CS 188: Artificial Intelligence
Spring 2007

Lecture 1: Welcome and Introduction
1/16/2007

Srini Narayanan– ICSI and UC Berkeley
Many slides over the course adapted from
Dan Klein, Stuart Russell or Andrew Moore

Administrivia

http://inst.cs.berkeley.edu/~cs188

Instructor Access

- Instructor : Srini Narayanan
  - Office Hours Thursday 11-1 739 Soda
  - Email snarayan@icsi.berkeley.edu
- TA: Sean Markan
  - Office Hours :
  - Email markan@cs.berkeley.edu
- TA: Jason Wolfe
  - Office Hours:
  - Email jwolfe@cs.berkeley.edu
- TA: Nuttapong Chentanez
  - Office Hours:
  - Email nchentan@cs.berkeley.edu

Course Details

- Book: Russell & Norvig, AI: A Modern Approach, 2nd Ed.
- Prerequisites:
  - (CS 61A or B) and (Math 55 or CS 70)
  - There will be a lot of statistics and programming
- Work and Grading:
  - 7-8 assignments (3-4 coding 4 written). Total 45%
    - Python, groups of 1-2, 5 late days
  - Mid-term and final (Midterm 20%, Final 30%)
  - Participation (5%)
  - Academic dishonesty policy

Announcements

- Important stuff:
  - No section this week
  - Python intro in section next week.
  - Tutorial intro to Python (1/24, 1/26) 3-5 pm
  - Get your account forms (in front after class)
  - First assignment on web on Thursday
- Questions?

Python

- Python is an open source scripting language.
- Developed by Guido van Rossum in the early 1990s
- Named after Monty Python
- Available for download from http://www.python.org

Why Python for CS 188?

- Easy to learn and expressive
  - Combines features from Scheme and Java.
- Textbook Code: Very Object Oriented
  - Python much less verbose than Java
- AI Processing: Symbolic
  - Python's built-in datatypes for strings, lists, and more.
- AI Processing: Statistical
  - Python has strong numeric processing capabilities: matrix operations, etc.
  - Suitable for probability and machine learning code.
- History
  - Used for the last two semesters

Today

- What is AI?
- Brief History of AI
- What can AI do?
- What is this course?

Sci-Fi AI?



A REAL Accomplishment: DARPA Grand Challenge

http://video.google.com/videoplay?docid=8594517128412883394

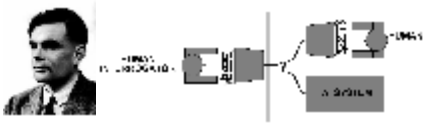
What is AI?

The science of making machines that:

Think like humans	Think rationally
Act like humans	Act rationally

Acting Like Humans?

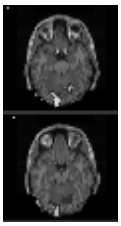
- Turing (1950) ``Computing machinery and intelligence''
  - ``Can machines think?'' → ``Can machines behave intelligently?''
  - Operational test for intelligent behavior: the *Imitation Game*



  - Predicted by 2000, 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 or amenable to mathematical analysis

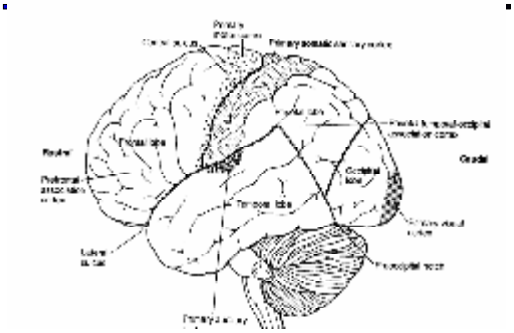
Thinking Like Humans?

