Intelligent Agents

Chapter 2

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/

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 & Mackwo…

something that acts in an environment; it does something.

Use "computation" to explain and traceback the actions

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

Courtesy Frank Ferraro

Possible AI approaches

	Like humans	Well
Think		
Act		

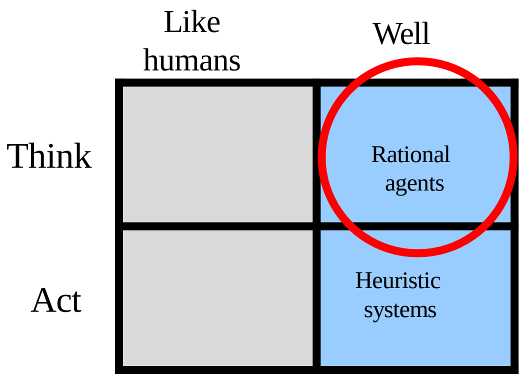
Possible approaches

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

AI tends to work mostly in this area

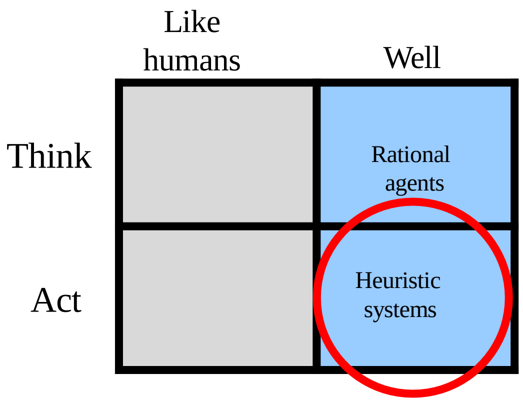
Think well

- 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



Courtesy Tim Finin

Agents and AI

- Much of our text is organized around the concept of **autonomous agents**
- It defines an agent as

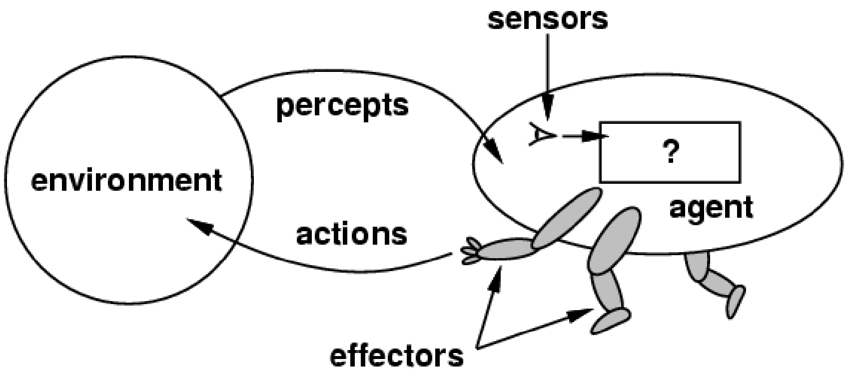
  "anything that can be viewed as perceiving its ***environment*** *through* ***sensors*** *and acting upon that environment through* ***actuators****"*

- Not all AI problems or tasks need or fit this concept, but it's quite general

How do you design an intelligent agent?

- **Intelligent agents** perceive environment via **sensors** and act *rationally* on them with their **effectors**

- General properties
  - **Reactive** to the environment
  - Pro-active or **goal-directed**
  - **Interacts** with other agents through communication or via the environment
  - **Autonomous**



Example: autonomous taxi

- **Percepts**: Video, sonar, speedometer, odometer, engine sensors, keyboard input, microphone, GPS, …
- **Actions**: Steer, accelerate, brake, horn, speak, …
- **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

- 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 $\rightarrow$ 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, money earned, …

Information and learning

- Rational agents aren’t expected to know everything
- But to use what they do know effectively
- Rationality includes ***information gathering*** -- If you don’t know something that might be useful, find out!
- Rationality also can exploit ***learning*** – making generalization from past experience to fill in missing information

