Intelligent Agents

Chapter 2

ICS 279 Fall 09

Agents

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

  Human agent:

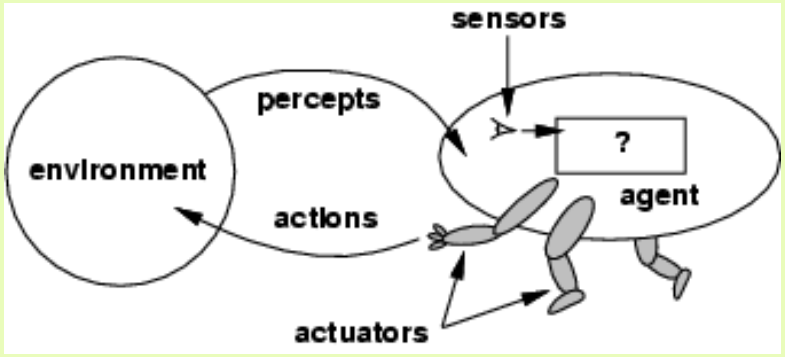
  eyes, ears, and other organs for sensors;

  hands, legs, mouth, and other body parts for actuators

- Robotic agent:

  cameras and infrared range finders for sensors;
  various motors for actuators

Agents and environments



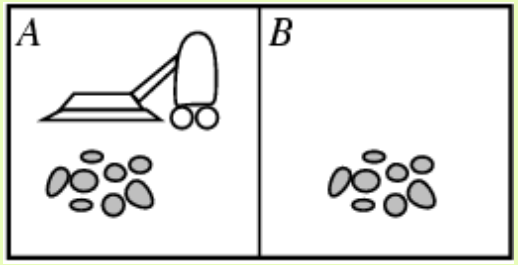
- The agent function maps from percept histories to actions:

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

- 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

- **Rational Agent: 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

- **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**

Discussion Items

- **An realistic agent has finite amount of computation and memory available. Assume an agent is killed because it did not have enough computation resources to calculate some rare eventually that ended up killing it. Can this agent still be rational?**

- **The Turing test was contested by Searle by using the "Chinese Room" argument. The Chinese Room agent needs an exponential large memory to work. Can we "save" the Turing test from the Chinese Room argument?**

Task Environment

- **Before we design an intelligent agent, we must specify its "task environment":**

PEAS:

Performance measure

Environment

Actuators

Sensors

PEAS

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

PEAS

- **Example: Agent = 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)

PEAS

- Example: Agent = 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

Environment types

- **Fully observable** (vs. **partially observable**): An agent's sensors give it access to the complete state of the environment at each point in time.

- **Deterministic** (vs. **stochastic**): The next state of the environment is completely determined by the current state and the action executed by the agent. (If the environment is deterministic except for the actions of other agents, then the environment is **strategic**)

- **Episodic** (vs. **sequential**): An agent's action is divided into atomic episodes. Decisions do not depend on previous decisions/actions.

Environment types

- **Static** (vs. **dynamic**): The environment is unchanged while an agent is deliberating. (The environment is **semidynamic** if the environment itself does not change with the passage of time but the agent's performance score does)

- **Discrete** (vs. **continuous**): A limited number of distinct, clearly defined percepts and actions.

  How do we **represent** or **abstract** or **model** the world?

- **Single agent** (vs. **multi-agent**): An agent operating by itself in an environment. Does the other agent interfere with my performance measure?

task environm.	observable	determ./ stochastic	episodic/ sequential	static/ dynamic	discrete/ continuous	agents
crossword puzzle	fully	determ.	sequential	static	discrete	single
chess with clock	fully	strategic	sequential	semi	discrete	multi
poker						
back gammon						
taxi driving	partial	stochastic	sequential	dynamic	continuous	multi
medical diagnosis	partial	stochastic	sequential	dynamic	continuous	single
image analysis	fully	determ.	episodic	semi	continuous	single
partpicking robot	partial	stochastic	episodic	dynamic	continuous	single
refinery controller	partial	stochastic	sequential	dynamic	continuous	single
interact. Eng. tutor	partial	stochastic	sequential	dynamic	discrete	multi