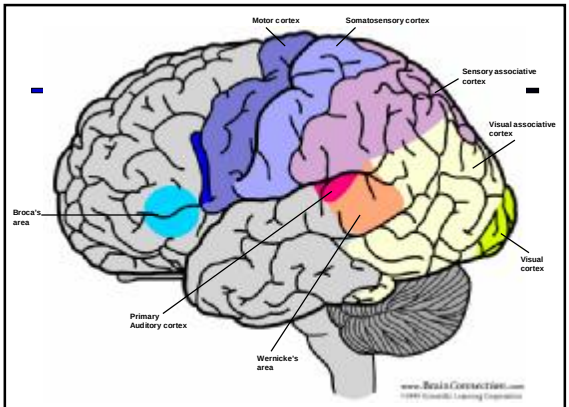
- The Cognitive Science approach:
  - 1960s ``cognitive revolution'': information-processing psychology replaced prevailing orthodoxy of behaviorism
- Scientific theories of internal activities of the brain
  - What level of abstraction? "Knowledge" or "circuits"?
  - **Cognitive science:** Predicting and testing behavior of human subjects (top-down)
  - **Cognitive neuroscience:** Direct identification from neurological data (bottom-up)
  - Both approaches 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!



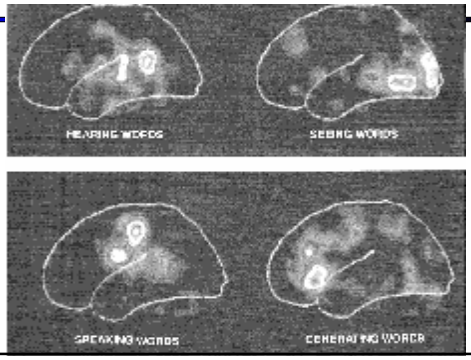
Images from Oxford fMRI center

BRAIN





Imaging the Brain



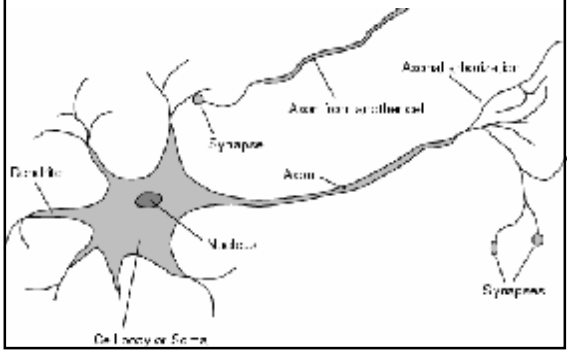
Sensory Systems

- Vision (nearly 30-50%)
- Audition (nearly 10%)
- Somatic
- Chemical
  - Taste
  - Olfaction

Motor Systems

- Locomotion
- Manipulation
- Speech

NEURON



Neural Basis of Intelligence

- How does a system of neurons with specific processes, connectivity, and functions support the ability to think, reason, and communicate?

Brain Like Computing

- Surge of research in recent years.
- Brain as a computing device is significantly different than modern computers.
  - How?
- This course will NOT tackle this kind of computing
  - 182 (ok, shameless plug) does.
- One lecture will identify the main points of convergence and divergence between AI and brain-based computation.

Brains ~ Computers

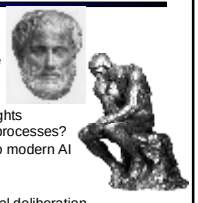
- 1000 operations/sec
- 100,000,000,000 units
- 10,000 connections/
- graded, stochastic
- embodied
- fault tolerant
- evolves, learns
- 1,000,000,000 ops/sec
- 1-100 processors
- ~ 4 connections
- binary, deterministic
- abstract
- crashes
- designed, programmed

What is AI?

The science of making machines that:

Think like humans	Think rationally
Act like humans	Act rationally

Thinking Rationally?



