

Artificial Intelligence: An overview

Thomas Trappenberg

January 4, 2009

Based on the slides provided by Russell and Norvig, Chapter 1 & 2

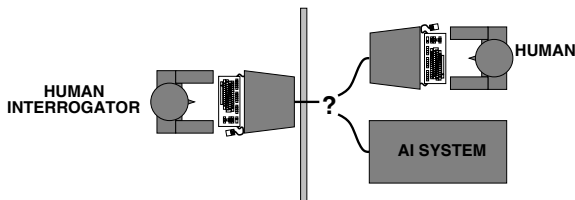
What is AI?

Systems that think like humans	Systems that think rationally
Systems that act like humans	Systems that act rationally

Acting humanly: The Turing test

Turing (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 following 50 years
- ▶ Suggested major components of AI: knowledge, reasoning, language understanding, learning

Problem: Turing test is not **reproducible**, **constructive**, or amenable to **mathematical analysis**

Thinking humanly: Cognitive Science

1960s “**cognitive revolution**”: information-processing psychology replaced prevailing **behaviorism**

- ◇ Requires scientific theories of internal activities of the brain
 - ▶ What level of abstraction? “Knowledge” or “circuits”?
 - ▶ How to validate? Requires
 - 1) Predicting and testing behavior of human subjects (top-down), or
 - 2) Direct identification from neurological data (bottom-up)
 - ▶ Both approaches (roughly, **Cognitive Science** and **Cognitive Neuroscience**) share with AI the search for theories to explain (or engender) anything resembling human-level general intelligence.

- ◇ Computational Neuroscience: **How the brain thinks!**

Thinking rationally: Laws of Thought

Aristotle: what are correct arguments/thought processes?

Several Greek schools developed various forms of **logic: notation** and **rules of derivation** for thoughts;
may or may not have proceeded to the idea of mechanization

Direct line through mathematics and philosophy to modern AI

Problems:

- 1) Not all intelligent behavior is mediated by logical deliberation
- 2) **What is the purpose of thinking?** What thoughts **should** I have out of all the thoughts (logical or otherwise) that I **could** have?

Acting rationally

Rational behavior: doing the right thing

The right thing: that which is expected to maximize goal achievement, given the available information. Doesn't necessarily involve thinking—e.g., blinking reflex—but thinking should be in the service of rational action.

Rational agent: An **agent** is an entity that perceives and acts

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

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

For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance

Caveat: **computational limitations make perfect rationality unachievable**

→ design best **program** for given machine resources

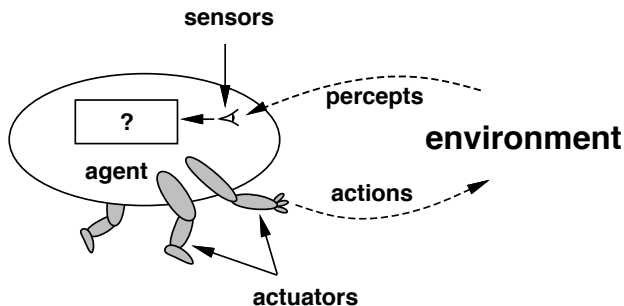
AI themes

- ▶ State graphs & search
 - ▶ Uninformed search
 - ▶ Heuristic search
- ▶ Knowledge representation & expert systems
 - ▶ Formal logic (propositional, first-order)
 - ▶ semantic nets
 - ▶ case-based reasoning
- ▶ Machine learning & probabilistic reasoning
 - ▶ Bayesian networks
 - ▶ Hidden Markov models & Kalman filters
 - ▶ Artificial Neural Networks & Support Vector Machines
- ▶ Common concepts & applications
 - ▶ Intelligent (rational) agent systems
 - ▶ Planning and decision making
 - ▶ Natural language processing
 - ▶ Games

Brief history of AI

- 1943** McCulloch & Pitts: Boolean circuit model of brain
- 1950** Turing's "Computing Machinery and Intelligence"
- 1952–69** Look, Ma, no hands!
- 1950s** Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1956** Dartmouth meeting: "Artificial Intelligence" adopted
- 1965** Robinson's complete algorithm for logical reasoning
- 1966–74** AI discovers computational complexity
Neural network research almost disappears
- 1969–79** Early development of knowledge-based systems
- 1980–88** Expert systems industry booms
- 1988–93** Expert systems industry busts: "AI Winter"
- 1985–95** Neural networks return to popularity
- 1988–** Resurgence of probability; general increase in technical depth
"Nouvelle AI": ALife, GAs, soft computing
- 1995–** Agents, agents, everywhere . . .
Machine learning comes to age, web intelligence, smart machines
- 2003–** Human-level AI back on the agenda

Agents and environments



Agents include humans, robots, softbots, thermostats, etc.

