# CMSC 471
# Artificial Intelligence
# Spring 2021

Frank Ferraro – ferraro@umbc.edu

# Recap from Last Class: What is AI?

Q. What is artificial intelligence?

A. It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable.

http://www-formal.stanford.edu/jmc/whatisai/

Courtesy Tim Finin

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." -- Poole & Mackworth

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." -- Poole & Mackworth

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

something that acts in an environment; it does something.

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

???

something that acts in an environment; it does something.

# Recap: intelligence?

Q. Yes, but what is intelligence?

A. Intelligence is the computational part of the ability to achieve goals in the world. Varying kinds and degrees of intelligence occur in people, many animals and some machines

http://www-formal.stanford.edu/jmc/whatisai/

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals

something that acts in an environment; it does something.

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals
- Balances short & long-term appropriately

something that acts in an environment; it does something.

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals
- Balances short & long-term appropriately
- Flexible & reactive

something that acts in an environment; it does something.

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals
- Balances short & long-term appropriately
- Flexible & reactive
- Learns/recognizes patterns
- Aware of computational/task budgets & limitations

something that acts in an environment; it does something.

# Okay, but really? What is AI?

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals
- Balances short & long-term appropriately
- Flexible & reactive
- Learns/recognizes patterns
- Aware of computational/task budgets & limitations

Use "computation" to explain and traceback the actions

something that acts in an environment; it does something.

# Possible AI approaches

| | Like humans | Well |
|---|---|---|
| Think | | |
| Act | | |

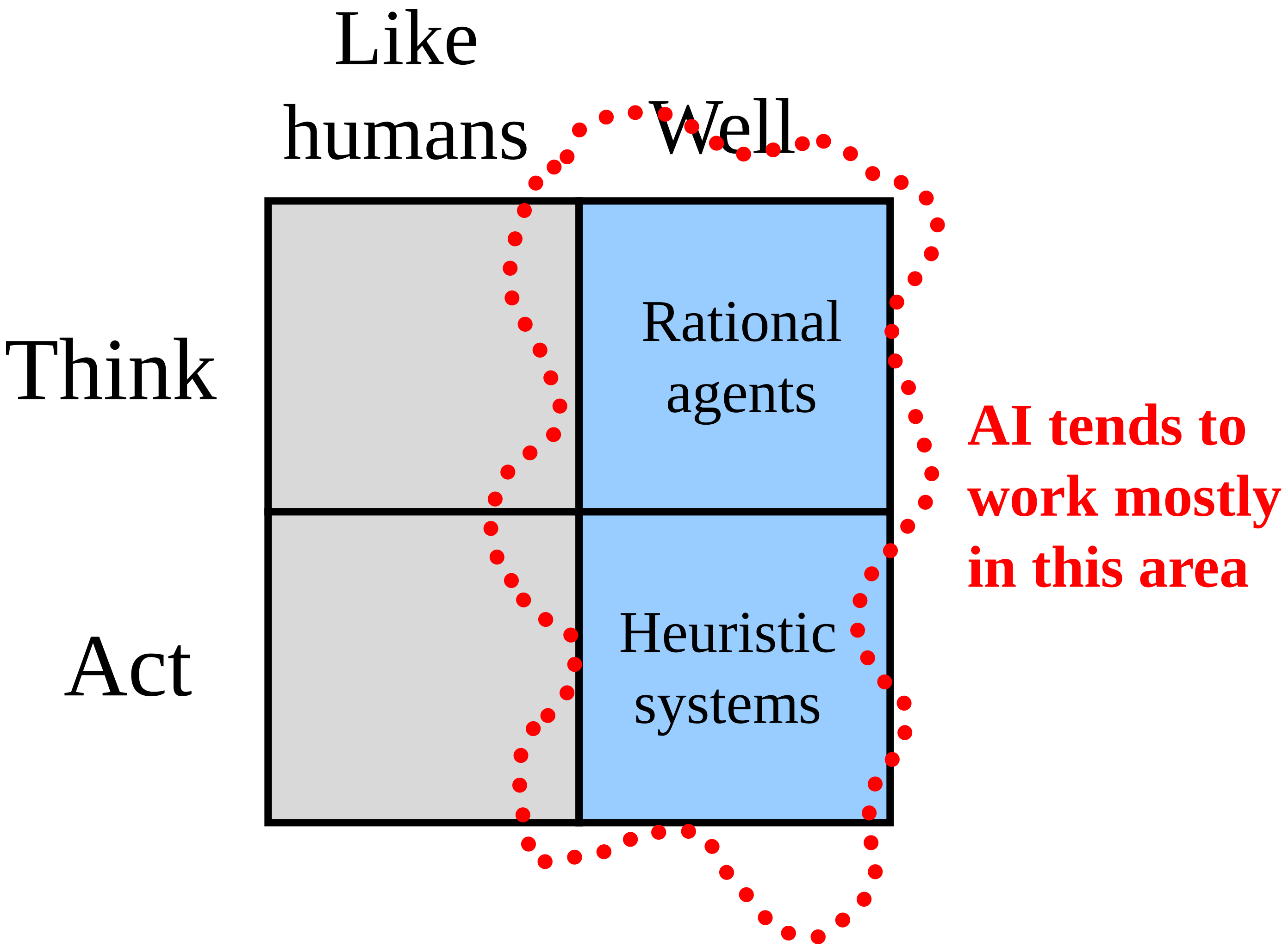
Courtesy Tim Finin

# Possible approaches

| | Like humans | Well |
|---|---|---|
| Think | | Rational agents |
| Act | | Heuristic systems |

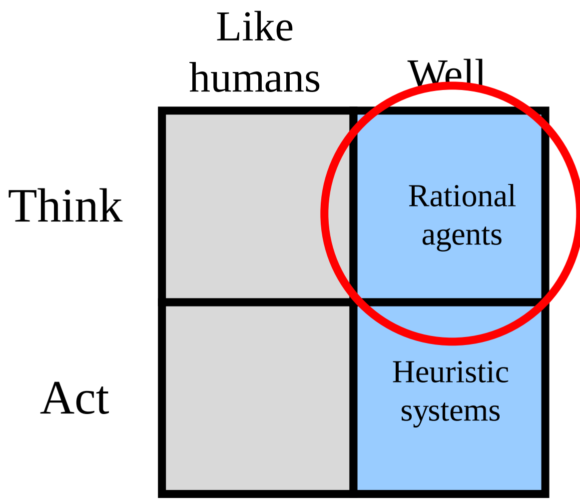
Courtesy Tim Finin

# Possible approaches



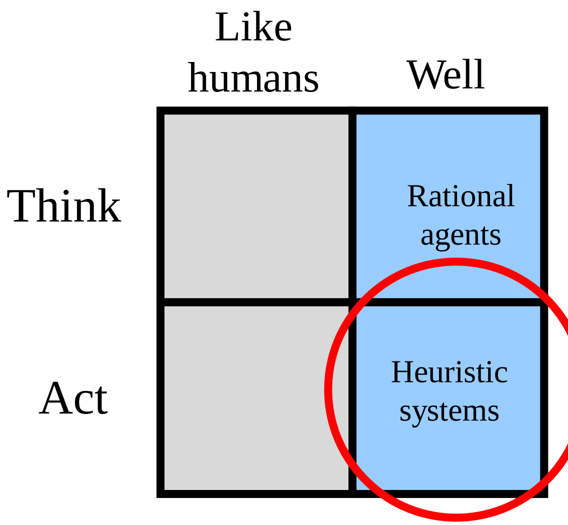

Courtesy Tim Finin

# Think well




- Develop formal models of knowledge representation, reasoning, learning, memory, problem solving, that can be rendered in algorithms
- Often an emphasis on systems that are provably correct, and guarantee finding an optimal solution

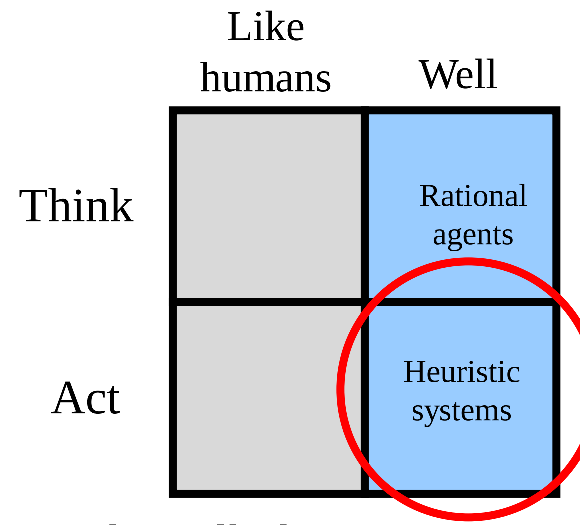
# Act well




- For a given set of inputs, generate output that's not necessarily correct but gets job done

