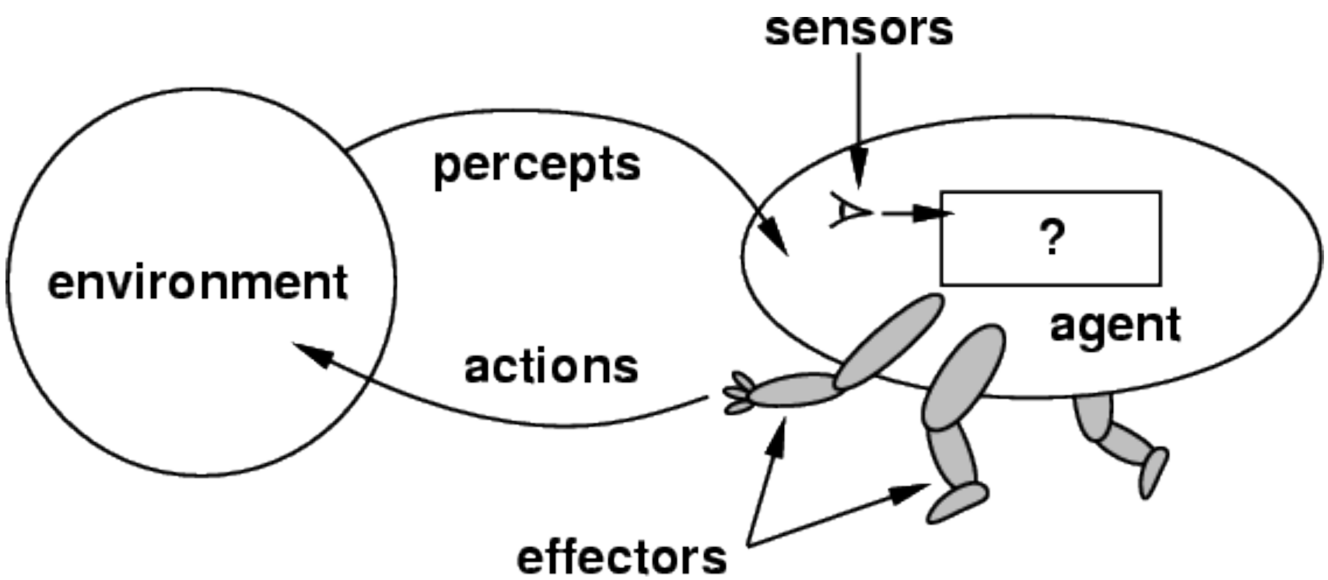
Artificial Intelligence
Class 2: Intelligent Agents



Slides adapted with thanks from: Dr. Marie desJardin

Bookkeeping

- Due last night:
  - Introduction survey
  - Academic integrity

  If you haven't done these, do!
- HW 1
  - Writing sections: 2 readings, 1 short essay, 6 questions
    - http://tiny.cc/mc-what-is-ai
    - http://ai100.stanford.edu/2016-report
  - Coding problems: out this afternoon
    - We will update on Piazza
- Due **11:59pm, 9/19**

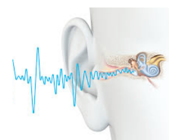
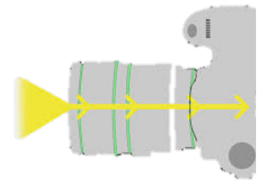
Today's Class

- What's an agent?
  - Definition of an agent
  - Rationality and autonomy
  - Types of agents
  - Properties of environments

Pre-Reading: Quiz

- What are sensors and percepts?



- What are actuators (aka effectors) and actions?