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

simple

Some agent types

complex

(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

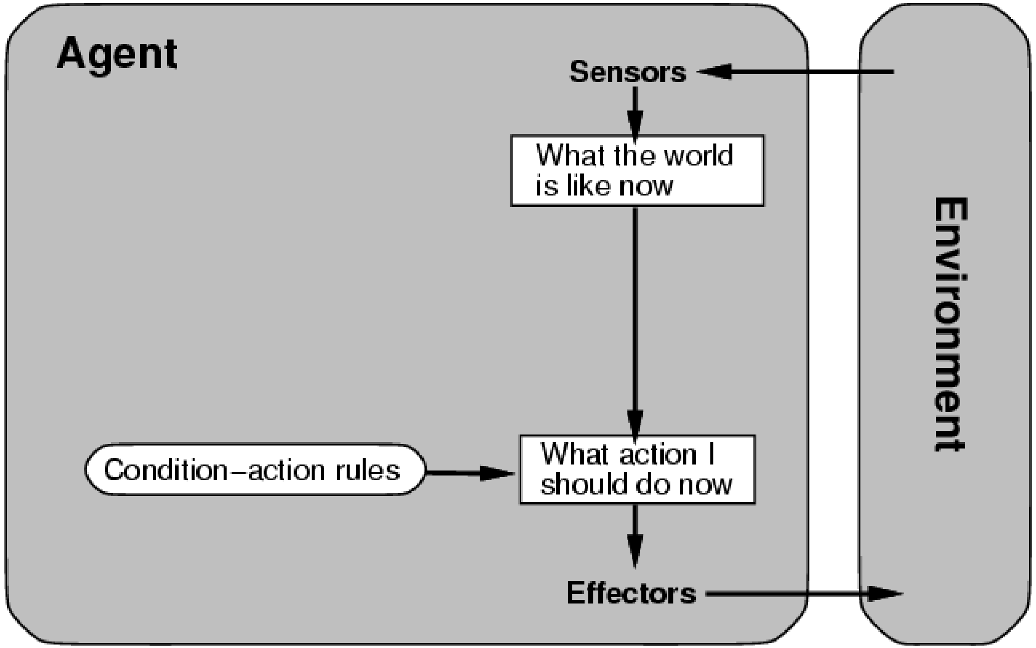
(4) Utility-based agents

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

Courtesy Tim Finin

(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

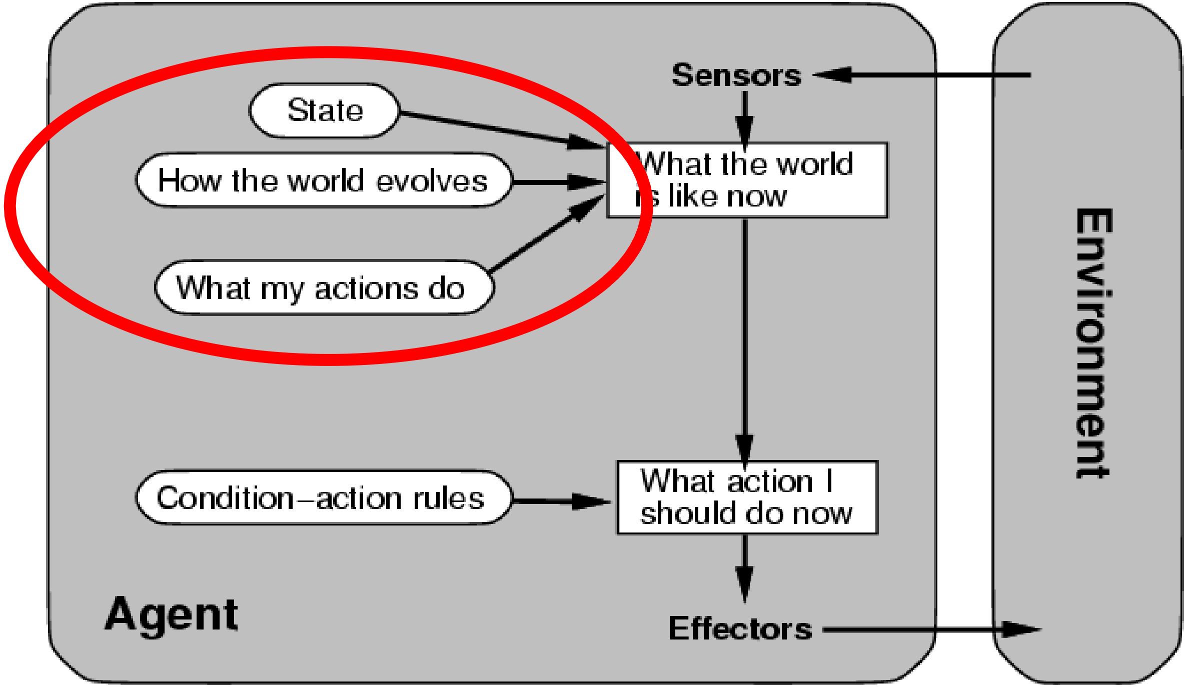
(1) Simple reflex agents

