EECS3401

Introduction to Artificial Intelligence & Logic Programming

Some of the material coverage in this course is based on previous EECS3401 offerings and include contributions of Yves Lesperance, Gunnar Gotshalks, Vitaliy Batusov, Parke Godfrey, Kamen Kanev, and others.

Introduction

Welcome to the *Introduction to Artificial Intelligence & Logic Programming*!

And the winner is

…*Théâtre D’opéra Spatial*

Jason M. Allen *via Midjourney* of Pueblo West, Colorado won the blue ribbon (*first place*) in this year's *contest for emerging digital artists* at the Colorado State Fair's *annual art competition*.

He won with his entry, “Théâtre D’opéra Spatial.”

“I’m not going to apologize for it. I won, and I didn’t break any rules.”

- Jason M. Allen, as quoted by the *New York Times* in the article, “An A.I.-Generated Picture Won an Art Prize. Artists Aren’t Happy,” by Kevin Roose, 2 September 2022.

“Théâtre D’opéra Spatial.”

What is controversial?

Mr. Allen did not “create” the picture.
Well, not directly, at least.

He generated it via an AI program (“agent”) made by *Midjourney. (See “Midjourney” at Wikipedia.)*

Questions

- What are the right of the digital artists that use *Midjourney*?
- *How* does *Midjourney* work?!
- What are the right of artists whose work, sometimes under copyright, was used to *train Midjourney*?

Johannes Vermeer, 1665

"DALL-E can now help you imagine what's outside the frame of famous paintings: OpenAI introduces native 'outpainting' for its AI image generator."

— *James Vincent*,
The Verge, 5 September 2022

Johannes Vermeer & DALL-E, 2022

“A protein’s function is determined by its 3D shape.”

Protein Folding

Proteins

- Proteins are molecular sequences of amino acids.
- A protein will "ball up" into a 3D shape, which is determined by its sequence.
- A protein's *function* is determined by
  - its sequence *and*
  - its shape.

Challenges

- Determining a given protein's sequence is nowadays quite doable.
- *Determining* a given protein's shape is doable, but *really* hard.
- *Predicting* a given protein's shape from its sequence was thought to be *next to impossible*.

The problem of predicting how a sequence of amino acids — a protein — will fold has been a major challenge in biology and medicine.

Deepmind's AlphaFold

"'It will change everything': DeepMind's AI makes gigantic leap in solving protein structures"

— *Ewen Callaway*, **Nature**, 30 November 2020

"Google's deep-learning program for determining the 3D shapes of proteins stands to transform biology, say scientists."

The LaMDA Chatbot Is a Sentient AI?

Former Google employee Blake Lemoine claims so!

LaMDA: *Language Model for Dialogue Applications*

Q. Is he right?!

A vast majority of AI researchers say no.

"Google Fires Engineer Who Claimed LaMDA Chatbot Is a Sentient AI"

— *Michael Kan*, **PC Magazine**, 25 July 2022

Sentient AI refers to **AI technologies that can process emotions and perceive the world like human beings**.

Currently, AI is not sentient; it doesn't understand or perceive the world in any way.

AI Successes

- In 2016, Google DeepMind *AlphaGo* beat decade's top player.
- In 2011, IBM's *Watson* beat the top Jeopardy winners.
- In 2005, Stanford team won DARPA Grand Challenge: 132 mile / 212 kilometre race in desert.
- In 1999, NASA's *Remote Agent* used AI planning to control a spacecraft.
- In 1997, IBM's *Deep Blue* beat chess world champion.

Artificial *General* Intelligence (AGI)

One Endeavor / Definition:

AI is the general pursuit to create a machine that is *intelligent* in a general way in the way that humans are.

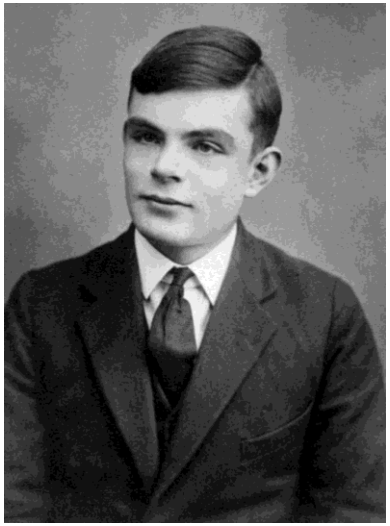
"First solve AI, then use AI to solve everything else."

— *Demis Hassabis*, CEO of Google DeepMind

Q. How would we know if we were successful?

The Turing Test

Alan Turing, perhaps the father of AI, proposed a *test* in 1950 — which we now call *the Turing Test* — to answer whether a machine can *think*.



"A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses come from a person or from a computer."

— **RN** p.2

Do we even want to pass the Turing Test?

