can be turned into learning agents.

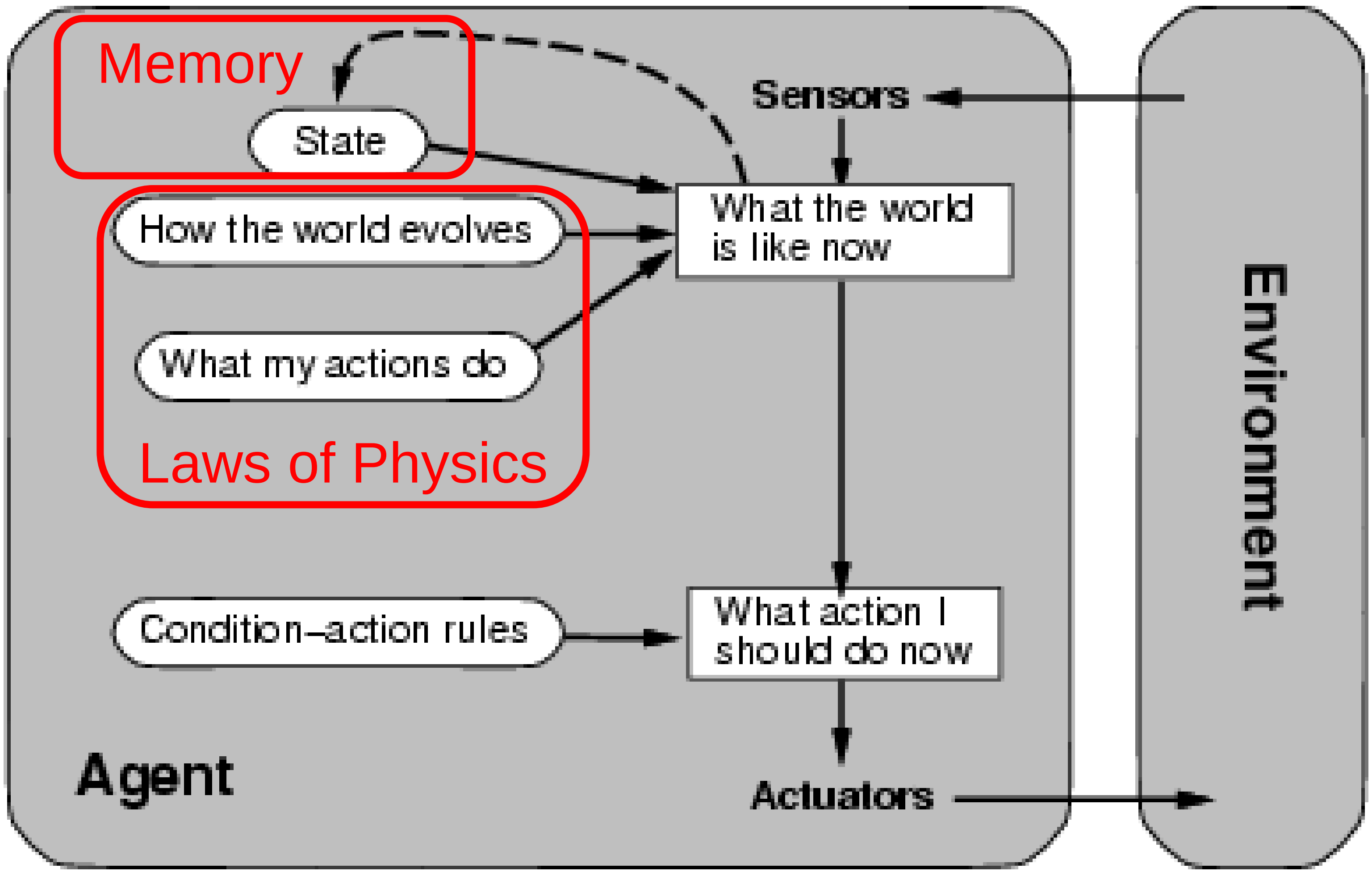
Simple Reflex Agent

Actions are only based on rules.



