Introduction to artificial intelligence

Introduction

Artificial intelligence, or AI, is one of the hottest topics at the moment — some go as far as to talk about a revolution. Some people believe AI applications will solve all of the human kind's problems while others are afraid of it. Most people, however, don't know enough about AI to form informed opinions on it, and in these cases it's easy to grow weary of AI. Understanding the basic principles of AI require some work, though: first of all, the term AI may refer to a variety of things from well-functioning video games to machines that are capable of learning. Depending on the way we define artificial intelligence, we could say that we're already surrounded by AI — or that we haven't seen even one truly intelligent application of it yet. The fact that most texts on AI are written from a technical point of view doesn't make it any easier to understand what we mean by AI, leaving the non-technically oriented people outside.

Nevertheless, the need for deeper understanding of AI has never been more urgent. The development of AI is no longer relevant to researchers only, as more and more people have access to it. Most importantly, AI has more and more impact on our everyday life.

This learning material is written as an easy-access introduction to artificial intelligence and its potential. The material will cover some technical aspects as well, but from a humane point of view. When you're familiar with the most important concepts of AI, you'll be able to find more information on it and follow the discussion around it. This is highly important: the development has taken huge leaps in a short period of time, so it's hard to say where it will go next. We hope this material provides you with the necessary tools to form your own opinions on the nature and future of artificial intelligence — that way, who knows, you'll be able to form the future yourself.

Contents of the material

In the first chapter of this material, we help you form an overall view of artificial intelligence by taking a look into the history of AI research and other closely related fields. After that, we dig deeper into a form of intelligence we're already more familiar with — the human intelligence. What is it, and how should we portray it in order to simulate it in machines? In this context, we also compare artificial and human intelligences, and talk about their interaction. Next, we take a closer look into the way computers operate, and what it takes to simulate intelligence in machines. The chapter on machine learning will provide more information on how all models don't have to be given to computers, as they are able to form their own models as well. In other words, computers can learn new things. This chapter is essential for understanding AI, as machine learning is one of the most important milestones AI has reached.

Machine learning requires learning material just like learners need their textbooks. Suitable material for machines can be found in databases, and therefore they are covered in their own chapter. After that, we introduce different aspects and applications of artificial intelligence in chapters dedicated to machine vision, natural language processing and robotics. In the robotics context, we also discuss the ethical issues concerning artificial intelligence.

Learning material (such as databases and input from the environment)
Models of language
Models of human mind and perception
Logic
Algorithms
Neural networks
Computing power

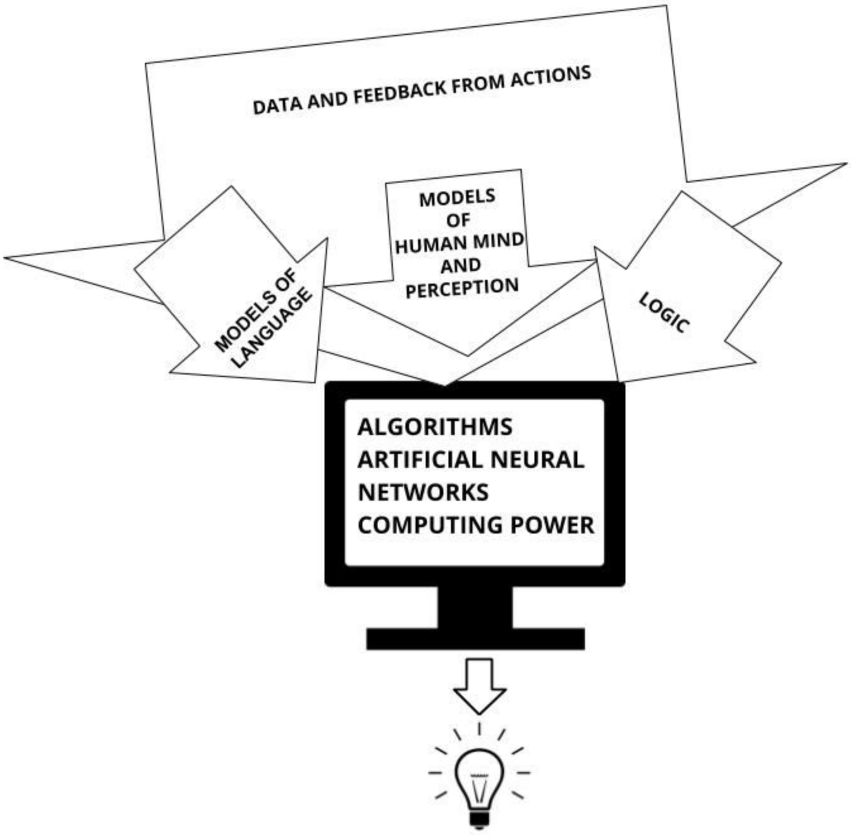

Image 1. The most important elements of AI

Keep in mind

This learning material will provide you with more information on:

- History of artificial intelligence
- Modeling human intelligence
- Interaction between humans and machines

- How computers operate
- Machine and deep learning
- Databases
- How machines process and produce natural language
- Robotics

Closer look

After each chapter, you'll find exercises related to the chapter. The exercises include discussion topics and information gathering assignments, and they state whether they are meant for primary schools, middle schools, secondary schools or adults.
They can be done independently or used for teaching.

Blue boxes will provide further information on different topics.

History of artificial intelligence

Artificial intelligence has roots in many sciences, such as philosophy, logic, mathematics, psychology, computer science and cognitive science. The starting point in the history of artificial intelligence can be pinpointed to the oldest of these sciences, philosophy, which was born out of curiosity towards existence. The most pressing questions related to artificial intelligence include the boundaries and essence of human knowledge. Philosophical discussions are still relevant when we're determining the boundaries and nature of artificial intelligence, and especially the ethical aspects related to them.

Additionally, philosophy sparked interest towards logic, and the first models of deductive reasoning were developed in the ancient Greece in the 4th century BC. The advancement of logic played an important role in the development of mathematics and, later on, the computer science. The history of artificial intelligence began in the 1940s when digital computers were invented. At that time, one of the most important figures in the development of computers was the British mathematician Alan Turing who devised logical and mathematical theories on what could be calculated and how to do it. The second World War created urgent need for developed computers, as they were highly needed for decrypting coded messages.

The development of actual AI applications saw light in the USA in the 1950s. Computers had evolved so rapidly that it cast hope on artificial intelligence doing the same. Additionally, the idea of human nature at that time supported the hopes that intelligence could be reproduced mechanically: human behaviour was seen as mechanical reactions to stimulus, which would have been easy to replicate. AI research was greeted with enthusiasm, and it received substantial funding. AI researchers predicted that a machine that could match human intelligence was only one generation away.

Algorithm is a detailed, step-to-step description or directive on how to solve a problem or perform a task.

The first applications of artificial intelligence were based on logic programming. The basic principle of that is that the programmer writes rules which determine how the machine interpretes predefined markers. Most applications were based on algorithms that had a specific task to perform. Performing an AI task could be compared to walking through a maze: you move forward step by step, and if you end up in a dead end, you go backwards. This kind of traditional AI is often referred to as the Good Old Fashioned AI (GOFAI). The problem was that for most tasks, the number of alternative solutions was enormous. This meant that the number of solutions had to be narrowed down by using shortcuts or rules of thumb that steered the program towards a solution even when it couldn't guarantee reaching one. For example a chess-playing program was given the rule of thumb “always guard the queen” as it was a useful tactic — even though sacrificing the queen could have been a better option in some cases.

The first logic-based artificial intelligence was successful in clearly restricted and defined tasks. It performed well in problems that were difficult for humans, such as calculations and geometric questions. However, tasks that even a small child could perform, such as detecting objects, were impossible for it to tackle. By the 1970s, it was clear that the wild expectations of artificial intelligence would not be met.

The problem with traditional artificial intelligence was that all the knowledge humans possess cannot be presented as rules. Even if we coded all the information found in dictionaries and thesauruses into a computer, it still wouldn’t know as much about the world as we do: humans know that a plate of porridge usually weighs more than a mitten, or that dogs don’t go to sleep in pickle jars.