- A **heuristic** (heuristic rule, heuristic method) is a rule of thumb, strategy, trick or simplification which drastically limits search for solutions in large problem spaces

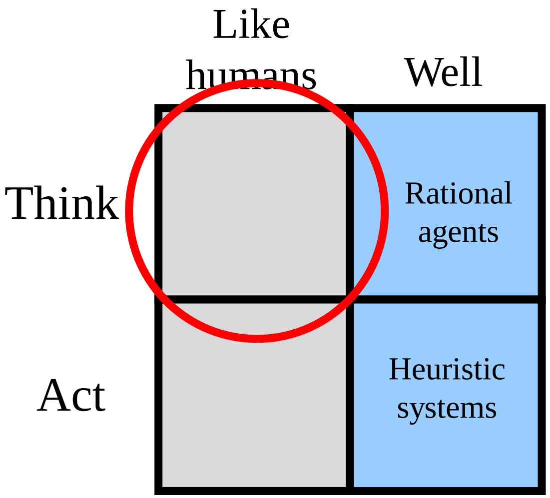
# Act well




- For a given set of inputs, generate output that's not necessarily correct but gets job done
- A **heuristic** (heuristic rule, heuristic method) is a rule of thumb, strategy, trick or simplification which drastically limits search for solutions in large problem spaces
- Heuristics don't guarantee optimal solutions or even any solution at all: "**all that can be said for a useful heuristic is that it offers solutions which are good enough most of the time**"
  - –Feigenbaum and Feldman, 1963, p. 6

Courtesy Tim Finin

# Think like humans



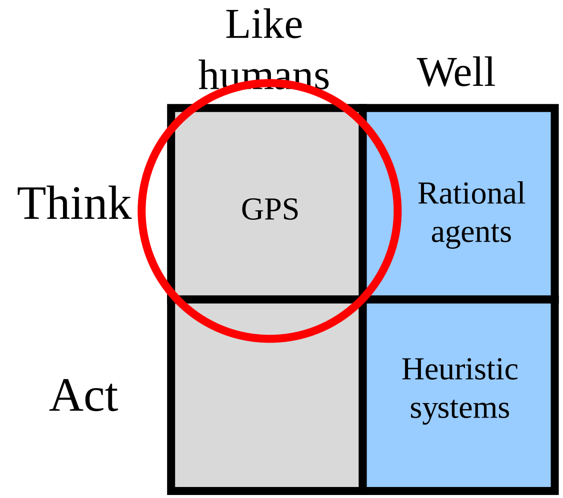

- Cognitive science approach

- Focus not just on behavior & I/O
  but also look at reasoning process

- Computational model should reflect "how" results were obtained

Courtesy Tim Finin

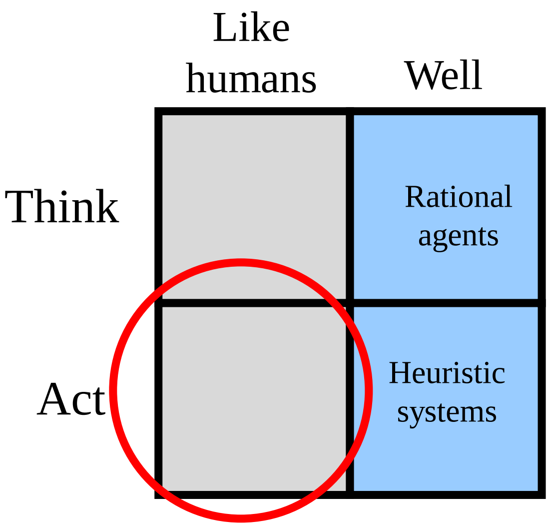
# Think like humans




- Cognitive science approach
- Focus not just on behavior & I/O but also look at reasoning process
- Computational model should reflect "how" results were obtained
- Provides new language for expressing cognitive theories & new mechanisms for evaluating them
- GPS (**General Problem Solver**): Goal not just to produce humanlike behavior, but to produce a sequence of steps of reasoning process that was similar to those followed by a person

Courtesy Tim Finin

# Act like humans




- Behaviorist approach
- Not interested in how you get results, just similarity to what human results are
- Exemplified by the Turing Test (Alan Turing, 1950)
- Has applications in interactive enter-tainment (e.g., computer games, CGI), virtual worlds and in modeling human intentions

# What’s easy and what’s hard?

- Easy: many high-level tasks usually associated with “intelligence” in people
  - e.g., symbolic integration, proving theorems, playing chess, medical diagnosis
- Hard: tasks many animals can do
  - walking around without running into things
  - catching prey and avoiding predators
  - Interpreting sensory info. (e.g., visual, aural, …)
  - modeling internal states of other from behavior
  - working as a team (e.g., with pack animals)
- Is there a fundamental difference between these?

Courtesy Tim Finin

# AI: Computational Agents Acting Intelligently

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." --Poole & Mackworth

- Makes appropriate actions for circumstances & goals
- Balances short & long-term appropriately
- Flexible & reactive
- Learns/recognizes patterns
- Aware of computational/task budgets & limitations

Use "computation" to explain and traceback the actions

something that acts in an environment; it does something.

