74.319
Introduction to Artificial Intelligence

Dr. John Anderson
Department of Computer Science
476 E2-EITC
474 8839
andersj@cs.umanitoba.ca
http://www.cs.umanitoba.ca/~andersj

Textbook

- G. Luger, *Artificial Intelligence,* Fifth Edition, Addison-Wesley, 2004
- Lots of old copies of the last edition around too, but you don't want to go back too far

Course Work

- 3-4 assignments (10%)
- A research project (15%)
  - You will choose the topic according to a set of guidelines I will provide
  - Ample room for programming/non-programming
  - Full project handout next week
- In-Class test (Monday, Oct. 23, 15%)
- Final exam (60%)

Academic Dishonesty

- The Faculty of Science and The University of Manitoba regard acts of academic dishonesty in quizzes, tests, examinations, laboratory reports or assignments as serious offences and may assess a variety of penalties depending on the nature of the offence
- Acts of academic dishonesty include, but are not limited to bringing unauthorized materials into a test or exam, copying from another individual, using answers provided by tutors, plagiarism, and examination personation
- ***Note:* cell phones, pagers, PDAs, MP3 units or electronic translators are explicitly listed as unauthorized materials, and must not be present during tests or examinations.**

Academic Dishonesty

- Penalties that may apply, as provided for under the University of Manitoba's Student Discipline By-Law, range from a grade of zero for the assignment or examination, failure in the course, to expulsion from the University. The Student Discipline By-Law may be accessed at: http://umanitoba.ca/admin/governance/policies/section_1200/1202.shtml
- Suggested minimum penalties assessed by the Faculty of Science for acts of academic dishonesty are available on the Faculty of Science web-page: http://umanitoba.ca/science/student/webdisciplinedocuments.html
- All Faculty members (and their teaching assistants) have been instructed to be vigilant and report all incidents of academic dishonesty to the Head of the Department.

Academic Dishonesty

- A declaration sheet, which states that the work being submitted is completely your own, is at http://www.cs.umanitoba.ca/newsite/honesty.html. This sheet must be printed out, filled in, signed, and submitted for every assignment.
- No grade will be recorded unless this declaration is provided.

General Class Info

- Class format will be slides like this
- These will not be made available in advance, but will be posted to the website approximately twice a week
- My advice re notes is to keep a scratchpad in class for taking side notes and note the number of the slide your notes refer to – you can transcribe them after or study them together
- All slides will be numbered

courses.cs

- Course materials and resources will be available via the department's courses page
- http://courses.cs.umanitoba.ca
- All materials except the course outline require login (Click login in the upper right)
- Your userid is your ccu id; click "forgot your password" to have an initial password mailed to yourself
- It make take a little while to get access to this, since additions to the registration list are not made in real time

IMPORTANT

- Your CCU account Needs to be working !!!
- i.e. if you have a .forward set to another account, you need to ensure that the account it's forwarded to works
- i.e. if you've gotten forwarded too much crap in hotmail and it's rejecting incoming messages, you'll never get your password
- Don't forward your ccu email elsewhere, keep your school account for school
  - and use it (please!), there's nothing more mockable than me getting questions from hot_wpg_dude_2006@hotmail.com or whatever….

Class Ground Rules

- I don't take attendance (though classes are both incredibly fun and a joy to be at, so do come – you'll find you miss things if you just rely on printed notes)
- …but I **DO** expect to you be fully interested and participatory when you are here
  - i.e. being here but reading the paper doing work on your computer once class has started isn't going to provoke a friendly reaction from me. Idle chatter, cell phones, and other non-class-related activities will annoy both me and your neighbours, and they may be harsher than me ☺

Class Ground Rules

- In a big room I may not see you; stop me anytime for a question or digression
- come see me anytime in my office (I will post official office hours where i am guaranteed to be there, but I am usually available for a few minutes at least if you find me outside those times)
- I am also the course director of 214 and have a number of graduate students I have to meet with regularly, so there may sometimes be a wait