- The "Laws of Thought" approach
  - What does it mean to "think rationally"?
  - Normative / prescriptive rather than descriptive
- Logicist tradition:
  - Logic: notation and rules of derivation for thoughts
  - Aristotle: what are correct arguments/thought processes?
  - Direct line through mathematics, philosophy, to modern AI
- Problems:
  - Not all intelligent behavior is mediated by logical deliberation
  - What is the purpose of thinking? What thoughts should I (bother to) have?
  - Logical systems tend to do the wrong thing in the presence of **uncertainty**

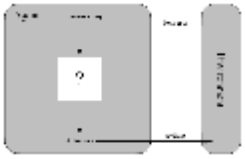
Acting Rationally

- 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
  - Thinking can be in the service of rational action
  - Entirely dependent on goals!
  - Irrational ≠ insane, irrationality is sub-optimal action
  - Rational ≠ successful
- Our focus here: rational agents
  - Systems which make the best possible decisions given goals, evidence, and constraints
  - In the real world, usually lots of uncertainty
    - ... and lots of complexity
  - Usually, we're just approximating rationality
- "Computational rationality" a better title for this course

Rational Agents

- An **agent** is an entity that perceives and acts (more examples later)
- This course is about designing rational agents
- Abstractly, an agent is a function from percept histories to actions:

$$\mathcal{P}^* \rightarrow \mathcal{A}$$



- For any given class of environments and tasks, we seek the agent (or class of agents) with the best performance
- Computational limitations make perfect rationality unachievable
- So we want the best program for given machine resources

AI-Adjacent Fields

- 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 and statistics
- Psychology
  - Adaptation
  - Phenomena of perception and motor control
  - Experimental techniques (psychophysics, etc.)
- Economics: formal theory of rational decisions
- Linguistics: knowledge representation, grammar
- Neuroscience: physical substrate for mental activity
- Control theory:
  - homeostatic systems, stability
  - simple optimal agent designs

Today

- What is AI?
- Brief History of AI
- What can AI do?
- What is this course?

A (Short) History of AI

- 1940-1950: Early days
  - 1943: McCulloch & Pitts: Boolean circuit model of brain
  - 1950: Turing's ``Computing Machinery and Intelligence''
- 1950—70: Excitement: 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
- 1970—88: Knowledge-based approaches
  - 1969—79: Early development of knowledge-based systems
  - 1980—88: Expert systems industry booms
  - 1988—93: Expert systems industry busts: "AI Winter"
- 1988—: Statistical approaches
  - Resurgence of probability, focus on uncertainty
  - General increase in technical depth
  - Agents, agents, everywhere... "AI Spring"?
- 2000—: Where are we now?

What Can AI Do?

Quiz: Which of the following can be done at present?

- ✓ Play a decent game of table tennis?
- ✓ Drive safely along a curving mountain road?
- ✗ Drive safely along Telegraph Avenue?
- ✓ Buy a week's worth of groceries on the web?
- ✗ Buy a week's worth of groceries at Berkeley Bowl?
- ? Discover and prove a new mathematical theorem?
- ✗ Converse successfully with another person for an hour?
- ? Perform a complex surgical operation?
- ✗ Unload a dishwasher and put everything away?
- ✓ Translate spoken English into spoken Swedish in real time?
- ✗ Write an intentionally funny story?

Unintentionally Funny (weird) Stories

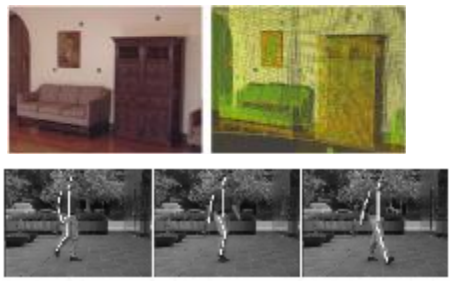
- One day Joe Bear was hungry. He asked his friend Irving Bird where some honey was. Irving told him there was a beehive in the oak tree. Joe walked to the oak tree. He ate the beehive. The End.
- Henry Squirrel was thirsty. He walked over to the river bank where his good friend Bill Bird was sitting. Henry slipped and fell in the river. Gravity drowned. The End.
- Once upon a time there was a dishonest fox and a vain crow. One day the crow was sitting in his tree, holding a piece of cheese in his mouth. He noticed that he was holding the piece of cheese. He became hungry, and swallowed the cheese. The fox walked over to the crow. The End.