# Agents

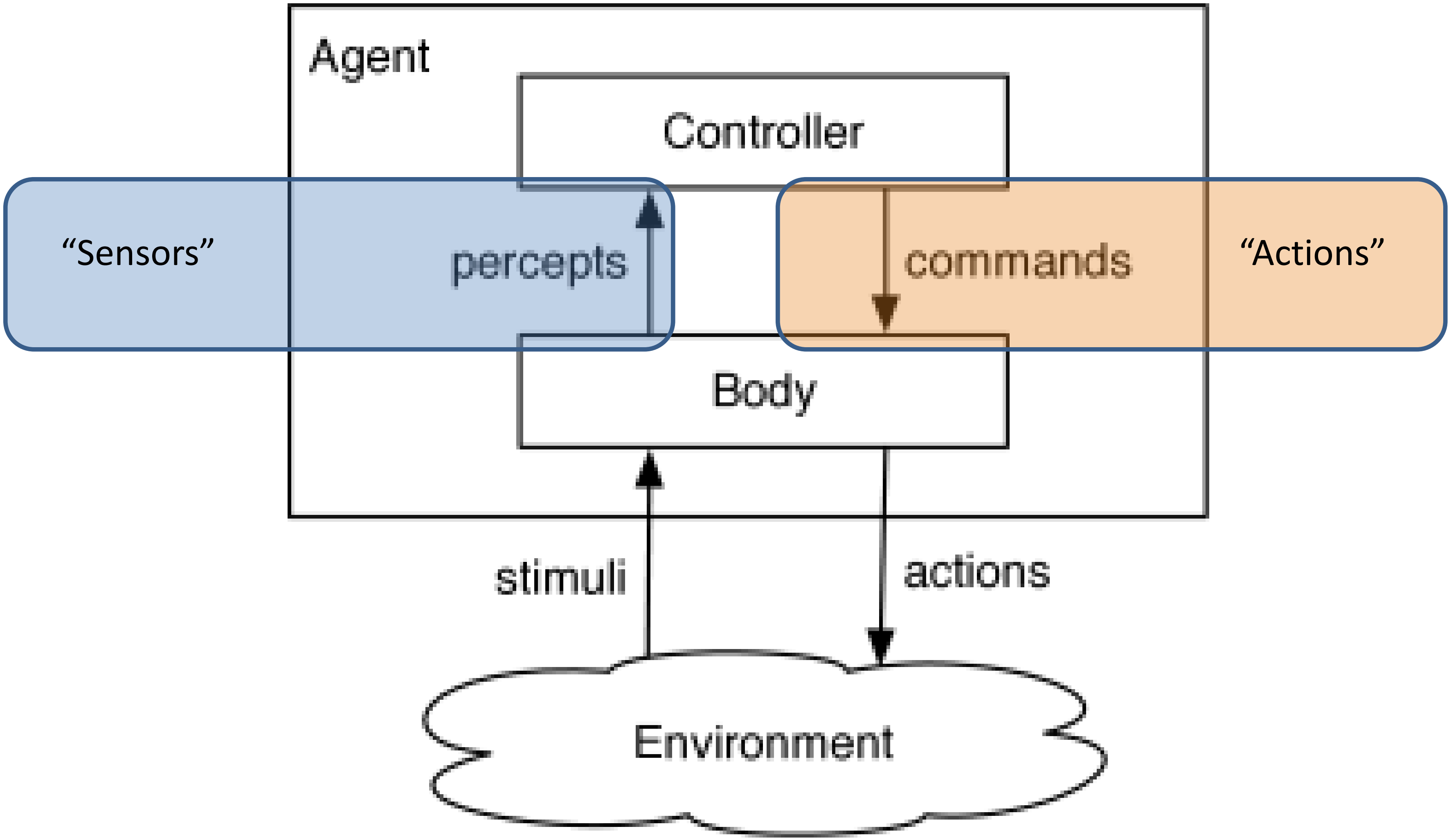



Fig. 2.1

# Agents

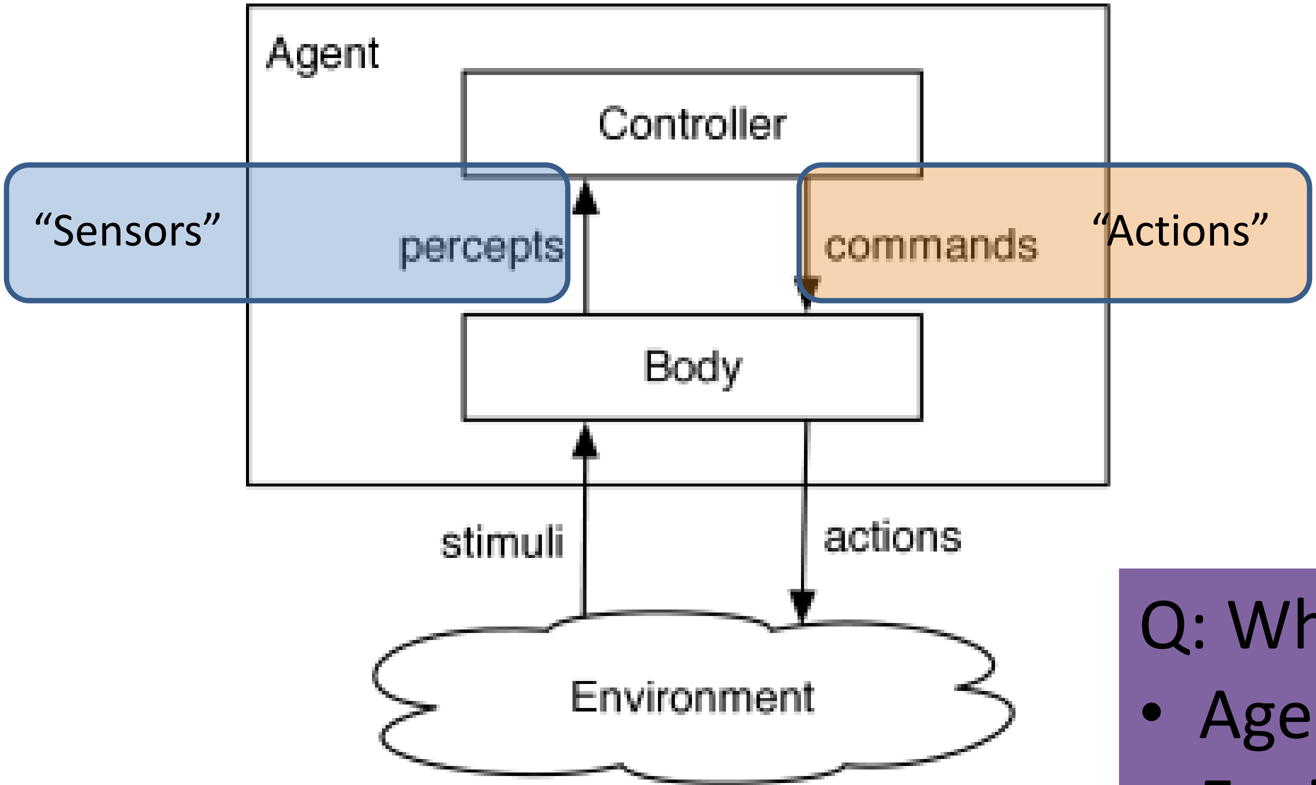



Fig. 2.1

Q: What's an example:
- Agent
- Environment
- Body
- Sensor(s)
- Action(s)

# Agent Desiderata

"Artificial intelligence, or AI, is the field that studies the synthesis and analysis of computational agents that act intelligently." -- Poole & Mackworth

- Rationality
- Autonomy
- Memory/Persistence
- Explainability
- … (any others?)

# Rationality

Ideal **rational agents** should, for each input, act
to maximize expected performance measure based on

(1) percept sequence, and

(2) its built-in and acquired knowledge

Courtesy Tim Finin

# Rationality

- Ideal **rational agents** should, for each input, act to maximize expected performance measure based on
  (1) percept sequence, and
  (2) its built-in and acquired knowledge

- Rationality includes *information gathering* -- If you don't know something, find out!

Courtesy Tim Finin

# Rationality

- Ideal **rational agents** should, for each input, act to maximize expected performance measure based on
  (1) percept sequence, and
  (2) its built-in and acquired knowledge

- Rationality includes *information gathering* -- If you don’t know something, find out!

- Rationality $\rightarrow$ Need a *performance measure* to say how well a task has been achieved

Courtesy Tim Finin

# Rationality

- Ideal **rational agents** should, for each input, act to maximize expected performance measure based on
  (1) percept sequence, and
  (2) its built-in and acquired knowledge
- Rationality includes *information gathering* -- If you don't know something, find out!
- Rationality → Need a *performance measure* to say how well a task has been achieved
- Types of performance measures: false alarm (false positive) & false dismissal (false negative) rates, speed, resources required, effect on environment, …

Courtesy Tim Finin

# Another View on Rationality: Qualitative **&** Quantitative Reasoning

- Quantitative Reasoning: Reasoning (drawing inferences and conclusions) based on numeric values
- Qualitative Reasoning: Reasoning (drawing inferences and conclusions) via logic and/or “fuzzy” values
  - Landmarks: what happens flipping an **empty** vs. **full** cup?
  - Order-of-magnitude reasoning: a cup 55% full vs. 45% can both be **partially full**
  - Qualitative derivatives: is a value increasing or decreasing?

# Autonomy

- A system is autonomous to extent that its behavior is determined by its experience

Courtesy Tim Finin

# Autonomy

- A system is autonomous to extent that its behavior is determined by its experience
- A system isn't autonomous if guided by its designer according to *a priori* *decisions*

Courtesy Tim Finin

# Autonomy

- A system is autonomous to extent that its behavior is determined by its experience
- A system isn't autonomous if guided by its designer according to *a priori decisions*
- An autonomous agent can always say "no"

Courtesy Tim Finin

# Autonomy

- A system is autonomous to extent that its behavior is determined by its experience
- A system isn't autonomous if guided by its designer according to *a priori decisions*
- An autonomous agent can always say "no"
- To survive, agents must have:
  - Enough built-in knowledge to survive
  - The ability to LEARN

Courtesy Tim Finin

## (0) Table-driven agents

Use percept sequence/action table to find next action. Implemented by a **lookup table**

## (1) Simple reflex agents

Based on **condition-action rules**, stateless devices with no memory of past world states

## (2) Agents with memory

have **represent states** and keep track of past world states

## (3) Agents with goals

Have a state and **goal information** describing desirable situations; can take future events into consideration