Online Discussions

- There's a discussion group on our class website
- Feel free to use it (questions that are intended to be answered by me are best sent directly to me - I'll see email much more quickly)
  - I'll post Q&A that are likely useful to others on the discussion group myself as well
- Basic ground rules for our forum fall under the heading of "Act Like Adults"
  - keep it relevant to course topics and treat others with respect

Forum Rules

- No sharing answers to assignments, however partial
- Complaints should be directed to me personally: I am happy to discuss anything you have problems with in person
- No abuse of others
- You must be identifiable, to me at least, in forum postings. That means don't change your name to "The Phantom" or whatever. If you wish privacy from your classmates, it's fine to abbreviate "Cindy K." or whatever (provided your name is indeed Cindy ☺).
- If you're really concerned about fellow students knowing who you are, by all means submit a posting to me, and I'll post it under my id
- The forum is subject to editing by me if these guidelines are not followed. Repeated abuse will result in being banned

Further To's

- I'm interested in this stuff – I hope that I can get some of you as interested as I am
- To that end, I'm going to be pointing out a lot of side issues and interesting systems and ideas
- These will be on slides marked *Further To*, and will not be examinable material. Every week there will be a **Stuff of the Week** file to point you in new directions: again, non-examinable
- Now without further ado, let's talk a little about what you and I do that's intelligent…

Solve the following problem

DONALD
+ **GERALD**

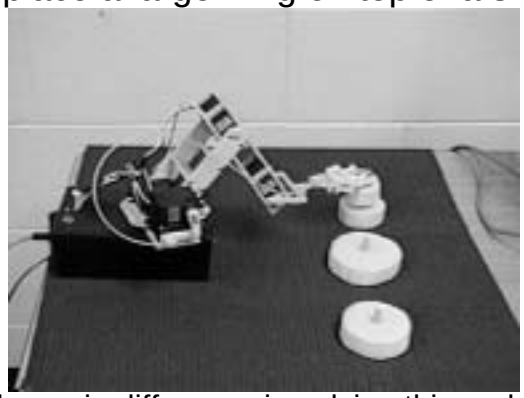
ROBERT

Hint: D=5
1:1 mapping

- D = 5, so T = 0, with carry of 1
- R = 2L + 1, so R is odd
- R = D + G = 5 + G, so R > 5 (R = 7 or 9)
- O + E = O ---> E = 0 or 9 (dep on carry in)
- Since T = 0, E must be 9 - so R = 7!
- Since E is 9, A + A = 9 with a carry in , so A = 4
- since there's a carry in where R=7, L+L=17. Since there's a carry in from D+D, L = 8
- 5 + G = 7, plus carry out from O+9, G=1
- This leaves N, B, O, and 6,3,2… just test all 3:N = 6, B = 3, O = 2 works!

Towers of Hanoi

- Move one ring at a time
- Do not place a larger ring on top of a smaller ring



- What's the main difference in solving this problem in this picture compared to the last example?

Where's John?

Where's John Part II

What can I do here?

What's the purpose of this?

What do you do when you see this?

What do you do when you see this?

What's In These Bottles?

What's In These Bottles?

What's In These Bottles?

What do you see here?

What does it take to do this?

What does it take to do this? Part II

Intelligence

- These are all situations that require intelligent information processing in various forms
- there are two broad categories here – "thinking" or exploring possibilities, and recognizing things and "knowing" what to do
  - Problem solving (thinking) skills are evident in puzzles like checkers, towers of hanoi, cryptarithmetic
  - Pattern recognition can be used to identify particular situations or useful features
    - Often a combination: that's why ordinary things from unfamiliar places are very interesting cognitively – you recognize elements immediately, but have to work to put things together

What Makes Something Intelligent?