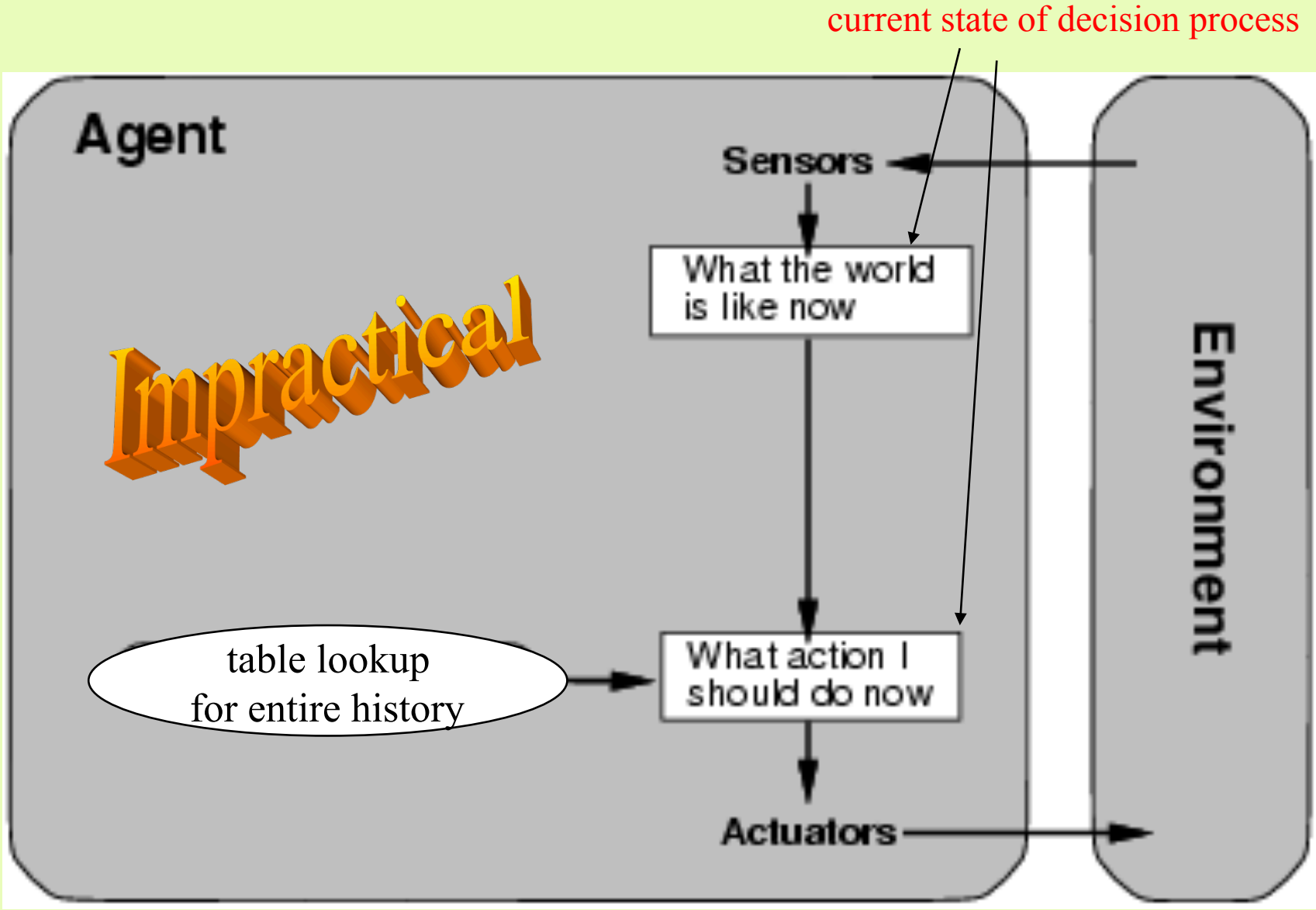
task environm.	observable	determ./ stochastic	episodic/ sequential	static/ dynamic	discrete/ continuous	agents
crossword puzzle	fully	determ.	sequential	static	discrete	single
chess with clock	fully	strategic	sequential	semi	discrete	multi
poker	partial	stochastic	sequential	static	discrete	multi
back gammon						
taxi driving	partial	stochastic	sequential	dynamic	continuous	multi
medical diagnosis	partial	stochastic	sequential	dynamic	continuous	single
image analysis	fully	determ.	episodic	semi	continuous	single
partpicking robot	partial	stochastic	episodic	dynamic	continuous	single
refinery controller	partial	stochastic	sequential	dynamic	continuous	single
interact. Eng. tutor	partial	stochastic	sequential	dynamic	discrete	multi

