CPS 170
Introduction

Ron Parr

Contact Information

- Professor
  - Ron Parr
  - D209 LSRC, parr at cs.duke.edu, 660-6537
  - Office hours: TBA
- TA
  - Yuqian Li
  - 006 North, yuqian at cs.duke.edu, 660-4006
  - Office hours: TBA

About Me

- My 12th year at Duke
- Bachelor's degree in philosophy (Princeton)
  - Philosophy of mind
- Ph.D. in computer science (Berkeley)
  - Hierarchical planning under uncertainty
- Current interests:
  - Planning under uncertainty
  - Probabilistic reasoning
  - Game theory
  - Reinforcement learning
  - Robotics
  - Sensing & Vision

Requirements

- Good programming skills:
  - C/C++, Java, Matlab or other high level language
  - Note: We may use python for some assignments
- Other expectations
  - Ability to do short proofs
  - Basic probability concepts
    (though we will review all of this)
  - Basic algorithmic concepts
    - Complexity - O()
    - Analysis of algorithms
  - Math
    - Basic calculus (partial derivatives)

Major Topics Covered

- Search
  - Uninformed search, informed search, CSPs
- Game Playing
  - minimax, alpha-beta search, introduction to game theory
- Logic and Knowledge Representation
  - Propositional Logic, First order logic, theorem proving
- Reasoning under uncertainty
  - Probability, Bayes nets, HMMs & tracking
- Planning
  - Classical planning, Decision theory, stochastic planning(MDPs)
- Introduction to robotics
- Introduction to machine learning

Major Topics *Not* Covered

- Natural Language
- Vision

Class Mechanics

- Textbook: *Artificial Intelligence, A Modern Approach*, Russell & Norvig (*third* edition)
  - Kindle edition is available
  - Not in the book store yet
- Homeworks: 40%
  - Discussion OK, write-up must be your own (see comments on next slide)
- Midterm: 30%
  - Closed book, in class, no collaboration
- Final: 30%
  - Closed book, finals week, no collaboration
- Homeworks will be a mix of short proofs, algorithm design/analysis, and small scale programming projects

Academic Honesty

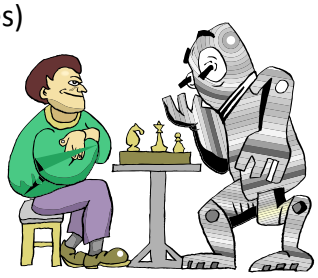
- Brainstorming with friend is encouraged!
- (But don't confuse brainstorming with letting your smart friends tell you the answers)
- You *must* write up solutions on your own
- Always ask before using code that is not your own
- Always give credit to original authors if you incorporate code that is not your own into your solutions

Attendance

- You are not required to attend class – though it's a ***good idea***
- You will not hurt my feelings if you skip class
- When you attend class:
  - Please be present in mind ***and*** body
  - Please don't:
    - Text/email
    - Surf the web
    - Play games
- About multitasking:
  - You might think you can do this w/o cost, but you can't
  - You aren't fooling anybody

Cool AI Applications

- Games (deep blue, Watson, solving checkers, video games)
- Handwriting recognition (PDAs, tablet PCs, post office)
- Speech recognition (my car, voice jail, Siri)
- MS Windows diagnostics
- E-commerce (collaborative filtering)
- Mobile robotics (grand/urban challenges)
- Space exploration
- Logistics planning
- Lots of Google tools
- Computer security



So, what is this AI stuff?

- Make machines *think* like humans
  - Is this enough?
  - Is this too much?
- Make machines *act* like humans
  - Is this sufficient?
  - Is this desirable?

Turing Test

- Computer must be indistinguishable from a human based upon written exchanges
  - Does this imply intelligence?
  - How could the computer cheat?
  - Does intelligence imply a certain type of computation?
  - Could an intelligent machine still fail the test?
- Does our notion of intelligence transcend our concept of humanity?

