CS486/686: Introduction to Artificial Intelligence
Lecture 1 - Introduction

Jesse Hoey & Victor Zhong

School of Computer Science, University of Waterloo

January 7, 2025

Readings: Poole & Mackworth 1.1

People, Books & Web

Instructors: Jesse Hoey (S3 & S2) & Victor Zhong (S1)

People, Books & Web

Instructors: Jesse Hoey (S3 & S2) & Victor Zhong (S1)

TAs:

- Negar Arabzadehghahyazi (narabzad)
- Arthur Chen (h559chen)
- Varun Dhanraj (vdhanraj)
- Yuxuan Li (y624li)

- Tien Dat Nguyen (tdnguyen)
- Ruoxi Ning (r2ning)
- Hala Sheta (hsheta)
- Shivani Upadhyay (sjupadhy)
- Dake Zhang (d346zhan)

People, Books & Web

Instructors: Jesse Hoey (S3 & S2) & Victor Zhong (S1)

TAs:

- Negar Arabzadehghahyazi (narabzad)
- Arthur Chen (h559chen)
- Varun Dhanraj (vdhanraj)
- Yuxuan Li (y624li)

- Tien Dat Nguyen (tdnguyen)
- Ruoxi Ning (r2ning)
- Hala Sheta (hsheta)
- Shivani Upadhyay (sjupadhy)
- Dake Zhang (d346zhan)

Lectures:

- Section 3: T/Th 1:00pm-2:20pm in MC 2038 (Jesse)
- Section 2: T/Th 8:30am-9:50am in MC 2038 (Jesse)
- Section 1: T/Th 2:30-3:50pm in MC 2017 (Victor)

People, Books & Web

Instructors: Jesse Hoey (S3 & S2) & Victor Zhong (S1)

TAs:

- Negar Arabzadehghahyazi (narabzad)
- Arthur Chen (h559chen)
- Varun Dhanraj (vdhanraj)
- Yuxuan Li (y624li)

- Tien Dat Nguyen (tdnguyen)
- Ruoxi Ning (r2ning)
- Hala Sheta (hsheta)
- Shivani Upadhyay (sjupadhy)
- Dake Zhang (d346zhan)

Lectures:

- Section 3: T/Th 1:00pm-2:20pm in MC 2038 (Jesse)
- Section 2: T/Th 8:30am-9:50am in MC 2038 (Jesse)
- Section 1: T/Th 2:30-3:50pm in MC 2017 (Victor)

Office hours:

Wednesday 2:00-3:00pm DC 3613/online (Jesse)
Tuesday 10:30-11:30am DC 2637 (Victor)

Office hours (TA): Notified near assignment due dates

Grade Breakdown

CS486 (Undergraduate Students)

- 4 Assignments (40%: 10% each) – approx deadlines: weeks 5, 7, 9, 12
- 1 midterm exam (15%) – Feb 13th 7:00 pm to 8:50 pm
- 1 final exam (45%) (must pass to pass the course)
- Optional project (5% bonus, proposal due at midterm)

CS686 (Graduate Students)

- 4 Assignments (25%: 6.25% each) – same deadlines
- 1 midterm exam (10%) – same time
- 1 final exam (35%)
- 1 course project report (30%, proposal due at midterm)

Grade Breakdown

CS486 (Undergraduate Students)

- 4 Assignments (40%: 10% each) – approx deadlines: weeks 5, 7, 9, 12
- 1 midterm exam (15%) – Feb 13th 7:00 pm to 8:50 pm
- 1 final exam (45%) (must pass to pass the course)
- Optional project (5% bonus, proposal due at midterm)

CS686 (Graduate Students)

- 4 Assignments (25%: 6.25% each) – same deadlines
- 1 midterm exam (10%) – same time
- 1 final exam (35%)
- 1 course project report (30%, proposal due at midterm)

Students wishing to write a project (and all CS686 students) **must** submit a project proposal

Grade Breakdown

CS486 (Undergraduate Students)
- 4 Assignments (40%: 10% each) – approx deadlines: weeks 5, 7, 9, 12
- 1 midterm exam (15%) – Feb 13th 7:00 pm to 8:50 pm
- 1 final exam (45%) (must pass to pass the course)
- Optional project (5% bonus, proposal due at midterm)

CS686 (Graduate Students)
- 4 Assignments (25%: 6.25% each) – same deadlines
- 1 midterm exam (10%) – same time
- 1 final exam (35%)
- 1 course project report (30%, proposal due at midterm)

Students wishing to write a project (and all CS686 students) **must** submit a project proposal

The final is cumulative (i.e., covering all course material) with a focus on the post-midterm material

Assignment Deadlines

- Assignment deadline times are at 11:59pm waterloo time on the day indicated.
- All assignment deadlines are fixed and will not be extended.
- Every assignment has a **grace period** of 72 hours.
- If you use university approved STA/VIF on assignment, you enter grace period: no late penalty.
- No assignment will be accepted after the grace period ends.
- No assignment questions will be answered during the grace period.
- Any student missing the midterm with university approved STA/VIF will have the midterm weight shifted to the final.