## (4) Utility-based agents

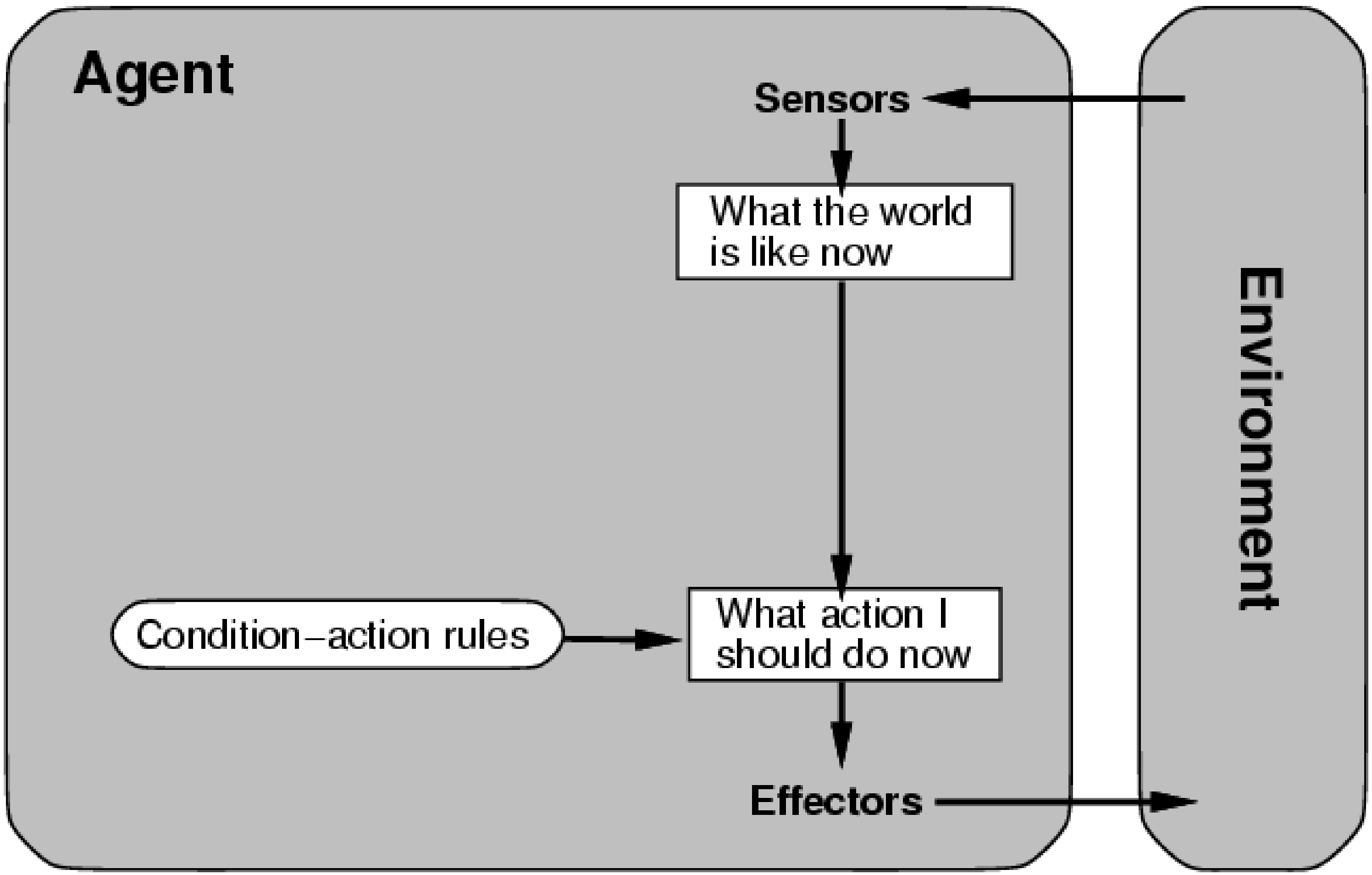
base decisions on **utility theory** in order to act rationally

simple

complex

# (0/1) Table-driven/reflex agent architecture

Use percept sequence/action table to find the next action.
Implemented by a (large) **lookup table**




Courtesy Tim Finin

# (0) Table-driven agents

**Table lookup** of percept-action pairs mapping from every possible perceived state to optimal action for it

**Problems:**

- Too big to generate and to store (e.g., chess has about $10^{120}$ states)
- No knowledge of non-perceptual parts of the current state (e.g., desirability)
- Not adaptive to changes in the environment; entire table must be updated if changes occur
- Looping: Can't make actions conditional on previous actions/states

Courtesy Tim Finin

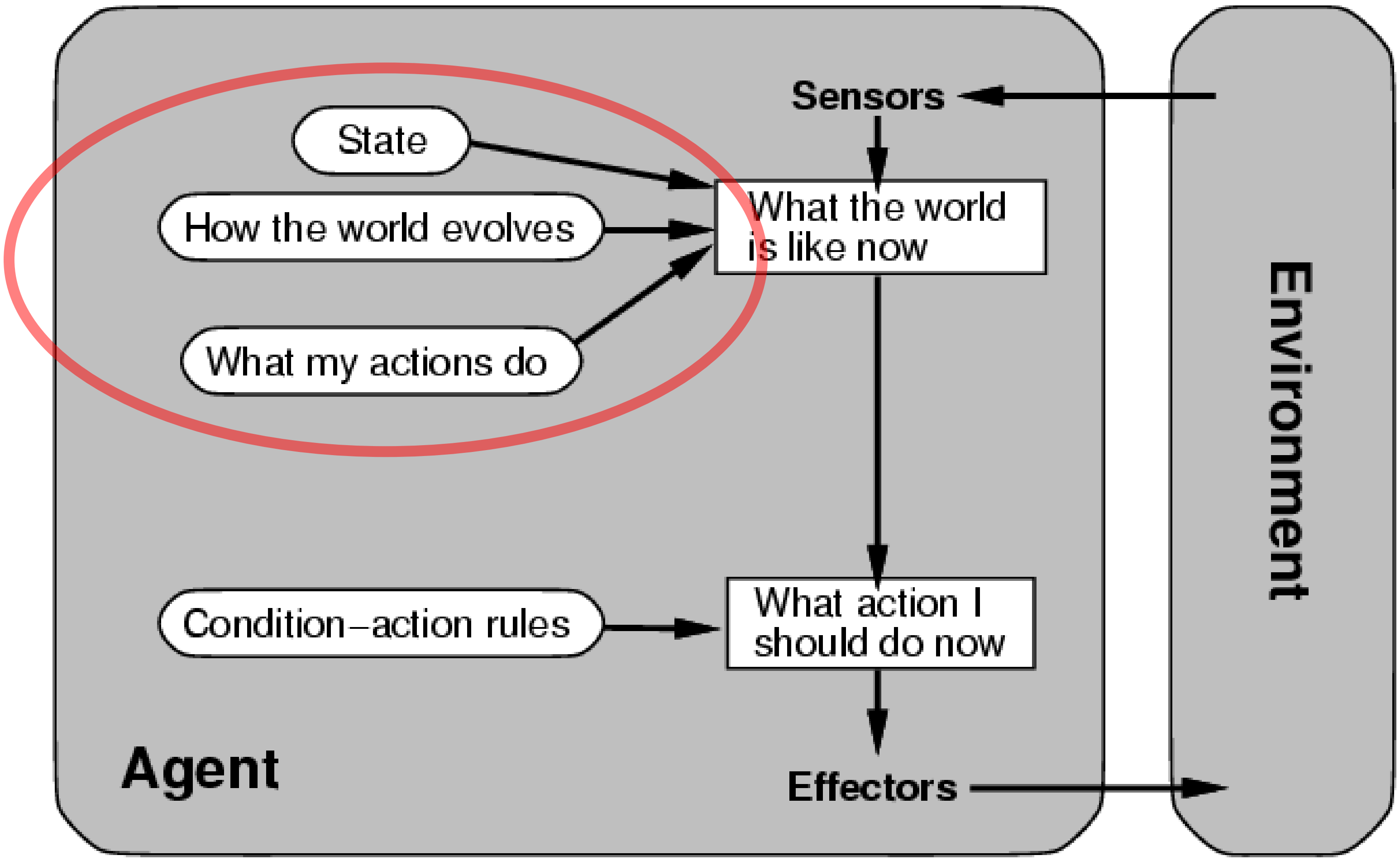
# (1) Simple reflex agents

- **Rule-based reasoning** maps percepts to optimal action; each rule handles collection of perceived states (aka reactive agents)
- **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 environment; collection of rules must be updated if changes occur
  - Still can’t condition actions on previous state
  - Difficult to engineer if the number of rules is large due to conflicts