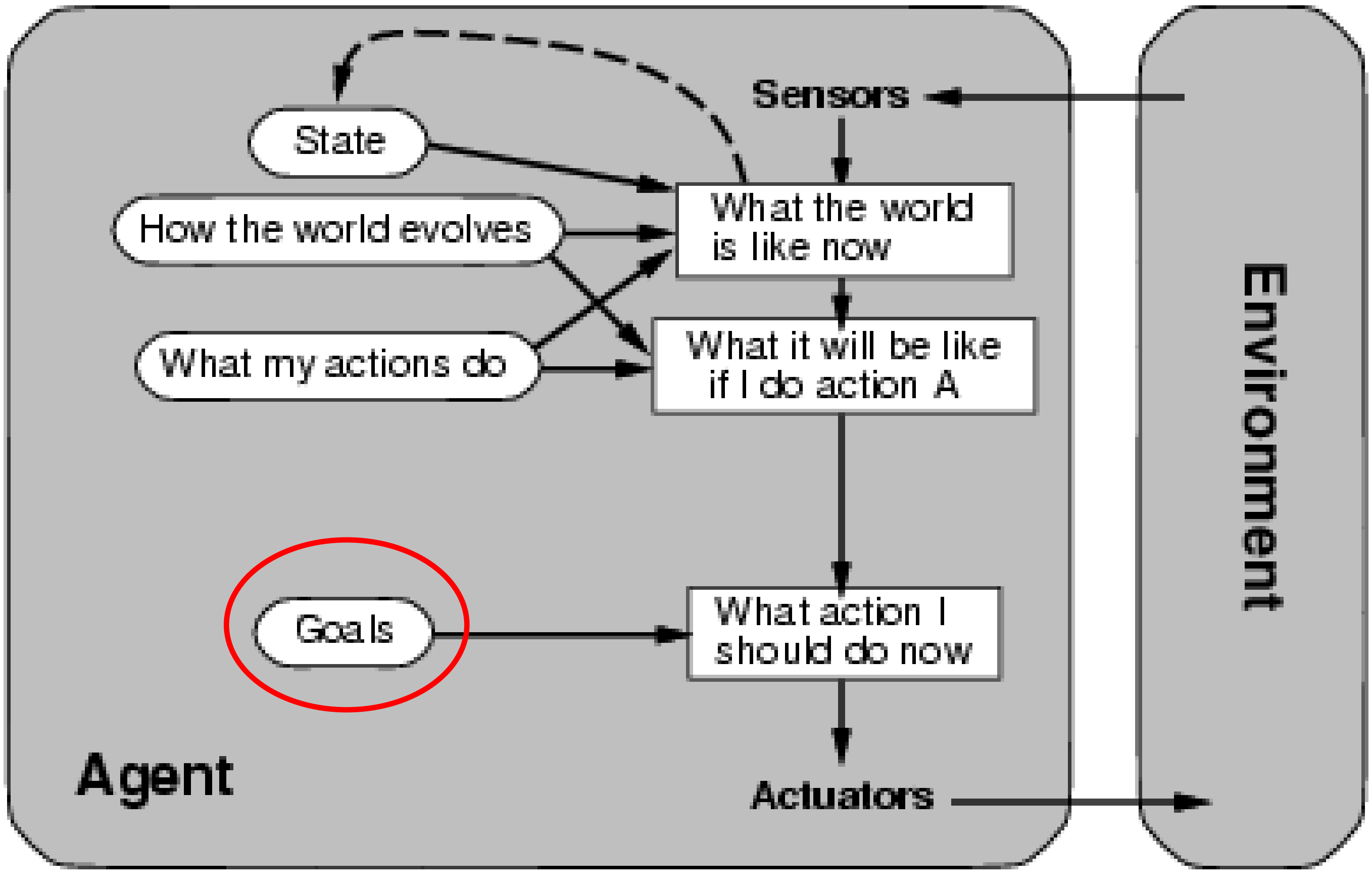
Model-Based Reflex Agent

The agent keeps track of the "**state**" of itself and the world. It knows how the state is changed.



