COMP111: Artificial Intelligence

Section 2. Views of AI and History of AI

Frank Wolter

Content

- Views of AI

Content

- Views of AI
- Brief history of AI

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.
- Theorem Proving: four colour theorem.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.
- Theorem Proving: four colour theorem.
- Speech recognition: smartphones.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.
- Theorem Proving: four colour theorem.
- Speech recognition: smartphones.
- Game playing: chess, GO, poker.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.
- Theorem Proving: four colour theorem.
- Speech recognition: smartphones.
- Game playing: chess, GO, poker.
- Semantic web search: structured search.

Typical AI Application Areas

- Natural language processing: understanding, generating, translating.
- Vision: object recognition, face recognition, scene recognition.
- Robotics: car driving, autonomous aerial vehicles.
- Theorem Proving: four colour theorem.
- Speech recognition: smartphones.
- Game playing: chess, GO, poker.
- Semantic web search: structured search.
- Diagnosis in healthcare

What is AI?

- The goal of work in AI is to build machines that perform tasks normally requiring human intelligence. (Nilsson 1971)

What is AI?

- The goal of work in AI is to build machines that perform tasks normally requiring human intelligence. (Nilsson 1971)
- AI is everything we can't do with today's computers. (Possibly McCarthy, 1980s)

What is AI?

- The goal of work in AI is to build machines that perform tasks normally requiring human intelligence. (Nilsson 1971)
- AI is everything we can't do with today's computers. (Possibly McCarthy, 1980s)
- My definition of AI is any algorithm that is new in computer science. Once the algorithm becomes accepted then it's not AI, it's just a boring algorithm.
  At one time windows, mouse, menus, scroll bars etc. were considered an AI technique for making computers understand natural language. (Possibly Keane, 1990s)

What is AI?

- The goal of work in AI is to build machines that perform tasks normally requiring human intelligence. (Nilsson 1971)
- AI is everything we can't do with today's computers. (Possibly McCarthy, 1980s)
- My definition of AI is any algorithm that is new in computer science. Once the algorithm becomes accepted then it's not AI, it's just a boring algorithm.
  At one time windows, mouse, menus, scroll bars etc. were considered an AI technique for making computers understand natural language. (Possibly Keane, 1990s)
- A branch of computer science dealing with the simulation of intelligent behavior in computers. The capability of a machine to imitate intelligent human behavior. (Merriam Webster Dictionary)

AI is both science and engineering

- the science of understanding intelligent entities: developing theories/models which attempt to explain and predict the nature of such entities.

AI is both science and engineering

- the science of understanding intelligent entities: developing theories/models which attempt to explain and predict the nature of such entities.
- the engineering of intelligent entities.

Basic Terminology

- AGI (artificial general intelligence, or strong AI): artificial system that can accomplish any intellectual task that humans can perform. Often sentience or consciousness are also required in definitions of AGI.

Basic Terminology

- AGI (artificial general intelligence, or strong AI): artificial system that can accomplish any intellectual task that humans can perform. Often sentience or consciousness are also required in definitions of AGI.
- Narrow (or weak) AI: artificial system that is able to solve a small number of specific intellectual problems. Does not experience consciousness.

Basic Terminology

- AGI (artificial general intelligence, or strong AI): artificial system that can accomplish any intellectual task that humans can perform. Often sentience or consciousness are also required in definitions of AGI.
- Narrow (or weak) AI: artificial system that is able to solve a small number of specific intellectual problems. Does not experience consciousness.

Before ChatGPT, other large language models (LLMs), and more generally generative AI, came up in 2022/23, almost all AI researchers and practitioners would have described their goal as narrow AI. That has changed a bit and some argue that generative AI could be a significant step towards AGI.

Four Views of AI

- Systems that think like humans

Four Views of AI

- Systems that think like humans
- Systems that act like humans

Four Views of AI

- Systems that think like humans
- Systems that act like humans
- Systems that think rationally (think in a correct way)

Four Views of AI

- Systems that think like humans
- Systems that act like humans
- Systems that think rationally (think in a correct way)
- Systems that act rationally (act optimal to achieve a goal)

Thinking Humanly vs Acting Humanly

- Cognitive science: to understand and build intelligent machines we should simulate human thinking.
- Computer science (often): to build intelligent machines simulate the intelligent behavior of humans. It is not necessary nor desirable that intelligent machine process information in the same way as humans (think humanly) but the aim is to create machines that are able to behave (e.g., take part in conversations) in the same way as humans.

