CSIS 4463: Artificial Intelligence

Introduction: Chapter 1

What is AI?

- **Strong AI:** Can machines really think?
  - The notion that the human mind is nothing more than a computational device, and thus in principle computers are capable of thought.
  - E.g., sentient or self-aware machines
  - A machine truly capable of reasoning and solving problems
- Philosophers have pondered this question for decades
- Most AI researchers focus attention elsewhere

What is AI?

- **Weak AI:** Can machines act intelligently?
  - The notion that machines can accomplish specific reasoning or problem solving tasks that do not fully encompass human cognitive abilities
  - E.g., A machine capable of solving a problem that would seem to require “intelligence”
- Has lead to a large body of algorithms that can solve problems at least as effectively as humans
  - E.g., Deep Blue’s win against Gary Kasparov

How to define artificial intelligence?

- AI defined differently by different people
- Topic is of interest within (and influenced by) diverse academic disciplines
- “Thinking” vs “Acting”
  - “Intelligent” thought processes / reasoning vs. “Intelligent” behavior
- Human-level performance vs Ideal performance
  - Does the system perform at the level of a human on a given task? Vs. Does the system perform a task rationally (ideal performance)?

What is AI?

Views of AI fall into four categories:

Thinking humanly	Thinking rationally
Acting humanly	Acting rationally

The textbook advocates "acting rationally"

A few definitions of AI

Thinking Humanly	**Thinking Rationally**
"The exciting new effort to make computers think … machines with minds, in the full and literal sense." (Haugeland, '85) "[Automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning…" (Bellman, '78)	"The study of mental faculties through the use of computational models." (Charniak & McDermott, '85) "The study of the computations that make it possible to perceive, reason, and act." (Winston, '92)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, '90)	"AI …is concerned with intelligent behavior in artifacts." (Nilsson, '98)

Thinking Humanly	**Thinking Rationally**
•Cognitive modeling •Field of cognitive science brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind. •Newell & Simon's "General Problem Solver" ('61) was an attempt to solve problems not simply correctly, but in the same way as human test subjects.	•The "laws of throught" •Aristotle's "right-thinking" (i.e., irrefutable reasoning processes) •Reasoning logically to a correct conclusion •This direction within AI is known as the "logicist tradition" and focuses on building on the work of 19th century logicians to create intelligent systems.
Acting Humanly	**Acting Rationally**
A machine passes the "Turing Test" for machine intelligence if a human is unable to determine which of 2 subjects is the human and which is the machine based on written responses to questions. (Turing, '50)	•"A rational agent is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome." (Russell & Norvig) •Draws much from economics

Thinking humanly: cognitive modeling

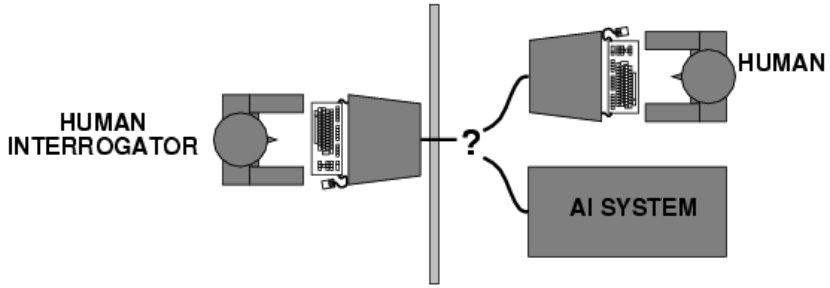
- 1960s "cognitive revolution": information-processing psychology
- Requires scientific theories of internal activities of the brain
  - How to validate? Requires

    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

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 70% 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 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. How do we state informal knowledge in the formal terms required for logical reasoning? Especially knowledge that is less than 100% certain?
  2. Big difference between solving a problem “in principle” vs “in practice”

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

Acting Rationally

- Perfect rationality (always doing the right thing) is not feasible
  - Too expensive computationally
- Perfect rationality is a useful working hypothesis
  - Often an underlying assumption of foundational elements of AI algorithms
    - E.g., Game playing search
- Bounded rationality: Acting appropriately when there is insufficient time to do all computations required for perfect rationality
- Related to Herb Simon's notion of a "satisficing" solution

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:

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

- 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's Foundations

- Philosophy — Logic, methods of reasoning, mind as physical system foundations of learning, language, rationality
  - Can formal rules be used to draw valid conclusions?
  - How does the mind arise from a physical brain?
  - Where does knowledge come from?
  - What is knowledge?
  - How does knowledge lead to action?
- Mathematics — Formal representation and proof algorithms, computation, (un)decidability, (in)tractability, probability
  - What are the formal rules with which to draw conclusions?
  - What can be computed?
  - What can't be computed?
  - How do we reason with uncertain information?

AI's Foundations

- Economics utility, decision theory, game theory
  - How should we make decisions to maximize profit?
  - How should we do something when others might not go along?
  - How should we do something for which the payoff might be far in the future?
  - What decisions should we make when interacting with others?
- Neuroscience physical substrate for mental activity
  - How do brains process information?
- Psychology phenomena of perception and motor control, experimental techniques
  - How do humans think and act?
  - How do animals think and act?
  - How do humans learn?
  - How do animals learn?

AI's Foundations

- Computer engineering
  - How do we build an efficient computer?
- Control theory & Cybernetics
  design systems that maximize an objective function over time
  - How can machines operate under their own control?
- Linguistics
  knowledge representation, grammar
  - How does language relate to thought?

