More Intro to AI

Dr. Melanie Martin

CS 4480

August 24, 2012

What is AI?

- "The practice of designing systems that possess and acquire knowledge and reason with knowledge." (Tanimoto 1987)

- "The design and study of computer programs that behave intelligently." (Dean, Allen, Aloimonos 1995)

- "The branch of computer science concerned with making computers behave like humans." (Webopedia)

What is AI?

- But then, what is intelligence???
  - "the capacity for learning, reasoning, understanding, and similar forms of mental activity; aptitude in grasping truths, relationships, facts, meanings, etc." (Webster's Encyclopedic Unabridged Dictionary of the English Language 1996)

What is AI?

Categories under AI on Cora (~1999-2001)
Domain Specific Search Engine for CS papers

- Agents
- Data Mining
- Expert Systems
- Games and Search
- Knowledge Representation
- Machine Learning
- Theory, Case-Based, Rule Learning, ...
- Natural Language Processing
- Planning
- Robotics
- Speech
- Theorem Proving
- Vision & Pattern Recognition

What is AI?

- Goals in AI
  - Engineering: Solve real-world problems. Build systems that exhibit intelligent behavior.
  - Scientific: Understand what kind of computational mechanisms and knowledge are needed for modeling intelligent behavior.

What is AI?

- A few recurring issues:
  - How important is cognitive modeling in our systems?
  - How do we balance scientific and engineering goals?
  - How do we evaluate our system?

How strong do you like your AI?

- Weak AI
  - Machines could act as if they were intelligent
- Strong AI
  - Machines that act intelligent are actually thinking

Inside Joke

- If it works it is not AI

What is AI?

Views of AI fall into four categories:

| Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

The textbook advocates "acting rationally"

What is AI?

- Do we really want to model humans?
  - Seem like our best example, but....
  - Should we build airplanes with wings that flap like birds?
- How do we know we did it?
  - Turing test?
    - Focus on behavior instead of internal algorithm
    - Defines success in terms of human intelligence

Acting humanly: Turing Test

- Turing (1950) "Computing machinery and intelligence":
- "Can machines think?" → "Can machines behave intelligently?"
- Operational test for intelligent behavior: the Imitation Game
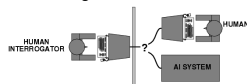
- Predicted that by 2000, a machine might have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI in following 50 years
- Suggested major components of AI: knowledge, reasoning, language understanding, learning

The Chinese Room

- Searle (1980) p. 1031

- Human: CPU
- Rule Book: Program
- Paper: Memory
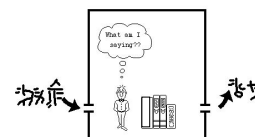
- Human understands only English
- Input symbols, output symbols based on rules
- Appears to have conversation in Chinese

- Syntax is not sufficient for Semantics

Thinking humanly: cognitive modeling

- 1960s "cognitive revolution": information-processing psychology
- Requires scientific theories of internal activities of the brain
- -- 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"

- 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 agent

- 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 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

- Economics   Utility, decision theory

- Neuroscience   Physical substrate for mental activity

- Psychology   Phenomena of perception and motor control, experimental techniques

- Computer engineering   Building fast computers

- Control theory   Design systems that maximize an objective function over time

- Linguistics   Knowledge representation, grammar

Bits of History

- 1956: The name "Artificial Intelligence" is coined
- 60's: Search and games, formal logic and theorem proving
- 70's: Robotics, perception, knowledge representation, expert systems
- 80's: More expert systems, AI becomes an industry
- 90's: Rational agents, probabilistic reasoning, machine learning
- 00's: Systems integrating many AI methods, machine learning, reasoning under uncertainty, robotics again

State of the art

- Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
- Proved a mathematical conjecture (Robbins conjecture) unsolved for decades
- No hands across America (driving autonomously 98% of the time from Pittsburgh to San Diego)
- During the 1991 Gulf War, US forces deployed an AI logistics planning and scheduling program that involved up to 50,000 vehicles, cargo, and people
- NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
- `Proverb` solves crossword puzzles better than most humans
- Watson defeats humans at Jeopardy

Outline for Chapter 2

- Agents and environments
- Rationality
- PEAS (Performance measure, Environment, Actuators, Sensors)
- Environment types
- Agent types

Agents

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

- Human agent: eyes, ears, and other organs for sensors; hands,
- legs, mouth, and other body parts for actuators

- Robotic agent: cameras and infrared range finders for sensors;
- various motors for actuators

Agents and environments
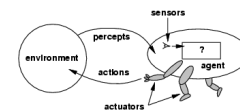
- The agent function maps from percept histories to actions:

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

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

- agent = architecture + program