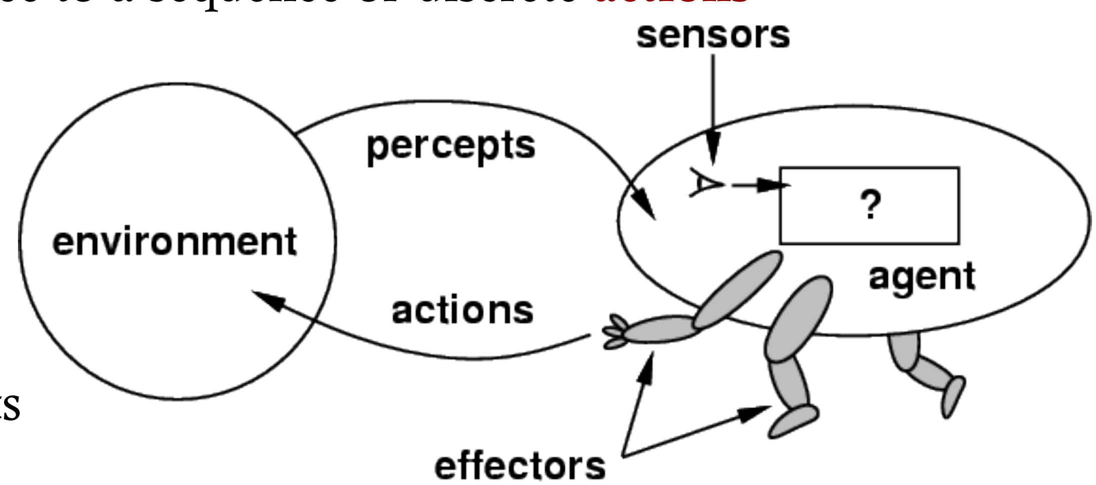


- What are the six environment characteristics that R&N use to characterize different problem spaces?

Observable	Deterministic	Static
# of Agents	Episodic	Discrete

How Do You Design an Agent?

- An intelligent agent:
  - Perceives its environment via **sensors**
  - Acts upon that environment with its **effectors** (or **actuators**)
- A discrete agent:
  - Receives **percepts** one at a time
  - Maps this percept sequence to a sequence of discrete **actions**
- Properties:
  - Autonomous
  - Reactive to the environment
  - Pro-active (goal-directed)
  - Interacts with other agents via the environment



Human Sensors/Percepts, Actuators/Actions

- Sensors:
  - Eyes (vision), ears (hearing), skin (touch), tongue (gustation), nose (olfaction), neuromuscular system (proprioception), …
- Percepts: "that which is perceived"
  - At the lowest level – electrical signals from these sensors
  - After preprocessing – objects in the visual field (location, textures, colors, …), auditory streams (pitch, loudness, direction), …
- Actuators/effectors:
  - Limbs, digits, eyes, tongue, …
- Actions:
  - Lift a finger, turn left, walk, run, carry an object, …

Human Sensors/Percepts, Actuators/Actions

- Sensors:
  - Eyes (vision), ears (hearing), skin (touch), tongue (gustation), nose (olfaction), neuromuscular system (proprioception), …
- Percepts: "that which is perceive
  - At the lowest level – electrical signa
  - After preprocessing – objects in the ,
    …), auditory streams (pitch, loudn
- Actuators/effectors:
  - Limbs, digits, eyes, tongue, …
- Actions:
  - Lift a finger, turn left, walk, run, carry an object, …

The Point:

- Percepts and actions need to be *carefully defined*
- Sometimes at different levels of abstraction!

E.g.: Automated Taxi

- **Percepts**: Video, sonar, speedometer, odometer, engine sensors, keyboard input, microphone, GPS, …
- **Actions**: Steer, accelerate, brake, horn, speak/display, …
- **Goals**: Maintain safety, reach destination, maximize profits (fuel, tire wear), obey laws, provide passenger comfort, …
- **Environment**: U.S. urban streets, freeways, traffic, pedestrians, weather, customers, …
- **Different aspects of driving may require different types of agent programs!**

