CS 486/686: Introduction to Artificial Intelligence

Introduction

Plan for Today

- Introductions
- Logistics for the course
- What is AI?
  - What will **this** course be about?

Who Am I?

My name is Kate Larson. I am a Professor and University Research Chair at U of Waterloo and a Research Scientist with Google DeepMind.

My work/education history:

- PhD, Computer Science, Carnegie Mellon University
- Undergraduate, Mathematics, Memorial University of Newfoundland

Research Interests: artificial intelligence, multiagent systems, game theory and social choice, reinforcement learning

Introductions: Meet Your Peers

- In the next two minutes introduce yourself to someone you don’t know.
- Talk about courses, co-op, summer activities, the headaches involved in finding somewhere to live, extra-curricular activities, graduation, jobs, etc,…

Course Logistics

- Instructor: Kate Larson (DC 2518)
- Lectures:
  - Tuesdays and Thursdays at 8:30-9:50 and 11:30-12:50
- Office Hours:
  - In person: Thursdays from 10:15-11:15 in DC 2518
  - Online: Mondays from 2:00-3:00 pm (on Teams, see course website for more info)
- Great team of TAs
  - Amin Bigdali, Jess Gano, Aryan Hadday, Liam Herbert, Yanting Miao, Kyle Tilbury, Dake Zhang

Course Logistics

- Website: We will be using Learn
  - Schedule
  - Lecture Notes
  - Assignments and Assignment Submission
- Communications:
  - Learn
  - Piazza: Details for signing up are in the syllabus
- Texts
  - Artificial Intelligence: A Modern Approach by S. Russell and P. Norvig (3rd or 4th edition)
  - Artificial Intelligence: Foundations of Computational Agents by D. Poole and A. Mackworth (available for free online – see syllabus)

Evaluation

CS 486

- 4 Assignments 40%
- Midterm (Oct 16) 20%
- Final Exam 40%

CS 686

- 4 Assignments 28%
- Midterm (Oct 16) 12%
- Final Exam 35%
- Project Proposal 0%
- Project 25%

Late Submissions

- The due-date for assignments will be clearly indicated on each assignment.
  - Late assignments will be accepted up to 48 hours after the due date without penalty.
  - After the 48-hour grace-period assignments are not accepted.
  - Course personnel do not provide help on assignments after the due date.

Generative AI and Assignments

- The recent advances in generative AI are exciting and have opened up many avenues for AI applications and research.

- While we encourage you to explore and experiment with generative AI, these models are **not** to be used in assignments **unless the assignment clearly states that they are allowed. If the use of generative AI is allowed on an assignment (or part of an assignment) then proper documentation, citation, and acknowledgement is required**

- **You are accountable for the content and accuracy of all work you submit in this class, including any supported by generative AI**

What is Artificial Intelligence?

What is AI? (Different Definitions from R&N)

The exciting new effort to make computers that think… machines with minds in the full and literal sense [Haugeland 85] [The automation of] activities that we associate with human thinking, such as decision making, problem solving, learning [Bellman 78]	The study of mental faculties through the use of computational models [Charniak & McDermott 85] The study of computations that make it possible to perceive, reason and act [Winston 92]
The art of creating machines that perform functions that require intelligence when performed by a human [Kurzweil 90] The study of how to make computers do things at which, at the moment, people are better [Rich&Knight 91]	A field of study that seeks to explain and emulate intelligent behavior in terms of computational processes [Schalkoff 90] The branch of computer science that is concerned with the automation of intelligent behavior [Luger&Stubblefield93]

My Favorite Definition

"Artificial Intelligence is that activity devoted to making machines intelligent, AND intelligence is that quality that enables an entity to function appropriately and with foresight in its environment."

Nils J. Nilsson

What is AI?

Definitions tend to differ along two dimensions:

- Reasoning vs behaviour
- Fidelity to humans vs rationality

Systems that **think** like **humans**	Systems that **think** **rationally**
Systems that **act** like **humans**	Systems that **act** **rationally**

What is AI?

Systems that **think** like **humans**	Systems that **think** **rationally**
Systems that **act** like **humans**	Systems that **act** **rationally**

Systems that think like humans

- Cognitive science
  - Brain imaging
  - Introspection
  - etc
- Fascinating area, but we will not be covering it in this course

What is AI?