The 1970s weren’t very eventful when it comes to artificial intelligence. In the 1980s, most of the development took place in expert-systems. Artificial intelligence was used to analyse for example medical samples. For the first time, AI applications were truly useful. Expert-systems worked because they focused on a very restricted task. The programs were fairly simple and easy to alter.

The sciences behind AI made progress during 1970–1990 as well:

- Computers and their computing power were improved.
- Internet was first created for universities, but from the 1990s onwards, it spread out to most households.
- More and more information was saved in digital form.
- Brain imaging developed, which helped gather more and more information on human brains.
- Cognitive psychology and cognitive science that study the intellectual functions of humans gained more prestige.

Moore’s law: Since the 1970s, the number of transistors used in computers has doubled approximately every two years. The number of transistors is one of the factors in increasing the computing power.

By the 21st century, these advancements had led to a new approach of creating artificial intelligence. The approach is based on machines learning new things by themselves, meaning that all of their knowledge doesn’t have to be programmed into them. Learning is possible with the help of either mathematical statistics or distributed, multilevel information processing. The inspiration for the model of neural computation, i.e. neural networks, was taken from the human brains where connected brain cells transmit signals and strengthen the most commonly used connections. This meant that the metaphor was turned upside down: humans were no longer regarded as machine-like operators, and computers were developed to resemble humans. We could well say that the modern artificial intelligence is more closely based on biology than logic.

The most important feature of machine learning is that it can process uncertainties and probabilities. Because of this, our uncertain and complex world is no longer impossible for it to understand — like it was for the traditional, logic-based AI. Machine learning AI is already more accurate than humans in many tasks, but all current AI applications still represent the so-called weak AI. However, weak doesn't mean defective, as the weakness is defined in contrast to strong AI or Artificial General Intelligence (AGI) that can perform broad intellectual tasks like a human can. Strong AI is still only a daydream some researchers have, but most AI researchers disregard these classifications and focus on creating well-functioning applications without paying too much attention to how their applications compare to human abilities. However, human intelligence is what we regard as the basis of intelligence, and therefore it is important that we stop to think about it and its modeling.

Important milestones and people in the history of AI

4th century BCE
Ancient Greeks contemplate deductive reasoning, and the development of logic begins.

15th century
Printing press is introduced, enabling fast reproduction of information.

17th century
Blaise Pascal creates a mechanical calculator that could perform addition, subtraction, multiplication and division. The machine operated on cogwheels.
Thomas Hobbes publishes his book The Leviathan, which included the theory on mechanical view of human thought.
Gottfried Leibniz creates the binary numbering system. He believed human thoughts could be traced back to arithmetic operations to solve disagreements.

19th century
Joseph-Marie Jacquard invents the first programmable machine, a power loom that utilised punched cards.
George Boole develops mathematics in a way that is still the basis of modern computers.
In 1818, Mary Shelley publishes the novel Frankenstein. In it Dr Frankenstein creates a thinking and feeling monster out of a dead body in a scientific experiment.
In 1837, Charles Babbage and Ada Byron design a mechanical, programmable machine called the Analytic Engine. The machine was never built.

1920s
In 1921, Karel Capek's play R.U.R (Rossum's universal Robots) premieres. The word "robot" was used for the first time in that play.

1930s

In 1936, Alan Turing publishes an article called "On Computable Numbers, with an Application to the Entscheidungsproblem" in which he outlines the operating principles of computers — before there were any computers at all.

1940s

In 1942, science fiction author Isaac Asimov introduces the three laws of robotics. When robots follow these laws, they are safe for humans. The laws are:

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.

In 1943, AI research pioneers Warren McCulloch and Walter Pitts publish ”A Logical Calculus of the Ideas Immanent in Nervous Activity” that formed the basis for neural networks.

In 1944, the first ever programmable digital computer, Colossus, is put to use in Great Britain. It was used to decrypt German coded messages during the Second World War.

In 1945, the computer ENIAC (Electronic Numerical Integrator and Computer) is developed in the USA for warfare purposes.

1950s

The development of AI applications begins in the University of Dartmouth in the USA.

John MacCarthy uses the term artificial intelligence for the first time.

Allen Newell, J.C. Shaw and Herbert Simon create the first AI programs called “the Logic Theorist” and “General Problem Solver”.

1960s

The first ever robotics company, Unimation, is founded.

ANALOGY, a program developed by Thomas Evans, proves that computers can solve problems used in intelligence tests.

J. Alan Robinson creates a mechanical proof method called The Resolution method.

In 1965, Gordon E. Moore makes the observation commonly known as the Moore’s law: the number of transistors in an integrated circuit doubles approximately every two years. In 1975, he revised the number to two.

In 1969, AI researchers Marvin Minsky and Seymour Papert publish the book Perceptron which highlighted the restrictions of single-layer neural networks. The book put neural network research temporarily on hold.

The first ever international conference on artificial intelligence was held in the USA.

1970s

Stanford University introduces ARPAnet, the forerunner of internet.

INTERNIST program for solving medical diagnostic problems is created.

1980s

Information retrieval methods based on neural networks are developed and become more commonly used.
Personal computers (PCs) become more commonly used, and more and more information is saved in digital form.

1990s
The machine learning era begins.
In 1996, The Deep Blue chess computer wins against a chess world champion.
Internet access becomes broadly available outside universities and to common people.

2000s
The development of self-driving cars begin.
AI research on detecting and expressing feelings takes steps forward.
Robot vacuum cleaners become available.
Human-like robots walking on two legs become better and better at walking and grabbing things.[1]

2010s
The deep learning era begins.
AlphaGo, a program developed by Google, wins a game of the difficult board game Go against a human.
Automatic translation systems and speech recognition take steps forward thanks to deep learning.[2]
Artificial intelligence outperforms humans in certain tasks related to image recognition.[3]
Ethical questions concerning artificial intelligence become a topic in public discussion.

Keep in mind

Artificial intelligence has roots in many sciences such as philosophy, logic, mathematics, psychology, computer science and cognitive science.

AI has taken huge leaps on the 21st century thanks to advanced machine learning and, most importantly, deep learning.

[1] See for example https://www.darpa.mil/program/darpa-robotics-challenge
[2] See for example https://translate.google.com/
[3] See for example https://aws.amazon.com/rekognition/

Closer look

Middle school, secondary school and adults

Choose a milestone or a person from the timeline and use a search engine to look up more information. The most influential individuals in the history of AI include Alan Turing and Marvin Minsky, among others.

Modeling intelligence

When discussing artificial intelligence, it's impossible to ignore human intelligence — after all, that's what we use as the benchmark for intelligence in the first place. There's no consensus on the definition of human intelligence, but we could say that the ability to learn, solve problems and adapt to changing circumstances are, among others, important markers for intelligence. We can also differentiate human intelligence into several categories, such as social or linguistic intelligence. In some categories, including logical-mathematical intelligence, machines have already outperformed humans. Human-like general intelligence is still only a daydream some ambitious AI researchers have, as most are content with an application that's capable of a single type of intelligence.

The underlying assumption behind artificial intelligence is that the intellectual functions of humans can be modeled. We need to make this assumption as the operation of computers is based on models; only the features that can be defined in detail can be simulated in machines. Modeling the human mind is essential for artificial intelligence, and especially cognitive science has made progress in that. It is an interdisciplinary science that combines findings from philosophy, psychology, computer science, linguistics and neuroscience. What sets cognitive science apart from psychology is that it concentrates on the cognitive functions including perception, learning, memory, attention, deduction, decision-making and language production. Unlike some branches of psychology, cognitive science strives for the precision of exact sciences in its

Modeling means representing the features and relations of a phenomenon or a system with the help of a model. Examples include maps, miniatures and mathematical models.
Some things are harder to model than others: for example consciousness is very hard to model.

findings. In other words, accurate quantitative research and modeling are important goals for cognitive science. Its research methods include brain imaging, psychological tests and computer simulations. A good example of a research interests is the distribution of attention when a driver uses their mobile phone behind the wheel. This could be examined with the help of eye movement detection cameras in a driving simulator.