- **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



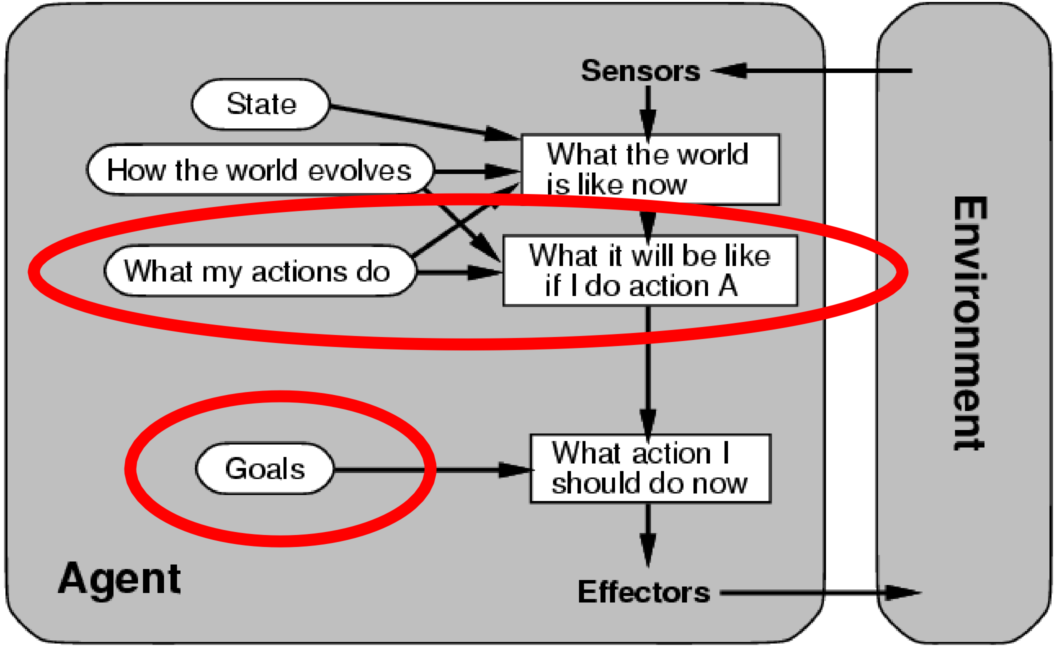
Courtesy Tim Finin

(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

(3) Architecture for goal-based agent

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



Courtesy Tim Finin

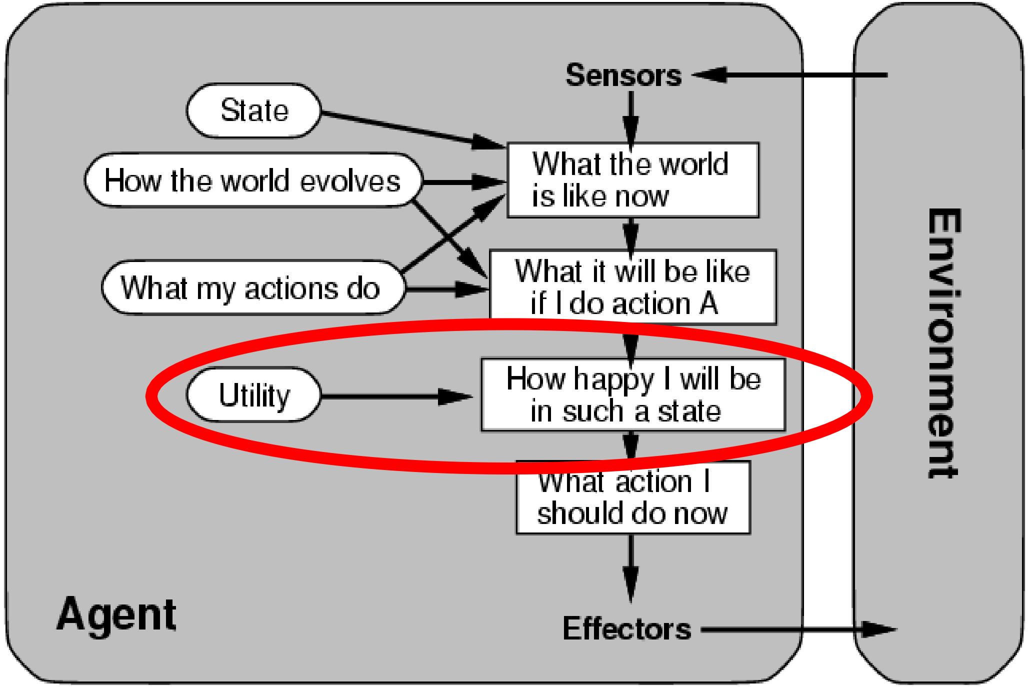
(3) Goal-based agents

- **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

Courtesy Tim Finin

Properties of Environments