From a philosophical point of view, The *Turing Test* is problematic in a number of ways.

Q. Should imitating humans be the goal?

- Should that be something we want to accomplish?
- Is mimicking *how* humans *think* and *act* a way forward to making *thinking*, and *autonomously acting*, machines?

Human Intelligence

Plus, we don't really know *how* human intelligence works!

We know little about how the human brain performs its higher level processes.

Hence, this point of view provides little information from which a scientific understanding of these processes can be built.

Neuroscience has been quite influential in some areas of AI, however. For example, in robotic sensing and computer vision.

How to pass the Turing Test?

The *Turing Test* does provide a guide, however, for what would be involved to be able to pass it.

- **natural language processing** to *communicate* successfully in a human language
- **knowledge representation** to *store* what it knows or hears
- **automated reasoning** to *answer* questions and to *draw new conclusions*
- **machine learning** to *adapt* to new circumstances and to *detect* and *extrapolate patterns*
- **computer vision** and **speech recognition** to *perceive* the world
- **robotics** to *manipulate* objects and move about.

Specific AI

AI is meant to “solve [the] kinds of problems now reserved for humans.”

— from the prospectus for the *1956 Dartmouth Summer Workshop on AI*

- This centres the endeavour of AI on problem solving.
- It is usually quite domain specific.
- The focus is on solving types of problems computationally that we could not solve before.

This is a bit of a *receding* definition: once we know how to do it, it no longer is AI!

Computational Intelligence

- AI tries to understand and model intelligence as a *computational process*.
- Thus, we try to construct systems for which computation achieves, or approximates, a desired notion of *rationality*.
- Hence, AI is part of *computer science*.

There are other areas interested in the study of intelligence; e.g., *cognitive science*, which focuses on human intelligence. Such areas are quite related, but their central focus tends to be different.

What would be a desired notion of *rationality*?

Rationality

- The alternative to mimicking human intelligence relies on the notion of rationality.
- Typically, this is a precise *mathematical notion* of what it means to “do the right thing” in any particular circumstance.

This provides:

- A precise mechanism for analyzing and understanding the properties of this ideal behavior we are trying to achieve.
- A precise benchmark against which we can measure the behavior of the systems we build.

Characterizations of Rationality

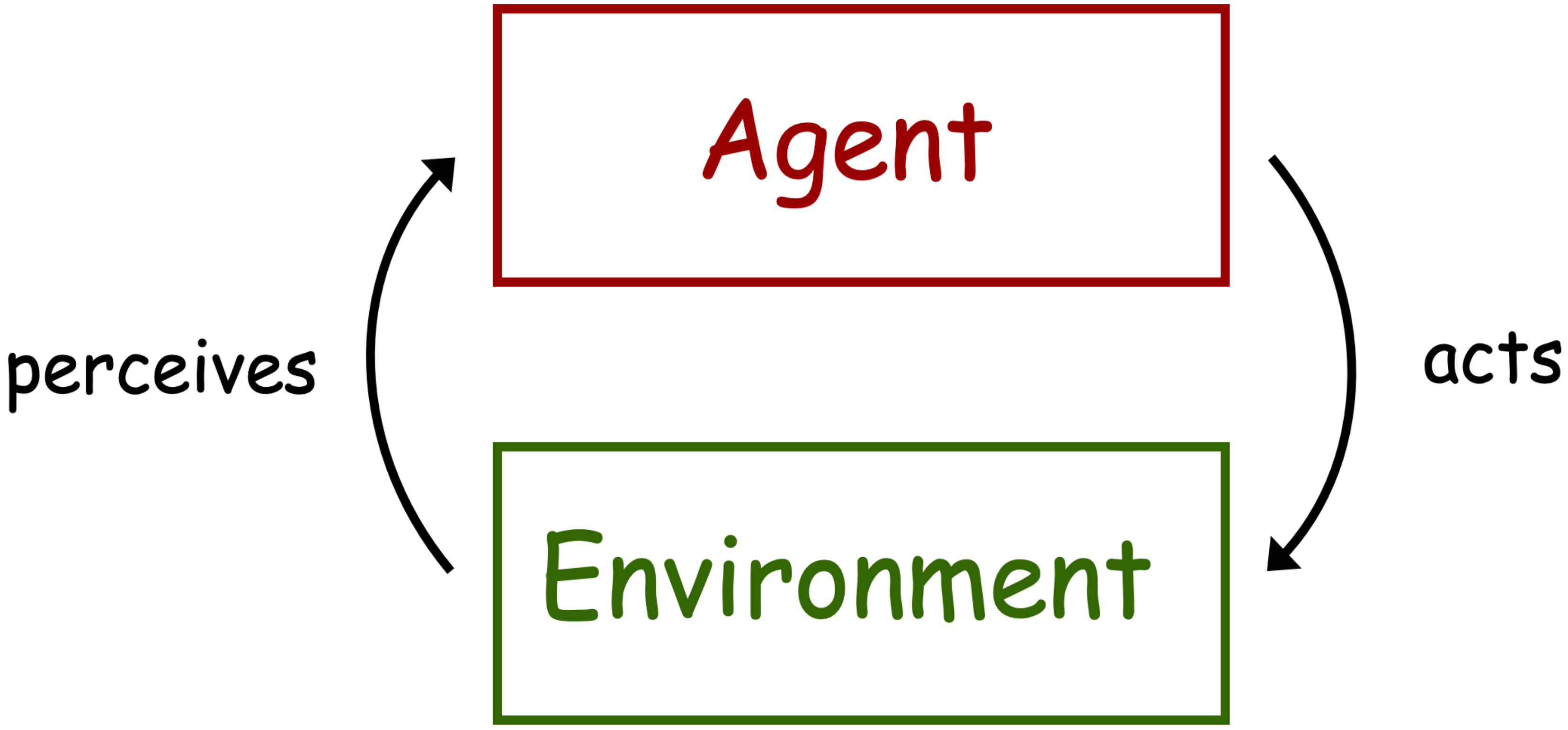
- Mathematical characterizations of rationality have come from diverse areas,
  - like *logic*
    - “laws of thought”
  - and *economics*
    - *utility theory*,
    - how best to act under *uncertainty*,
    - *game theory*, and
    - how self-interested, agents interact.
- There is no universal agreement about which notion of rationality is best. But since these notions are precise, we can study them, and give exact characterizations of their properties, good and bad.