Cognitive science aims at discovering models of how the human brains represent, process and transform information. Intelligence is seen as the ability to create memory and understanding, detect patterns, and learn from previous experiences. The underlying idea of human is that humans are active information processors who learn effectively by processing things by themselves, and creating their own data structures. The main focus is on the human cognition, but cognitive science is also interested in animal cognition and data processing in artificial systems.

The findings of cognitive science research are also utilised in AI research, but cognitive science in itself doesn't strive for creating artificial intelligence. Cognitive science is all about analysing, describing and predicting cognitive functions of the human mind. The models created in this way are important for AI research.

The difference between human and artificial intelligence

In many categories of intelligence, the models of artificial intelligence have already outperformed humans. For years now, AI has been able to beat humans in tasks that require calculations, such as a variety of games. Nowadays, machines already make less mistakes than humans in certain tasks related to language usage and recognition.

Famous matches between humans and artificial intelligence
Deep blue program won a game of chess against a human in 1997. AlphaGo won a game of the notoriously complicated Go against a human in 2015.

However, humans are still the only ones capable of general intelligence that is flexible and multi-functioning. In other words, human intelligence can (more or less) tackle all sorts of tasks, whereas artificial intelligence only knows how to perform the exact task it has been given, and its ability to make generalisations from new information is limited.

What makes it so hard to compare human and artificial intelligence is that we are still far from knowing everything about the human mind. For example, there is no consensus on how human consciousness works. It is especially hard to explain how all the sensory information from the sensory receptors come together with memory and the frame of mind to create a coherent sense of consciousness. The sense of selfhood is also closely related to this.

Additionally, the human sense and strive for esthetics is still not fully determined. Some say that only humans are capable of creativity. Whether this is true or not depends on what we mean by creativity. If we define it as the ability to create new combinations of existing elements, machines are perfectly capable of it.

AI has already written poetry and composed music.

Use a search engine to look up *AI poetry* or *AI music*.

Interaction between humans and artificial intelligence

Another reason for modeling human functions is that it improves the interaction between humans and machines. Ever since the time of punched card computers, we have have the need to think about the way we use machines. As computers have evolved, the means of interaction have changed. Nowadays, we use keyboards, mouses and touch screens. Additionally, a whole new branch of science has evolved: user interface design. It aims at enabling smoother interaction in order to make using machines as easy as possible.

AI applications have revolutionised the interaction between humans and machines. Nowadays, it's not just about the user knowing how to use a machine: machines can learn more and more about the user as well. The fact that machines have gotten better and better at understanding both spoken and written language has made a huge difference in the way we interact with machines. This applies to machine vision as well: machines can for example assess personal traits and the alertness level of a driver by just "looking" at the driving.

At the moment, one of the most important research topics related to the interaction between humans and machines is emotional intelligence. Improved detection of facial expressions and gestures makes AI better and better at it. Recognising emotions is highly important to many AI applications, such as automated customer service and care robots, as interacting with applications and robots that are able to take note of our current mood is much more comfortable for us.

Keep in mind

Modeling intellectual functions is key to reproducing them in machines, only the features that can be defined in detail can be simulated in machines.

Cognitive science is interested in modeling the human mind which is essential for artificial intelligence.

Modeling human functions is also essential for improving the interaction between humans and machines.

Closer look

Middle school, secondary school and adults

1.
How would you model mobile phone addiction or cheerfulness so that a computer would be able to process them? (Think about the features that make them what they are, and a way to measure those features.)

2.
If there was an AI application that could recognise emotions, what kinds of practical applications could you create to make people's life a little easier?

3.
How does the creativity of AI differ from human creativity — or are there any differences in the first place? Using a search engine, look up artworks, poems or songs created by AI, and see if your friends can identify whether they were created by AI or a human.

4.
Use an arts lesson to create art mechanically. You can try to differentiate features of a painting, and then combine them according to set rules, for instance.
Here are some examples of possible rules:

- Divide the original painting into 10 equal sections and number them from 1 to 10.
- In your own artwork, use only the colors you can see in section 5.
- In your own artwork, use only the shapes you can see in section 8.
- In your own artwork, copy one random section from the original painting as it is.

Afterwards, think about the creativity behind your artwork and the whole process of doing it. How was that different from seeking inspiration from existing artworks?

Logic behind artificial intelligence

Logic is the study of valid reasoning, and reasoning is an important part of intellectual functions. That's why logic is essential for both computer science and artificial intelligence. It forms the theoretical basis for computer science by defining what can and cannot be calculated and classified.

We can define logic as an exact and precise method of looking into the validity of a statement. The exactness and precision are achieved by looking only at the formal content of the reasoning without paying attention to the actual meaning. Therefore, reasoning can be formally valid regardless of what the statements mean, or even if they make no sense at all, as you can see from the following examples:

> If all humans are mortal
> and Socrates is a human,
> Socrates is mortal.
>
> If all boys are apples
> and Oliver is a boy,
> Oliver is an apple.

In computers, logic is implemented with the help of algorithms. To understand the basic principles of how computers and artificial intelligence work, it's essential that we know what algorithms are. An algorithm is a detailed, step-to-step description or directive on how to solve a problem or perform a task. We could say that a recipe in a cookbook is an algorithm for preparing food. The outcome is a result of what ingredients have been added and how they were handled. In computers, the ingredients as well as their combinations and outcomes follow logic, and they are executed with the help of a programming language.

Programming languages are translated into a form that a computer can read: sequences of ones and zeros. After that, we move from the software level to the physical level of the hardware, which also follows the rules of logic. The whole transferring of signals is based on so-called logic gates that either allow or block signals from moving forward according to the rules of logic. Their operating principle is based on the logic of AND-, OR- and NOT-operations, the most used of which is the AND-operation. In computers, AND-gate receives two (or more) input signals, and generates a single output signal. The output signal is true only if all input signals are true. The following chart illustrates the explicit definition of the AND-, OR- and NOT-operations.

AND

A	B	A AND B
true	true	true
true	false	false
false	true	false
false	false	false

OR

A	B	A OR B
true	true	true
true	false	true
false	true	true
false	false	false

NOT

A	NOT A
true	false
false	true

All of digital technology is based on these simple AND-, OR- and NOT-operations. They also illustrate the basic principles of how computers work, i.e. how precise segments are connected using explicit rules.

Logic and artificial intelligence or machine learning

Traditional AI applications were purely based on logic. They could for example have instructions such as “if a green block is next to a red square, move a yellow triangle to the left”. This kind of a conditional statement sets preconditions, and the command will be executed only if these preconditions are met. Logic based applications require that all factors are clearly defined, and all rules are explicit.

In real life, there aren’t many intellectual tasks that are so explicit that they can be modeled with logic. At which stage is a red square no longer red if we mix it with white to make a lighter shade? When does a square become a circle if we round its edges?

These are the kind of problems that made traditional AI fall through. Ever-changing and diverse world just cannot be divided into segments that would fit logic-based systems. However, there are many intellectual tasks that don’t require logical perfection: in many cases, probable or close-enough answers are all we need. In real life, practical reasoning is more important than valid reasoning. As machine learning is more flexible than traditional logic-based AI, it suits practical reasoning better.

Critique towards modeling intelligence

In the previous section, we established that the logic-based operating principles of computers require us to model intelligence in a way that determines the factors of intelligence explicitly. Now, let’s imagine we have succeeded in that, and computers are able to simulate all forms of human intelligence perfectly. Would the intelligence of computers be identical to human intelligence?

One of the most famous arguments against artificial intelligence is a thought experiment called Chinese room[4], developed by philosopher John Searle in 1980. According to it, machines cannot understand or be aware of what they're doing. In the experiment, you're asked to imagine a room that has two doors. From one of the doors, someone delivers questions in Chinese, and through the other one, someone in the room delivers answers to the same questions in perfect Chinese. An outsider might be fooled to think that someone in the room must know Chinese, but in closer inspection, we see that the room is equipped with dictionaries that have entries on all of the questions and answers. Someone in the room simply uses the dictionary to look up the questions and then pairs them with their answers — without understanding a word of Chinese. This kind of mechanical pairing of characters without any understanding of their meaning is exactly what computers do. In a broader sense, the thought experiment suggests that meanings exist regardless of the mechanical processing rules applied to them, which would make it impossible for machines to achieve human-like intelligence.