Course Project

Optional for CS486 students (5% bonus)

Mandatory for CS686 students (30% of grade)

You **must** submit a correctly constructed and formatted proposal by the midterm – will be pass/fail with **no mark**

Final project dues before the final exam

Course Project

Optional for CS486 students (5% bonus)

Mandatory for CS686 students (30% of grade)

You **must** submit a correctly constructed and formatted proposal by the midterm – will be pass/fail with **no mark**

Final project dues before the final exam

Individual project (CS686)

Group project (CS486, up to 3 members):

- **Must be substantially** more involved than individual projects
- Each team member's contribution **to the project must be clearly and specifically described in both the proposal and report**
- There must be more papers referenced and discussed for team projects

Course Project

Optional for CS486 students (5% bonus)
Mandatory for CS686 students (30% of grade)

You **must** submit a correctly constructed and formatted proposal
by the midterm – will be pass/fail with **no mark**

Final project dues before the final exam

Individual project (CS686)

Group project (CS486, up to 3 members):

- **Must be substantially** more involved than individual projects
- Each team member's contribution **to the project must be clearly and specifically described in both the proposal and report**
- There must be more papers referenced and discussed for team projects

```
https:
//cs.uwaterloo.ca/~jhoey/teaching/cs486/projects.html
```

GenAI Policy

GenAI (ChatGPT, etc.) are known to be unreliable
(https://arxiv.org/pdf/2309.01219), and their usage in this
class is **strongly discouraged**.

Your ability to solve problems on a final exam will be severely
impacted if you don't solve assignment problems yourself using the
methods we learn in class. **Any usage of these tools must be
declared on all submitted work. Usage of GenAI without
declaring exactly what it was used for will result in a reduced
grade for the work.** Further, for research projects, these tools will
generate the uniform content across users, and projects using
GenAI may also lose marks for a lack of originality.

Textbooks

Textbook:

- David Poole and Alan Mackworth
  Artificial Intelligence: Foundations of Computational Agents
  (3rd edition) https://artint.info

Secondary textbooks:

- Russell and Norvig
  Artificial Intelligence https://aima.cs.berkeley.edu
- Ian Goodfellow, Yoshua Bengio and Aaron Courville
  Deep Learning - https://deeplearningbook.org

Course materials: LEARN

Discussion forum: Piazza,
https://piazza.com/uwaterloo.ca/winter2025/cs486686

Assignments handed in and returned, grades: LEARN

Overview of the Course

Overview of the Course

Introduction, Agents and AI

Overview of the Course

Introduction, Agents and AI

Deterministic AI

- Representation and reasoning
  - States and searching
  - Features and constraints
- Supervised learning
  - Regression
- Deterministic planning

  $\approx$ end of midterm material

Overview of the Course

Introduction, Agents and AI

Deterministic AI

- Representation and reasoning
  - States and searching
  - Features and constraints
- Supervised learning
  - Regression
- Deterministic planning

  $\approx$ end of midterm material

Uncertainty in AI

- Representation and reasoning
  - Uncertainty and probability
  - Neural networks
- Learning
  - Bayesian learning
  - Unsupervised learning
  - Deep learning
- Planning with uncertainty
  - Markov decision processes
  - Reinforcement learning
- Topics on language models
  - N-Gram language models
  - Neural language models

Academic Integrity and Intellectual Property

See also official course outline on LEARN Property of UW:

- Lecture content, spoken and written (and any audio/video recording thereof)
- Lecture handouts, presentations, and other materials prepared for the course (e.g., PowerPoint slides)
- Questions or solution sets from various types of assessments (e.g., assignments, quizzes, tests, final exams)
- Work protected by copyright (e.g., any work authored by the instructor or TA or used by the instructor or TA with permission of the copyright owner)

Academic Integrity and Intellectual Property

See also official course outline on LEARN Property of UW:

- Lecture content, spoken and written (and any audio/video recording thereof)
- Lecture handouts, presentations, and other materials prepared for the course (e.g., PowerPoint slides)
- Questions or solution sets from various types of assessments (e.g., assignments, quizzes, tests, final exams)
- Work protected by copyright (e.g., any work authored by the instructor or TA or used by the instructor or TA with permission of the copyright owner)