Rationality

- An ideal **rational agent**, in every possible world state, does action(s) that **maximize its expected performance**
- Based on:
  - The percept sequence (world state)
  - Its knowledge (built-in and acquired)
- Rationality includes information gathering
  - If you don't know something, find out!
  - No "rational ignorance"
- Need a **performance measure**
  - False alarm (false positive) and false dismissal (false negative) rates, speed, resources required, effect on environment, constraints met, user satisfaction, …

Autonomy

- An autonomous system **determines its own behavior**
- But not if all its decisions are included in its design
  - I.e., all decisions are made by its designer according to *a priori* decisions
- Good autonomous agents need:
  - Enough built-in knowledge to survive
  - The ability to learn
- In practice this can be a bit slippery

Some Types of Agent (1)

1. Table-driven agents
   - Use a percept sequence/action table to find the next action
   - Implemented by a (large) **lookup table**
2. Simple reflex agents
   - Based on **condition-action rules**
   - Implemented with a **production system**
   - Stateless devices which do not have memory of past world states
3. Agents with memory
   - Have **internal state**
   - Used to keep track of past states of the world

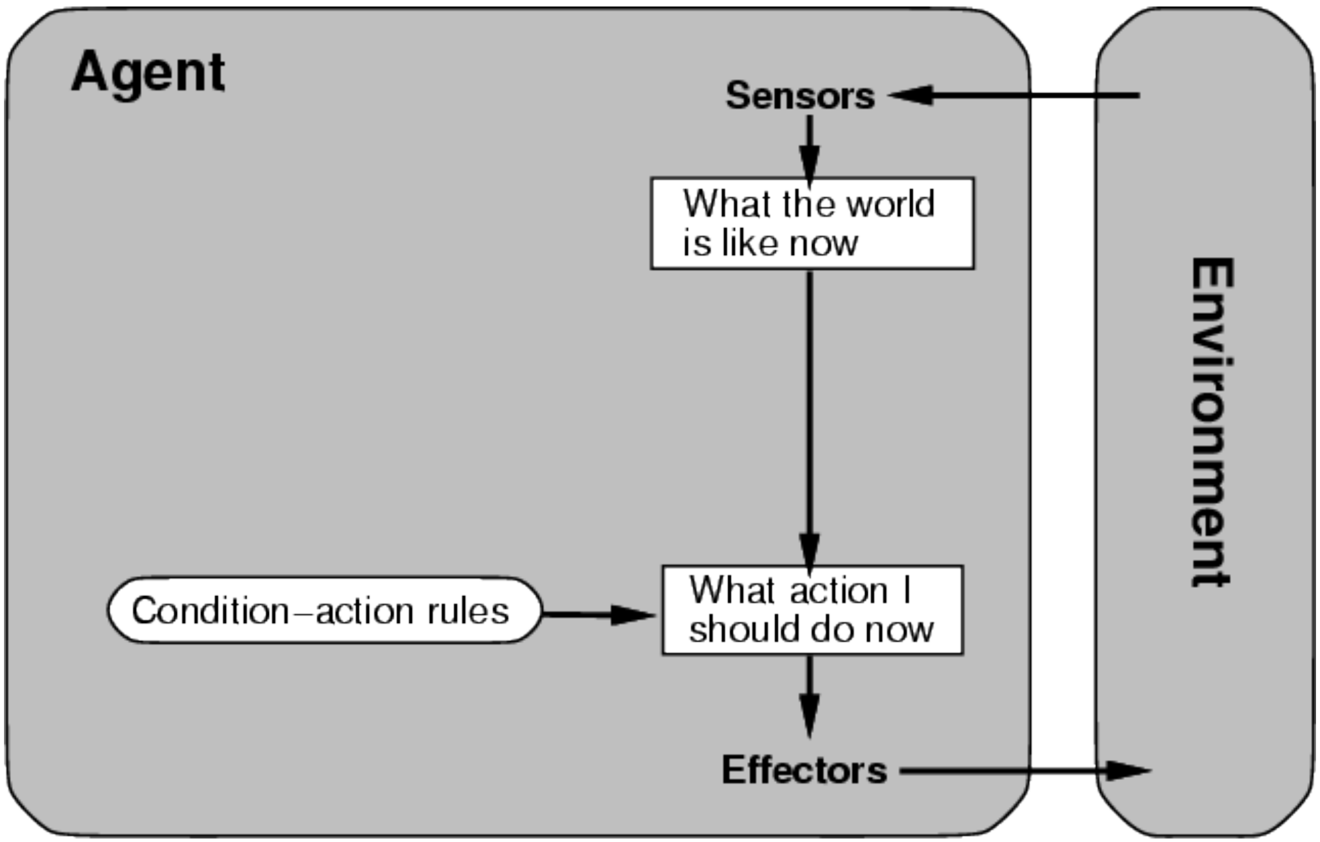
Some Types of Agent

4. Agents with goals
   - Have internal state information, plus
   - **Goal information** about desirable situations
   - Agents of this kind can **take future events** into consideration
5. Utility-based agents
   - Base their decisions on classic **axiomatic utility theory**
   - In order to **act rationally**

(1) Table-Driven Agents

- Table lookup of:
  - Percept-action pairs mapping
  - Every possible perceived state ←→ optimal action for that state
- Problems:
  - **Too big** to generate and store
    - Chess has about 10^{120} states, for example
  - No knowledge of **non-perceptual** parts of the current state
    - E.g., background knowledge
  - Not **adaptive to changes** in the environment
    - Change by updating entire table
  - No **looping**
    - Can't make actions conditional on previous actions/states