task environm.	observable	determ./ stochastic	episodic/ sequential	static/ dynamic	discrete/ continuous	agents
crossword puzzle	fully	determ.	sequential	static	discrete	single
chess with clock	fully	strategic	sequential	semi	discrete	multi
poker	partial	stochastic	sequential	static	discrete	multi
back gammon	fully	stochastic	sequential	static	discrete	multi
taxi driving	partial	stochastic	sequential	dynamic	continuous	multi
medical diagnosis	partial	stochastic	sequential	dynamic	continuous	single
image analysis	fully	determ.	episodic	semi	continuous	single
partpicking robot	partial	stochastic	episodic	dynamic	continuous	single
refinery controller	partial	stochastic	sequential	dynamic	continuous	single
interact. Eng. tutor	partial	stochastic	sequential	dynamic	discrete	multi

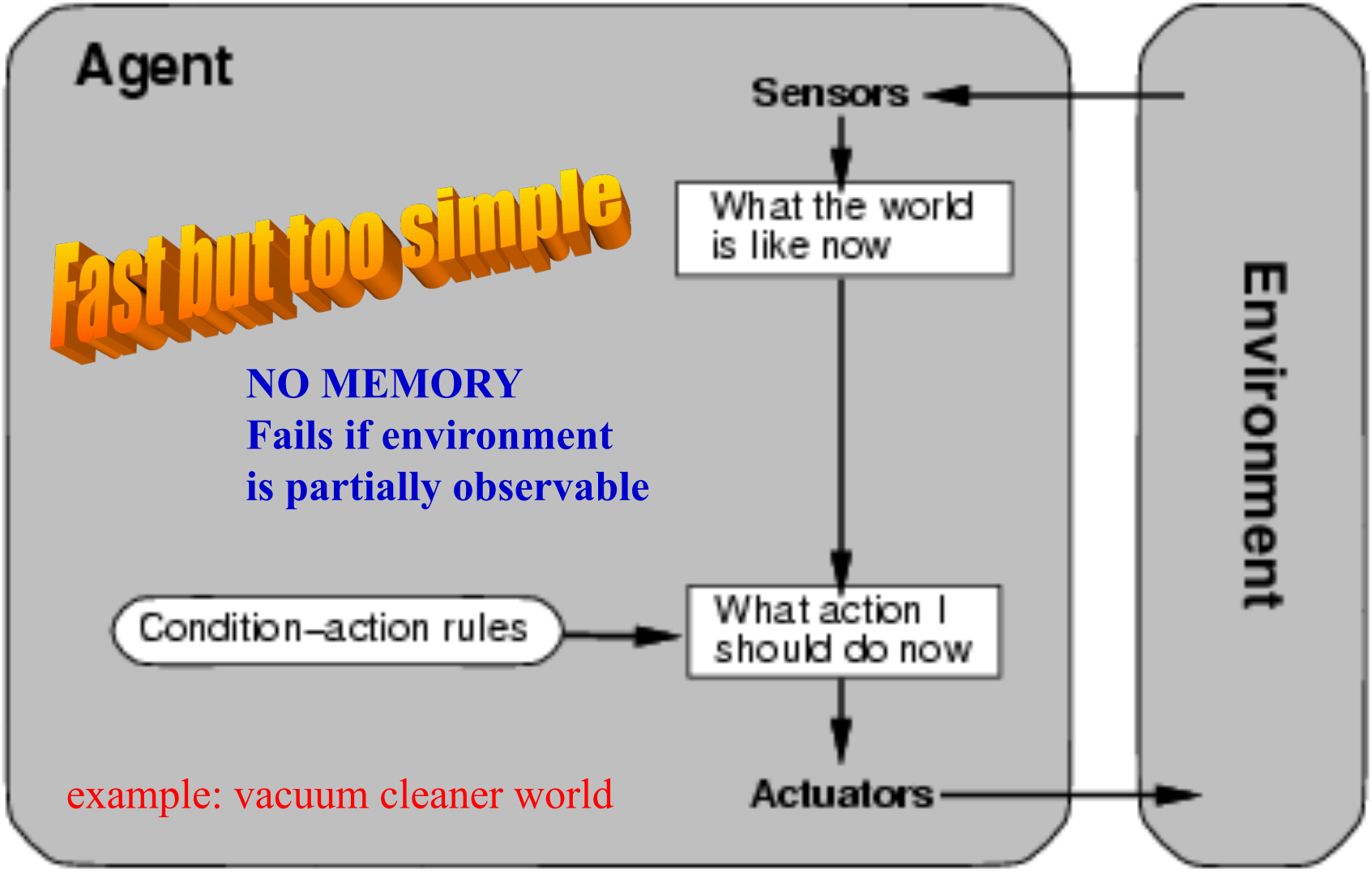
Agent types

- 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 Agent.