- **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

- **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

Courtesy Tim Finin

Properties of Environments

- **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

Courtesy Tim Finin

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

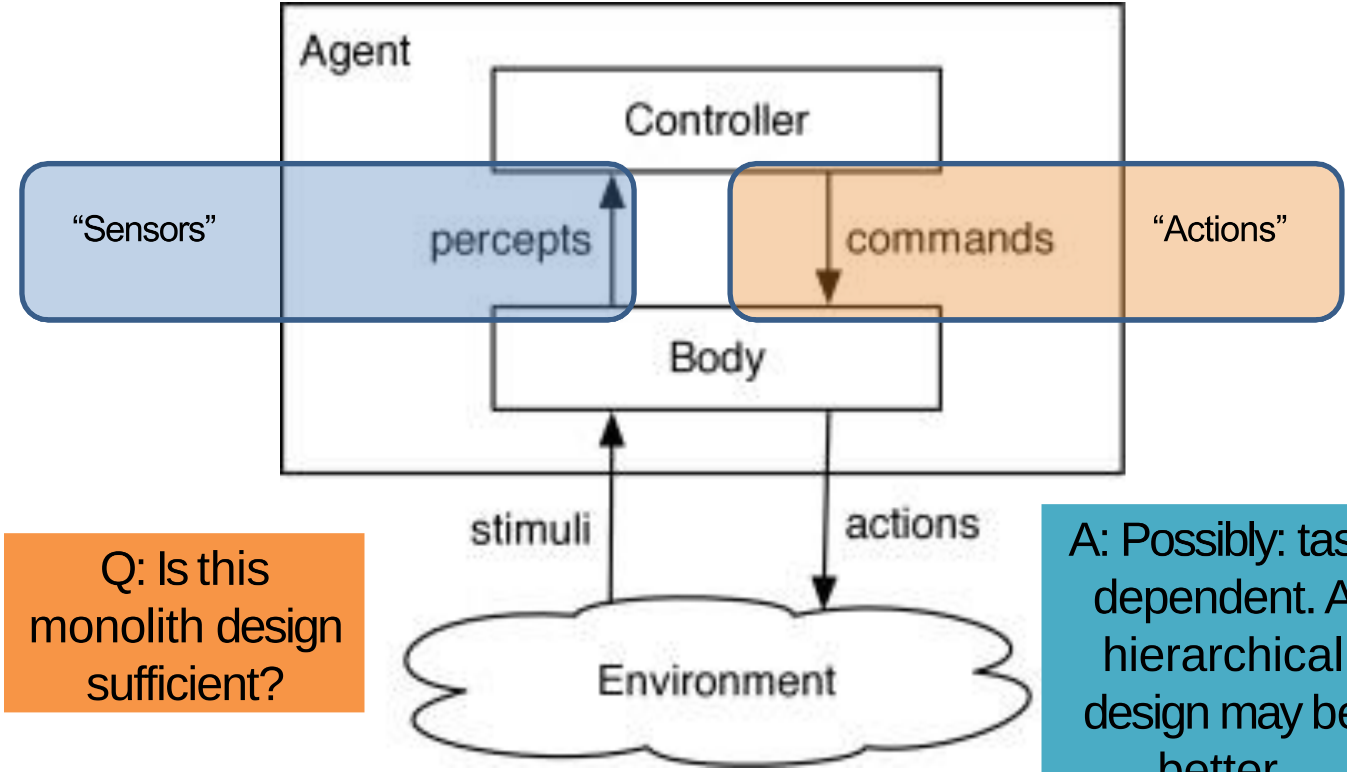
Characteristics of environments

	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 sufficient?