(1) Table-Driven/Reflex Agent



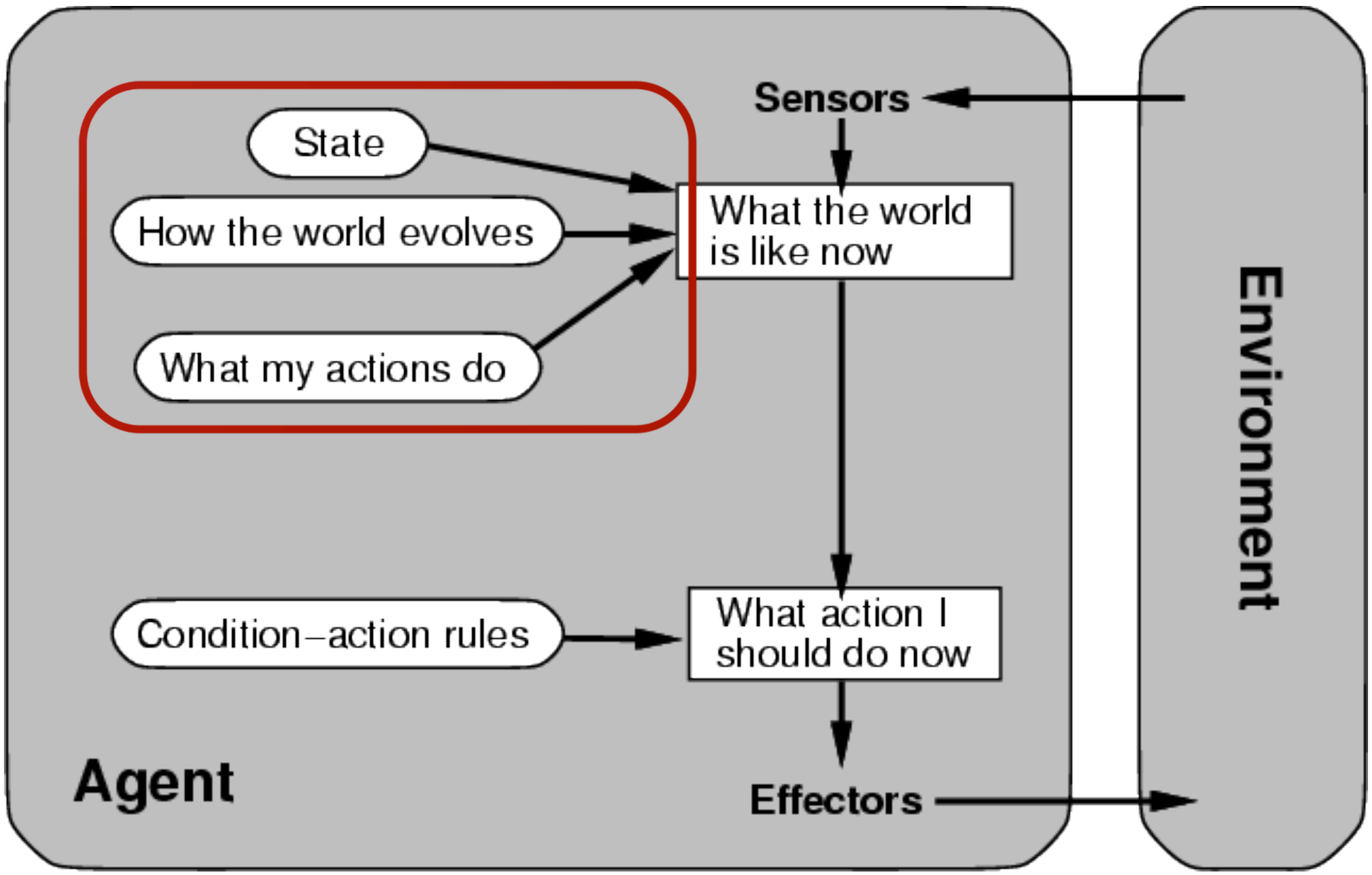
(2) Simple Reflex Agents

- **Rule-based reasoning**
  - To map from percepts to optimal action
  - Each rule handles a collection of perceived states
- Problems
  - Still usually too big to generate and to store
  - Still no knowledge of non-perceptual parts of state
  - Still not adaptive to changes in the environment
    - Change by updating collection of rules
  - Actions still not conditional on previous state

(3) Agents With Memory

- Encode "internal state" of the world
  - Used to remember the past (earlier percepts)
- Why?
  - Sensors rarely give the whole state of the world at each input
  - So, must build up environment model over time
  - "State" is used to encode different "world states"
  - Different worlds generate the same (immediate) percepts
- Requires ability to represent **change** in the world
  - Could represent just the latest state
  - But then can't reason about hypothetical courses of action
- Example: Rodney Brooks' Subsumption Architecture.