Courtesy Tim Finin

# (2) Architecture for an agent with memory

**internal state** used to keep track of past states of the world




Courtesy Tim Finin

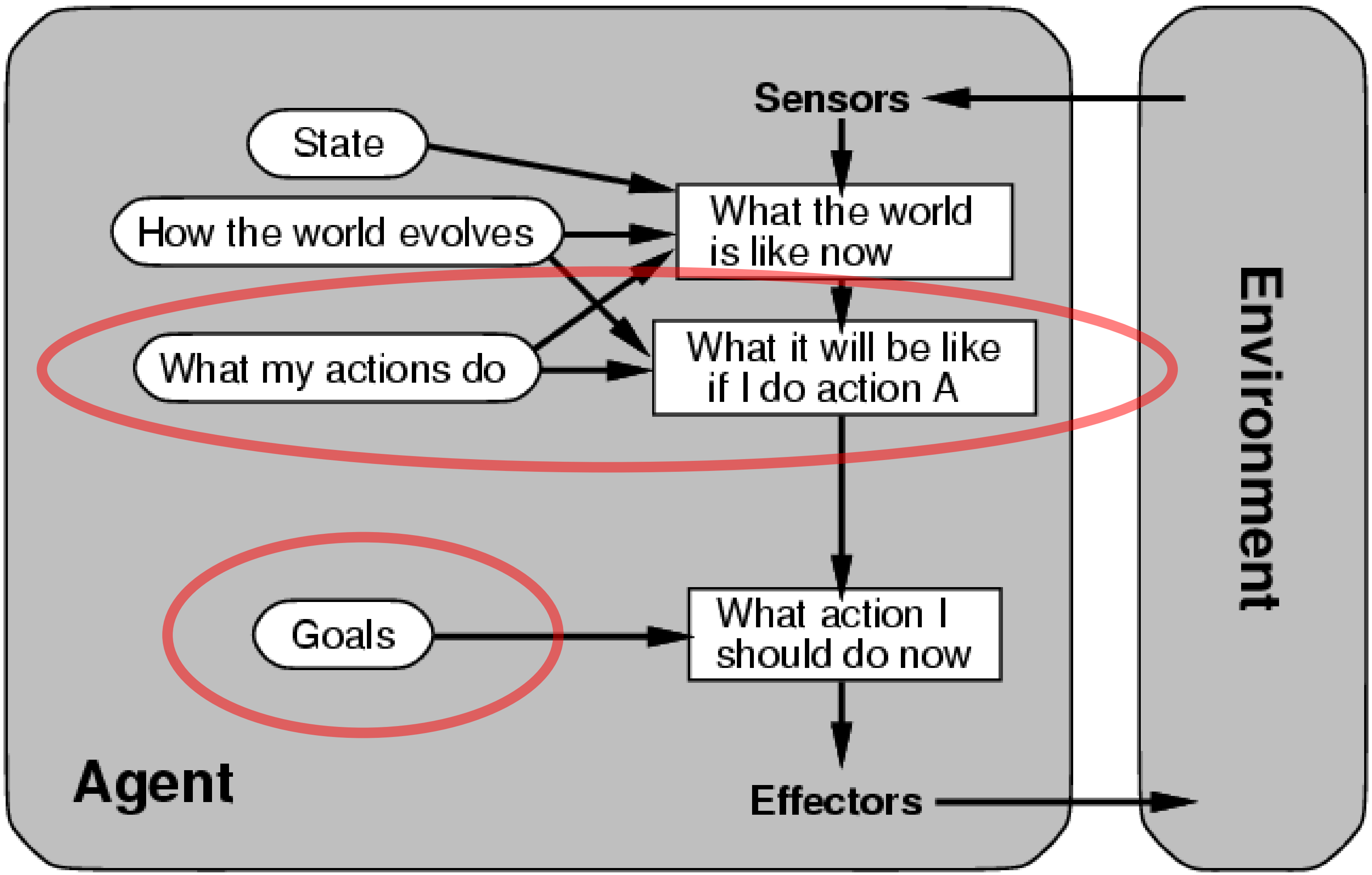
# (2) Agents with memory

- Encode *internal state* of world to remember past as contained in earlier percepts
  - Note: sensors don't usually give entire world state at each input, so environment perception is *captured over time*
  - *State* used to encode different "world states" that generate the same immediate percept
- Requires *representing change* in the world
  - Might represent just latest state, but then can't reason about hypothetical courses of action

Courtesy Tim Finin

# (3) Architecture for goal-based agent

state and **goal information** describe desirable situations allowing agent to take future events into consideration




Courtesy Tim Finin

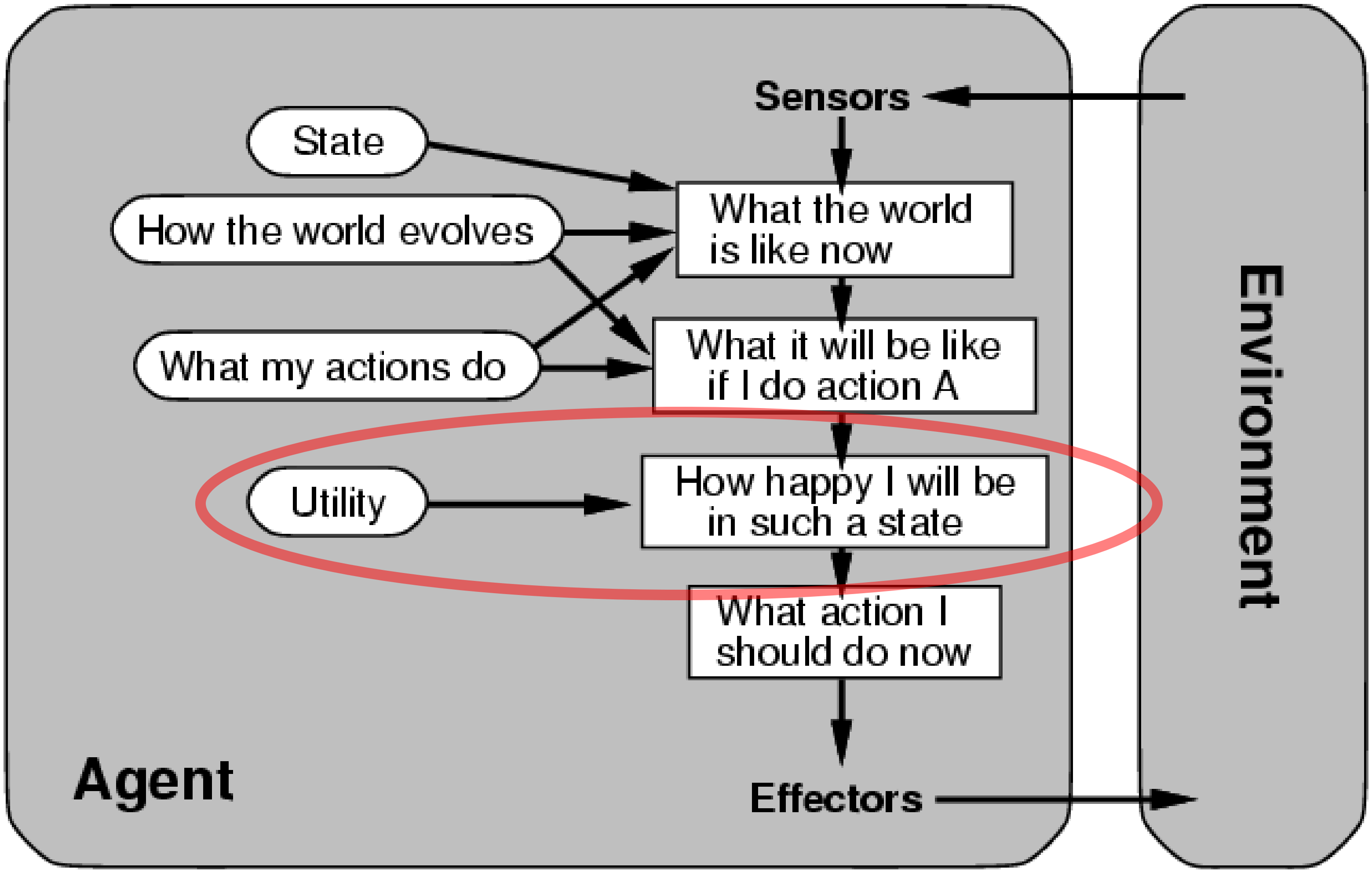
# (3) Goal-based agents


- **Deliberative** instead of **reactive**
- Choose actions to achieve a goal
- Goal: description of a desirable situation
- Keeping track of current state often not enough; must add goals to decide which situations are good
- Achieving goal may require an action sequence
  – Model action consequences: *"what happens if I do...?"*
  – Use *planning* algorithms to produce action sequences

Courtesy Tim Finin

# (4) complete utility-based agent

base decisions on utility theory in order to act rationally




Courtesy Tim Finin

# (4) Utility-based agents

- For multiple possible alternatives, how to decide which is best?
- Goals give a crude distinction between happy and unhappy states, but often need a performance measure for *degree*
- Utility function **U: State→Reals** gives measure of success/happiness for given state
- Allows decisions comparing choices between conflicting goals and likelihood of success and importance of goal (if achievement **uncertain**)