Abridged history of AI

- 1943 McCulloch & Pitts: Boolean circuit model of brain
- 1950 Turing's "Computing Machinery and Intelligence"
- 1956 Dartmouth meeting: "Artificial Intelligence" adopted
- 1952—69 Look, Ma, no hands!
- 1950s Early AI programs, including Samuel's checkers program, Newell & Simon's Logic Theorist, Gelernter's Geometry Engine
- 1965 Robinson's complete algorithm for logical reasoning
- 1966—73 AI discovers computational complexity
  Neural network research almost disappears
- 1969—79 Early development of knowledge-based systems
- 1980-- AI becomes an industry
- 1986-- Neural networks return to popularity
- 1987-- AI becomes a science
- 1995-- The emergence of intelligent agents

A brief history of AI

- What happened after WWII?
  - 1943: Warren McCulloch and Walter Pitts: a model of artificial boolean neurons to perform computations.
    - First steps toward connectionist computation and learning (Hebbian learning).
    - Marvin Minsky and Dann Edmonds (1951) constructed the first neural network computer
  - 1950: Alan Turing's article "Computing Machinery and Intelligence"
    - First complete vision of AI.

A brief history of AI

- The birth of AI (1956)
  - Dartmouth Workshop bringing together top minds on automata theory, neural nets and the study of intelligence.
    - Allen Newell and Herbert Simon:
      - The logic theorist (first nonnumerical thinking program used for theorem proving)
    - For the next 20 years the field was dominated by these participants.
  - Great expectations (1952-1969)
    - Newell and Simon introduced the General Problem Solver.
      - Imitation of human problem-solving
    - Arthur Samuel (1952-)investigated game playing (checkers) with great success.
    - John McCarthy(1958-) :
      - Inventor of Lisp (second-oldest high-level language)
      - Logic oriented, Advice Taker (separation between knowledge and reasoning)

A brief history of AI

- The birth of AI (1956)
  - Great expectations continued ..
    - Marvin Minsky (1958 -)
      - Introduction of microworlds that appear to require intelligence to solve: e.g. blocks-world.
      - Anti-logic orientation, society of the mind.
- Collapse in AI research (1966 - 1973)
  - Progress was slower than expected.
    - Unrealistic predictions.
  - Some systems lacked scalability.
    - Combinatorial explosion in search.
  - Fundamental limitations on techniques and representations.
    - Minsky and Papert (1969) Perceptrons.

A brief history of AI

- AI revival through knowledge-based systems (1969-1979)
  - General-purpose vs. domain specific
    - E.g. the DENDRAL project (Buchanan et al. 1969)
      - First successful knowledge intensive system.
      - Inferring molecular structure from mass spectrometer data
  - Expert systems
    - MYCIN to diagnose blood infections (Feigenbaum et al.)
      - Introduction of uncertainty in reasoning.
  - Increase in knowledge representation research.
    - Logic, frames, semantic nets, …

A brief history of AI

- AI becomes an industry (1980 - present)
  - R1 at DEC (McDermott, 1982)
    - 1st successful commercial Expert System
  - Fifth generation project in Japan (1981)
  - American response MCC corporation
- Puts an end to the AI winter
  - Boomed from a few million dollars in 1980 to billions of dollars in 1988
  - Companies specializing in expert systems, vision systems, robotics, etc
- Connectionist revival (1986 - present)
  - Parallel distributed processing (RumelHart and McClelland, 1986); backprop.

A brief history of AI

- AI becomes a science (1987 - present)
  - Neats vs. scruffies
    - In speech recognition: hidden markov models
    - In neural networks
    - In uncertain reasoning and expert systems: Bayesian network formalism
    - …
  - “Neats”: those who think AI should be grounded in mathematical rigor
  - “Scruffies”: those who prefer to try out lots of stuff, write some programs, and assess which things seem to work
- The emergence of intelligent agents (1995 - present)
  - The whole agent problem:

    “How does an agent act/behave embedded in real environments with continuous sensory inputs”

State of the art

- Game playing: Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
- Robotic vehicles:
  - No hands across America (driving autonomously 98% of the time from Pittsburgh to San Diego) in mid 90s
  - More recently: 2005 DARPA Grand Challenge, Stanley drove 132mile dessert course
  - 2006 Urban Challenge, CMU's Boss drove in traffic through streets obeying traffic rules and avoiding pedstrians and vehicles (on a closed Air Force base)
- Logistics Planning:
  - 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
- Autonomous Planning & Scheduling
  - NASA's on-board autonomous planning program controlled the scheduling of operations for a spacecraft
- `Proverb` solves crossword puzzles better than most humans
- Spam Fighting: Learning algorithms classify over a billion messages a day as spam
- Robotics: companies like iRobot bringing AI and Robotics into households (e.g., the Roomba vacuum)

Many Subfields have Developed

- Artificial Life
- Autonomous Agents
- Biologically-Inspired Computing
- Computational Intelligence
- Constraint Programming
- Evolutionary Computation
- Knowledge-Based Systems
- Machine Learning
- Machine Vision
- Multi-Agent Systems
- Natural Language Processing
- Neural Networks
- Pattern Recognition
- Planning Systems
- Robotics
- Stochastic Search
- Swarm Intelligence

Just to name a few….