The emphasis on simulating behavior rather than internal processes has consequences for the question: *when can we count a machine as being intelligent*?

Thinking Humanly vs Acting Humanly

- Cognitive science: to understand and build intelligent machines we should simulate human thinking.
- Computer science (often): to build intelligent machines simulate the intelligent behavior of humans. It is not necessary nor desirable that intelligent machine process information in the same way as humans (think humanly) but the aim is to create machines that are able to behave (e.g., take part in conversations) in the same way as humans.

The emphasis on simulating behavior rather than internal processes has consequences for the question: *when can we count a machine as being intelligent*?

"Can machines think?" $\longrightarrow$ "Can machines behave intelligently?"

Turing test (1950)

- A human interrogator asks questions to a system, without knowing whether it is human or a machine.

Turing test (1950)

- A human interrogator asks questions to a system, without knowing whether it is human or a machine.
- Systems passes as AI system if the questioner cannot tell the difference between the answers given by a human and the answers given by a machine.

Turing test (1950)

- A human interrogator asks questions to a system, without knowing whether it is human or a machine.
- Systems passes as AI system if the questioner cannot tell the difference between the answers given by a human and the answers given by a machine.
- Of great theoretical importance for discussion about AI, but not meaningful as a practical test.

Thinking and Acting Rationally

- Investigate how we actually think and act is one route to AI. But how about how we should think and act.

Thinking and Acting Rationally

- Investigate how we actually think and act is one route to AI. But how about how we should think and act.
- Use techniques from logic and probability theory to investigate and create machines that can reason correctly.

Thinking and Acting Rationally

- Investigate how we actually think and act is one route to AI. But how about how we should think and act.
- Use techniques from logic and probability theory to investigate and create machines that can reason correctly.
- Result is idealised reasoning

Thinking and Acting Rationally

- Investigate how we actually think and act is one route to AI. But how about how we should think and act.
- Use techniques from logic and probability theory to investigate and create machines that can reason correctly.
- Result is idealised reasoning
- Acting rationally = acting to achieve one's goals, given one's beliefs.

Thinking and Acting Rationally

- Investigate how we actually think and act is one route to AI. But how about how we should think and act.
- Use techniques from logic and probability theory to investigate and create machines that can reason correctly.
- Result is idealised reasoning
- Acting rationally = acting to achieve one's goals, given one's beliefs.
- Also use techniques from economics/game theory to investigate and create machines that act rationally.

A brief history of AI

First steps towards AI: 1943–55

- Based on first insights into the functioning of neurons in the brain and Turing's theory of computation McCulloch & Pitts (1943) suggest a model of artificial neural networks.

First steps towards AI: 1943–55

- Based on first insights into the functioning of neurons in the brain and Turing's theory of computation McCulloch & Pitts (1943) suggest a model of artificial neural networks.
- Minsky and Edmonds (1951) build the first neural network computer (SNARC) using vacuum tubes to simulate a network of 40 neurons.

First steps towards AI: 1943–55

- Based on first insights into the functioning of neurons in the brain and Turing's theory of computation McCulloch & Pitts (1943) suggest a model of artificial neural networks.
- Minsky and Edmonds (1951) build the first neural network computer (SNARC) using vacuum tubes to simulate a network of 40 neurons.
- Shannon, Turing (1950): the first chess playing programs.

Dartmouth Conference: 1956

- A two months conference. 10 attendees. Main figures: McCarthy, Minsky, Shannon, Samuel, Newell, Simon.
- Newell & Simon presented LOGIC THEORIST program. Simon: "We have invented a computer program capable of thinking non-numerically, and thereby solved the venerable mind-body problem."
- Corresponding article was rejected by Journal of Symbolic Logic.
- Name 'Artificial Intelligence' coined by John McCarthy was adopted.
- No technical breakthroughs.

