Artificial Intelligence
Class 1: Course Overview

Dr Cynthia Matuszek (Dr M)
cmat@umbc.edu

Slides adapted with thanks from: Dr. Marie desJardin

Course Staff

- Professor: Dr. M
  - cmat@umbc.edu
  - ITE 331
  - Office hours: M 11-12, W 9:15-10:15, or by appointment
- TA: Nikhil Mengani,
  - mnikhil1@umbc.edu
  - ITE 353H
  - Office hours: TBD

My Research

- Robotics
  - How can we go from industrial robots to useful robots in human environments? (Schools, cars, homes…)
- Natural Language Processing
  - How can computers learn to understand and speak human languages (English)?
- Artificial intelligence
  - How to get computers to behave in ways that we would consider to be "intelligent"
- Human-Robot Interaction (HRI)

Today: Intro & Overview

- Review of syllabus and schedule
  - Academic honesty
  - Expectations
- Brief history of AI
  - What is AI? (and why is it so cool?)
  - What's the state of AI now?
- Topics we'll cover
- What is 'intelligence'?

From handout:
http://tiny.cc/ai-schedule
http://tiny.cc/ai-class
http://tiny.cc/ai-piazza

Classroom Policies

- Be courteous to classmates and instructors.
- No devices in use except when specified.
  - You don't learn as much.
  - People around you don't learn as much.
  - http://tiny.cc/devices-in-class
- No food or drink in this classroom.
  - Water is fine.

5

Grading

- Grades in Blackboard
  - Know your grades
    but also
  - Keep track of what's left
- Grade questions:
  - 24-hour "cooling" period
- Grade changes/regrades:
  - Requests to professor and TA
  - TA cannot change grades!

Class participation	5%
Midterm	15%
Homework	30%
Quizzes and surveys	5%
Project	25%
Final exam	20%

Pop quiz: Can Dr M add?

6

Participation

- Attend class.
- Speak up.
  - Answer questions
  - Ask questions
  - Tell us your thoughts
  - There are lots of opportunities to talk here!
- Be active on Piazza.
  - Ask and answer questions.
  - Post links to interesting material.

7

~6 Homework Assignments

- Written, problem set, and programming
  - Due at 11:59pm the day before class
  - Late: 25% off /day
- Assignments will be turned in electronically
  - Blackboard / online forms / email
  - Assignment will specify
  - 10% penalty for not following turn-in instructions
  - Example: Wrong file type
- Questions? **Piazza, *then* TA**

8

Time Management

- Some things can be rescheduled
  - E.g., overlapping exams
- Individual extensions *may* be given:
  1. With reasonable cause
  2. When made in advance
- Please talk to me!

9

Reading

- Pre-readings: Do these before that class
  - It will be hard to follow if you don't
- Readings: Do these after class
  - More detail on concepts

Before class | After class

CLASS	DATE	TOPIC	PRE-READINGS	READINGS	HOMEWORK	HANDOUTS / NOTES
1	9/1	Introduction and overview	Class web page Integrity policy	Ch. 1	Intro survey out Policies signature out HW1 out	Intro & URLs
2	9/6	Agents	2.1, 2.2 intro, 2.2.1; skim 2.3.1-2.3.2	Ch. 2	Survey due Policies signature due *By 11:59 9/5—see above*	
3	9/8	Problem-solving as search	3.1 intro, 3.1.1,			

10

Academic Integrity

- Instructor's responsibilities:
  - Be respectful
  - Be fair
  - Be available
  - Tell the students what they need to know and how they will be graded
- Students' responsibilities:
  - Be respectful
  - Do not cheat, plagiarize, or lie, or help anyone else do so
  - Do not interfere with other students' academic activities

11

Academic Integrity Policy

- *"By enrolling in this course, each student assumes the responsibilities of an active participant in UMBC's scholarly community, in which everyone's academic work and behavior are held to the highest standards of honesty.* ***Cheating, fabrication, plagiarism, and helping others to commit these acts are all forms of academic dishonesty, and they are wrong.*** *Academic misconduct could result in disciplinary action that may include, but is not limited to, suspension or dismissal."*

[Statement adopted by UMBC's Undergraduate Council and Provost's Office]

12

Integrity: Plagiarism

- Representing someone else's work as your own is plagiarism.
  - **What if the reference is in the bibliography?**
    - If you didn't explicitly quote the text you used *and* cite the source *where* you used the text, it is plagiarism.
  - **What if I only use some of the words?**
    - Scattering some of your own words and rephrasing isn't enough. If the ideas are not restated entirely in your own words, it is plagiarism.