Systems that **think** like **humans**	Systems that **think** **rationally**
Systems that **act** like **humans**	Systems that **act** **rationally**

Systems that think rationally

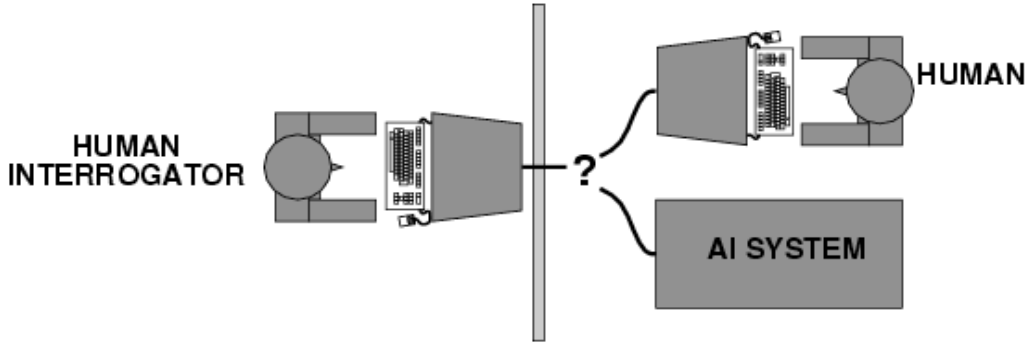
- What are the correct thought processes? (Aristotle)
- Systems that reason in a logical manner
- Systems that do inference correctly

What is AI?

Systems that **think** like **humans**	Systems that **think** **rationally**
Systems that **act** like **humans**	Systems that **act** **rationally**

Systems that act like humans

- Turing Test: Introduced by Alan Turing in the 1950 paper "Computing machinery and intelligence"
- Predicted that by 2000 a computer would have a 30% chance of fooling a lay person for 5 minutes
- Anticipated all major arguments against AI
- Suggested the major components of AI
  - Knowledge, Reasoning, Language Understanding, Learning



What is AI?

Systems that **think** like **humans**	Systems that **think** **rationally**
Systems that **act** like **humans**	Systems that **act** **rationally**

Systems that act rationally

- Rational behaviour: “doing the right thing”
  - Does not require “intelligence” or “thinking” but it often helps!
- Rational Agent paradigm (This course)
  - Agent: entity that perceives and acts
  - Rational agent: acts so as to achieve best outcome

What is AI?