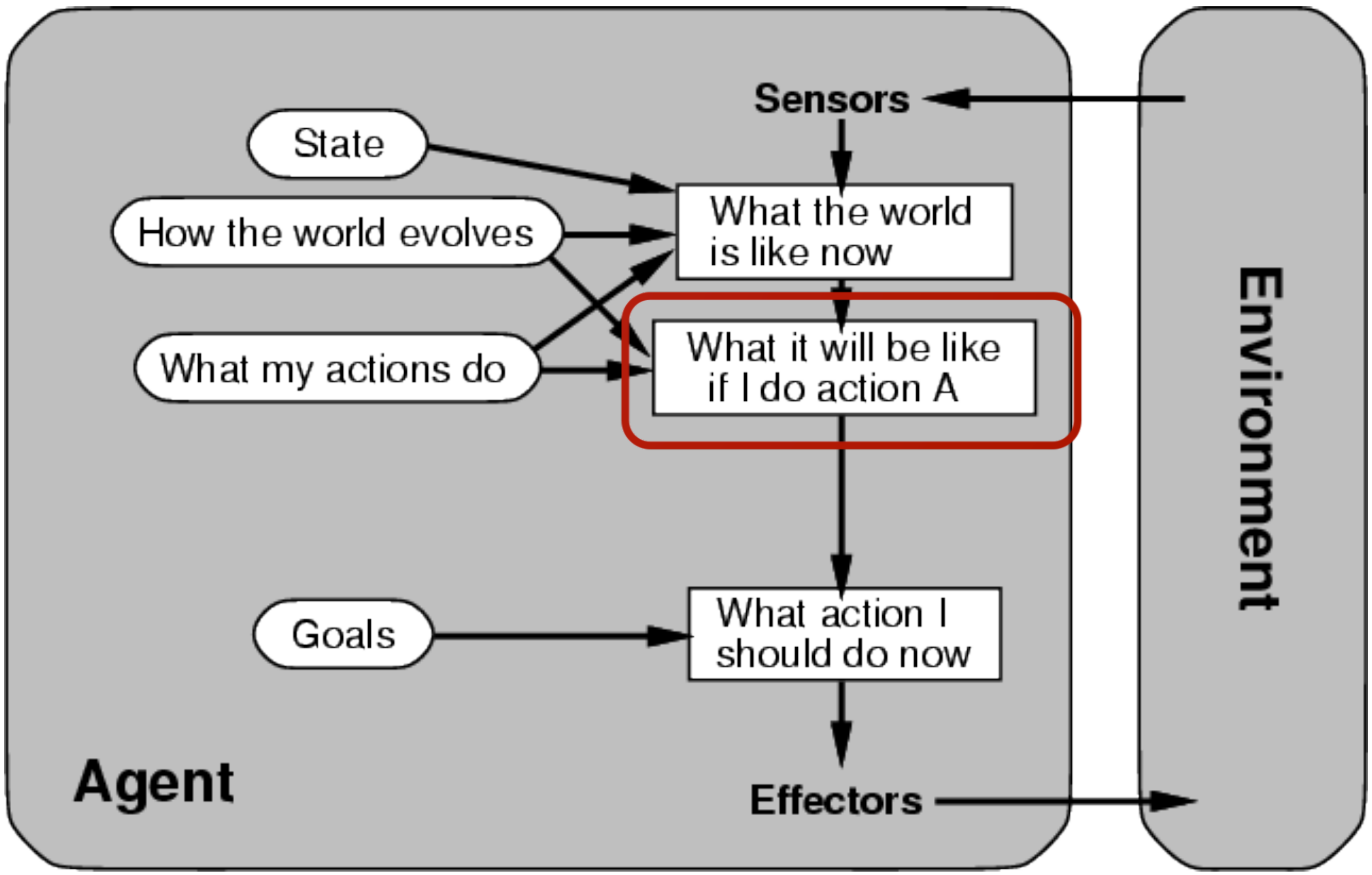
(3) Architecture for an Agent with Memory



(4) Goal-Based Agents

- Choose actions that achieve a goal
  - Which may be given, or computed by the agent
- A goal is a **description of a desirable state**
  - Need goals to decide what situations are "good"
  - Keeping track of the current state is often not enough
- Deliberative instead of reactive
  - Must consider sequences of actions to get to goal
  - Involves thinking about the future
  - *"What will happen if I do...?"*