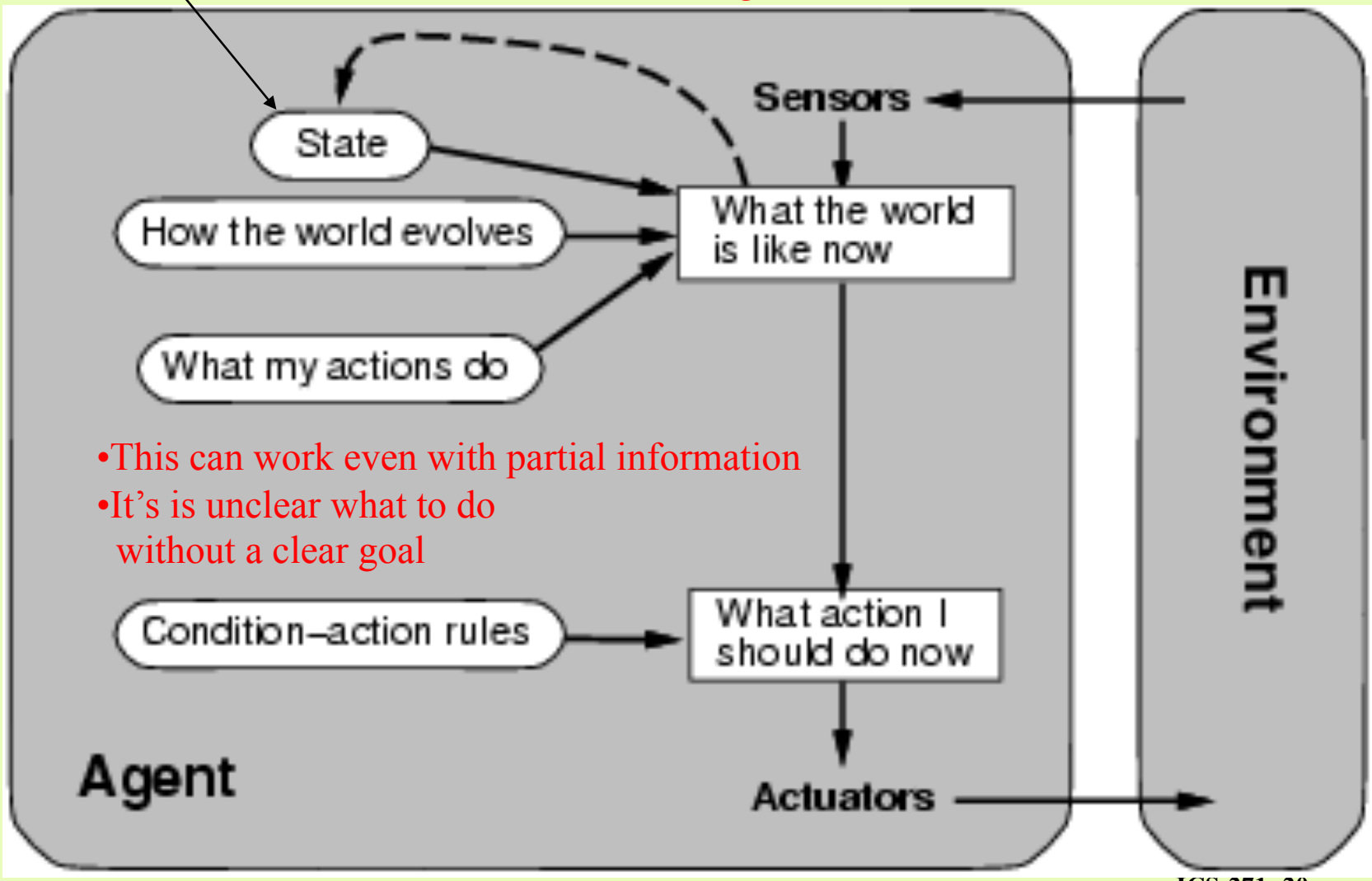
Simple reflex agents



Model-based reflex agents

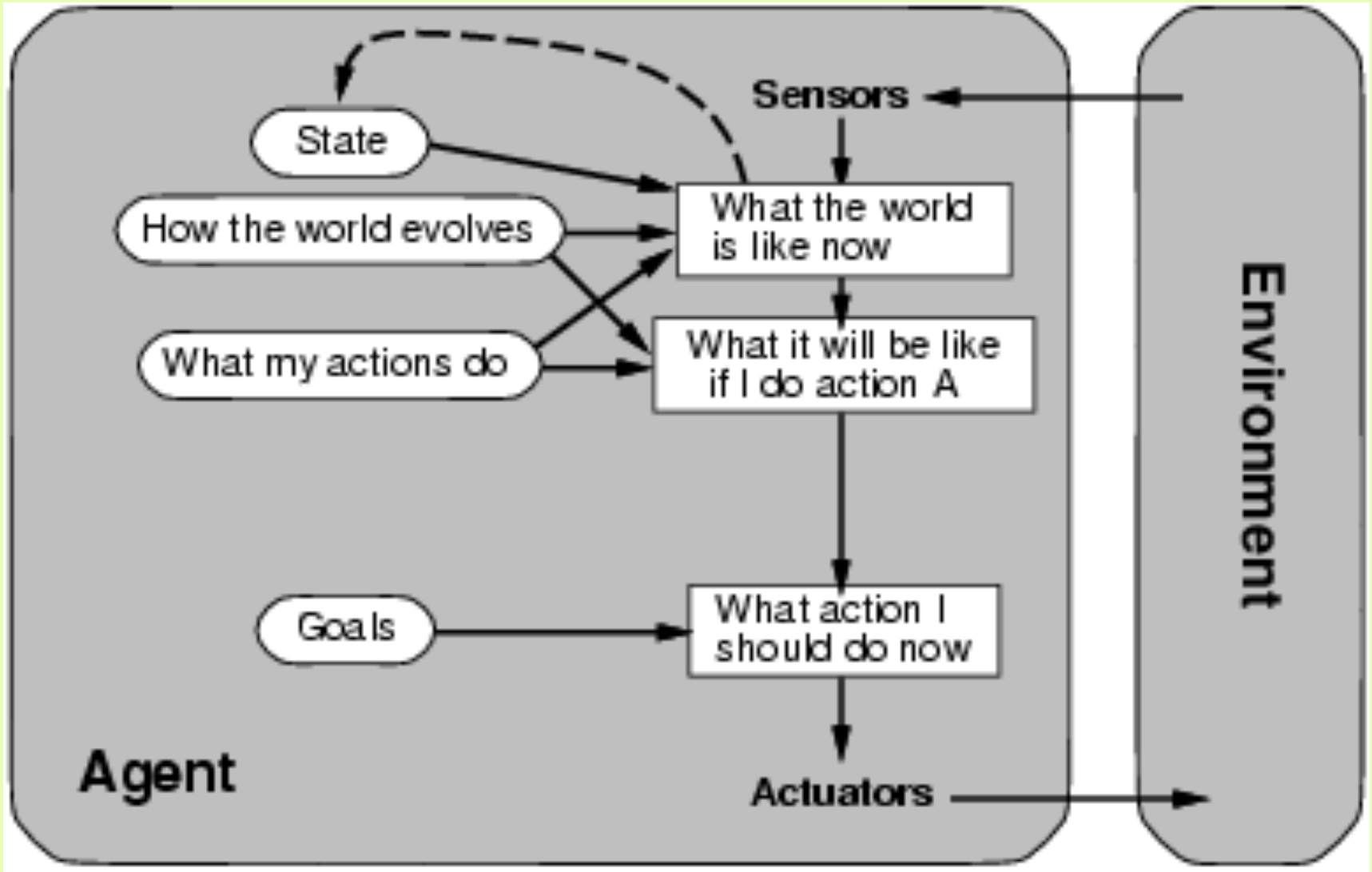
description of
current world state

Model the state of the world by:
modeling how the world chances
how it's actions change the world