The argument was made against strong AI. Weak AI applications aren't even supposed to be capable of human-like intelligence, which means that the argument doesn't apply to them. Additionally, the thought experiment was developed in a time when AI was still based on logic alone. On the other hand, even the modern and machine learning AI applications are based on combining characters mechanically, regardless of the fact that the characters used in machine learning are more subtle, and the rules of combining aren't always known to us.

Strong or general AI (Artificial general intelligence AGI) is capable of human-like, all-round and flexible intelligence, and can apply its intelligence to new kinds of tasks. Some say strong AI is capable of experiencing consciousness, but there is no consensus on this matter. So far, no one has been able to create a strong AI.

[4] "Minds, Brains, and Programs," by John R. Searle, from *The Behavioral and Brain Sciences*, vol. 3. Copyright 1980 Cambridge University Press.

Keep in mind

Logic is central in both computer science and artificial intelligence. Logic is the study of valid reasoning. In computers, logic is carried out by algorithms.

Algorithms are detailed, step-to-step descriptions or directives on how to solve a problem or perform a task. We could say that a recipe in a cookbook is an algorithm for preparing food.

Closer look

Middle school, secondary school, adults

1.
Use an internet-based test to familiarise yourself with logical reasoning. Use a search engine to look up *logical reasoning exercises* or *mensa*.

2.
What aspects of humanity are hard or downright impossible to express in logical form (for example intuition, soul, love or beauty)?
How closely are they related to intellectual processing? In other words, is it possible to model intellectual processing without them?

Machine learning

Machine learning refers to computers being able to learn new things without humans having to give them detailed directions on what to do in each situation. The idea is that computers search inside their collection of data in accordance with prewritten algorithms to discover patterns in the data. With their help, the computer can then make estimations and predictions as well as answer questions on the topic at hand.

The development of machine learning started in the 1950s hand in hand with AI applications. However, it took until the 2000s to take machine learning to the level where it could outperform humans in certain tasks. The goal couldn't have been reached without the increased amount of learning material, i.e. databases, improved computing power, and new kinds of advanced algorithms.

One of the most used applications of machine learning is housing price prediction. The price of an apartment cannot be determined based on its size alone, as a variety of other factors play an important part in the price as well: location, condition, floor, number of bathrooms, and so on. This is exactly the type of complex task where machine learning is at its best.

The precondition for machine learning is that we have enough learning material on the subject at hand. In order to predict housing prices, we need to give the computer information on previously sold apartments, including selling prices and detailed descriptions of the apartments. In the learning phase, computer tries to find possible connections within the data, and determine how much weight each factor is given. If we think about a studio apartment, the number of bathrooms doesn't play any role in the selling price, but if a large apartment has only one bathroom, it will affect the price. Some factors have stronger connections than others, and some have no connections at all. These connections and their weights are used to form patterns that will help the computer predict the prices of apartments entering the housing market.

Different types of machine learning

Machine learning can be divided into three most important types: supervised learning, unsupervised learning and reinforcement learning. All of these have their own applications, but the basic principle remains the same: computers are provided with large datasets that they use to learn how to detect patterns in the subject area.

Supervised learning

In supervised learning, computers are provided with examples on what kinds of results have been reached previously. One application of this is housing price prediction: the computer is given information on available housing options as well as the prices that have been paid previously. The idea is that computers are taught how to recognise and form patterns that they can use to make predictions based on similar new data.

Unsupervised learning

In unsupervised learning, computers aren't given the classification of what kinds of results are needed. Spam email filtering can be carried out with the help of unsupervised learning: in time, email programs can learn how to tell spam and normal messages apart based on the contents, sender address, and previous cases. Compared to the problem-oriented nature of supervised learning, unsupervised learning aims at making classifications without any help.

Reinforcement learning

The third type of machine learning is reinforcement learning. In it, learning computers are called agents. These agents then scout their environment and adapt their actions based on the feedback they get from the environment, while striving to reach the best possible outcome.

As an example of reinforcement learning, we could teach a computer to play backgammon. As is the case with all games, there are numerous possible states and outcomes for each step of the game. Programming all of the steps one by one would take ages, but with the help of reinforcement learning, we can teach the computer to react to all of the steps. We can teach the computer by letting it play against an actual human or against another reinforcement learning algorithm.

At first, the agent performs random actions. If an action improves the agent's premises, it is rewarded. The agent always strives to achieve the best possible outcome. The actions it performs are improved step by step, and after several trials, it will learn ways to develop them even further. After a number of attempts, the agent will be able to form a sequence of actions that lead to high points.

Examples on algorithms used for machine learning

All machine learning models are put into practice with the help of algorithms that can be based on mathematical statistics or neural networks. The housing price program we discussed earlier is an example of a statistical algorithm called linear regression. It is well suited for depicting complex events as it allows analysing the simultaneous effect of several variables.

Genetic algorithms

Another type of statistical algorithms used for machine learning is genetic algorithms that are based on the evolution theory. The algorithms mimic mechanisms of natural selection, such as survival of the fittest and the inheritance of traits from the fittest.

With the help of supervised learning, we can teach a genetic algorithm to write a sentence "programming is fun", for example. First we define the goal, which in this case would be writing the sentence. After that, we give the computer a rough description of how to reach that goal. The description could include the number of characters (18 in this case) and a clarification that the characters we're looking for are letters of the alphabet, numbers and punctuation marks such as a full stop or a space. Then it's time to specify the initial population, i.e. a bunch of random attempts to reach the goal, and the size of the population, which can be anything from ten to thousands. In our example, a desired trait could be "how many characters are in their right place", and the more right characters there are, the fitter that individual is.

After the population is specified and gone through fitness analysis, it's time to create rules for forming the next generation. We want the fittest individuals to have the best chances to reproduce and spread their genes, so we create an environment where the fittest ones get most picks. For the purpose of our sentence-building example, we could use a simple method of inheriting the first part of the sentence from one parent, and the latter part from another.

The last step of the algorithm is mutation. In humans, the most simple mutations are ones where a single nucleotide base is changed in a sequence of DNA. Mutations in genetic algorithms follow the same principles: some traits don't correspond to the inheritance from parents. Mutations maintain the diversity of a population, and keep it from converging prematurely.

Genetic algorithms can represent either supervised or reinforcement learning. If fitness analysis is based on comparing results with the desired outcome, it's supervised learning. If the analysis is carried out by examining how different individuals perform, we're talking about reinforcement learning.

Neural networks

The idea of networked information processing in artificial intelligence has been around since the dawn of AI research, but it wasn't until the 1990s that the neural networks started to take notable steps forward. Their model is based on human brains where information processing is carried out by a vast network of neurons. A neuron transfers a signal from e.g. a visual perception from one neuron to another through their network in accordance with the strength and type of the stimulus.

An artificial neural network (ANN) consists of an input layer, one or more hidden layers, and an output layer. Each layer consists of so-called knots, or nodes, that perform calculations and other operations, and are connected to the nodes on the previous and the next levels. Signals go through knots in accordance with certain rules, starting from the input layer and through the hidden layers onto the output layer. The connections differ in strength, and the operations can alter that strength. Information processing in neural networks is based on these connections.

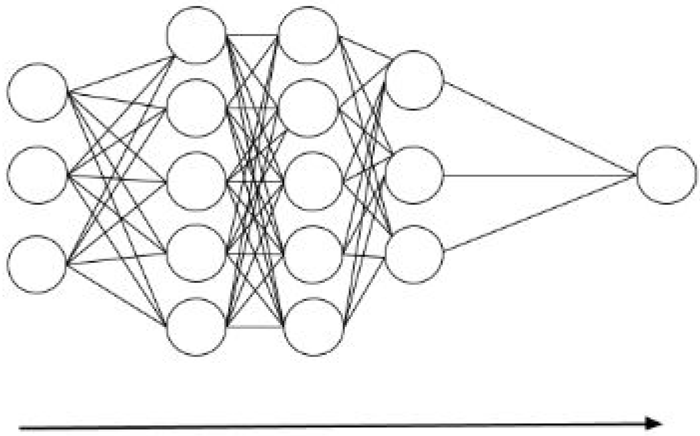

Image 2. On the left side, you see an input layer that's followed by three hidden layers, and on the right side, an output layer. As input, we can give the network for example sentences that the network analyses to form patterns on the relationships between words, and learn to identify sentence elements.