- What would it take for me to convince you that I'm intelligent?
- There are common features that we associate with intelligence:
  - ability to respond flexibly to one's environment (adaptive behaviour)
  - ability to learn from previous situations or a teacher
  - ability to achieve goals (goal-directed behaviour)
  - ability to communicate and share ideas
  - ability to be creative
  - self-awareness

Subjectivity

- These criteria are all subjective – we wouldn't likely have identical degrees of acceptance for what it means to learn, for example
- Further, suppose I could demonstrate these things to you in a physical setting, e.g. you interacting with me via a visual or non-visual network connection
- Would that convince you I'm intelligent?
- How would your biases toward or away from machines affect this?
- What else might be necessary?

Alan Mathison Turing

- 1912-1954
  One of the greatest mathematicians in history
- Huge breadth: father of computability theory, major figure in cryptography
- Namesake of the Turing Machine
- Also interested in how computability affected the design of the human mind & intelligence in general

The Turing Test

- In his later years Turing was concerned with how one might judge an intelligent machine
- He rationalized that since we judge ourselves as intelligent, true comparison must be made between people and machine
- To use a cliché from the same year as his paper, if it looks like a duck, walks like a duck and quacks like a duck -- it's a duck…
- He rationalized that the only way to accurately judge this was to put controls in place to ensure no bias for human or machine

Turing Test

- Based on an Edwardian party game
- Involves a setting where a human and a machine are connected via teletype to a judge
- The judge can question each, and it's essentially the job of the computer to fool the judge into thinking it is human
- Turing argued that the skills necessary to do this would be enough to demonstrate true intelligence in a system achieving this goal
  - really though we use a variant each time we say performance alone is what counts

Eliza

- The Eliza (aka Doctor) program was written by Joseph Weizenbaum in 1966
- The program was designed simple linguistic parsing, and responding to keywords, to act as a Rogerian psychotherapist
  - advantageous, as program reflects the questions back to the interrogator, and generally seeming mystic at times is ok
- A conversation was accidentally carried out on the computer system at MIT between an administrator and Eliza:

Eliza

Can I use the system now?

Why do you want to use the system now?

Because I have work to do.

What work to you have to do?

. . .

- The administrator was convinced that he was talking to a real person
- This may be more of a comment on the intelligence level of University Administrators as opposed to the intelligence level of Eliza
- What other factors come into play here?
- see http://fury.com/aoliza/ for fun examples

Our Mascot Julia

- Just to show you what you can do today with this, I've set up a local chatbot, Julia
- She is not always robust when there are lots of users (and i have not updated any of her knowledge in the last year, so she is sort of stuck in 2005), but you can find her at http://gold.cs.umanitoba.ca:10047
- You can also run your own (code downloadable, see stuff of the week)
- There's a competition called the Loebner prize given out every year in a Turing test setting. A version of this bot is the current champion (which shows you how poor this is)

Julia

- has some breadth, but horribly little depth
- DOES do some intelligent processing – following snippets of dialog, manipulating pronouns
- but gives the *illusion* of intelligent responses by being able to say a little about a lot of things and guessing well on context
  - Suspension of disbelief
  - 25,000 patterns gives you a lot of possibilities
- Most people would say there's really very little intelligence going on here, even when it is behaving well
- So what does this say about judging intelligence?

Searle and the Chinese Room

- The insides matter too?
- John Searle tried to illustrate this with his Chinese Room Argument
- A room where Chinese characters are pushed in one side, someone looks in a big index file, then pulls out a matching English (or other language) response and pushes it out the other side
- Searle argued there's no understanding here, and therefore no deep intelligence (strong ai vs. weak ai - see http://www.utm.edu/research/iep/c/chineser.htm)
- No concrete proof we don't work this way…in fact some argue we do (see Brooks)

Disciplines in AI

- AI is normally considered part of computer science – however, it's easily the most multidisciplinary field in Comp Sci
- Numerous Fields beyond computer science all play a role Engineering, Psychology, Philosophy, Biology, Complex Systems (Economics, Ecology, Evolutionary Science)
- Because of this, the goals of AI are also diverse…