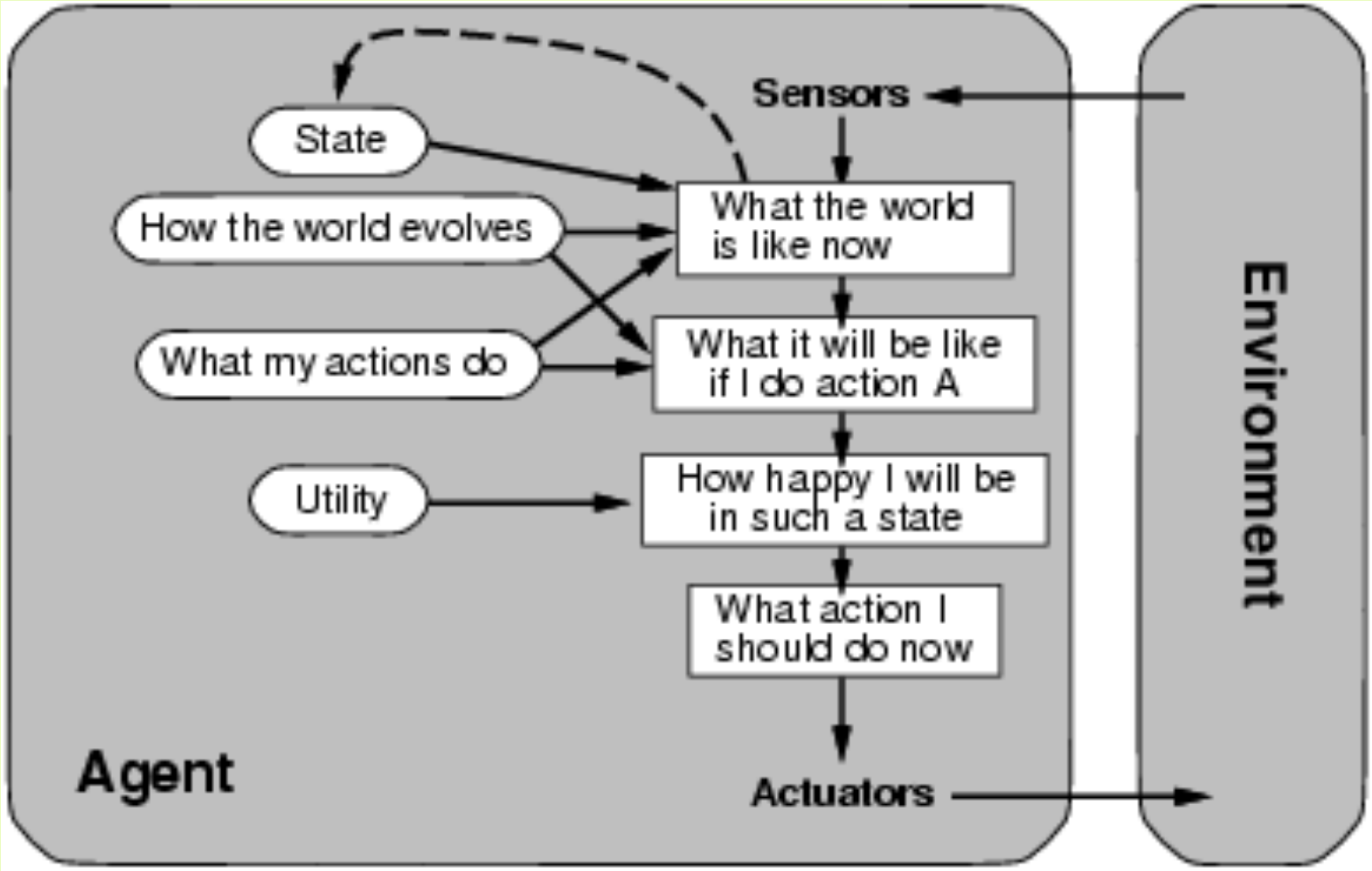
Goal-based agents

Goals provide reason to prefer one action over the other.
We need to predict the future: we need to plan & search



Utility-based agents

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