The subject of learning is deconstructed into components, and the connections between them are then presented based on their strength. As input, we can give things like sentences from a novel. When this input is run through the network several times, the network learns to recognise word classes based on where the words are located in sentences. First, the network learns to break the text into sentences, then words, and ultimately, words into their stems and inflectional affixes. For example nouns occur with certain grammatical cases, and can be identified as nouns that way.

The most important feature of neural networks is that they are self-organising, meaning they can alter their own connections. Input can be run through the network several times, and each time alters the connections. This makes the networks capable of learning. Their goal is to detect different kinds of connections in their input, and use that to form patterns.

Deep learning

Neural networks that consist of several layers are called deep learning. Deep learning methods have developed rapidly in the 2010s, thanks to algorithms, increased computing power, and large databases. Several layers enable several interphases, meaning they can perform more and more complicated tasks. Thanks to deep learning, artificial intelligence has learned to identify objects in images, and understand and produce natural language. In restricted areas, deep learning can even outperform humans in these tasks.

Deep learning models don't depend on the alterations humans make to their algorithms. Instead, they are self-teaching, meaning that they can make their own alterations to their information processing methods – just like the human brains. Of all the types of machine learning, deep learning is the closest equivalent to human learning. Thanks to its self-teaching nature, deep learning is flexible and able to work with uncertain information. This enables it to perform well in tasks presented by our complex world.

One application of unsupervised deep learning is a program that classifies video material. As input, we have millions of videos and text-based material related to them, such as descriptions and comments. A deep-learning system learns to detect connections between the videos and text-based descriptions, and ultimately recognise characters and events in the video material. The system can learn to recognise e.g. giraffes even if we don't ask it to do so. The system forms a sort of theory that covers everything it needs to know about recognising giraffes.

Table: Differences between traditional machine learning and deep learning

	Traditional machine learning	Deep learning
Operating principle	Different kinds of automated algorithms that learn to predict future outcomes by analysing historical data from databases.	Analyses the properties and connections of the database with the help of neural networks that feed relevant information through several layers.
Control	Algorithms are created and controlled by analysts.	Algorithms are self-adaptive.
Result /Output	Usually a numerical value, such as a result or a classification.	Whatever; text, pictures, sounds or other elements.

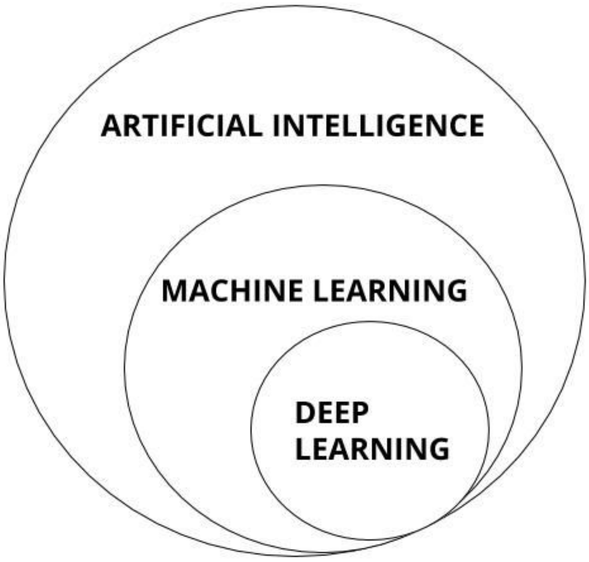

Image 3. The relationship between artificial intelligence, machine learning and deep learning

Keep in mind

Computers search inside their collection of data in accordance with prewritten algorithms to discover patterns in the data.

The advancements made in machine learning have been made possible by the growing number of databases that the computers need for learning, as well as increased computing power, and new kinds of developed algorithms.

The most effective form of machine learning used in AI applications is called deep learning. It means that the data is processed in a multi-layer network.

Closer look

Middle school, secondary school, adults
Compare your own learning to machine learning. What similarities and differences can you find? What strengths do humans and machine learning have?

Databases

Database is a collection of data on a specific subject. Databases are formed when a company, authority, or another organisation needs to store and retrieve information. Usually, the information saved in a database is related to a goal: if a supermarket chain wants to keep track of consumer behaviour, they can collect information on the loyalty card number, purchases, dates and times, and then save it in their database.

Achievements in machine learning and the expansion of databases have taken place simultaneously. Machine learning requires a massive amount of information, and the increased availability of collected data has advanced its development. Nowadays, there are databases on nearly everything we do. Trade and industry have their own interests, but more and more information on our private lives can also be found in databases. This information can consist of data on our phone calls, computer usage, bus rides, or even on how dirty our laundry is. If we use phone apps, the amount of data being collected increases: how many steps we take during a day, who we know, where we are (or where our phones are, to be more precise), what kinds of music we like...

Databases can be open or proprietary. Information in proprietary databases is non-public and only accessible to the administrator of the database. Such databases are maintained by authorities and health care providers, for example.

As an example of an open database, visit the Statistics Finland's Paavo database where you can retrieve information on residential areas and their population: https://www.stat.fi/tup/paavo/index_en.html

Where are the databases?

Databases can be stored either locally or in a cloud server. Local databases are stored on the same computer as the program using the database, whereas databases stored "in a cloud" refer to the servers of companies offering cloud storage services. Cloud storage can be accessed via internet anywhere and anytime. A cloud is basically a network of servers, i.e. computers. So after all, even the information in clouds is stored in a physical place somewhere.

Databases can be updated automatically by a device that's connected to the Internet, such as an electricity meter or an activity bracelet. In these cases, we talk about the Internet of Things (IoT). Its applications also include cars that send information to the manufacturer's database automatically. In case of a car crash, the information collected by the car's sensors is instantly and automatically sent to the manufacturer's database.

Nowadays, IoT produces more and more information, and transferring this data to a cloud requires faster and faster internet connections. For this reason, network equipment manufacturers have started selling 5G technology, which includes new and improved cellular mobile networks and wireless data transfer systems.

Many of the new AI applications, such as self-driving cars and industrial robots, need to be able to make decisions in microseconds. In these situations, transferring the data to a cloud takes too much time, which means that the data processing needs to take place closer to the data collection site. This technology is called edge computing, and it has made real-time data processing possible. However, edge computing isn't widely used yet, so the most important databases are still stored in clouds.

How is the data organised?

When it comes to machine learning, it's important that we don't focus only on how extensive a database is, but also on the quality of the data. When the database that machine learning uses for learning is extensive and of high quality, it's easier to find the right answers. If the database that machine learning uses is small and the data itself is poor, the answers can be pure nonsense. One of the factors affecting the quality is the way the data has been organised.

One common way of organising data is called the relational model, meaning that the information is represented as a collection of relations between concepts. When a database is built according to the relational model, the first step is to think about concepts that have to do with the subject, and then investigate how they are related to each other. If our subject was a supermarket, possible concepts would include a customer, shopping basket, and the prices and quantities of the things inside the shopping basket. Of all the possible concepts, we choose only the most crucial ones. After that, we try to find the connections between the concepts, and define them further by adding more attributes, such as the date of purchase, or the location of the store. Additionally, we try to organise the database in the most logical way possible, as this will help us avoid unnecessary copies, and make the queries more simple.

Customer: Customer number, Name, Address, Email address
Purchases: Purchase number, Customer number, Store number, Date
Store: Store number, Address, Manager

Image 4. An example of the connections between information in a relational database.

How are databases used?

When analysing a large amount of data, we can use traditional statistical methods or new machine learning algorithms. Traditional data analysis is based on models that have been drawn by humans based on things like averages and percentages. This kind of data analysis is descriptive and based on data that has been collected from previous events. It's possible to make predictions based on it, but the predictions are always made by humans.

In machine learning, the models, predictions, and evaluations of the models are all carried out without a human having to take part in it. Forming models of future events is also possible. All of this happens so quickly, extensively and accurately that a human could never match that.