Goals of Artificial Intelligence

- Several Levels
- To Create Intelligent Artificial Entities (variants on this – Strong AI)
- Much more practically: To create systems and approaches to solve problems in ways that would generally be considered intelligent if those problems were solved by humans
- To find ways of making computers more effective at solving problems that currently demand human intelligence (currently??)
- To create a better understanding of the nature of intelligence

From a Computer Science Standpoint

- To find PRACTICAL solutions to HARD problems
- To find good solutions to NP-Complete problems in not only finite but usually highly-bounded amounts of time
- Most real-life problems we encounter are NP-Complete: driving to work, making a meal, analyzing a visual scene…in fact all the problems you looked at in examples earlier are theoretically very very difficult
- YET we do them functionally, and easily
- From a CS standpoint, AI is about functioning well despite the complexity!

Brief History of AI

- AI has been studied actively since the 1950's, but has been postulated and debated much earlier
- Plato discussed the idea of artificial beings
- Very complex mechanical automatons existed even in medieval times

1840: Ada Byron, Countess of Lovelace

- Took Babbage's designs for the analytical engine, and studied how they WOULD work if he could get it working
- Wrote the first theoretical programs for Babbage's computing machine, including arguments against intellectual capabilities (1843)

1940: Alan Mathison Turing

- Already mentioned, but huge in terms of his work on computability and intelligence
- And in how we evaluate and compare intelligent systems

1950's

- Era of incredibly weak computing power but lots of scientific idealism – you'll see many predictions in this period (even in its films) about where we'd be 20 years ago relative to our current time
- First neural networks (perceptrons)
- LISP
- First examination of logic in a computational sense

Early 1960's

- Search for generality
- Belief that there are general computational mechanisms for implementing all intellectual capabilities
- Poor mathematical results on perceptrons led to abandonment of most neural net research
- Newell and Simon, Physical Symbol System Hypothesis
- Newell and Simon, General Problem Solver
- Lots of theory, extremely little success in real world problems!

1960's: Herb Simon

1916-2001

- Complex Systems Theorist, 1st Nobel prize winner working in AI
- Much work on rational decision-making
- Emphasis on the interaction of simple components, forerunnner to multi-agent systems
- Emphasis on how the environment shapes intelligence

Late 60's – Early 70's

- Era of specialization
- Realization that general techniques will only work in very simple situations (weak techniques!)
- Beginning to develop specialized techniques for special-purpose domains (strong techniques)
- Still largely lab-based research work
- Specialized techniques help fragment the field

70's – 80's

- Era of Knowledge-Based Systems
- Realization that much of what we consider high performance comes from expert knowledge gathered over time
- Techniques for extracting knowledge, major advances in representing knowledge
- AI research begins to move out of labs and into the real world
  - first commercial systems, significant intelligent industrial robotic systems
- Connectionist approaches appear once more

90's to the 21st Century

- Idea that intelligence exists and is specialized for a particular environment
- AGENT perspective: building whole systems for specialized environments (embedded systems, further specialization)
  - still specialized, but more holistic!
- Major breakthroughs in machine learning
- we use intelligent systems every day, usually in a very transparent fashion – everything from speech recognition over the phone to biometrics,to games, to analysis of credit card transaction patterns

Basic Concepts in AI

- Think back to the examples of intelligent information processing we did (the first few slides of the course)
- We differentiated two "kinds" or "levels" of intelligent processing that were going on
- On one hand we reasoned consciously about many problems
- Only the other, things can often be immediately recognized and dealt with "without thinking"
- This is the most basic division in AI

Differentiating These Two Levels

- Natural boundary between these two occurs at approx. the 100ms mark in human cognition
- This is the time that it takes to recognize a face
- Formal (conscious) reasoning ("thinking") typically requires more than 100ms
- Subconscious reasoning ("pattern recognition") occurs in 100ms or less
- When things "suddenly appear" without logical connection, you're usually doing pattern recognition
  - Things we think about now may be dealt with by recognition later, if we get lots of experience!