Courtesy Tim Finin

# Properties of Environments

**Observability**: Full vs. Partial

**Certainty**: Deterministic vs. Stochastic

**Atomicity**: Episodic vs. Sequential

**Malleability**: Static vs. Dynamic

**Percept & Action Type**: Discrete vs. Continuous

**Number of Participants**: Single agent vs. multi-agent

# Properties of Environments (I)

- **Fully/Partially observable**
  - Agent’s sensors give complete state of environment needed to choose action: environment is **fully observable**
  - Such environments are convenient, freeing agents from keeping track of the environment’s changes
- **Deterministic/Stochastic**

Courtesy Tim Finin

# Properties of Environments (I)

- **Fully/Partially observable**
  - Agent's sensors give complete state of environment needed to choose action: environment is **fully observable**
  - Such environments are convenient, freeing agents from keeping track of the environment's changes
- **Deterministic/Stochastic**
  - Environment is **deterministic** if next state is completely determined by current state and agent's action
  - **Stochastic** (i.e., non-deterministic) environments have multiple, unpredictable outcomes

Courtesy Tim Finin

# Properties of Environments (I)

- **Fully/Partially observable**
  - Agent's sensors give complete state of environment needed to choose action: environment is **fully observable**
  - Such environments are convenient, freeing agents from keeping track of the environment's changes
- **Deterministic/Stochastic**
  - Environment is **deterministic** if next state is completely determined by current state and agent's action
  - **Stochastic** (i.e., non-deterministic) environments have multiple, unpredictable outcomes
- In fully observable, deterministic environments agents need not deal with uncertainty

Courtesy Tim Finin

# Properties of Environments (II)

- **Episodic/Sequential**
  - In **episodic** environments subsequent episodes don't depend on actions in previous episodes
  - In **sequential** environments agent engages in a series of connected episodes
  - Episodic environments don't require agent to plan ahead
- **Static/Dynamic**

# Properties of Environments (II)

- **Episodic/Sequential**
  - In **episodic** environments subsequent episodes don't depend on actions in previous episodes
  - In **sequential** environments agent engages in a series of connected episodes
  - Episodic environments don't require agent to plan ahead
- **Static/Dynamic**
  - **Static** environments doesn't change as agent is thinking
  - The passage of time as agent deliberates is irrelevant
  - The agent needn't observe world during deliberation

# Properties of Environments III

- **Discrete/Continuous**
  - If number of distinct percepts and actions is limited (or representable by an integer), environment is **discrete**, otherwise it's **continuous**
- **Single agent/Multiagent**

Courtesy Tim Finin

# Properties of Environments III

- **Discrete/Continuous**
  - If number of distinct percepts and actions is limited (or representable by an integer), environment is **discrete**, otherwise it's **continuous**
- **Single agent/Multiagent**
  - In environments with other agents, agent must consider strategic, game-theoretic aspects of environment (for either cooperative *or* competitive agents)
  - Many engineering environments don't have multiagent properties, whereas most social and economic systems get their complexity from 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 | | | | | | |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

Courtesy Tim Finin

# 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 | | | | | | |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

Courtesy Tim Finin

# 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 | | | | | | |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

Courtesy Tim Finin

# 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 | | | | | | |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

Courtesy Tim Finin

# 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 | | | | | | |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

Courtesy Tim Finin

# Characteristics of environments

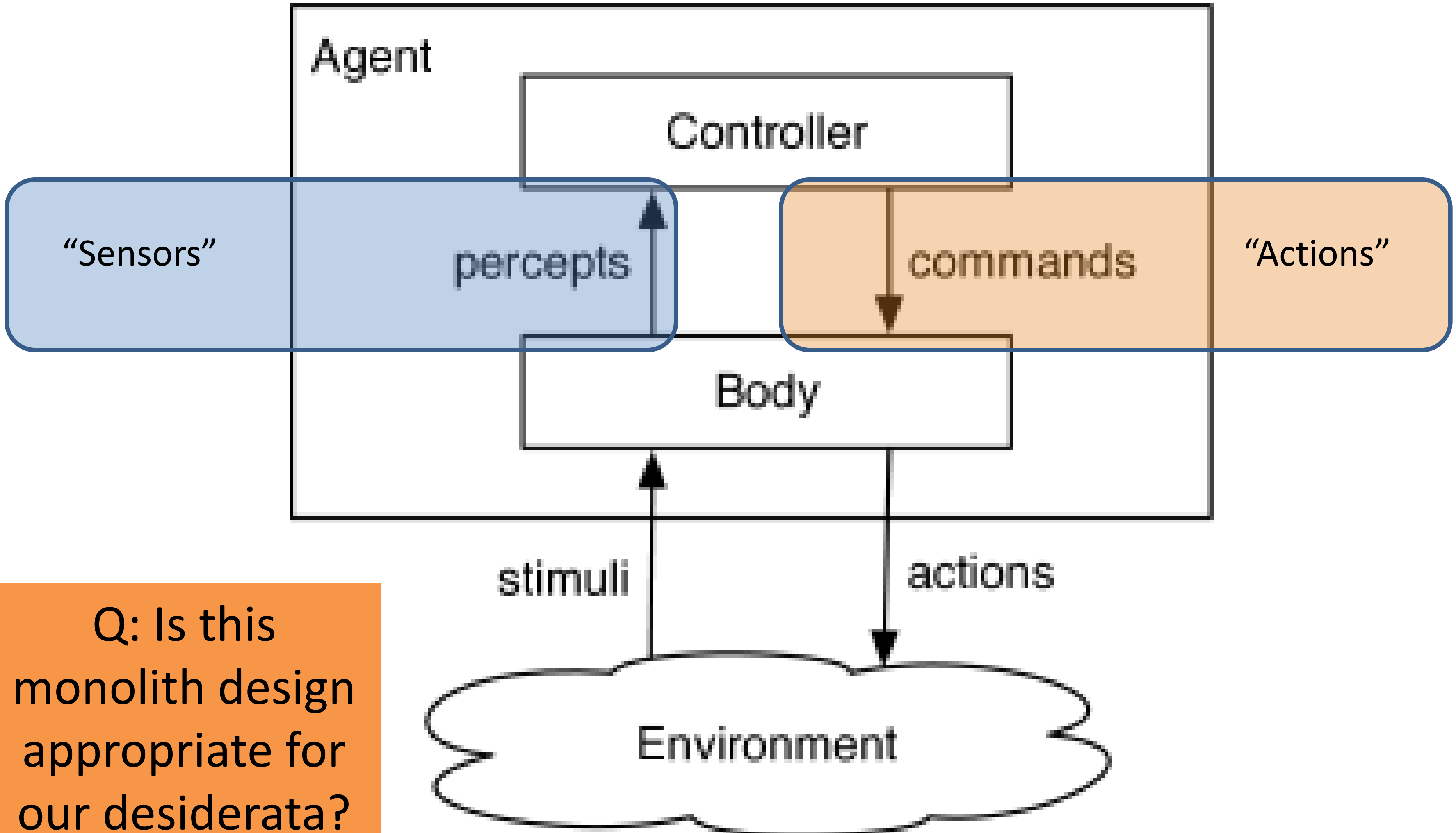
→ Lots of real-world domains fall into the hardest case!