13

Integrity: Plagiarism

- More Examples
  - **The introduction and background material are borrowed; all of the research is original.**
    - If somebody else's words appear in any document that you claim is written by you, it is plagiarism.
  - **It was a draft or not an official assignment**
    - If you represented somebody else's words as your own, even in an informal context, it is plagiarism.
  - **"But the professor *told* me to use that source!"**
    - Unless you are explicitly told to copy a quote from a source, you must write your answers *in your own words.*

14

Integrity: Abetting

- This includes putting someone's name on something when they didn't work on it.
  - "This is just everyone on our team" is wrong.
- Know what your project partners are doing.
  - Their cheating can hurt you.
- Helping another student to cheat, falsify, or plagiarize will result in you receiving **the same penalty**.

15

Integrity: What To Do

- You can **always** bring it to me
- Cheating from you / in your group / etc:
  - You **may** talk to them about it
    - Unless it's too late (it's been turned in, the test is over)
      - Then you are abetting unless you report
    - Some people may get sneakier instead of improving
- You **do not have to** talk to anyone but me

16

Integrity: Penalties

- Penalties depend on the offense and whether it recurs
- The **minimum** penalties are:
  - Receiving a zero on an assignment
  - Being required to redo the assignment, without credit, in order to pass the class
- Additional penalties may include:
  - Receiving a full grade reduction in the class
  - Failing the class without possibility of dropping it
  - Suspension or expulsion from the university

17

About Groupwork

- Study groups are encouraged!
  - Talking about the homework is completely acceptable
  - Don't share code
- Programming must be done individually
  - Programs must be written entirely by you
  - Copying another person's code is never acceptable
  - You can help debug
- Some homework is for 2-3 students working together
  - The assignment will say so; otherwise, it's individual.

18

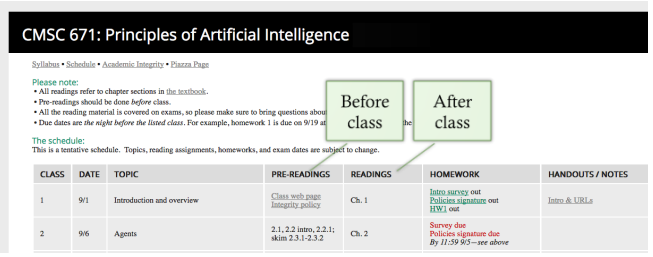
Availability & Communication

- **Post all questions to Piazza** (unless it violates integrity)
  - We will try to respond to Piazza posts immediately
  - Email takes 24-48 hours
- Always send email to professor ***and*** *TA*
  - Piazza, *then* TA, *then* prof+TA
- Office hours
- Drop by when my door is open
  - If I'm busy (often), we'll make an appointment
  - I will remain after class when I can

19

Schedule

- You will check this pretty much every class



20

What is AI?

Google

DeepMind

21

Artificial Intelligence

- Key types
  - Strong AI: mental/thought capabilities equal to (or better than) human
  - Weak (bounded) AI: intelligent actions or reasoning in some limited situations
- "Human-level" intelligence
  - In what situation?
  - Internally?
- Self-awareness / Consciousness

22

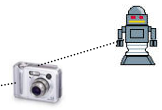
Intelligence

- These are problematic.
  - How do we measure it?
  - What's an 'intelligent action'?
    - In practice, 'previously human only'
- Is there something ineffable missing?
  - What?
- How do we test?

23

AI: A Vision

- Could an intelligent agent living on your home computer…
  - **Manage your email**
  - Coordinate your **work and social activities**
  - Help **plan your vacations**
  - **Watch your house** while you take those vacations?