We shall focus on acting rationally. This has implications for “thinking” / “reasoning”.

Agency

- It is also useful to think of intelligent systems as being *agents*, either:
  - with their own goals, or
  - or that act on behalf of someone (a “user”).
- An agent is an entity that exists in an *environment* and that *acts* on that environment based on its *perceptions* of the environment.
- An *intelligent agent* acts to further its own “interests” (or those of a user).
- An *autonomous agent* can make decisions without the user’s intervention, possibly based on what it has learned.

Agent Schematic (I)

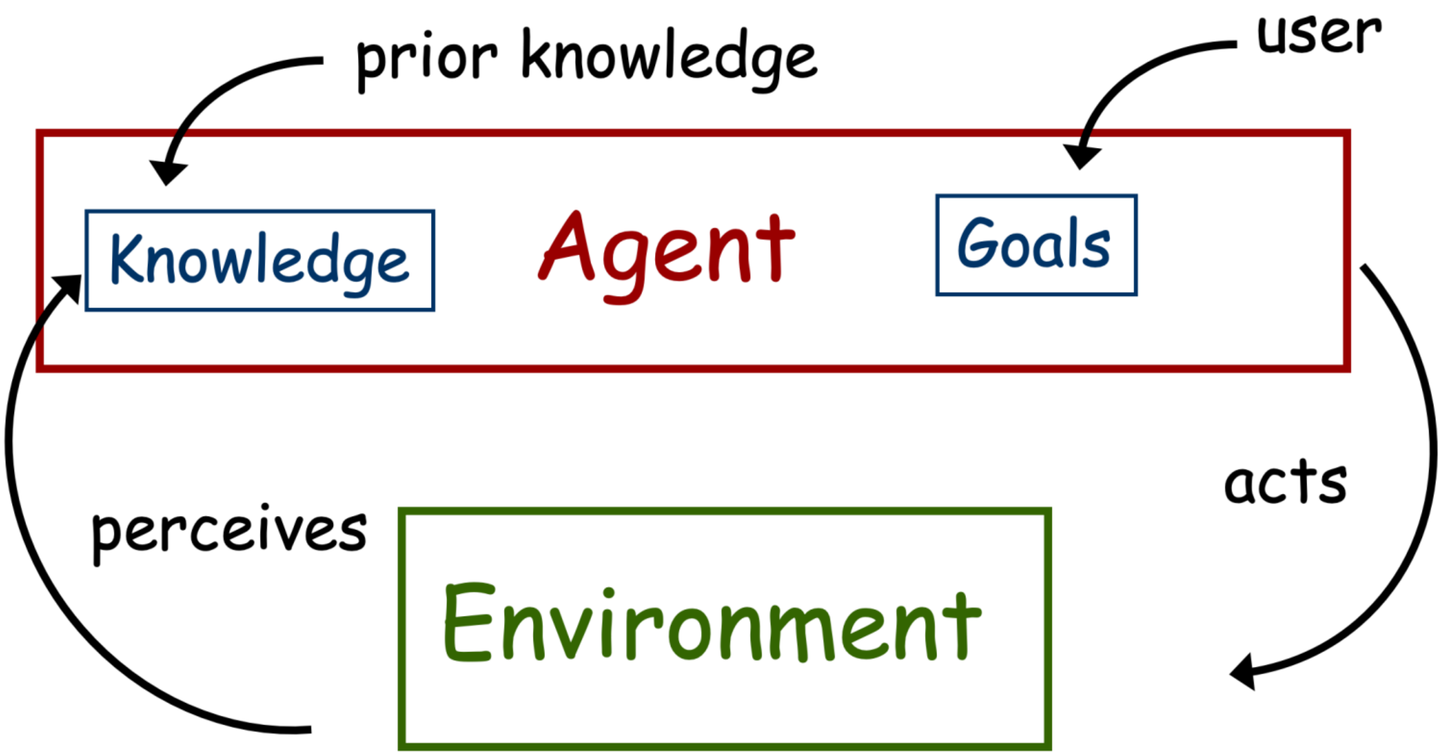


This diagram ignores the internal structure of the agent. It is a *black box*.

Types of Agents (I)

- **Simple reflex agents**:
  apply simple condition-action rules to decide the next action based on current percepts.
- **Model-based reflex agents**:
  maintain a model of the world, apply rules to decide next action based on the current state of the world model.
- **Goal-based agents**:
  decide next action based on current state of the world model and the current goal(s); may do planning. This is more flexible!

Agent Schematic (II)



Supports more flexible interaction with the environment, the ability to modify one's goals, and knowledge that is applied flexibly to different situations.

Types of Agents (II)

- **Utility-based agents**:
  choose actions to maximize their expected utility in *uncertain* worlds.

Learning

For all types of agents, one could add a *learning mechanism*: *explore* the space of possible rules/actions/models, evaluate performance, and *modify* agent to improve.

Types of Environments

- fully observable vs partially observable
- deterministic vs stochastic
- episodic vs sequential
- static vs dynamic
- discrete vs continuous
- single-agent vs multi-agent
- known dynamics vs unknown dynamics

See **RN** §2.3 for details.

Logic Programming

E.g., Prolog

- *Logic programming* is a programming language based on *first-order logic* and the *predicate calculus*.
- This paradigm offers a formal approach to *knowledge representation* and *reasoning*.
- A logic program acts much like a *simple* agent:
  - One asks the program a *question*.
  - The program uses logical inference over its “knowledge” to infer *answer(s)* to the question.

We spend a good deal of time in this course on the logic-programming paradigm and on **Prolog**. So, we shall be returning to this!