[Shank, Tale-Spin System, 1984]

Natural Language

- Speech technologies
  - Automatic speech recognition (ASR)
  - Text-to-speech synthesis (TTS)
  - Dialog systems
- Language processing technologies
  - Machine translation:
    F-E:
    Aux dires de son président, la commission serait en mesure de le faire .
    According to the president, the commission would be able to do so .
    E-R-E:
    The spirit is willing but the flesh is weak.
    The vodka is good but the meat is rotten.
  - Information extraction
  - Information retrieval, question answering
  - Text classification, spam filtering, etc…

Vision (Perception)



Images from Jitendra Malik

Robotics

- Robotics
  - Part mech. eng.
  - Part AI
  - Reality much harder than simulations!
- Technologies
  - Vehicles
  - Rescue
  - Soccer!
  - Lots of automation…
- In this class:
  - We ignore mechanical aspects
  - Methods for planning
  - Methods for control

Images from stanfordracing.org, CMU RoboCup, Honda ASIMO sites

Logic

- Logical systems
  - Theorem provers
  - NASA fault diagnosis
  - Question answering
- Methods:
  - Deduction systems
  - Constraint satisfaction
  - Satisfiability solvers (huge advances here!)

Image from Bart Selman

Game Playing

- May, '97: Deep Blue vs. Kasparov
  - First match won against world-champion
  - ``Intelligent creative'' play
  - 200 million board positions per second!
  - Humans understood 99.9 of Deep Blue's moves
  - Can do about the same now with a big PC cluster
- Open question:
  - How does human cognition deal with the search space explosion of chess?
  - Or: how can humans compete with computers at all??
- 1996: Kasparov Beats Deep Blue
  "I could feel --- I could smell --- a new kind of intelligence across the table."
- 1997: Deep Blue Beats Kasparov
  "Deep Blue hasn't proven anything."

Text from Bart Selman, image from IBM's Deep Blue pages

Decision Making

- Many applications of AI: decision making
  - Scheduling, e.g. airline routing, military
  - Route planning, e.g. mapquest
  - Medical diagnosis, e.g. Pathfinder system
  - Automated help desks
  - Fraud detection
- … the list goes on.

Some Real Accomplishments of AI

- DARPA Grand Challenge – 123 miles through the desert
- Deep Space 1 – Remote Agent Experiment
- Deep Blue defeated the reigning world chess champion Garry Kasparov in 1997
- Proved a mathematical conjecture (Robbins conjecture) unsolved for decades
- Logistics Planning for 1991 Gulf War
- Computer Algebra Systems
- Credit Evaluation
- Fraud Detection

Today

- What is AI?
- Brief History of AI
- What can AI do?
- What is this course?

Course Topics

- Part 0: Introduction: Agents and Rationality (Week 1)
- Part 1: Problem Solving and Search (Week 2 – Week 4)
  - Fast search
  - Constraint satisfaction
  - Adversarial and uncertain search
- Part 2: Logical Agents (Week 4 – Week 6)
  - Propositional Logic
  - First Order Logic
  - Ontologies and Inference
- Part 3: Uncertainty and Beliefs (Week 6 – Week 9)
  - Probability
  - Bayes' nets
  - Decision theory
- Part 4: Learning Agents (Week 9 – Week 13)
  - Classification
  - MDPs and Reinforcement Learning
  - Neural Networks
- Throughout: Applications
  - Natural language
  - Vision
  - Robotics
  - Games

Course Projects

- Search and game playing
- Bayes nets
- Spam/digit recognition
- Robot learning

Some Hard Questions…

- Who is liable if a robot driver has an accident?
- Will machines surpass human intelligence?
- What will we do with superintelligent machines?
- Would such machines have conscious existence? Rights?
- Can human minds exist indefinitely within machines (in principle)?