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

Hierarchical Design

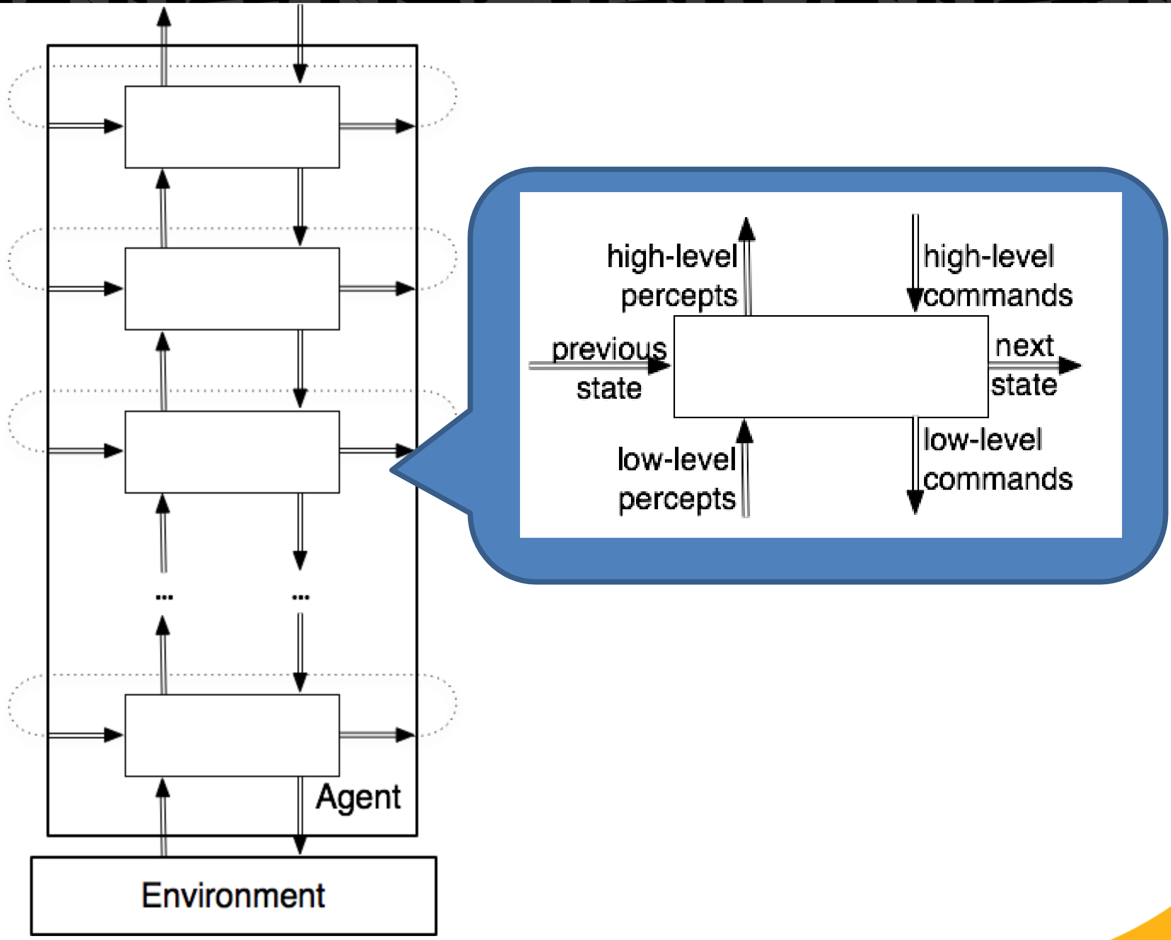


Fig. 2.4

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