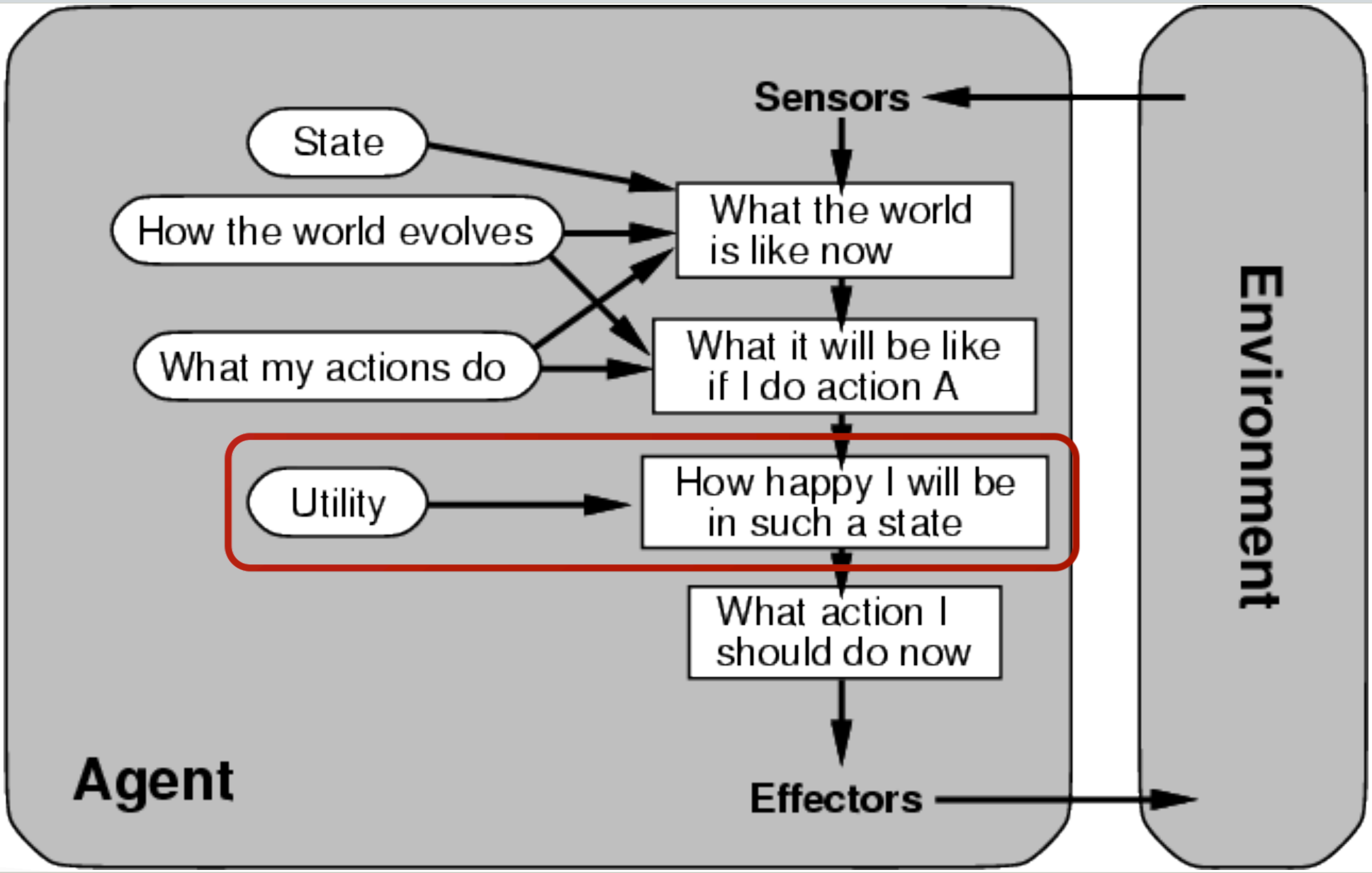
(4) Architecture for Goal-Based Agent



(5) Utility-Based Agents

- How to choose from multiple alternatives?
  - What action is best?
- What state is best?
  - Goals → crude distinction between "happy" / "unhappy" states
  - Often need a more general performance measure (how "happy"?)
- Utility function gives success or happiness at a given state
- Can compare choice between:
  - Conflicting goals
  - Likelihood of success
  - Importance of goal (if achievement is uncertain)

(4) Architecture for a complete utility-based agent



Properties of Environments

These should be familiar!