| | 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** |

A **Yes** in a cell means that aspect is simpler; a **No** more complex

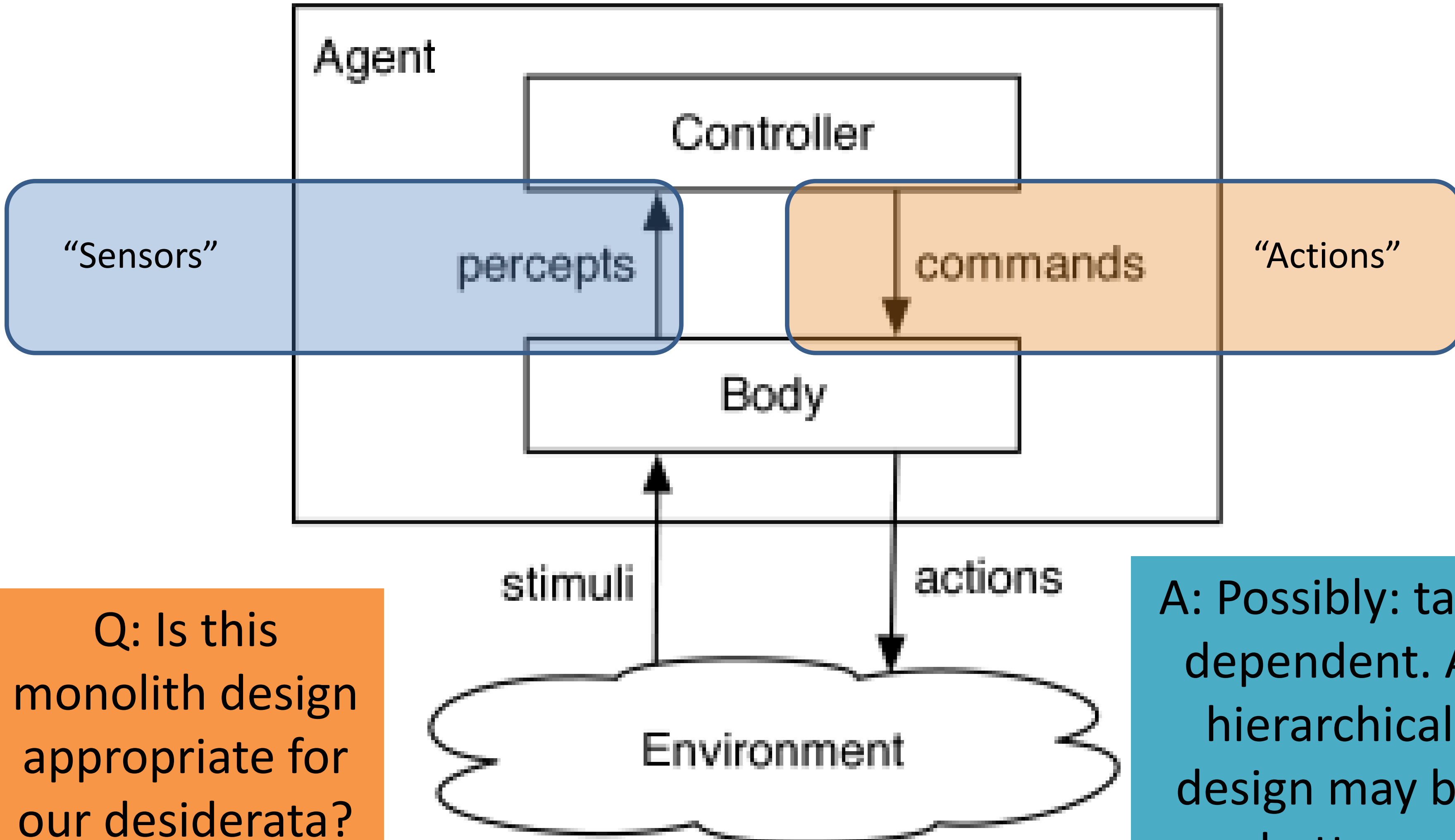
Courtesy Tim Finin

# Let’s Re-examine Our Agent Design




Q: Is this monolith design appropriate for our desiderata?

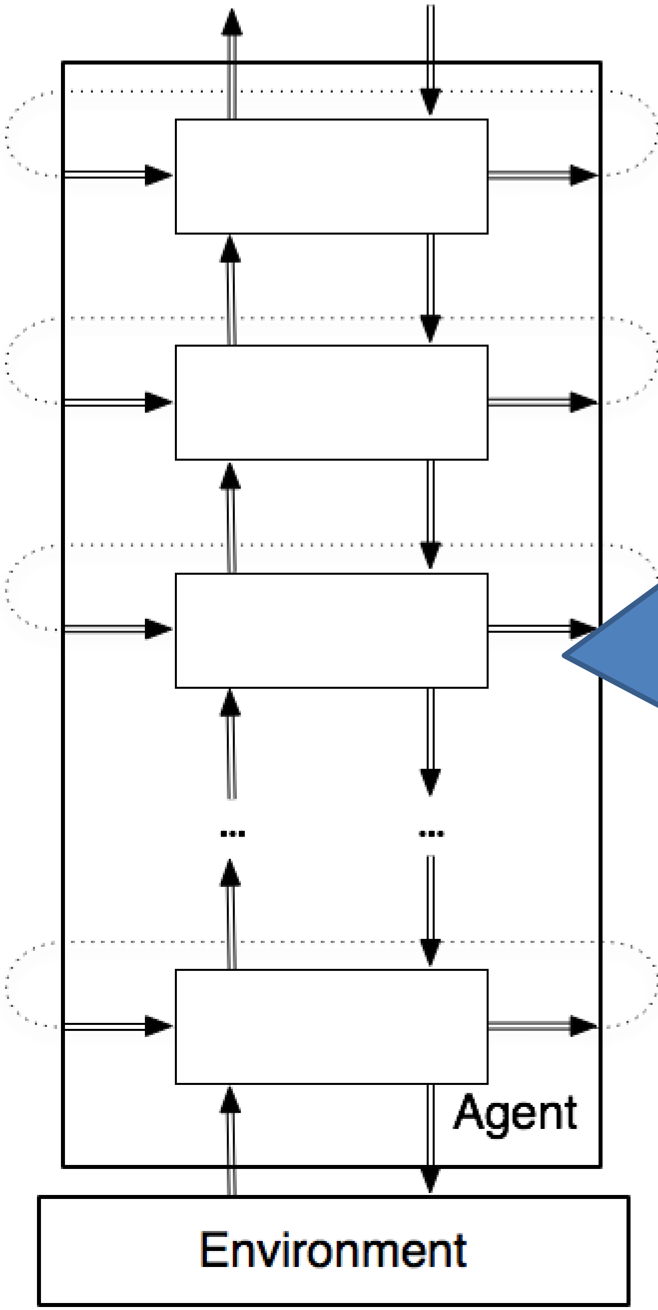
# Let's Re-examine Our Agent Design




Q: Is this monolith design appropriate for our desiderata?

A: Possibly: task dependent. A hierarchical design may be better.

# A Hierarchical Design



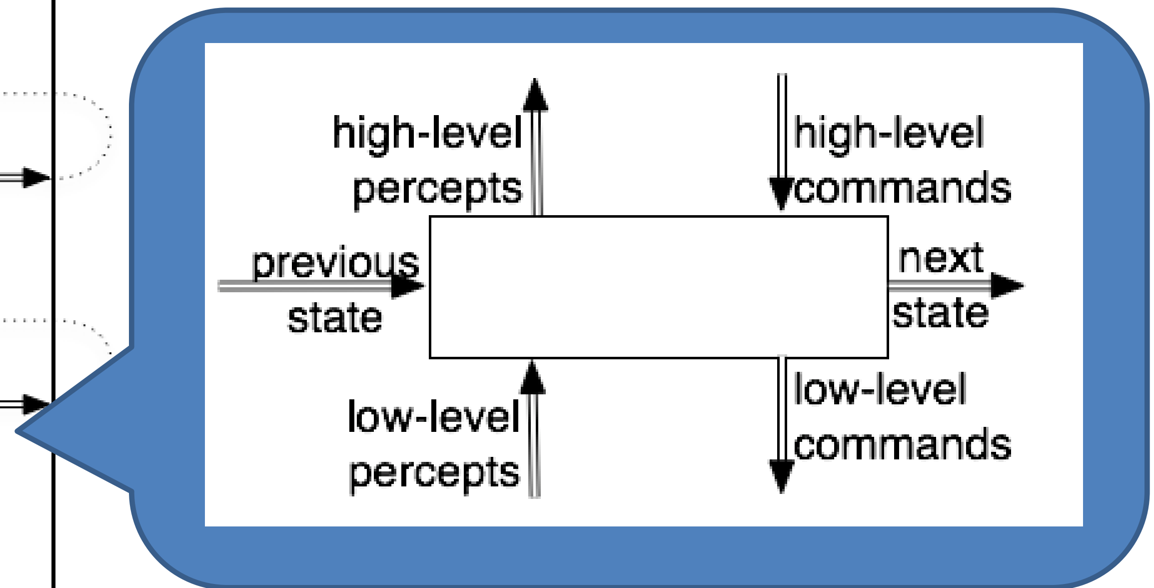
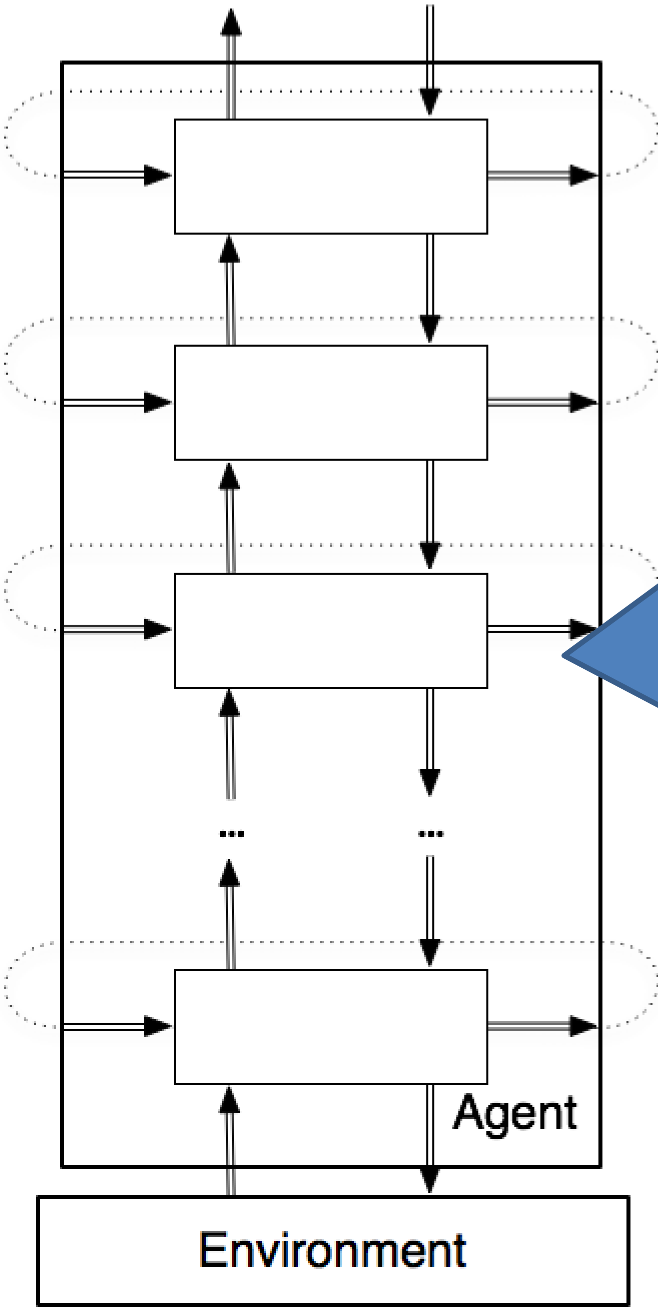