Data protection issues

A lot of data may be collected, but it cannot be done without limitations. Legislation has been slowly catching up to secure the citizens' rights related to personal data. The EU General Data Protection Regulation (GDPR) that came into force in 2018 guaranteed all EU citizens the right to review the data that's been collected on them, get information on how the data has been collected and how it is handled, and who has access to it. Additionally, citizens now have the right to rectify incorrect information, and remove their personal data from a register. Respectively, the keepers of the registers are obliged to work towards fulfilling these rights.

The term personal data covers basically all information related to a private person, such as names, pictures uploaded into social media, email addresses and health records. Even the IP address of your computer could be regarded as personal data even though it cannot be used to identify anyone alone. There are very few companies who aren't affected by the this regulation. However, many countries outside of Europe don't regulate the use of personal data as strictly.

For most people, the only concrete impact of the GDPR has been one extra click to accept the collection of data by accepting cookies on websites, for example. Most people don't even know how much data is collected on them, or how it is used to steer their behaviour.

Cookie is information that the server saves on a user's device. The browser then sends the information back to said server. Cookies can be saved on the user's device, or deleted after the user is done using the service.

Cookies collect information on things like:

- User's IP address
- Time
- Previously visited websites

Social media and search engines tailor their contents and advertising based on the user data collected from the internet. The advertisements can then be targeted at a specific group of people that is more likely to buy the advertised products than the general public. Because of personally tailored contents, users may end up living in a bubble where they only see contents that support their own views, and interact with like-minded people. This can make it harder to find new information that could challenge the user's personal views. We call that phenomenon the filter bubble.

The organisation of databases can also improve privacy. This is especially true in the case of blockchain technology which allows strangers to produce and manage distributed databases. The idea of distributed databases is that they are divided into blocks between the participants, and all of the blocks are chained. This means that the database can be verified and compiled from multiple sources. All users have the same, updating version of the database. As all alterations remain visible and they can't be changed retroactively, all members can trust each others. Blockchain technology has enabled digital currencies such as Bitcoin, intelligent contracts, micropayments, and voting systems.

Keep in mind

Database is a collection of data on a specific subject.

Databases are essential for the development of artificial intelligence, as machine learning requires a massive amount of learning material.

Closer look

1.
What kinds of databases can you create about your family or classmates? What kinds of databases are useful? Collect information on your classmates (favourite song, right- or left-handedness, number of coloured pencils in their pencil case etc.) and collect it in a chart to create a database. Can you find new information by analysing the database? (Such as a connection between the number of coloured pencils and a good arts grade.)

2.
What kinds of human activity is left outside databases?

Machine vision and object detection

Machine vision can be defined as automated information collection from digital pictures and videos. Object detection and classification tasks related to machine vision aim at modeling and automating functions that would match the human visual system, and such applications are developed within AI research. However, usually the goal is simply an automated system. These kinds of system can be used for quality control in industrial robots, for example.

When it comes to AI applications, machine vision has proved to be one of the biggest success stories in recent years. If we measure the intelligence of an AI application by comparing it to the skills of a human, object detection is one of the fields of AI that has reached the level of humans in a restricted area. With the help of artificial intelligence, computers can even outperform humans in object detection. AI applications have for example beaten humans in identifying different species of monkeys and orchids. Additionally, programs are able to identify objects from video material and even predict what's going to happen next in that video.

If given enough input, AI applications can learn to recognise basically anything. The program creates a set of rules for itself and uses them to identify the object. Deep learning makes the process easier.

Giving rules on how to detect an object is hard to do in advance. You get the idea by looking at a coin in your hands: depending on the angle, distance and the way your finger covers it, the coin can look circular, oval, linear, or like a dot or a crescent. A human is still able to say it's a coin, as we have come across a coin so often that our brains have saved its concept on the cerebral cortex and the network of brain cells. In other words, these connections between the brain cells activate each time we see a coin. The same principle applies to machines that tweak the connections based on the input they receive.

Applications

The oldest application of machine vision is sorting and quality control in factories. A machine vision that oversees an assembly line can check the measurements and structures of products, read the barcodes, and count the number of products. Machines are both more accurate and faster than humans in these kinds of tasks. The machine vision applications used in manufacturing are good examples of tasks that can and should be assigned to machines. It's relatively quick to start utilising machine vision on an assembly line, provided that the deep learning image recognition program gets to see enough samples in the beginning. That way it can learn to detect anomalies in the production.

Another important application of machine vision is medical image analysis. The amount of medical imaging and information has increased while the examinations have become more and more complicated. Artificial intelligence can help medical professionals spot relevant areas from X-ray images, or enhance the quality of images. AI can help humans reach a diagnosis faster, and it also makes the process more accurate and a lot less expensive. AI still hasn't replaced the role of humans in diagnostics, but it's a tool doctors can use in their work.

From the AI's point of view, object detection offers a lot of interesting applications in the form of self-driving cars and robotics. Moving and operating in a space requires constant observations of the surroundings. A moving robot can use infrared sensors to spot obstacles, but self-driving cars need to also identify what the obstacles are. Is the obstacle on the road a plastic bag that can ran over, or is it a large stone? This level of object detection requires deep learning. The speed of information processing is essential: the faster the robot or car moves, the less time there is for object detection, which means that real-time processing is essential.

Other new and promising AI research interests include social perception. Machines have learned to detect micro expressions and use them for emotion recognition, but the next challenge is to make machines understand why humans experience those emotions and why they react to situations like they do. This understanding of human behaviour would be especially important to AI applications that interact with humans.

Another interesting application of object detection is augmented reality (AR). AR means that you can see computer-generated perceptual information that appear on top of the real-life environment on the screen of a phone, for example. This information can consist of images, sounds, videos, texts or information on the current location. In order to place the additional content in the right part of the natural environment, the program needs to monitor the environment through the phone's camera, and recognise objects in the natural environment.

Machines still haven't reached human perception

Machines may well be more accurate than humans in classifying their sightings in a restricted environment, but real life is far too complex and ever-changing for machines to reach the potential of human perception. Everything we see cannot be turned into learning material for machines. Additionally, machines aren't yet able to give reasons for why things are the way they are in pictures.

Another difference between machines and humans is that humans can detect and recognise things they encounter rarely or for the first time. In other words, humans can learn new things the very first time they see something, whereas a deep learning neural networks always need several inputs of the same thing.

Keep in mind

AI applications of machine vision and object detection have taken huge leaps in the 2010s mostly due to deep learning.

The most important application of machine vision is quality control in factories. Machine vision is faster and more accurate than humans in detecting faulty products on the production line.

Closer look

Middle school, secondary school and adults
Find out what CAPTCHA stands for and how it’s related to AI applications of machine vision.

Processing natural language

For many AI applications, the ability to understand and generate the language we use, the so-called natural language, is very much needed nowadays. However, if we look at natural language from the computer science's viewpoint, we notice that it's by no means precise. It's not only a very wide and diverse system, as it's also a very open-ended one.

Image 5. Example on language ambiguity.

Before a machine can understand and generate natural language, it needs to be able to process it. Natural Language Processing (NLP) refers to using computer programs to analyse and generate natural text and speech.

Natural language processing is a branch of language technology which diverged from linguistics and became its own science in the 1960s as computers were evolving. In the 1980s, the need for language processing became even more urgent as personal computers brought text editing programs to all desktops. That's when the first spell and grammar checkers as well as hyphenators were introduced for the first time. More and more information was produced and saved in electronic form.

Processing natural language was still clumsy, and for example translations were generated word for word. Traditional, logic-based programming didn't make it easy to model natural language.

It was a piece of cake. = "It was a slice of a baked good."
It's raining cats and dogs. = "Kittens and puppies are falling from the sky."

By the 21st century, deep learning had taken natural language processing to a whole new level. Neural networks have made modeling complex meanings possible, as they learn from examples: now we don't even have to try forming rules for inexact phrases. Neural network algorithms have evolved, and the number of available databases has increased, which has meant huge leaps for natural language processing.