Optimism regarding AGI (1956-70)

Promising performance of early AI systems on simple examples (integration problems, algebra problems, moving blocks) let to great expectations (Simon (1957)):

> *It is not my aim to surprise or shock you - but the simplest way I can summarize is to say that there are now in the world machines that can think, that can learn and that can create. Moreover, their ability to do these things is going to increase rapidly until - in a visible future - the range of problems they can handle will be coextensive with the range to which the human mind has been applied.*

It does not work... (1970s)

The systems from the 1960s almost always failed completely when tried out on a wider selection of or more difficult problems.
Examples:

It does not work... (1970s)

The systems from the 1960s almost always failed completely when tried out on a wider selection of or more difficult problems. Examples:

- Translation of Russian scientific papers into English (space race between USSR and US) by simple syntactic transformations failed completely because of lack of stored general knowledge to resolve ambiguities (and for other reasons).

It does not work... (1970s)

The systems from the 1960s almost always failed completely when tried out on a wider selection of or more difficult problems. Examples:

- Translation of Russian scientific papers into English (space race between USSR and US) by simple syntactic transformations failed completely because of lack of stored general knowledge to resolve ambiguities (and for other reasons).
- Game playing programs failed because of inherent intractability (combinatorial explosion). Techniques developed for toy examples did not scale.

Combinatorial Explosion (Chess)

In principle, it is easy to write a program playing chess by computing all possible sequences of moves (say, up to 40 moves) and then select an optimal move. In chess, however, there are approximately

- 400 different positions after one move apiece;
- 72,084 different positions after two moves apiece;
- 9×10^6 positions after three moves apiece;
- 288×10^9 different possible positions after four moves apiece;
- and so on.

Combinatorial Explosion (Chess)

In principle, it is easy to write a program playing chess by computing all possible sequences of moves (say, up to 40 moves) and then select an optimal move. In chess, however, there are approximately

- 400 different positions after one move apiece;
- 72,084 different positions after two moves apiece;
- 9×10^6 positions after three moves apiece;
- 288×10^9 different possible positions after four moves apiece;
- and so on.

In practice, it simply is not possible (and never will be possible) to compute all relevant sequences of moves!

Combinatorial Explosion (Rubik's Cube)

For a Rubik's Cube there are

- 43,252,003,274,489,856,000 combinations
- Up to 481,229,803,398,374,426,442,198,455,156,736 brute-force solution attempts

(For comparison: The 2019 Summit machine at the Oak Ridge National Lab in Tennessee executes roughly 10^{18} instructions per second.)

Combinatorial Explosion (Rubik's Cube)

For a Rubik's Cube there are

- 43,252,003,274,489,856,000 combinations
- Up to 481,229,803,398,374,426,442,198,455,156,736 brute-force solution attempts

(For comparison: The 2019 Summit machine at the Oak Ridge National Lab in Tennessee executes roughly 10^{18} instructions per second.)

Insight: The fact that a program can find a solution in principle does not mean that the program will be able to find it in practice.

Combinatorial Explosion (Rubik's Cube)

For a Rubik's Cube there are

- 43,252,003,274,489,856,000 combinations
- Up to 481,229,803,398,374,426,442,198,455,156,736 brute-force solution attempts

(For comparison: The 2019 Summit machine at the Oak Ridge National Lab in Tennessee executes roughly 10^{18} instructions per second.)

Insight: The fact that a program can find a solution in principle does not mean that the program will be able to find it in practice.

The Lighthill Report (1973) on AI for the British Research Council came to the conclusion that AI researchers had failed to address the combinatorial explosion problem and was pessimistic regarding research in, for example, general robotics and natural language processing. It formed the basis for the decision of the UK government to essentially end support for AI research.

Expert Systems (Narrow AI)

- General purpose, brute force techniques don't work, so use knowledge rich solutions.
- Early 1980s saw emergence of expert systems as systems capable of exploiting knowledge about tightly focused domains to solve problems normally considered the domain of experts.