Fig. 2.4

# A Hierarchical Design



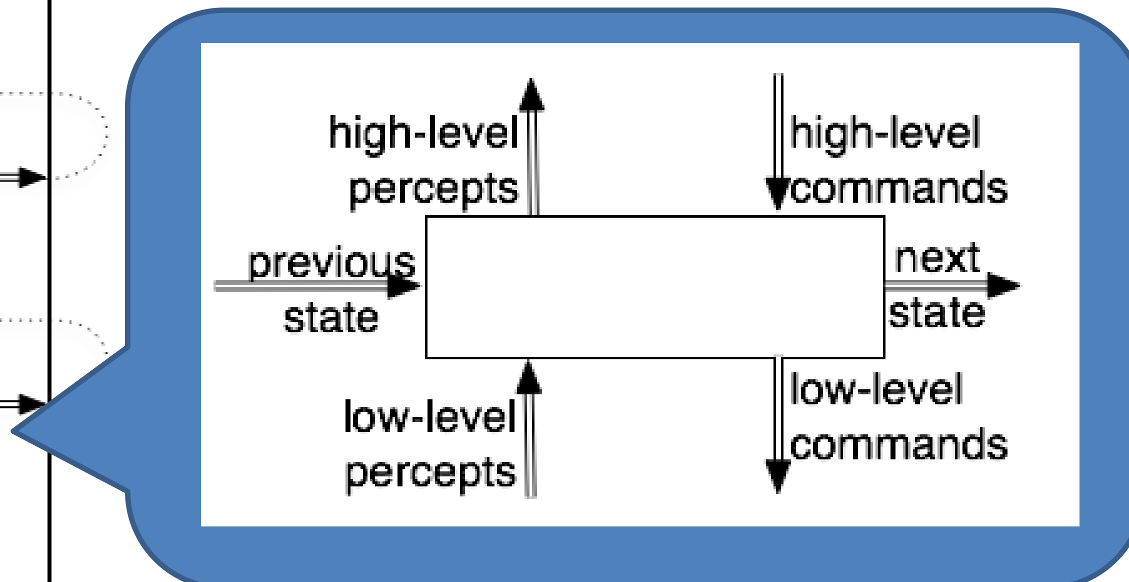




Fig. 2.4

Each lower-level layer can be seen as a “virtual body”

Q: How might this setup aid our desiderata (e.g., rationality, explainability, etc.?)

# A Delivery Robot: An Example Hierarchical Design

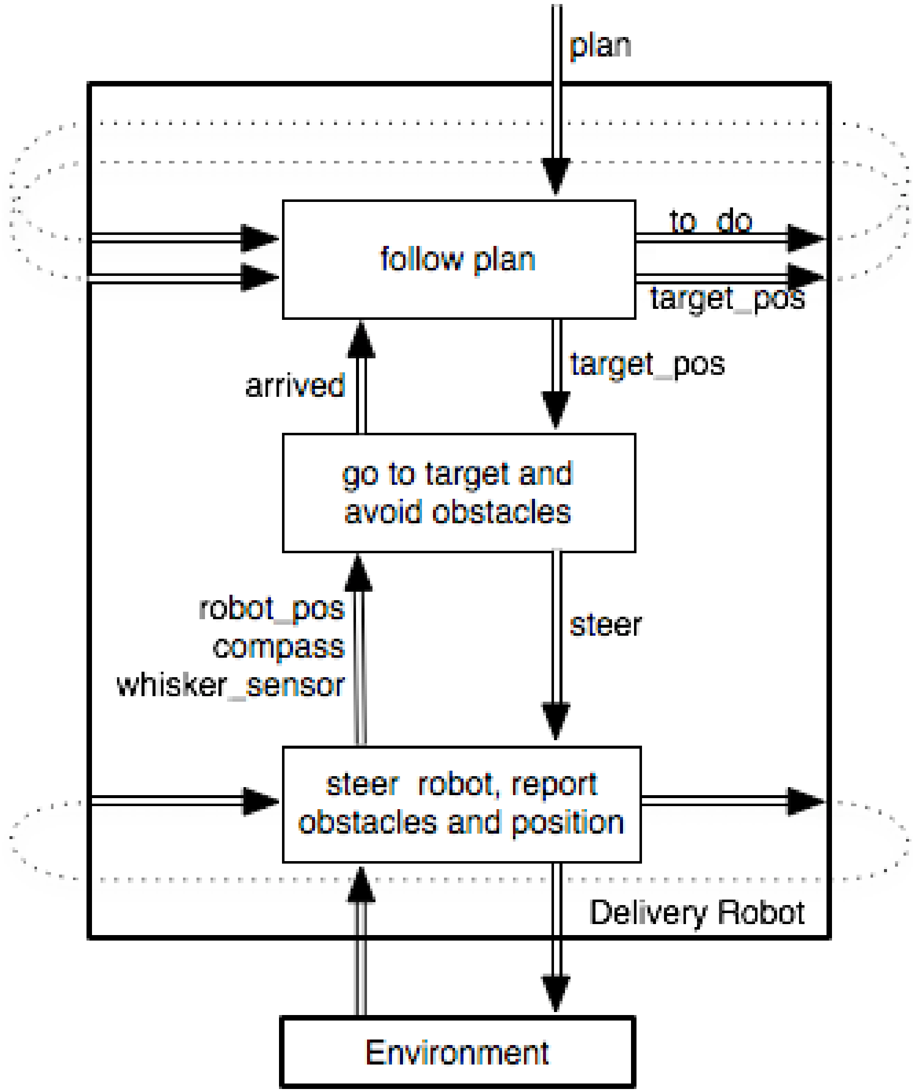



Fig. 2.5

# Summary

- **Agent programs** map percepts to actions and update their internal state
  - **Reflex** agents respond immediately to percepts
  - **Goal-based** agents act to achieve their goal(s)
  - **Utility-based** agents maximize their utility function
- **Representing knowledge** is important for good agent design
- Most challenging environments are partially observable, stochastic, sequential, dynamic, and continuous and contain multiple agents

Courtesy Tim Finin