We are interested in intelligent systems where

a system is intelligent if and only if it acts rationally

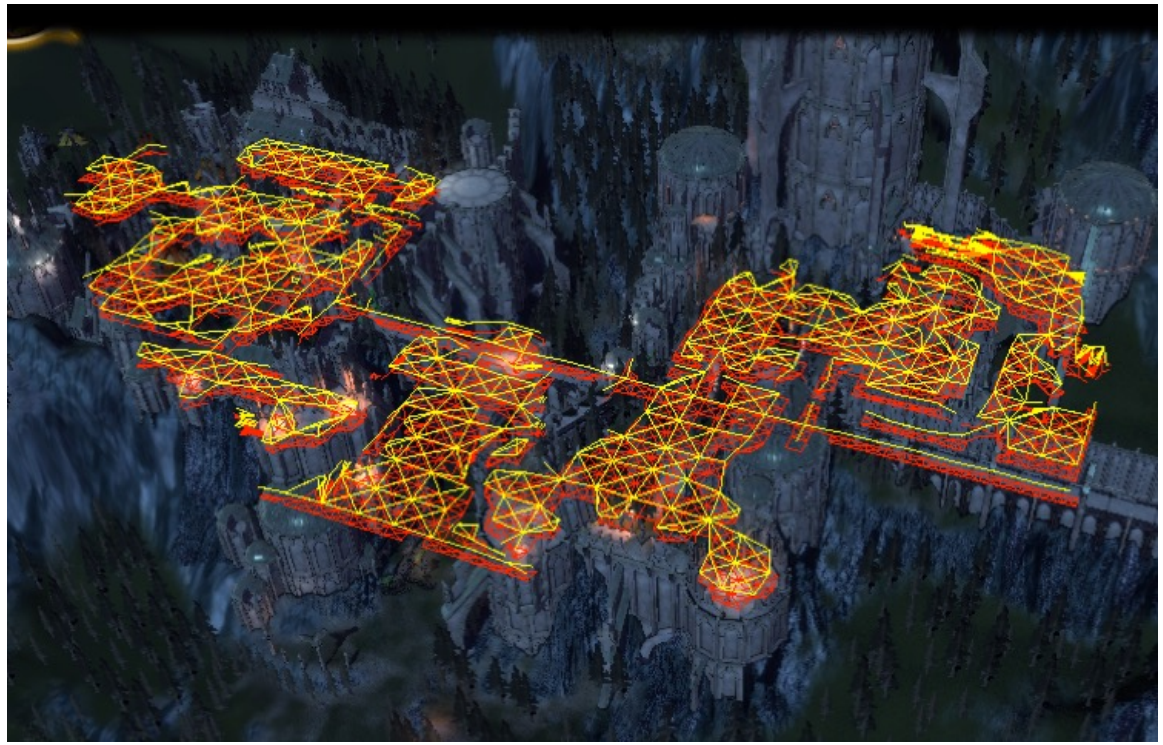
Questions to think about

- Why do we care about behaviour instead of thought processes and reasoning?
- Why do we measure success against rationality instead of against humans?

Topics We Will Cover

- **Search**
  - Uninformed and heuristic search
  - Constraint Satisfaction
- **Reasoning Under Uncertainty**
  - Probability Theory and Decision Theory
  - Probablistic Inference, Causal Inference
  - Bayesian Networks, Markov Decision Processes
- **Learning**
  - Decision Trees, Statistical Learning, Neural Networks
  - Bandits, Reinforcement Learning
- **Multiagent Systems**
  - Game-tree search, Game Theory, Multiagent Reinforcement Learning

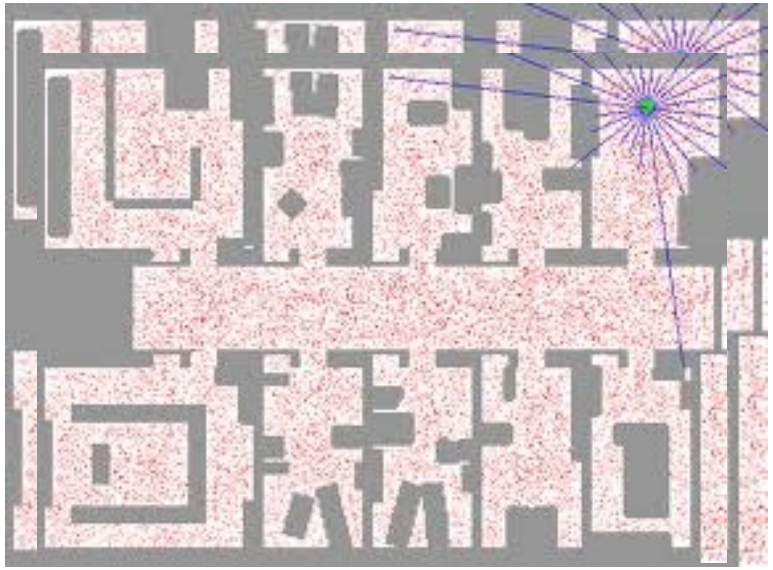
Search



From Fener, Koenig, Sturtevant, AAAI 2020

7		1			3		6	8
	6	3	2	5				
8					6	5		3
			8		9		7	
2			1		4			9
	9		5		7			
1		8	4					6
				8	2	1	4	
5	3		6			9		2

Reasoning Under Uncertainty

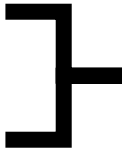


CoBots, CORAL at CMU

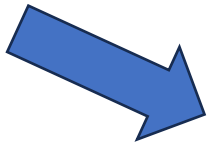
Machine Learning

- Provide examples to the machine
- Machine learns to accomplish tasks based on examples

- Three Categories
  - Supervised Learning
  - Unsupervised Learning
  - Reinforcement Learning

