CS451/CS551/EE565
Artificial Intelligence
Fall 2013

Prof. Janice T. Searleman
jets@clarkson.edu
Outline

- Administrivia
- The Turing Test
- Intelligent Agents (IAs)

Announcements: Activity Fair on Tuesday, August 27th, 7–8 pm, Cheel Arena
Course Description

This course is a comprehensive introduction to core concepts in artificial intelligence using an agent based approach, and surveys active research areas in AI.

Topics include:

- Intelligent Agents
- Search & heuristics, game playing, video game AI
- Knowledge representation: first-order logic, semantic networks, neural nets, frame-based systems, and rules
- Automated reasoning: reasoning under uncertainty
- Problem-solving, expert systems, planning
- Natural language understanding
- Learning, Evolutionary Computation, Connectionism
Administrivia

- course webpage: http://www.clarkson.edu/~jets/cs451/
- contact info:
  office: SC375 & Applied CS Labs:VR (SC336)
  phone: 268–2377
  email: jets@clarkson.edu

- AFS directory
  /afs/cu/class/cs451
Textbook & Software


- Companion Website
  aima.cs.berkeley.edu/
# Grading Policy

<table>
<thead>
<tr>
<th></th>
<th>CS451</th>
<th>CS551/EE565</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>2 Midterm Exams</strong></td>
<td>30%</td>
<td>30%</td>
</tr>
<tr>
<td>(tentatively 10/8 &amp; 11/14)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Final Exam</strong></td>
<td>35%</td>
<td>35%</td>
</tr>
<tr>
<td><strong>HW &amp; class participation</strong></td>
<td>15%</td>
<td>10%</td>
</tr>
<tr>
<td><strong>Research Paper Presentation</strong></td>
<td>n/a</td>
<td>5%</td>
</tr>
<tr>
<td><strong>Project</strong></td>
<td>20%</td>
<td>20%</td>
</tr>
</tbody>
</table>
Grading Policy (continued)

- **Attendance:** Participation in class discussions are an essential part of this course, so attendance is required and you must be prepared for class (i.e. have done the reading in advance).

- Written HW and programming assignments are expected to be individual efforts – do not copy or allow your work to be copied. Late HW will be accepted only if negotiated with the instructor.

- Each student in CS551/EE565 will also present a research-level paper to the class. All graduate students will write a reaction paper to it (details to follow later).
Project

The course project allows you to explore some area of AI in more depth. It consists of a written research paper coupled with a program that illustrates the concepts. You may work on any area of AI that you find interesting, upon approval of the instructor; including (but not limited to) computer vision, learning, neural nets, genetic algorithms, intelligent tutoring systems, game AI, robotics, and so on.

A proposal describing your project is due on Wednesday, 9/25/13, and the final project is due on Wednesday, 12/04/13. All projects will be demonstrated to the instructor no later than 12/09/13.
What is Artificial Intelligence?

"It's beginning to show some human characteristics - faulty reasoning, tactfulness and repetition."
Outlook

- AI is a very exciting area right now.
- This course will teach you the foundations.

Getting started:
- Read AIMA: Chapters 1 & 2
<table>
<thead>
<tr>
<th>What is AI?</th>
</tr>
</thead>
<tbody>
<tr>
<td>“[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning …” (Bellman, 1978)</td>
</tr>
<tr>
<td>“The study of mental faculties through the use of computational models” (Charniak &amp; McDermott, 1985)</td>
</tr>
<tr>
<td>“The art of creating machines that perform functions that require intelligence when performed by people” (Kurzweil, 1990)</td>
</tr>
<tr>
<td>“The branch of computer science that is concerned with the automation of intelligent behavior” (Luger &amp; Stubblefield, 1993)</td>
</tr>
</tbody>
</table>
Views of AI fall into four categories:

<table>
<thead>
<tr>
<th>Thinking humanly</th>
<th>Thinking rationally</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acting humanly</td>
<td>Acting rationally</td>
</tr>
</tbody>
</table>

Russell & Norvig opt for “Acting rationally”
Acting humanly: Turing Test

Alan Turing (1950)

- "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
Thinking humanly: cognitive modeling

- 1960s "cognitive revolution": information-processing psychology replaces behaviorism.
- Cognitive science brings together theories and experimental evidence to model internal activities of the brain.
  - What level of abstraction? “Knowledge” or “Circuits”?
  - How to validate models?
    - Predicting and testing behavior of human subjects (top-down)
    - Direct identification from neurological data (bottom-up)
Thinking rationally: "laws of thought"

Aristotle (~450 BCE) attempted to codify "right thinking"; what are correct arguments/thought processes?

e.g. “Socrates is a man. All men are mortal. Therefore Socrates is mortal.”

Several Greek schools developed various forms of logic: notation and rules of derivation for thoughts.

Problems:

- Uncertainty – not all “facts” are certain (e.g. the flight might be delayed).
- Resource limitations: not enough time to compute, insufficient memory, etc.
Acting rationally: rational agent

- **Rational** behavior: Doing the right thing!
- The right thing: that which is expected to maximize goal achievement, given the available information
- Provides the most general view of AI because it includes:
  - correct inference ("laws of thought")
  - uncertainty handling
  - resource limitation considerations (e.g. reflex vs. deliberation)
  - cognitive skills (NLP, AR, knowledge representation, etc.)
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
  \[ \rightarrow \text{design best } \text{program} \text{ for given machine resource} \]
What is an (Intelligent) Agent?

- An **agent** is anything that can be viewed as **perceiving** its **environment** through **sensors** and **acting** upon that environment through **actuators**

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

- An agent can perceive its own actions, but not always its effects