EECS-3401

It is an *introduction* to AI and Logic Programming.
And it is about *symbolic* AI!

Our focus will be on

- knowledge representation
- reasoning (*logical* & *probabilistic*)
- decision making (*search*, *planning*, & *decision theory*)

Logistics

More about the course:

- textbooks
- tentative schedule
- requirements
    - prerequisites
    - deliverables
    - evaluation
- and so on…

Check the course syllabus.

not database systems

Database Systems, specifically *relational*, share a common origin with symbolic AI.

Graph database systems underly present development of *knowledge graphs*.

- EECS-3421
- EECS-4411

not natural language processing

aka *computational linguistics*

Natural language processing (NLP) seeks to have computers “understand” and “speak” human (*natural*) languages (e.g., English).

Is this now solved with transformer-based *large language models* (LLMs)?!

- To crack the problem of *grammar*, the answer seems to be *yes*!
- For true *actionable understanding*, the answer is much murkier.

not pattern recognition (per se)

The field of *pattern recognition* explores / explored *stochastic* methods to recognize symbols and objects from, say, images.

This has been largely subsumed under *computer vision*. And the techniques replaced by neural-network methods.

This is a “bottom up” approach rather than a “top down” one. We *train* a *model* over data (examples).

not knowledge discovery in databases (KDD)

aka *data mining*

Data mining extracts patterns / knowledge from data.

This was considered initially a branch of AI.

Until we knew good techniques … then it was not “AI” anymore.

not machine learning

Machine learning is about techniques to *learn* / train from data to predict classes of new instances.

Deep neural networks dominate approaches presently.

We call these the “dark arts”.

- EECS-3404
- EECS-3405

Why *this* course?

Is it still worth studying *symbolic* AI? Of course!

- It is a systematic way of thinking about general “reasoning” tasks.
- Many AI researchers are starting to conclude that the *dark arts* — er, machine learning — alone does not suffice.

  That somehow, perhaps, we have to couple *symbolic* and *stochastic* approaches, to couple *top-down* and *bottom-up*.

For this lecture:

- Russell & Norvig (“RN”): Chapters 1 & 2.

For next time:

- RN: Chapter 8; review Chapter 7
- Review your background in first-order logic (e.g., **EECS-1090**).