Levels of natural language

Natural language can be divided into different levels, some of which are easier to model than others. For example phonetics, the study of physical properties of speech sounds, deals with physical, measurable study interest that is therefore easy to model: sound frequency can be measured and expressed in hertz, and all vowels have their own frequency.

The level of syntax is also easy to model. Syntax refers to the study of the rules of sentence structures which can be illustrated with the help of syntactic trees.

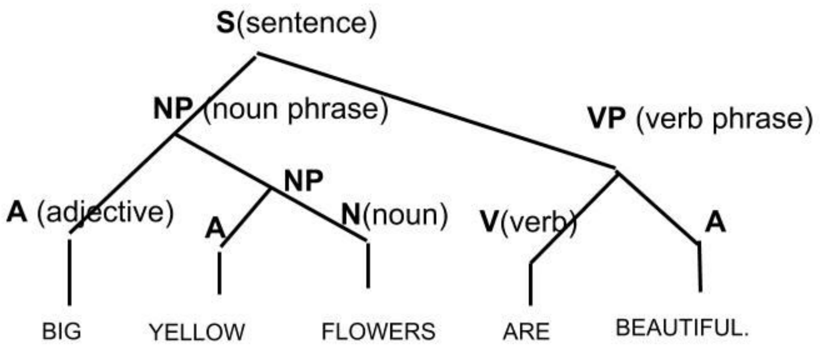

Image 6. Syntactic tree

Of all the levels, semantics and especially pragmatics are the hardest to model. Semantics deals with meanings and relationships between words and what they stand for. Machine learning models are already able to learn meanings and relationships between constructs.

Examples on the relationships between meanings: homonyms look or sound the same but have different meanings (She'll park the car near the park), and synonyms have the same meaning (quick, fast, rapid).

Pragmatics, however, studies the use of language in a wider context. People often communicate much more than what is actually said. Most sentences in any language are ambiguous, but we rarely notice it as the context makes the meaning clear to us. If someone says "I saw bats", you'll most likely know whether they have looked at animals or baseball bats — or whether they are using a saw to cut them. Pragmatics also deals with the unspoken rules related to the use of language, and how the rules can be broken in the name of humor or irony. This wider context of language is still impossible for computers to truly understand.

Steps in natural language processing

The process of turning natural language into a form that a computer can understand takes multiple steps. Its starts with preparing the textual or voice-based material by removing irrelevant information, such as headers and footers, email address bars, and page numbers. The actual language processing can be carried out using various methods, such as recognising pairs and clusters of words that often occur together, and different kinds of data compression techniques that downsize the problem.

Depending on the task at hand, the next step is to form models with the help of machine learning. The final step is to evaluate the models and use them for an application.

Natural language processing in AI applications

AI applications need to be able to process natural language especially if they are meant for communicating with humans or need to understand communication between humans. For example discussion forum moderation, machine translations, automatic speech recognition, automatic text generation and speech to text -applications all require natural language processing.

One good example of a relatively easy language-related task that AI applications deal with is spam email filtering. Directing all messages that contain the word “casino” or “bonus” is simple enough, but the next step is to process language like search engines do, as they need to recognise the inflected forms of words as well.

The most difficult tasks related to natural language processing include translating, summarising messages, and answering questions. One way to test the quality of AI is the Turing test in which a human tries to identify whether their questions are answered by another human or a computer program. If the program is able to fool the tester, we can say that the AI has reached the level of humans. The same test can be applied to chatbots as well, but it cannot be used to determine intelligence or the level of thinking on a broader level. First of all, many of the highly developed AI applications don’t communicate with humans in the first place, so they wouldn’t be able to pass the test even if they were considered highly intelligent. On the other hand, even human intelligence isn’t always articulated well in linguistic form.

When it comes to challenging translation and summarising tasks, a lot has happened in the 2010s. Programs can understand the semantics and contextual meanings of sentences better and better. This means that for example machine translations have improved substantially, and automated customer service understands its customers better than ever. This progress can be seen in both textual and spoken forms. For humans, speech is the primary form of communication, and the number of speech recognition based applications has increased dramatically. The most famous ones include the voice-activated virtual assistants Siri and Alexa created by Apple and Amazon. Voice-activated applications have proved to be useful to many user groups, such as people with visual impairment.

Keep in mind

AI applications need to be able to process natural language especially if they are meant for communicating with humans or need to understand communication between humans. From the computer science's viewpoint, natural language is not precise at all, and that's why it needs to be processed. This kind of processing is the study interest of language technology.

Closer look

Primary school

Test how well a machine understands your speech or writing. (Use a search engine to look up *chatbot* or use the assistant on your phone.) Try to come up with a question only a human could answer. What makes the way we use language so hard for machines to process?

Middle school, secondary school, adults

Use a search engine to look up more information and videos on neural networks that can process language. You can type in *neural network learns to speak*, for example.

Notice how the network learns to speak like a child would. First, you can only hear random uttering, but as the network processes the input, you'll be able to understand more and more of what it's trying to say.

Robotics

Robots may well be the most widely recognised form of artificial intelligence, all thanks to the entertainment industry. Human-like intelligence is present in Data from Star Trek, Terminator, and R2-D2 and C-3PO from Star Wars. A number of movie screen robots have thought about their own consciousness and the differences between robots. Even the word “robot” itself comes from a play: it was used for the first time in the play R.U.R (Rossum’s universal Robots) written by the Czech playwright Karel Capek in 1920. The play deals with themes that still sound topical: robots have taken over people’s work, and humans need to think about the feelings and rights of robots. Most of the robots we use nowadays are industrial robots operating in factories, which the play also predicted.

When asked to think about a robot, most people immediately think of humanoid robots that resemble humans. However, industrial robots don’t look anything like humans: their appearance has been tailored to fit their task. In the industrial context, robots are used for tasks that are hazardous, hard, or repetitive for humans.

Automation and robotics are closely related terms. Both automation and robots can be defined as devices that perform a mechanic task. The difference between them is that robots are always controlled by a computer, but automation can function mechanically without one. An example of this would be a thermostat of a radiator, as it doesn’t require a computer to perform its task. When a computer controls a system, we call it an embedded system. In other words, the computer is embedded in the device. This means that all robots are embedded systems.

In the Finnish national core curriculum for basic education 2014, the following learning objectives were outlined for middle school crafts: “Implement embedded systems in crafts by using programming for designs and products”. (Finnish National Agency for Education 2014: grades 7-9, crafts S3). In other words, all pupils should create an embedded system such as a robot or any other product that can be controlled by a programmable computer. You’ll find more information on robotics and embedded systems in the *Introduction to teaching robotics* learning material.

Humanoid robotics

Humanoid robotics is a narrow, highly demanding branch of robotics. It covers basically all the core elements of artificial intelligence, i.e. deduction, predicting future events, natural language processing, and machine vision. On top of that, a humanoid robot should be able to move as well as recognise and express emotions. Without machine learning, these functions would not be possible. As machine learning has taken steps forward, so have robots that move or speak like a human.

Modeling movements that come naturally to humans is a challenge. That's why developers have turned to less complicated creatures, such as insects and snakes. The easiest way to move a robot is to give it wheels, but wheels lose their power in the face of an obstacle.

We could well say that robotics forms its own branch of artificial intelligence, the embodied AI, that is based on the idea that the intelligence of operations is created in interaction with environment. The role of bodily form in human intelligence has gained more prestige, and nowadays we don't consider intelligence being created solely by the brains. It's more and more common to see the human body as a one big entity that's affected by interaction between physical and social surroundings, emotions, and even gut flora.

Robots interact with their surroundings with the help of sensors and actuators. Sensors provide robots with input that the processors in their computers can analyse. Actuators then turn the results into a form we can see or hear. We could well compare this to the way humans operate: our senses provide us with information about our surroundings, we process the information in our brains, and based on the results, the actuators, meaning our body and language systems, make us perform certain tasks.

Philosophical questions concerning robotics

Many of the philosophical questions become more pressing when artificial intelligence takes steps from the virtual world towards the physical world and interaction with humans. Some of the questions have to do with the ethics of AI. Machine learning has made artificial intelligence more and more autonomous, which means that it is capable of deciding how and when to work. This may cause fear or at least spark questions on who's responsible for what a robot does. Could it decide to go against the original plan? In a moral sense, would it be responsible for its own actions?