Symbolic & Subsymbolic Models

- These two levels are also referred to as the symbolic and subsymbolic levels
- When we reason, we make connections between symbols – designations having meaning to us
  - may be conscious names or internal designations
- e.g. Socrates is a Man, All Men are Mortal, Socrates is Mortal
- or "I see a green button, Green means Go, that's probably the one to open the door!"
- When we can't illustrate this symbolic connection, we're typically working at a lower level

Example of Subsymbolic Reasoning

- Identification of gender by facial photograph
- Here there are some factors (e.g. jewelry, hairstyle, jawline) that may give some indication, but we usually don't have to reach that level, unless we're very uncertain
- We also can't typically explain why we think a face is of one gender or another either – unless there's some obvious element like the above (there are no logical connections to minor features, we just get the conclusion)
  - again, unless we've hit that logical level due to being really uncertain

Symbolic AI

- To carry out formal reasoning, knowledge is represented by a collection of symbols
- Reasoning is performed by manipulating the symbols

```
Bird(Tweety)
If Bird(x) Then HasWings(x)
∴HasWings(Tweety)
```

- Symbolic AI systems can be programmed to carry out high-level cognitive functions at an expert level of performance
- We'll be spending a lot of time on Symbolic AI

Physical Symbol Systems

- A physical symbol system consists of a collection of symbols and a set of processes that operate on the symbols
  - A system in Prolog or Lisp for example
  - Or our own inner reasoning, as implemented in wetware!
- From the idea of Physical Symbol Systems, Newell and Simon arrived at the Physical Symbol System Hypothesis:

PSS Hypothesis: Newell and Simon

- A physical symbol system has the **necessary** and **sufficient** means for general intelligent action
- Allen Newell and Herbert Simon, *Comm. of the ACM,* Volume 19, Number 3, 1976 – interesting paper
- Think of what this means…
- *NECESSARY*: We can't have general intelligent action without it
- *SUFFICIENT*: It's all we need for general intelligent action
- Remember we're talking about general intelligence as opposed to some special purpose system

Subsymbolic AI

- As the name implies, reasoning below the symbol level : somehow associating a pattern with a solution (e.g. face recognition)
- No internal designation of symbols. Instead, knowledge is represented by a collection of nodes and links, and we can't actually identify a specific purpose/fact associated with each node
- This paradigm has been used very successfully to develop intelligent systems that perform low-level cognitive tasks such as speech recognition
- very difficult to use this approach at a high level (e.g. maintaining context in dialogue!)

Subsymbolic AI

- Subsymbolic processing is often referred to as “connectionism”. Neural networks are one form of implementation of a connectionist approach (symbolic connectionism also poss.)
- Artificial neural networks perform their processing in a manner that is significantly simpler than the processing performed by the brain
  - Simply because they're both much smaller and geared to specialized tasks!
  - Brain: 100 billion neurons, a quadrillion connections (1 to 10,000 conections per neuron!)

Symbolic/Subsymbolic Models of Intelligence

Property	Conscious/Symbolic	Unconscious/Subsymbolic
engaged by	novelty	repetition
	emergencies	expected events
	danger	safety
used in	new circumstances	routine situations
can handle	decisions	nonbranching tasks
accepts	logical propositions	logic or inconsistencies
operates	sequentially	simultaneously
controls	volition	habits
capacity	relatively small	enormous
persists for	tenths of seconds	decades (lifelong)

J. Raskin, *The Humane Interface*, p. 16

Humans as Symbolic/Subsymbolic Creatures

- Considering what can be demonstrated scientifically (ignoring the mind-body problem), our knowledge is implemented in a connectionist fashion – by physical neurons
- More sophisticated than any neural net, but sharing the same general concepts
- Yet, we DO do symbolic reasoning (our inner monologue)
- At some point, we have a symbolic system implemented using subsymbolic techniques