- **Fully observable/Partially observable**
  - If an agent's sensors give it access to the **complete state of the environment**, the environment is **fully observable**
  - Such environments are convenient
    - No need to keep track of the changes in the environment
    - No need to guess or reason about non-observed things
  - Such environments are also rare in practice

Properties of Environments

- **Deterministic/Stochastic**.
  - An environment is **deterministic** if:
    - The next state of the environment is completely determined by
      - The current state of the environment
      - The action of the agent
  - In a **stochastic** environment, there are multiple, unpredictable outcomes.
  - In a fully observable, deterministic environment, the agent need not deal with uncertainty.

Properties of Environments

- **Episodic/Sequential**
  - An **episodic** environment means that subsequent episodes do not depend on what actions occurred in previous episodes.
  - In a **sequential** environment, the agent engages in a series of connected episodes.
  - Such environments do not require the agent to plan ahead.
- **Static/Dynamic**
  - A **static** environment does not change while the agent is thinking.
  - The passage of time as an agent deliberates is irrelevant.
  - The agent doesn't need to observe the world during deliberation.

Properties of Environments III

- **Discrete/Continuous**
  - If the number of distinct percepts and actions is limited, the environment is **discrete**, otherwise it is **continuous**.
- **Single agent/Multi-agent**
  - If the environment contains other intelligent agents, the agent needs to be concerned about strategic, game-theoretic aspects of the environment (for either cooperative *or* competitive agents)
  - Most engineering environments don't have multi-agent properties, whereas most social and economic systems get their complexity from the interactions of (more or less) rational agents.

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire						
Backgammon						
Taxi driving						
Internet shopping						
Medical diagnosis						

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	**No**	**Yes**	**Yes**	**Yes**	**Yes**	**Yes**
Backgammon						
Taxi driving						
Internet shopping						
Medical diagnosis						

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	No	Yes	Yes	Yes	Yes	Yes
Backgammon	**Yes**	**No**	**No**	**Yes**	**Yes**	**No**
Taxi driving						
Internet shopping						
Medical diagnosis						

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	No	Yes	Yes	Yes	Yes	Yes
Backgammon	Yes	No	No	Yes	Yes	No
Taxi driving	**No**	**No**	**No**	**No**	**No**	**No**
Internet shopping						
Medical diagnosis						

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	No	Yes	Yes	Yes	Yes	Yes
Backgammon	Yes	No	No	Yes	Yes	No
Taxi driving	No	No	No	No	No	No
Internet shopping	**No**	**No**	**No**	**No**	**Yes**	**No**
Medical diagnosis						

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	No	Yes	Yes	Yes	Yes	Yes
Backgammon	Yes	No	No	Yes	Yes	No
Taxi driving	No	No	No	No	No	No
Internet shopping	No	No	No	No	Yes	No
Medical diagnosis	**No**	**No**	**No**	**No**	**No**	**Yes**

Characteristics of Environments

	Fully observable?	Deterministic?	Episodic?	Static?	Discrete?	Single agent?
Solitaire	No	Yes	Yes	Yes	Yes	Yes
Backgammon	Yes	No	No	Yes	Yes	No
Taxi driving	No	No	No	No	No	No
Internet shopping	No	No	No	No	Yes	No
Medical diagnosis	**No**	**No**	**No**	**No**	**No**	**Yes**

→ Lots of (most?) real-world domains fall into the hardest case! ←

Summary: Agents

- An **agent:**
  - Perceives and acts in an environment
  - Has an architecture
  - Is implemented by an agent program(s)
- An **ideal agent:**
  - Always chooses the "right" action
    - Which is, that which maximizes its expected performance
  - Given its percept sequence so far!
- An **autonomous agent:**
  - Uses its *own experience* to learn and make decisions
  - Rather than built-in knowledge
  - I.e., *a priori* world knowledge by the designer

Summary: Agents

- **Representing knowledge** is important for successful agent design
  - Percepts, actions and their effects, constraints, …
- The most challenging environments are:
  - Partially observable
  - Stochastic
  - Sequential
  - Dynamic
  - Continuous
  - Contain multiple intelligent agents