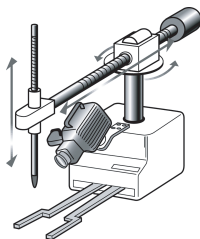
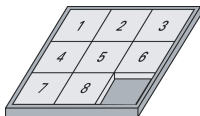
The **agent function** maps from percept histories to actions:

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

The **agent program** runs on the physical **architecture** to produce f

Example: 8-puzzle solving machine

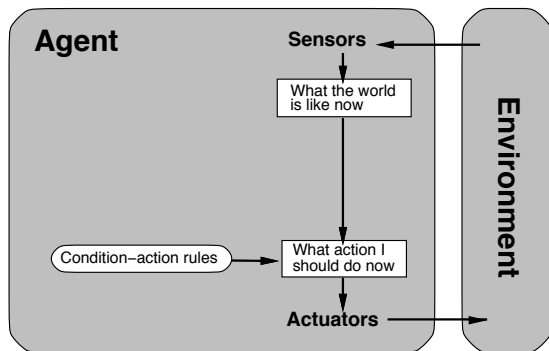
Build a machine that can take an 8-puzzle and solves it.



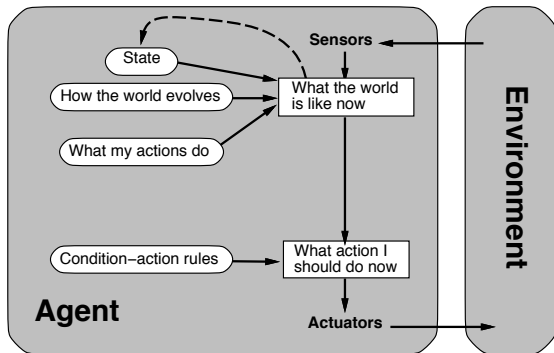
Challenges:

- ▶ image understanding
- ▶ motor control
- ▶ solving complex computational task

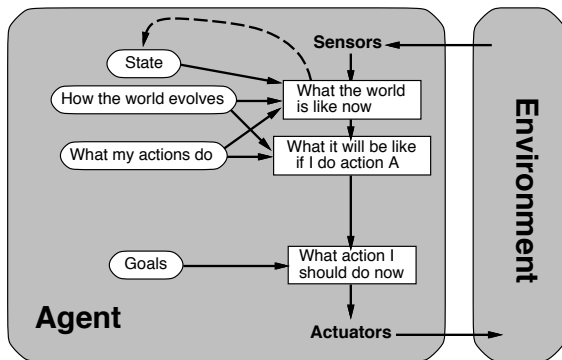
Simple reflex agents



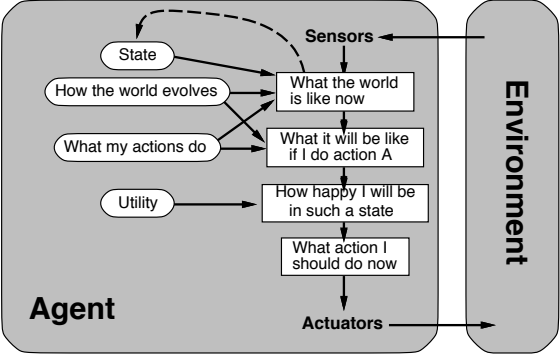
Reflex agents with state



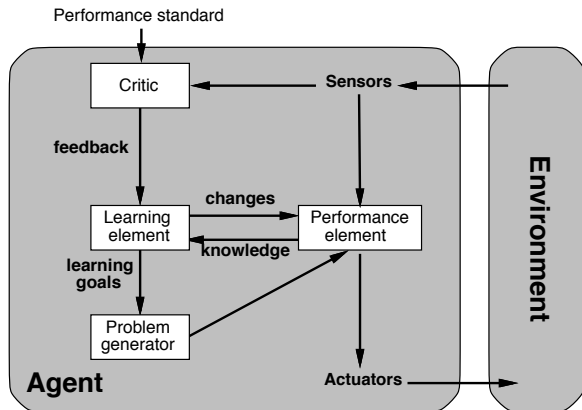
Goal-based agents



Utility-based agents



Learning agents



Anticipating agents: Generative world models