Another category of philosophical questions deals with the consciousness and "human rights" of robots. People have the tendency to humanize artificial intelligence applications and especially humanoid robots. This tendency is called the ELIZA effect. When an ATM tells us to "take the receipt", we may unconsciously think that the ATM is communicating with us, even though it's actually just showing us a sequence of characters as it has been programmed to do. This effect has us thinking that robots are smarter than they actually are, which makes imagining they have a consciousness even easier. As we established in the chapter on modeling intelligence, we don't even know everything about the human consciousness yet. That's why it's way too early to make any assumptions on robot consciousness either.

Keep in mind

Most of the robots we use nowadays are industrial robots operating in factories.

Automation and robotics are closely related terms. Both automation and robots can be defined as devices that perform a mechanic task. The difference between them is that robots are always controlled by a computer, but automation can function mechanically without one.

Humanoid robotics is a narrow, highly demanding branch of robotics. It covers basically all the core elements of artificial intelligence, i.e. deduction, predicting future events, natural language processing, and machine vision. On top of that, a humanoid robot should be able to move as well as recognise and express emotions.

Closer look

Primary school

1.

Play a robot game where you're only allowed to move following predetermined commands and repetition. You can use commands such as:

Step forward
Turn right
Turn left
Repeat ___________ X times

This game will introduce and showcase the basic principles of programming, i.e. giving commands and progressing step by step.

Middle school

2.

Search further information on Isaac Asimov's Three laws of robotics. Are these rules devised for science fiction sufficient to make robotics safe?

Secondary school and adults

3.

Compare the autonomy of AI applications and the free will of humans. Are humans fully autonomous, considering our behaviour is affected by genes, environment and previous experiences?

Can an immobile robot be autonomous?

Closing words

Artificial intelligence has taken giant leaps, and it has already brought changes to people's lives all around the Western world. It's hard to predict where the development will lead to, but we can and should steer its course. The least we can do is minimise the risks and harms, but we can also set our goals at building artificial intelligence that works for the greater good of humankind. It's good to think about these things now that we are still forming policies on the new and more effective applications of AI that have been made possible by new technology, such as deep learning.

The most important values behind rules and policies concerning AI research have been transparency, responsibility and societal benefits. Transparency has to do with the stability of AI applications, meaning that they operate the way they were intended. The only way to gain certainty of this is to make sure the applications operate in a way that is transparent.

The call for transparency is also relevant for personal data privacy. The information in databases used by AI must be open to corrections, and the grounds for decision making need to be traceable. This ensures AI-based systems operate in a responsible way.

The future and values of AI is important to many operators, such as:
https://ai100.stanford.edu/
https://www.tekoalyaika.fi/en

To ensure that AI development can benefit the society as a whole, public government needs to be able to influence the course of development. At the moment, most of the AI development is in the hands of market forces and business interests of companies. This has raised concerns over the uneven distribution of the benefits AI has brought to us. This kind of uneven distribution can present itself in growing income and wealth gaps that may endanger the social cohesion of a society. The way public government can take part in preventing this is through education, research and public spending. It's essential that common people can take part in the discussion as well, and that's why it's important to know the operating principles and possibilities of artificial intelligence.

When you look into the possibilities of artificial intelligence, the course of development may feel concerning. However, it's important to keep in mind that artificial intelligence has been subject to high expectations ever since the concept was born, and some of them have never been realised. That's why critical thinking is needed even during the current AI craze. No matter where the development leads us, humans will still play an important part in complex systems.

Glossary

5G New cellular mobile networks and wireless data transfer systems that have been developed to meet the needs of the Internet of Things.

Artificial general intelligence (AGI) AGI is capable of human-like, all-round and flexible intelligence, and can apply its intelligence to new kinds of tasks. Some say strong AI is capable of experiencing consciousness, but there is no consensus on this matter. So far, no one has been able to create a strong AI.

Algorithm is a detailed, step-to-step description or directive on how to solve a problem or perform a task. We could say that a recipe in a cookbook is an algorithm for preparing food.

Augmented reality Computer-generated perceptual information that appear on top of the real-life environment on the screen of a phone, for example. The information can consist of images, sounds, videos, texts or information on the current location.

Blockchain Data storage system where all users have the same, updating version of the database. As all alterations remain visible, and they can't be changed retroactively, all members can trust each others. Blockchain technology has enabled digital currencies such as Bitcoin, intelligent contracts, micropayments and voting systems.

CAPTCHA (Completely Automated Public Turing test to tell Computers and Humans Apart) An image recognition task to keep malwares from accessing computer services. They can consist of hard-to-read letters set in distorted images, or image recognition tasks, to name a few. Machine learning systems have already learned how to crack certain types of text-based Captchas.

Chatbot A computer program that can hold a conversation with its user with the help of written or spoken messages. Commonly used for customer service purposes on websites.

Cloud service A storage for databases that can be accessed via internet, managed by a hosting company.

Cookies Information that the internet server downloads onto a user's computer. The browser then sends the information back to the same server. A cookie can be saved on the device either for good or for the duration of the session, and they can include information such as the user's IP address, and which web pages the user has visited in the past.

Data Bits of information that aren't necessarily meaningful but can be used as material for meaningful information. It consists of digital raw stock such as datasets, statistics, financial information, maps, pictures, videos and 3D models.

Data mining Data analysis methods that are used to find relevant information from large datasets.

Deep learning Neural networks that consist of several layers are called deep learning. Several layers enable several interphases, meaning they can perform more and more complicated tasks. Thanks to deep learning, artificial intelligence has learned to identify objects in images, and understand and produce natural language.

Edge Computing A new technology that enables devices to process data instead of sending it into a cloud first. This makes real-time data processing possible.

ELIZA effect The tendency to unconsciously assume computers have human-like features. The effect is named after the chatbot-like computer program ELIZA created in the 1960s.

Emotional artificial intelligence Detection and utilisation of emotions in interaction between humans and machines.

Filter bubble Internet users only see content that favours their own values and opinions as the content has been tailored to them according to their user information.

GDPR (General Data Protection Regulation) A regulation in EU law that sets boundaries for the collection and processing of personal data.

Humanoid Anything that has been built to resemble the human body.

IoT (Internet of Things) Internet of things refers to machines and appliances that are connected to internet and able to gather data and send it to a database. Electricity meters and activity trackers are good examples of such appliances.

IP address Number sequence that is used for network interface identification. IP does not identify the actual user.

Moore's law Since the 1970s, the number of transistors used in computers has doubled approximately every two years. The number of transistors is one of the factors in increasing the computing power.

Neural network A method of information processing that takes its inspiration from the human brains. A neural network consists of an input layer, one or more hidden layers, and an output layer. Each layer consists of so-called knots, or nodes, that perform calculations and other operations, and are connected to the nodes on the previous and the next levels. Signals go through knots in accordance with certain rules, starting from the input layer and through the hidden layers onto the output layer. The connections differ in strength, and the operations can alter that strength.

Open data Public data collected or created by the public sector, organisations, or private companies, that is open for access and free of charge for everyone.

Programming Giving instructions to a computer with the help of a programming language. Widely used programming languages include Python, Lisp, Prolog, Java, and C++.

Singularity A hypothesis stating that artificial superintelligence (ASI) will trigger technological growth and social changes that happen so quickly that humans no longer can fathom or predict the future.

Turing test A quality test for AI, developed by Alan Turing in 1950. Based on the Imitation Game, the test includes a human trying to identify whether their questions are answered by another human or a computer program. If the program is able to fool the tester, we can say that the AI has reached the level of humans. The same test can be applied to chatbots as well, but it cannot be used to determine intelligence or the level of thinking on a broader level.

Virtual reality (VR) A computer-generated experience taking place within a simulated environment. The environment can either simulate a real-life environment or aim at creating purely fictional surroundings.

Weak or narrow AI An AI that's meant for handling restricted tasks and doesn't strive for a human-like intelligence. All current applications of AI still represent weak AI. However, weakness does not mean that the applications are defective or that they perform poorly, as weakness is defined in contrast to strong AI or Artificial General Intelligence (AGI).