What Intelligence Isn't

- It's not about fooling people
- Fooling people is (in some cases) easy, e.g., eliza: http://chayden.net/eliza/Eliza.html
- More recent efforts: http://chatbots.org/

The Moving Target

- What is human intelligence?
  - At one time, calculating ability was prized
    - Now it is deprecated
    - Calculators permitted earlier and earlier in school
  - Chess was once viewed as an intelligent task
    - Now, massively parallel computers use not very intelligent search procedures to beat grand masters
    - Some say Deep Blue wasn't AI
  - Learning once thought uniquely human
    - Now it's a well-developed theory
    - Best backgammon player is a learning program
  - Vagaries of natural language once considered impenetrable
    - Now Watson wins at Jeopardy
    - Siri understands a lot of what you say
- Biological chauvinism at work?

Artificial Flight

- Even seemingly unambiguous terms such as "flight" were subject to biological chauvinism.
- Problem: Flight was largely irreducible (no easier subproblems)

- Demonstrable, unambiguous success ended chauvinism – could the same be true for AI?

Why is it hard: Ideal Intelligence

- Intelligence means making optimal choices
- Is anything truly intelligent?
- How do we define optimality?

- It took decades for people to realize that this was a thorny issue. Let's see how this played out:

Early Efforts: General (top down)

- Good news:
  - Many problems can be formalized as instances of
    - Search
    - Logical deduction
  - The space of all proofs is a (somewhat) searchable space
  - Knowledge base + theorem proving provide a satisfying picture or reasoning, knowledge and learning
    - Tell PC:
      - All men are mortal
      - Socrates is a man
    - Ask:
      - Is Socrates mortal?

Bad news for general methods

- Searching in proof space is hard
- Representing knowledge is hard
- Knowledge interconnected in strange ways
  - Chairs
  - People
  - Gravity
  - Customs…
- Early efforts were too general, ambitious
  - In most cases, could not solve the general knowledge representation problem
  - Even if KR problem was solved, theorem proving problem was intractable
- But then how does Watson work?

Early Efforts: Special Purpose (bottom up) Methods

- Neural networks
  - Attempted to reproduce function of human neurons
  - Highly abstracted from actual "wetware"
- Proverbial wing-flapping flying machine?
- Success at reproducing low-level tasks
  - Pattern recognition, associative memory
- Nearly became a religion
- Huge gap between low level and high level
- Early efforts were too specific

Overpromising and the AI Winter

- Years of
  - Naïve optimistism
  - Unrealistic assessments of challenges
  - Poor scientific/academic discipline
- Lead to (early 90's)
  - Backlash
  - Reduced government funding
  - Reduced investment from industry
  - The "AI Winter"

Intelligence: A web of abilities

- Intelligence is hard to define in isolation
- Mixture of special purpose and general purpose hardware
  - Special purpose
    - Recognizing visual patterns
    - Learning and reproducing language
  - General Purpose
    - Theorem proving
    - Learning and excelling at new tasks
- Seamless integration

- Solving pieces of the puzzle isn't enough, but it is *measurable*

AI Moving Forward

- More science/engineering
- Less philosophy

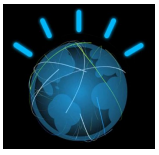
- Study broad classes of problems that would traditionally require human intelligence (but not intelligence itself)
- Restrict problem somewhat:
  - Develop a crisp input specification
  - Develop a well-defined success criterion
- Develop results with
  - Provable properties
  - Broad applicability

- Extract and study underlying principles behind successful methods

Eye on the prize

- AI's narrower focus has earned the field credibility and practical successes, yet
- Some senior researchers complain that we have taken our eye off the prize:
  - Too much focus on specific problems
  - Lack of interest in general intelligence
- Are we ready to tackle general intelligence?

Scaling Up



- Some have proposed "grand challenges" as a way to build on our successes with smaller problems
- DARPA grand challenge, CALO (precursor to Siri), Watson challenge researchers to integrate, build bigger systems
- Pros:
  - Some impressive successes
  - Some useful lessons learned about integration
- Cons:
  - Some non-general solutions
  - Siphons off a lot of resources, talent

Conclusion

- We want to solve hard problems that would traditionally require human-level intelligence. (Most we consider are at least NP-hard.)
- We want to be good computer scientists, so we force ourselves to use well-defined input/output specifications.
- We aim high, but we let ourselves simplify things if it allows us to produce a general-purpose tool with well-understood properties.
- We hope/expect/try to make these tools the building blocks of more advanced and "intelligent" systems