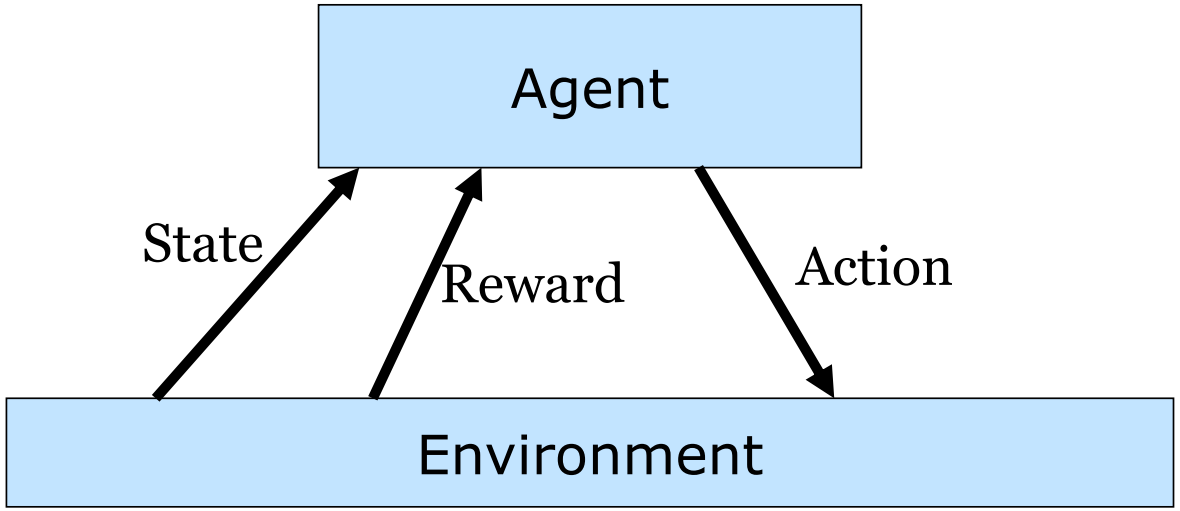


Covered in depth in CS 480



Covered in CS 486

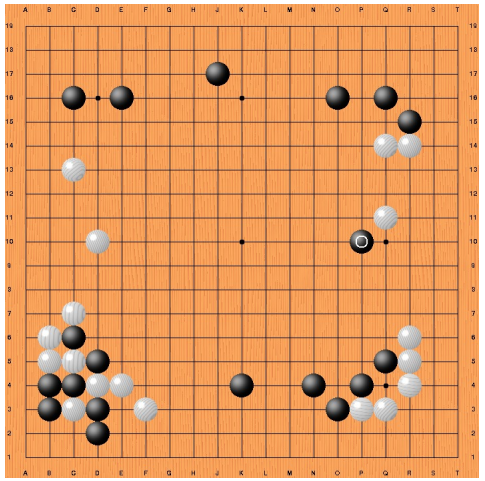
Reinforcement Learning Problem



Goal: Learn to choose actions that maximize rewards

Multiagent Systems

How should an agent reason and behave when there are other agents



2016: AlphaGo defeats top player Lee Sedol (4-1)
Game 2 move 37: AlphaGo plays unexpected move (odds 1/10,000)

Pluribus achieves superhuman performance in multiplayer poker, 2019

Brief History of AI

- 1943-1955: Initial Work in AI
  - McCulloch and Pitts Boolean circuit of the brain
- 1950's: Early AI Programs
  - Samuel's Checker Program, Newell and Simon's Logic Theorist, Gerlenter's Geometry Engine
- 1956: Dartmouth Meeting: "Artificial Intelligence"

Brief History of AI

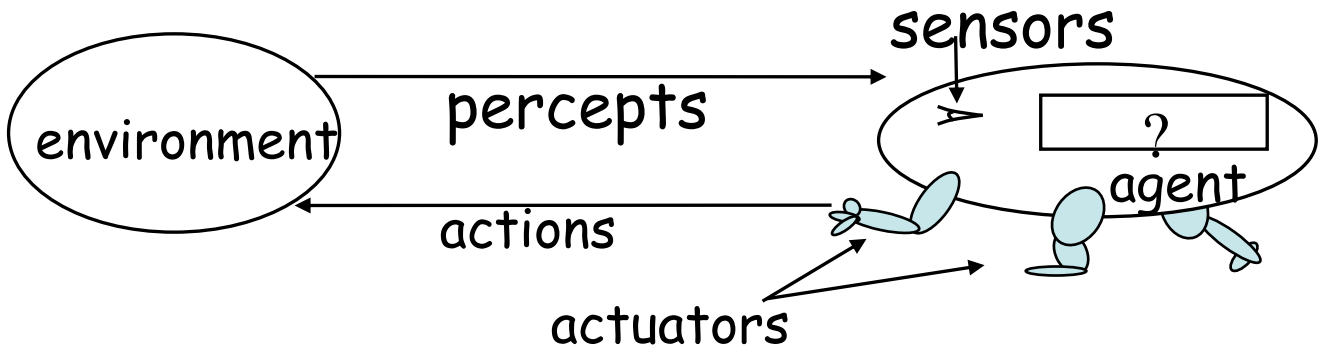
- 1950s-1969: Enthusiasm and Expectations
  - Many successes
  - LISP, time-sharing, resolution, neural networks, planning, learning theory ,vision, Shakey, machine translation
- 1966-1973: Reality Hits
  - Early programs had little knowledge of their subject matter
  - Computational complexity
  - Perceptrons paper (negative result about neural networks)