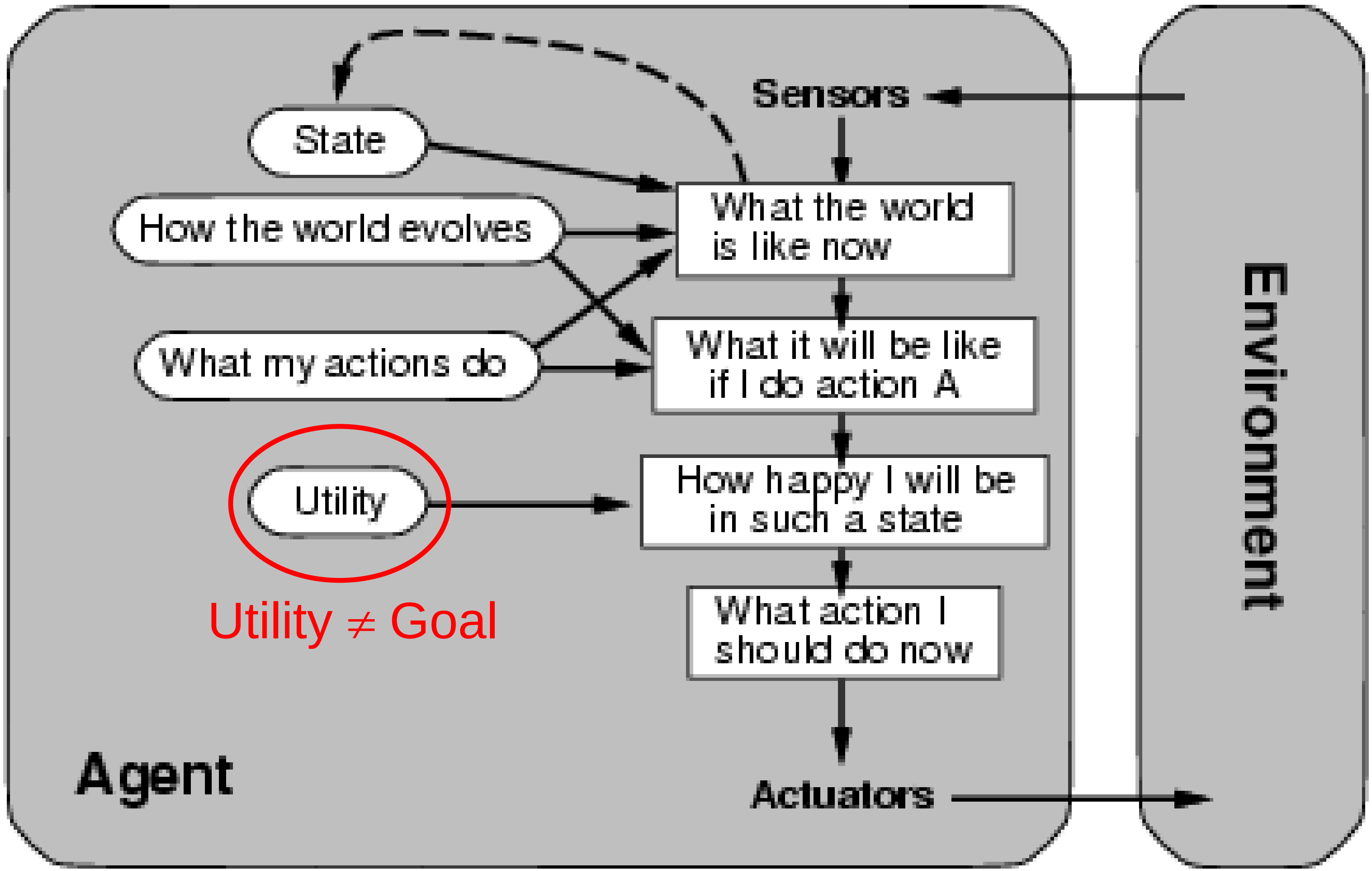
Goal-Based Agent

The agent aims to reach a "goal" state.
This affects how it selects the action.

