

# Artificial Intelligence

## CHAPTER 1

# Outline

- ◇ Course overview
- ◇ What is AI?
- ◇ A brief history
- ◇ The state of the art

## Administrivia

Class home page: <http://www-inst.eecs.berkeley.edu/~cs188>  
for lecture notes, assignments, exams, grading, office hours, etc.

Assignment 0 (lisp refresher) due 8/31

Book: Russell and Norvig Artificial Intelligence: A Modern Approach  
Read Chapters 1 and 2 for this week's material

Code: integrated lisp implementation for AIMA algorithms at  
<http://www-inst.eecs.berkeley.edu/~cs188/code/>

## Course overview

- ◇ intelligent agents
- ◇ search and game-playing
- ◇ logical systems
- ◇ planning systems
- ◇ uncertainty—probability and decision theory
- ◇ learning
- ◇ language
- ◇ perception
- ◇ robotics
- ◇ philosophical issues

# What is AI?

“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning ...” (Bellman, 1978)	“The study of mental faculties through the use of computational models” (Charniak+McDermott, 1985)
“The study of how to make computers do things at which, at the moment, people are better” (Rich+Knight, 1991)	“The branch of computer science that is concerned with the automation of intelligent behavior” (Luger+Stubblefield, 1993)

Views of AI fall into four categories:

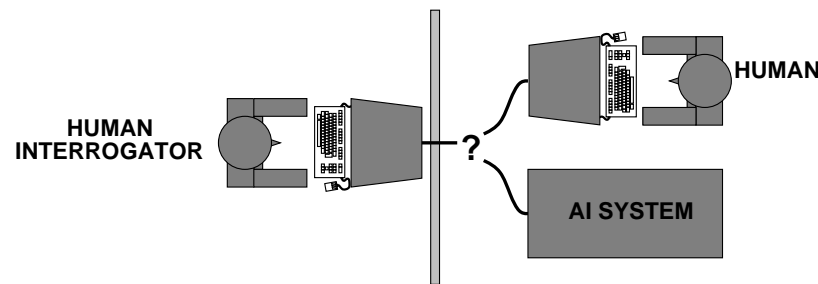
Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

Examining these, we will plump for acting rationally (sort of)

# 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 orthodoxy of 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) are now distinct from AI

# Thinking rationally: Laws of Thought

Normative (or prescriptive) rather than descriptive

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?



# 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

Aristotle (Nicomachean Ethics):

*Every art and every inquiry, and similarly every action and pursuit, is thought to aim at some good*

# Rational agents

An agent is an entity that perceives and acts

This course is about designing rational agents

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 prehistory

Philosophy	logic, methods of reasoning mind as physical system foundations of learning, language, rationality
Mathematics	formal representation and proof algorithms computation, (un)decidability, (in)tractability probability
Psychology	adaptation phenomena of perception and motor control experimental techniques (psychophysics, etc.)
Linguistics	knowledge representation grammar
Neuroscience	physical substrate for mental activity
Control theory	homeostatic systems, stability simple optimal agent designs

# Potted 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 probabilistic and decision-theoretic methods  
Rapid increase in technical depth of mainstream AI  
"Nouvelle AI": ALife, GAs, soft computing

## State of the art

Which of the following can be done at present?

- ◇ Play a decent game of table tennis
- ◇ Drive along a curving mountain road
- ◇ Drive in the center of Cairo
- ◇ Play a decent game of bridge
- ◇ Discover and prove a new mathematical theorem
- ◇ Write an intentionally funny story
- ◇ Give competent legal advice in a specialized area of law
- ◇ Translate spoken English into spoken Swedish in real time