



Welcome!

CS5811 Advanced Artificial Intelligence

Michigan Technological University

Information about me

- Nilufer Onder
- Official office hours: MWF, 10:00-11:00
- Research interests: Planning, planning under uncertainty, decision making under uncertainty, temporal reasoning

How about you?

Administrivia

- Textbook: Russell and Norvig's "AI A Modern Approach (AIMA)". 2nd edition, 2003.
- Both paper and programming assignments (50%)
- One midterm exam (seventh week) (25%)
- Final exam (twelveth week, before Thanksgiving break) (25%)
- Prerequisite: CS4811

Course overview

- Ch. 01: Introduction
- Ch. 02: Intelligent agents
- Ch. 03: Solving problems by searching (short)
- Ch. 04: Informed search and exploration
- Ch. 05: Constraint satisfaction problems
- Temporal Constraint Networks
- (Ch. 06: Adversarial search (skip))

Course overview (cont'd)

- (Ch. 07: Logical agents (skip))
- (Ch. 08: First-order logic (skip))
- (Ch. 09: Inference in first-order logic (skip))
- (Ch. 10: Knowledge representation (skip))
- Ch. 11: Planning
- Ch. 12: Planning and acting in the real world

Course overview (cont'd)

- Ch. 13: Uncertainty
- Ch. 14: Probabilistic reasoning
- Ch. 15: Probabilistic reasoning over time
- Ch. 16: Making Simple Decisions
- Ch. 17: Making Complex Decisions
- Ch. 18: Learning from observations (time permitting)

Course overview (cont'd)

- (Ch. 19: Knowledge in Learning (skip))
- (Ch. 20: Statistical Learning Methods (skip))
- Weeks 13,14: Student presentations

What is AI?

Systems that:

think like humans	think rationally
act like humans	act rationally

Thinking humanly

- Need to know how the human mind works (cognitive modeling)
- Introspection: catch your own thoughts, remember how you solved a problem or learned something
- Psychological experiments
- If a sufficiently precise theory of the mind is available, it might be possible to convert it to a computer program
- 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
 - Predicting and testing behavior of human subjects (top-down)
 - Direct identification from neurological data (bottom-up)

Cognitive Science (cont'd)

- Both approaches (roughly, *Cognitive Science* and *Cognitive Neuroscience*) are now distinct from AI
- Both share with AI the following characteristic:
the available theories do not explain (or engender) anything resembling human-level general intelligence
- Hence, all three fields share one principal direction!

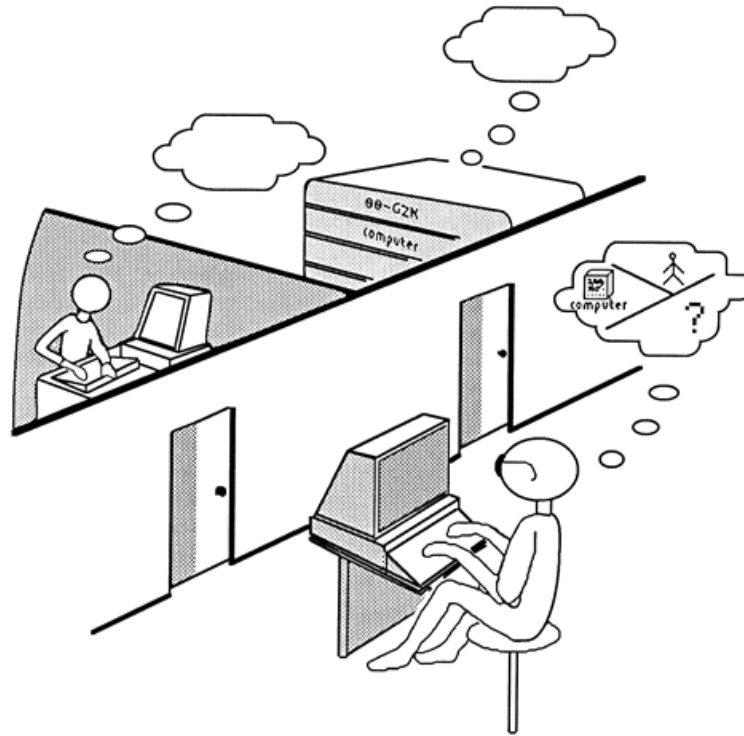
Acting humanly

- Turing (1950) “Computing machinery and intelligence”:
- “Can machines think?” → “Can machines behave intelligently?”
- 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

Acting humanly

- Suggested major components of AI: knowledge, reasoning, language understanding, learning
- Operational test for intelligent behavior: the Imitation Game (aka Turing Test)

The Turing test



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

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

Thinking rationally (cont'd)

- Direct line through mathematics and philosophy to modern AI
- Problems:
 - Not all intelligent behavior is mediated by logical deliberation
 - 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 emphasizes 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

Rational agents (cont'd)

- Caveat: *computational limitations make perfect rationality unachievable* → design best *program* for given machine resources
- *Limited rationality*

What can AI do?

- For you
- For the society