Brief History of AI

- 1969-1979: Knowledge-based systems
- 1980-1988: Expert system industry booms
- 1988-1993: Expert system busts, AI Winter
- 1986-2000: Return of neural networks
- 2000-present: Increase in technical depth
  - Probability theory, statistics, optimization, utility theory, game theory, learning theory
- 2010-present: Big data, deep neural networks

Back to Rationality: Rational Agent Paradigm

- A rational agent is an entity that perceives and acts:
  - Function from percepts to actions



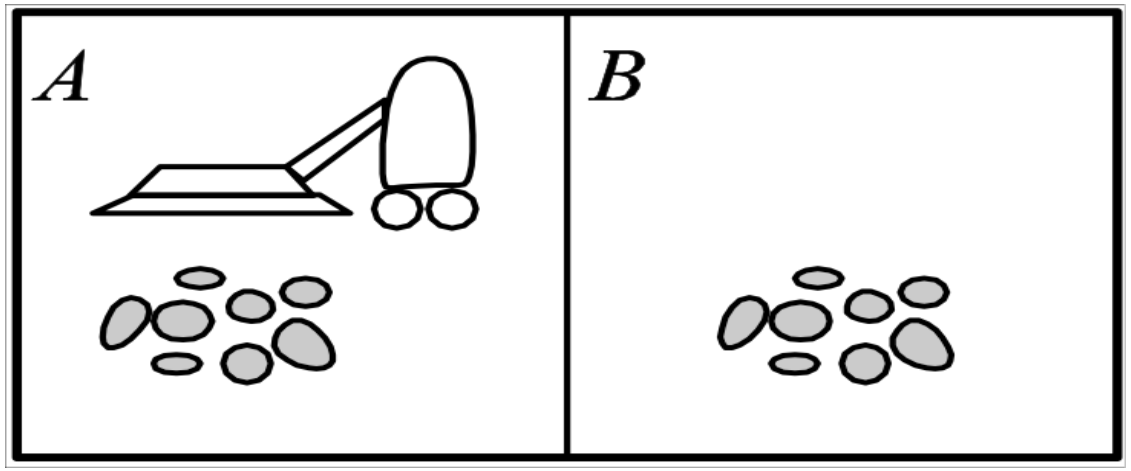
- Performance measures
  - Goal achievement, resource consumption,…
- Caveat: Computational limitations and environmental constraints means we do not have perfect rationality

Task Environments

To design a rational agent the task environment must be defined

- Performance measures
- Environment
- Actuators
- Sensors

Performance Measures Can Be Tricky



Percepts: [Location, Dirty or Clean]

Actions: Right, Left, Vacuum, NoOp, Dump

Function: ([A,Clean],Right), ([A, Dirty], Vacuum), ([B, Dirty], Vacuum), ([B, Clean], Left)...

Properties of the Task Environment

Properties of the task environment determine what AI approach is most suitable

- Fully Observable vs **Partially Observable**:
  - Does the agent's sensors give it access to the complete state of the environment?
- Deterministic vs **Stochastic**
  - Is the next state completely determined by the current state and action executed?
- Episodic vs **Dynamic**
  - Does the current decision or action influence future decisions or actions?
- Discrete vs **Continuous**
  - How are states, time, actions modelled?
- Static vs **Dynamic**
  - Is the environment changing as the agent is planning what to do next?
- Single Agent vs **Multiagent**

Properties of the Task Environment

Solitaire	**Computer Go**	**Recommender Systems**	**Autonomous Driving**
Fully Observable	Fully Observable	Partially Observable	Partially Observable
Deterministic	Deterministic	Stochastic	Stochastic
Sequential	Sequential	Episodic	Sequential
Static	Static	Dynamic	Dynamic
Discrete	Discrete	Discrete	Continuous
Single agent	Multi agent	Multi agent	Multi agent