Expert Systems (Narrow AI)

- Expert systems (combining basic logic and probability based reasoning):
  - MYCIN: diagnosis infections, recommend antibiotics, diagnosis of blood clotting diseases.
  - PROSPECTOR: aid geologists in mineral exploration, finding sites for drilling
- 1990s: Most companies set up to commercialise expert systems technology did not survive. Problems:
  - the knowledge elicitation bottleneck;
  - Lack of trust in recommendations given by expert systems.

Major steps in the past 20 years

1997: IBM Deep Blue defeated world chess champion Garry Kasparow.

Reasons for success: much faster machines with more storage. Heuristic search algorithms.

IBM Watson wins Jeopardy (2011)

- Jeopardy! is an American television quiz show: contestants are presented with general knowledge questions they must answer:

IBM Watson wins Jeopardy (2011)

- Jeopardy! is an American television quiz show: contestants are presented with general knowledge questions they must answer:
- Example:
  - Question: In the 2004 opening ceremonies of the Olympic Games a sole member of this team opened the Parade of Nations; the rest of his team closed it.
  - Answer: Greece.

IBM Watson wins Jeopardy (2011)

- Jeopardy! is an American television quiz show: contestants are presented with general knowledge questions they must answer:
- Example:
  - Question: In the 2004 opening ceremonies of the Olympic Games a sole member of this team opened the Parade of Nations; the rest of his team closed it.
  - Answer: Greece.
- Watson is a question answering computer system capable of answering questions posed in natural language.

IBM Watson wins Jeopardy (2011)

- Jeopardy! is an American television quiz show: contestants are presented with general knowledge questions they must answer:
- Example:
  - Question: In the 2004 opening ceremonies of the Olympic Games a sole member of this team opened the Parade of Nations; the rest of his team closed it.
  - Answer: Greece.
- Watson is a question answering computer system capable of answering questions posed in natural language.
- In 2011, Watson competed on Jeopardy! against former winners Brad Rutter and Ken Jennings. Watson received the first place prize.

IBM Watson wins Jeopardy (2011)

- Jeopardy! is an American television quiz show: contestants are presented with general knowledge questions they must answer:
- Example:
  - Question: In the 2004 opening ceremonies of the Olympic Games a sole member of this team opened the Parade of Nations; the rest of his team closed it.
  - Answer: Greece.
- Watson is a question answering computer system capable of answering questions posed in natural language.
- In 2011, Watson competed on Jeopardy! against former winners Brad Rutter and Ken Jennings. Watson received the first place prize.
- Reasons for success: progress in natural language processing and extracting knowledge from structured and unstructured content (required four terabytes of disk storage).

AlphaGo wins against 9-dan professional (2016)

- Go is Strategy board game for two players. The players take turns placing playing pieces in vacant intersections of a board with a 19×19 grid of lines.
- The objective of Go is to fully surround a larger total area of the board than the opponent.
- AlphaGo computer program by Google DeepMind in London.
- 2016, AlphaGo beats Lee Sedol in a five-game match.
- Reasons for success: heuristic search not enough. Progress in machine learning.

Libratus (CMU) wins in Poker (2017)

- 2017: Libratus beats four of the world's best poker players in a 20-day tournament.

Libratus (CMU) wins in Poker (2017)

- 2017: Libratus beats four of the world's best poker players in a 20-day tournament.
- Poker, in contrast to chess or GO, is a game with imperfect information. With chess and Go, each player can see the entire board, but with poker, players do not get to see each other's hands.
- the AI is required to bluff and correctly interpret misleading information in order to win.
- Reasons for success: progress in machine learning and applying probabilistic reasoning.

AlphaFold: Predict 3D models of protein structures (2020)

- Proteins underpin the biological processes of all living beings. They are the building blocks of life.
- There are more than 200 million known proteins, with many more found every year. Each has a unique 3D shape that determines how it works and what it does.
- Predicting the 3D shape of a protein from its sequence of amino acids has been an unsolved challenge for over 50 years.
- Solved by Google Deepmind in 2020.
- Reason for Success: progress in machine learning, in particular deep neural networks.

Self Driving Cars

- A ground vehicle capable of sensing its environment and navigating without human input.