Sharing intellectual property without the intellectual property owner's permission is a violation of intellectual property rights

Current Research in AI

Organizations

- Association for the Advancement of Artificial Intelligence (AAAI)
- International Machine Learning Society

Current Research in AI

Organizations

- Association for the Advancement of Artificial Intelligence (AAAI)
- International Machine Learning Society

Journals

- Artificial Intelligence
- Journal of AI Research
- Journal of Machine Learning Research
- arXiv categories cs.AI, cs.LG, cs.CL, cs.CV, cs.RO

Current Research in AI

Organizations

- Association for the Advancement of Artificial Intelligence (AAAI)
- International Machine Learning Society

Journals

- Artificial Intelligence
- Journal of AI Research
- Journal of Machine Learning Research
- arXiv categories cs.AI, cs.LG, cs.CL, cs.CV, cs.RO

Conferences

- International Joint Conferences on A.I. (IJCAI)
- AAAI Conference on Artificial Intelligence (AAAI)
- Neural Information Processing Systems (NeurIPS)
- International Conference on Machine Learning (ICML)
- International Conference on Learning Representations (ICLR)
- Conferences on special topics such as NLP, Vision, Robotics, etc.

What is Artificial Intelligence (AI)?

What is Artificial Intelligence (AI)?

The synthesis and analysis of computational agents that act *intelligently*

What is Artificial Intelligence (AI)?

The synthesis and analysis of computational agents that act *intelligently*

Intelligence is a matter of **degree**

The aspects that go into an agent acting intelligently include

- What it does is appropriate for its **circumstances**, its **goals**, and its **perceptual and computational limitations**
- It takes into account the **short-term and long-term consequences** of its actions, including the effects on society and the environment
- It **learns** from experience
- It is **flexible** to changing environments and changing goals

Goals of Artificial Intelligence Research

Goals of Artificial Intelligence Research

Scientific: Understanding the principles that making intelligent behaviour possible in natural or artificial systems

- **Analysis** of natural and artificial agents
- **Testing hypotheses** about the construction of intelligent agents
- **Building and experimenting with** intelligent agents

Goals of Artificial Intelligence Research

Scientific: Understanding the principles that making intelligent behaviour possible in natural or artificial systems

- **Analysis** of natural and artificial agents
- **Testing hypotheses** about the construction of intelligent agents
- **Building and experimenting with** intelligent agents

Engineering: Designing and building useful intelligent artifacts

Subareas of AI

Machine learning: Designing algorithms that learn from data

Deep learning: Machine learning, with specific focus on neural networks

Natural language processing and **computational linguistics**: Understanding and generating human language

Computer vision: Interpreting and generating visual information

Subareas of AI

Machine learning: Designing algorithms that learn from data

Deep learning: Machine learning, with specific focus on neural networks

Natural language processing and **computational linguistics**: Understanding and generating human language

Computer vision: Interpreting and generating visual information

Robotics: Building robots that can act autonomously

Subareas of AI

Machine learning: Designing algorithms that learn from data

Deep learning: Machine learning, with specific focus on neural networks

Natural language processing and **computational linguistics**: Understanding and generating human language

Computer vision: Interpreting and generating visual information

Robotics: Building robots that can act autonomously

Cognitive science: Understanding how humans and animals think, learn, and act

History of AI: The Turing Test

Alan Turing: 1912-1954 was primariliy known for codebreaking in WW11 (https://www.imdb.com/title/tt2084970/)

He also asked '*can machines think?*' and proposed the **imitation game** in 1950

An interrogator is connected to one person and one machine via a terminal, therefore can't see her counterparts. Her task is to find out which of the two candidates is the machine, and which is the human only by asking them questions. If the machine can "fool" the interrogator, it is intelligent.

Conversations with ELIZA - (Weizembaum 1966, Lisp)

ME: Hi
ELIZA: How do you do? What brings you to see me?
ME: I'm thinking of taking a course in artificial intelligence.
ELIZA: Why do you say that?
ME: I'm worried about it
ELIZA: Are you worried often?
ME: when I take AI courses
ELIZA: What happened when you take ai courses?
ME: I usually spend long nights doing assignments
ELIZA: Earlier you said you felt worried?
ME: about AI courses
ELIZA: Maybe your plans have something to do with this.
ME: no
ELIZA: Why not?
ME: because I think you are a computer
ELIZA: Is it because you think I am a computer that you came to me?
ME: no
ELIZA: Does it bother you that I ask that?
ME: yes

Loebner Prize: A Turing Test Competition

Fool all judges: $1M prize

No one has ever won

Example interactions:
CONFEDERATE: Hello, my name is Sam and I am the human.
JUDGE: Hi Sam, tell me which is larger, a TV screen or Belgium.
CONFEDERATE: Belgium is a little bit larger

JUDGE: Wljso lwjejdo wleje elwjeo wjeol, isn't it?
CONFEDERATE: Is that Polish?