24

Main Goals of AI

- **Represent** and **store** knowledge
- **Retrieve** and **reason** about knowledge
- **Behave** intelligently in complex environments
- **Learn** from environment and interactions
- Develop interesting and useful **applications**
- **Interact** with people, agents, and environment

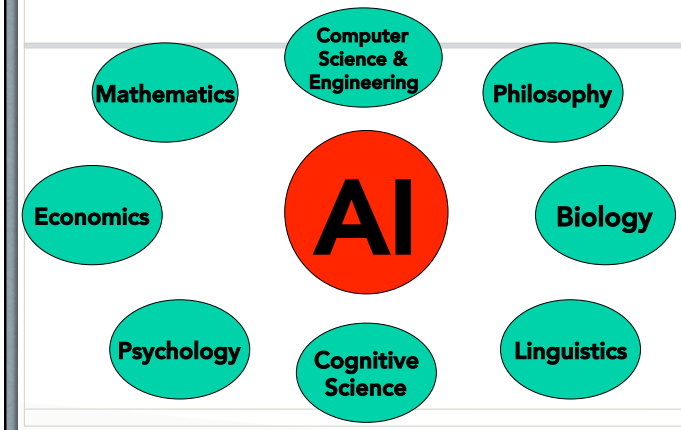
25

Why AI?

- **Engineering**
  - To get machines to do a wider variety of useful things
    - Understand spoken natural language
    - Recognize individual people in visual scenes
    - Find the best travel plan for your vacation
- **Cognitive Science**
  - Help understand how natural minds work
    - Visual perception, memory, learning, language, etc.
- **Philosophy**
  - As a way to explore interesting (and important) philosophical questions

26

Foundations of AI

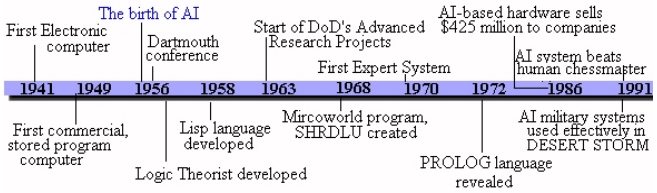


"AI hasn't accomplished anything"

- 1997: Deep Blue beats Garry Kasparov (world champion)
- 1998: Founding of Google
- 2000: Consumer robots vacuum while autonomously navigating and avoiding obstacles
- 2004: First DARPA Grand Challenge robot race
- 2007: Checkers is solved
- 2011: An AI named Watson beats the top Jeopardy! champions
- 2016: AlphaGo beats world champion at the game Go
- 2016: Self-driving car brings man with blood clot to hospital

28

A Bit of History



29

Eliza

- ELIZA: A program that simulated a psychotherapist
  - Successfully passed the Turing Test!
- Coded at MIT during 1964-1966 by Joel Weizenbaum.
- First script was DOCTOR.
  - The script was a simple collection of syntactic patterns
  - Each pattern had a reply which might include bits of the input
- Weizenbaum was shocked at reactions:
  - Psychiatrists thought it had potential.
  - People unequivocally anthropomorphized.
  - Many thought it solved the NL problem.

30

- **I am the psychotherapist. Please, describe your problems. Each time you are finished talking, type RET twice.**
- **i'm having a hard time with my students**
- **Why do you say that?**
- **well, they keep asking me for money**
- **Maybe your plans have something to do with this.**
- **what plans?**
- **Is it because of your life that you say what plans?**
- **No, it's because I didn't know what you were talking about**
- **Is it because you didn't know what I was talking about that you came to me?**
- **no, it's because i wanted to see if you could pass the turing test**
- **Does it bother you that you wanted to see if I could really pass the turing test?**
- **.....**
- **How do you reconcile problems at school?**
- **This is a waste of time. Good bye**
- **My secretary will send you a bill.**

31

PARRY

- Kenneth Colby modeled a paranoid using the same techniques circa 1968.
- PARRY has basic emotions. If it gets angry, its replies become more hostile.
- In a Turing test, expert psychiatrists were unable to distinguish PARRY's ramblings from those of real paranoids.