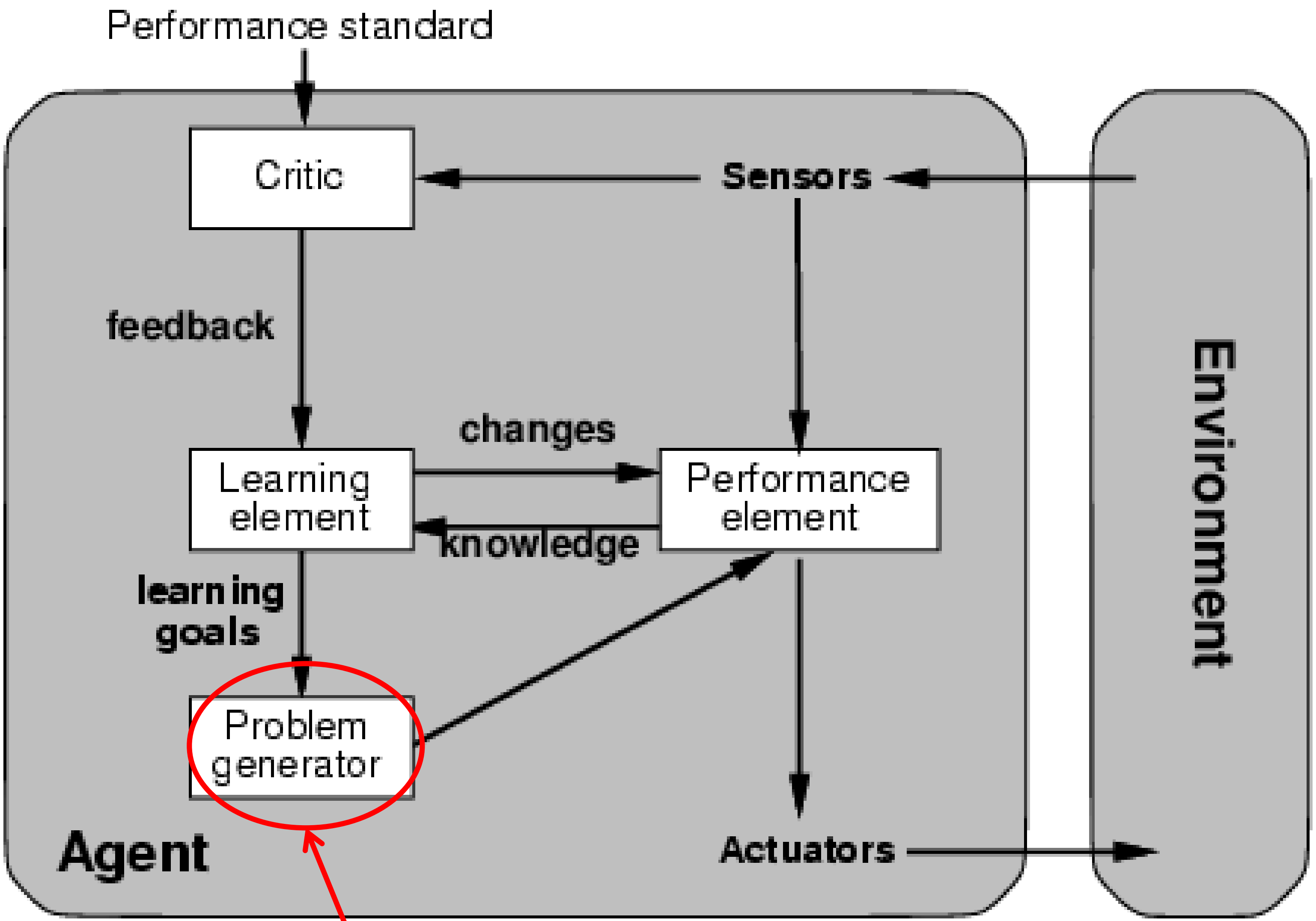


Utility-Based Agent

Utility is more general than goals.



Learning Agent



Exploration (not part of all learning agents)