JUDGE: Which is further away from this apartment, Queens or the Queen of England
CONFEDERATE: If the Queen is visiting Manhattan, she is closer.

Would ChatGPT win the Loebner Prize?

SHRDLU: A Program for Blocks World (Winograd, 1968)

First autonomous system supporting physical interaction

Blocks world only

Example conversation:

Pick up a big red block

OK

grasp the pyramid

I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN

What does the box contain?

THE BLUE PYRAMID AND THE BLUE BLOCK.

What is the pyramid supported by?

THE BOX.
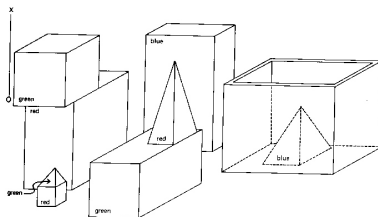
Game Playing

Two-player, zero-sum games

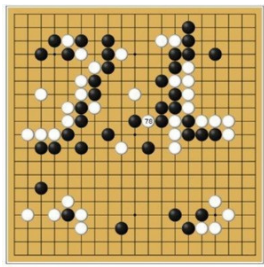Game tree for tic-tac-toe:
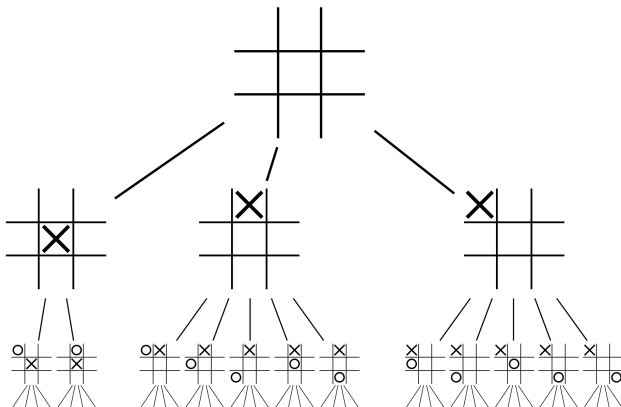
Go

The game tree has more than 3^{361} nodes

- 3: Possible state for a single cell
- 361: Number of cells (19×19) on the board

Google DeepMind: AlphaGo (2016)
https://www.nature.com/articles/nature16961
March 2016: AlphaGo beats Lee Sedol in 4/5 games
May 2017: AlphaGo beats Ke Jie in 3/3 games

StarCraft

A multi-agent problem with

- Imperfect information (partially observed map)
- Large action space (10^8 possibilities)
- Delayed credit assignment
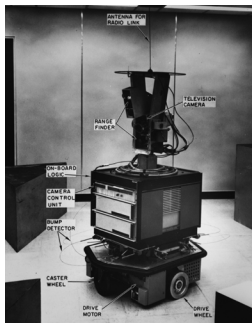
StarCraft

A multi-agent problem with

- Imperfect information (partially observed map)
- Large action space (10^8 possibilities)
- Delayed credit assignment

Google DeepMind:
https://www.nature.com/articles/s41586-019-1724-z

Robotics

Shakey
SRI
1970

José
UBC
2000

HRP-4C
AIST - Hiroshi Ishiguro
2010

Robocup

Soccer playing robots

Robocup 2023:
https://www.youtube.com/watch?v=Ub1ZO2dVKXM

Autonomous Cars

"Stanley" (team led by Sebastian Thrun from Stanford) won the
2005 DARPA Grand Challenge

Autonomous Cars: 20 Years

Waymo-operated Jaguar
I-Pace in SF (2023)

Autopilot from Tesla

Are Self-Driving Cars (or AlphaGo etc.) Intelligent?

Are Self-Driving Cars (or AlphaGo etc.) Intelligent?

Recap: AI cares about the synthesis and analysis of computational agents that act *intelligently*

Intelligence is a matter of **degree**

The aspects that go into an agent acting intelligently include

- What it does is appropriate for its **circumstances**, its **goals**, and its **perceptual and computational limitations**
- It takes into account the **short-term and long-term consequences** of its actions, including the effects on society and the environment
- It **learns** from experience
- It is **flexible** to changing environments and changing goals

Next

- Agents (Poole & Mackworth chapter 1.3-1.10)
- Search (Poole & Mackworth chapter 3)