32

H: Good afternoon. My name is Doctor Jones.
P: Pleased to meet you.
H: How are you?
P: All right I guess.
H: Could you tell me your name?
P: Pat Smith.
H: I though you might have some problems.
P: I've been upset lately.
H: By what?
P: People get on my nerves sometimes.
H: Tell me more.
P: Do you know anything about bookies?

Parry Meets Eliza

- P: I don't understand your motives.
- E: Why don't you understand my motives?
- P: Shouldn't I be?
- E: Why do you say that just now?
- P: That's just the way it is.
- E: Suppose that's the way it were not.
- P: So?
- E: I am not sure I understand you fully.
- P: You harp on the same questions too much.
- E: What does that similarity suggest to you?
- P: I don't get you.
- E: Do you wish that you did get I?

33

We've Come a Long Way (or have we?...)

34

Big Questions

- Can machines think?
- If so, how?
- If not, why not?
- What does this say about human beings?
- What does this say about the mind?

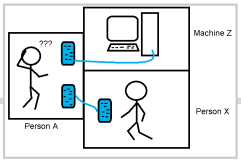
35

What's Easy and What's Hard?

- It's easi(er) to mechanize high-level tasks
  - Symbolic integration
  - Proving theorems
  - Playing chess
  - Medical diagnosis
- It's hard to mechanize tasks that lots of animals can do
  - Walking around without running into things
  - Catching prey and avoiding predators
  - Interpreting complex sensory information (e.g., visual, aural, …)
  - Modeling the internal states of other animals from their behavior
  - Working as a team (e.g., with pack animals)
- Is there a fundamental difference?

36

Turing Test



- Three rooms:
- 1 person, 1 computer, and 1 interrogator
  - The interrogator can communicate with the other two
  - The interrogator tries to decide which is the person
  - Both try to convince the interrogator they are the person
- If the machine succeeds, the machine can think

…Right? (no)

Image: filipinofreethinkers.org/2012/06/23/turings-tremendous-talent-and-trenchant-test/turing-test

The Loebner Contest

- A modern version of the Turing Test, held annually
  - $100,000 cash prize.
  - Hugh Loebner was once director of UMBC's Academic Computing Services (née UCS)
- Restricted topic (removed in 1995) and limited time.
- Participants: set of humans, set of computers, set of judges.
- Scoring
  - Rank from least human to most human.
  - Highest median rank wins $2000

What Can AI Systems Do Now?

- **Computer vision:** face recognition from a large set
- **Natural language processing:** machine translation
- **Expert systems:** medical diagnosis in a narrow domain
- **Spoken language systems:** ~1000 word continuous speech
- **Planning and scheduling:** Hubble Telescope experiments
- **Robotics:** autonomous (mostly) automobile
- **User modeling:** Bayesian reasoning in Windows help (the infamous paper clip…)
- **Games:** Grand Master level in chess (world champion), perfect play in checkers, Go
- **Search:** You've used Google.
- **Learning:** So much learning.

What Can't AI Systems Do Yet?

- Understand natural language robustly
- Learn a natural language
- Surf the web
- Interpret an arbitrary visual scene
- ~~Play Go as well as the best human players~~
- Construct plans in dynamic real-time domains
- Refocus attention in complex environments
- Perform life-long learning

Exhibit true autonomy and intelligence?

Who Does AI?

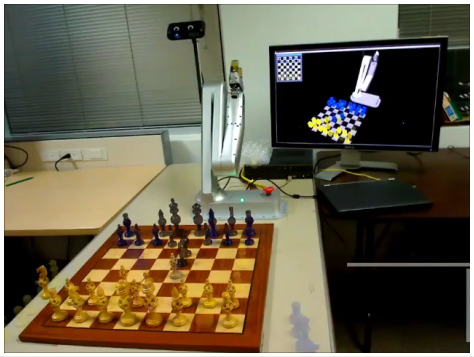
- Academic researchers (perhaps the most Ph.D.-generating area of computer science in recent years)
  - Some top schools: CMU, Stanford, Berkeley, MIT, UW, UMd, U Alberta, UT Austin, … (and, actually, UMBC!)
- Government and private research labs
  - NASA, NRL, NIST, IBM, AT&T, SRI, ISI, MERL, …
- Lots of companies!
  - Google/Alphabet, Microsoft, Amazon, Honeywell, Teknowledge, SAIC, MITRE, Fujitsu, Global InfoTek, BodyMedia, …