Self Driving Cars

- A ground vehicle capable of sensing its environment and navigating without human input.
- Major milestone 1995: CMU's NavLab 5 completed the first long distance drive (2,849 miles, 98.2 percent autonomous).

Self Driving Cars

- A ground vehicle capable of sensing its environment and navigating without human input.
- Major milestone 1995: CMU's NavLab 5 completed the first long distance drive (2,849 miles, 98.2 percent autonomous).
- Not so impressive (2004 DARPA Grand Challenge): Task of robot vehicles was to drive a 150-mile route in the Mojave Desert region of the United States. None of the vehicles finished the route. CMU's Red Team traveled the farthest distance, completing 7.32 miles.

Self Driving Cars

- A ground vehicle capable of sensing its environment and navigating without human input.
- Major milestone 1995: CMU's NavLab 5 completed the first long distance drive (2,849 miles, 98.2 percent autonomous).
- Not so impressive (2004 DARPA Grand Challenge): Task of robot vehicles was to drive a 150-mile route in the Mojave Desert region of the United States. None of the vehicles finished the route. CMU's Red Team traveled the farthest distance, completing 7.32 miles.
- Great progress over the past 15 years. Many car producers claim to essentially have the technology. However, trust and legal issues are major challenges.

Self Driving Cars

- A ground vehicle capable of sensing its environment and navigating without human input.
- Major milestone 1995: CMU's NavLab 5 completed the first long distance drive (2,849 miles, 98.2 percent autonomous).
- Not so impressive (2004 DARPA Grand Challenge): Task of robot vehicles was to drive a 150-mile route in the Mojave Desert region of the United States. None of the vehicles finished the route. CMU's Red Team traveled the farthest distance, completing 7.32 miles.
- Great progress over the past 15 years. Many car producers claim to essentially have the technology. However, trust and legal issues are major challenges.
- Progress: better sensors, progress in machine learning for mapping environment and self localisation.

Apparently much harder: RoboCup

- RoboCup is an annual football competition for robots founded in 1997.

Apparently much harder: RoboCup

- RoboCup is an annual football competition for robots founded in 1997.
- Offcial aim: By the middle of the 21st century, a team of fully autonomous humanoid robot soccer players shall win a soccer game, complying with the official rules of FIFA, against the winner of the most recent World Cup.

Apparently much harder: RoboCup

- RoboCup is an annual football competition for robots founded in 1997.
- Offcial aim: By the middle of the 21st century, a team of fully autonomous humanoid robot soccer players shall win a soccer game, complying with the official rules of FIFA, against the winner of the most recent World Cup.
- To explore progress visit `https://en.wikipedia.org/wiki/RoboCup`.

Generative AI (roughly 2022)

Systems that generate text, images, etc by responding to prompts.

- ChatGPT, released in 2022, an example of a chatbot based on large language model GPT3 (developed by openAI);
- now many more systems such as GPT4 and LaMDA (Google).

Generative AI (roughly 2022)

Systems that generate text, images, etc by responding to prompts.

- ChatGPT, released in 2022, an example of a chatbot based on large language model GPT3 (developed by openAI);
- now many more systems such as GPT4 and LaMDA (Google).
- All based on the transformer deep learning architecture. Underpinning research published in article "Attention Is All You Need" in 2017.

Generative AI (roughly 2022)

Systems that generate text, images, etc by responding to prompts.

- ChatGPT, released in 2022, an example of a chatbot based on large language model GPT3 (developed by openAI);
- now many more systems such as GPT4 and LaMDA (Google).
- All based on the transformer deep learning architecture. Underpinning research published in article "Attention Is All You Need" in 2017.
- trained using in the order of a trillion words using networks with in the order of 100 billion parameters (2023).

Major Challenge: Explainable AI

- In many domains it is important to understand the decisions and predictions of AI algorithms. Examples: clinical decision support systems, interview support systems, financial advice systems.
- Many AI algorithms (in particular deep learning) are naturally opaque and typically even experts cannot explain their decisions/predictions.
- Lots of research currently into trying to explain to humans the decisions/predictions of such AI algorithms.