

Syllabus

Course Homepage: [HLT Link](#)

Homework Submission: on eLearning

Piazza: [Piazza Link](#)

Optional Textbook: *Artificial Intelligence: A Modern Approach* Russell and Norvig, 3rd edition.

Miscellaneous Notes:

1. Honorlock will be used for exams.
2. Don't email for class-related questions, use Piazza private posts.
3. Textbook is not required, only what is presented in lecture will be tested.

Prerequisites:

1. Python for projects, maybe C/C++ and Java
2. Big-O notation
3. Propositional and first-order logic (will have review lecture)
4. Elementary knowledge of probability theory (will have review lecture)

Collaboration Policy:

1. Assignments can be done in group of two
2. Term projects can be done in group of two
3. Only need to turn in one solution/program per group

Late Policy:

1. 6 free late days (≤ 2 late days can be used per homework)
2. One day late is 10% penalty, two days late is 30%

Day 1

The main goal of the course is **problem solving and decision making**. There are two main types of problems:

- Problems do not require *sophisticated knowledge*

One example is the **Path Finding** problem. One way to solve the problem is to search through all possible paths and find the minimum. However, that is often not feasible because there are too many possible paths.

Another example is the **Map Coloring** problem. One way to solve the problem is to perform a complete (or exhaustive) search, which would take a very long time.

The focus for these search problems is to find efficient solutions to solving these difficult problems. Note that hardware has improved in the last 20 or 30 years, we do not care about those improvements because algorithmic improvements almost always outperform hardware improvements.

- Problems that do require *sophisticated knowledge*

One example is asking a computer “what is the Capital of California.” Although this is easy for a human who had learned this knowledge before to answer, it is hard for a computer to know even if they are given the Wikipedia article for the state of California. Humans are able to process contextual information based on prior knowledge as well as comprehension level; since computers are not able to do that, humans must turn natural language into some kind of **knowledge representation** so that the resulting information can be easily understood by a machine.

The second challenge is to use **knowledge reasoning** to derive new knowledge from existing knowledge. However, deriving information from some statements that are only partially true requires **Probabilistic Reasoning** to understand the final probability of a different statement.

AI Goal

Intelligence

“the capacity to learn and solve problems”

Views of AI falls into fall different perspectives:

1. Thinking: being able to think
2. Acting: being able to do something that can be observed
3. Human: emotion might be involved in decision-making
4. Rational: decisions are always rational and ideal

1. Acting Humanly (Turing Test)

Turing proposed the question of “can machines act intelligently?” The Test consists of a test subject who is chatting with either a real human or a machine. After a certain period of time, if the test subject can not tell if they’re talking to the machine or the human, then the machine would have passed by the Turing Test.

Key components from the Turing Test:

1. Natural Language Processing: processing the user's input and respond intelligently
2. Knowledge Representation: store and manipulate information from knowledge base
3. Automated reasoning: used store info to draw conclusion to answer questions
4. Machine Learning: to adapt and extrapolate from existing data

One example of a “human impersonator” is *ELIZA*, which attempts to impersonate a psychotherapist. It was built in the 1960s, and used basic strategies to decide on what response is the most fitting. *ALICE*, which was built 30 years later, is a lot smarter and can intelligently answer questions with additional information from a knowledge base.

2. Thinking humanly: modeling cognitive processes

Cognitive Science studies the internal activities of the brain, and *Cognitive Neuroscience* attempts to mirror the human brain.

3. Thinking rationally: formalizing the “laws of thought”

Using logic to make the right inferences. The classic example is if Socrates is a man, and all men are mortal, then the logical conclusion is that “Socrates is mortal”.

However, since not all decisions can be made by direct logical reasoning, logic has limited usage in AI.

4. Acting rationally: rational agent

Rational behavior means doing the right thing, or the best thing to **maximize goal achievement given the available information**, which might mean using a math textbook to score well on a math test.

Building Intelligent Machines

Two approaches

1. One approach is building an exact model of human cognition (main view from psychology and cognitive sciences)
2. Developing methods to match/exceed human performance, not necessary through emulating human behavior (main focus of this course as well as recent AI progress)

AI has many domains, and it also leverages the knowledge from a lot of different disciplines (philosophy, CS theory, mathematics/physics).

Historical Perspective

The goal is to obtain an understanding of the human mind. Major people include:

1. Formalizing the laws of human thought: George Boole, Gottlob Frege, and Alfred Tarski
2. Thinking as computation: Alan Turing, John von Neumann, and Claude Shannon

Direct founders;

1. John McCarthy, Marvin Minsky, Herbert, Simon, and Allen Newell

History of AI:

- 1943: 40-neuron network, model of artificial neurons
- 1950: Turing's "Computing machinery and intelligence"
- 1960s: Dartmouth meeting with founders
- 1950s: Early enthusiasm with great expectations for its future
- 1958: John McCarthy's LISP
- 1965: The **resolution principle**, which was a basis for automated theorem proving
 - At this time, the world is simplified and limited to very small number of states
- 1960s: "dose of reality", *NP-Completeness* (intractability of problems)
- 1970s: Expert systems that require a knowledge base (correlate symptoms with diseases)
- 1980s: Neural nets, LISP
- 2000-: focus on integration of reasoning

Achievements:

1. Microsoft '97, Answer Wizard
2. Deep Blue beating Russian chess grandmaster
3. 1999: *Proverb* Solving crossword puzzles
4. Robocup, Autonomous Vehicle

The recent progress is from the advancement of **machine perception**, which means that machines can see and hear from the outside world. Techniques that were developed by *assuming* outside input are now able to be applied into technology like autonomous driving. There is also advancement in both processing speed as well as memory space, which is getting close to humans' capabilities.

Other progress comes from crowd-sourced human data for training data, engineering teams at big companies, as well as research activity at many institutions. Recent achievements include Watson winning *Jeopardy!*, Google's AlphaGo, as well as Watson automating insurance claim workers' jobs.