Humans and the PSS Hypothesis

- The PSS hypothesis isn't at all harmed by the fact that we're implemented neurally
- It doesn't say that we can't have connectionism, only that it's the SYMBOLIC processing that's necessary for general intelligence – the implementation is irrelevant
- It does state essentally that if a connectionist system is generally intelligent, that it must at some point use symbolic reasoning
- It's NOT so clear on what role subsymbolic systems play in general intelligence

Modern Research and the PSS

- The power and utility of subsymbolic systems is unquestionable
- Few people would argue that they are not likely necessary for some tasks (or so much so that implementing these tasks symbolically would take multiple universe lifetimes)
- Very few deny the necessity of symbols for general intelligence, it's the sufficiency part that's questioned

Each Model has it's own Strengths

- Subsymbolic AI has been more successful in systems that perform low-level cognitive functions, while Symbolic AI is almost always used for high-level reasoning. There are situations where we see the two together
- e.g. robotic control –typically use symbolic systems to do higher level reasoning (should we turn right? Why?) and subsymbolic systems to handle low level elements (motor control)
- The crossover is obvious in many of our problem solving techniques, and is really intriguing!!!

Can Computers Think?

- Really part of the mind-body problem : If you don't believe philosophically that something is necessary beyond our physical neural implementation, then the potential is certainly there
  - very convenient really: you can just deny that whatever behaviour is exhibited involves "thought"
- It also has a lot to do with what we mean by "think"! Questions involving intelligence (even intelligence itself!) involve many subjective concepts like this
- If by "think" we mean associate concepts in useful and even novel ways, then yes!

Intelligence and Subjectivity

- As we mentioned, Intelligence, thought, all these things are subjective concepts
- We'll actually come up with a quantitative measure of intelligence for particular tasks in a bit
- But scientifically, we'd like to avoid the subjectivity
- Modern AI researchers typically don't aim for thought, they aim for RATIONALITY
- What does it mean to be rational?

Rationality

- To be rational means to do the right thing in a situation in which you find yourself
- What's the right thing? Well, that depends on your environment and your goals
- it's rational for me to go right in certain situations, left in others.
- Are we rational creatures?

Rationality

- Debatable, but generally an acceptable thought: we don't do the *wrong* thing deliberately, do we?
- Our definition of "most successful" requires some consideration - immediate or long term, for example
- Killing somebody for money might not seem irrational if the likelihood of getting caught is small

Rationality

- Rationality is also subject to the limits of knowledge
  - It's irrational to stick your fingers in an electrical socket
    - but not if you know the electricity's off
    - and not if you don't know what an electrical socket is
      - though you could argue that its irrational to stick your fingers into anything that might seem vaguely dangerous
- *When* we choose to judge is also important

Rationality and Success

- We generally define success in terms of some kind of performance measure
- For ourselves, this is a highly complex combination of things: love, money, toys, respect, power, happiness, honesty…
  - weighted very differently for different people, some very qualitative
- For the design of less broad agents than ourselves, an acceptable measure of some sort is usually quantifiable
- Rationality, then, has 4 dependencies:

Rationality

1 - the performance measure used.

2 - what the agent knows about the immediate world around it
  - Really, everything perceived so far: the percept sequence
  - Everything perceived so far: e.g. whether we've seen if the circuit beaker is on or off to the outlet we might want to stick our fingers into!

Rationality

3 - non-immediate knowledge of the world
- facts about the world in general beyond immediate perception - e.g. not knowing a shock is painful/dangerous

4. the repertoire of actions we have available
- if I'm in my office on the 5th floor and want to go to the parking lot, a fly action would be rational, unless of course I can't physically do that!

Further-To's

- Reading:
  - Turing, Alan M: Computing Machinery and Intelligence, Mind 59:433-560 (1950)
    - Online at http://www.loebner.net/Prizef/TuringArticle.html
  - Moravec, Hans P: When will Computer Hardware Match the Human Brain?
    - Online at http://www.transhumanist.com/volume1/moravec.htm
- AliceBot:
  - http://www.alicebot.org (warning – you'll have to hack code a lot, unless you want to use their version that requires a specific host)