41

Applications

42

Game Playing



43

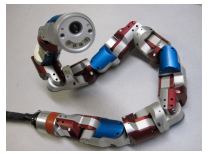
Text/Sketch Recognition

44

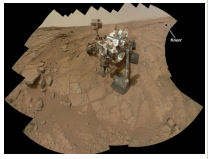
User Modeling & NLP

45

Robotics



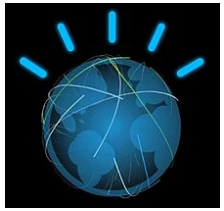


46

Knowledge Representation

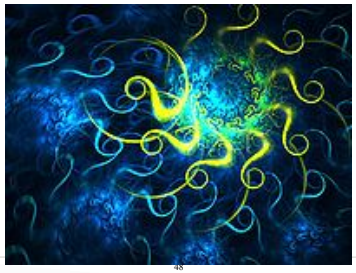
Watson



47

47

Evolutionary Art

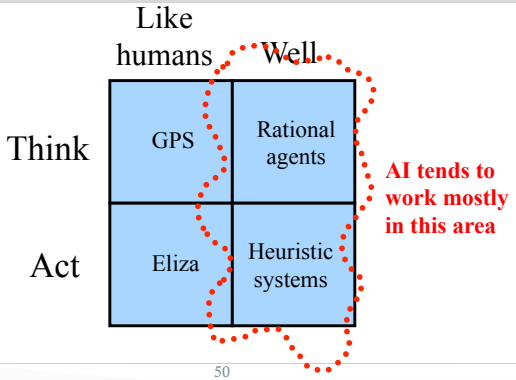


48

Computer Vision

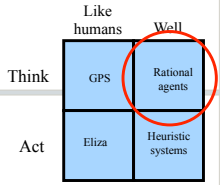
49

Possible Approaches



50

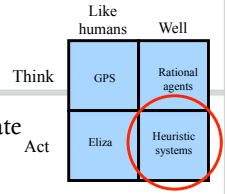
Thinking Well



- Develop formal models of knowledge representation, reasoning, learning, memory, and problem solving, that can be rendered in algorithms.
- There is often an emphasis on systems that are provably correct, and guarantee finding an optimal solution.

51

Acting Well



- For a set of inputs, generate an appropriate output that is not necessarily correct but gets the job done.
- A **heuristic** (**heuristic rule**, **heuristic method**) is a rule of thumb, strategy, trick, or any other kind of device which drastically limits search for solutions in large problem spaces.
- Heuristics do not guarantee optimal solutions; in fact, they do not guarantee any solution at all: **all that can be said for a useful heuristic is that it offers solutions which are good enough most of the time.**

  – Feigenbaum and Feldman, 1963, p. 6

52

Thinking Like Humans

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- Cognitive science approach
- Focus not just on behavior and I/O
  - Also look at reasoning process.
- Computational model reflects "how" results were obtained
- Providea a new language for expressing cognitive theories and new mechanisms for evaluating them
- GPS (General Problem Solver):
  - Not just to produce humanlike behavior, but to produce a sequence of steps of the reasoning process similar to the steps followed by a person

53

Acting Like Humans

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

- Behaviorist approach.
- Not about how you get results, just the similarity to what human results are.
- Exemplified by the Turing Test

54

What about Statistical Methods?

	Like humans	Well
Think	GPS	Rational agents
Act	Eliza	Heuristic systems

55

For Next Time

- Due at 11:59pm before next class:
  - Fill out the survey
  - Read academic integrity statement
  - Sign up for Piazza and join this class
- Look at the reading lists
- Do pre-reading for next time

56