Will we ever create true AI?

The Meaning of Human Intelligence

Hadar Ákos Tibor

Content

Introduction

Of course, we've all heard of AI. This is the bases for dozens of movies, series, books and other areas of pop-culture. And, as we've made our way towards a time completely dominated by technology and science, those stories are becoming more-and-more real each day.

A few decades ago, if you were to tell someone that in the near future a cylindrical 'box' (Alexa) will be able to play their favourite songs, tell them the weather forecast, read the news, tell jokes and much more all on verbal command, they would have considered you either incredibly foolish or insane. And yet, that is the reality we live in today. However, as much as we like to unify these devices under one category and call them Artificial Intelligence, I think we can all agree that they are still far from real intelligence. If I asked you whether or not you can have a genuine conversation with Alexa, Siri, Cortana or any other AI like the conversations you have with other people such as your friends, you'd probably say no.

So, the question comes naturally: what is intelligence? Once we have answered this, we can determine if we'll be able to create true AI. Of course, this is far from a simple question and psychologists are split among different answers to it, however it's simply too fascinating not to talk about. We will explore what defines an intelligent being, what types of intelligences exist, and the technology that is necessary to replicate intelligent behaviour.

As you can see from the title page, in this project we'll be diving into what intelligence actually is and what it means to be an intelligent being. We'll explore the limitations and, at the same time, advances of technology in the field of AI and machine learning. The goal here is to gain a better understanding of the human mind and to glance at the future to see what possibilities lay ahead in the realm of Artificial Intelligence; and, of course, to answer the question: will we ever create true AI? This EPQ will be broken up into two main parts: in the first part, we'll try to come up with a definitive model of intelligence, and in part two we'll discuss the presence of intelligence in AI based on that model.

Part I: Defining Intelligence

Spearman's Two-Factor Theory

Many people argue that intelligence is a single general ability, whereas others say that it's a more complex phenomenon which is made up of other, more specific skills and abilities. From this, it's not at all surprising that over the years, different psychologists have developed multiple and often contrasting theories about intelligence and created several tests in the attempt to measure it. The most famous and widely used is, of course, the IQ test. However, this is a relatively new type of test, so let's not jump ahead of ourselves.

One of the first people to study intelligence was Sir Francis Galton in the 19th century. He was fascinated by the concept of a gifted individual. So, he created a lab to measure reaction time and other physical characteristics to test his hypothesis which states that intelligence is a mental ability and is a product of evolution. It's interesting that his theory would state this as he was the cousin of the world-renowned biologist: Charles Darwin.

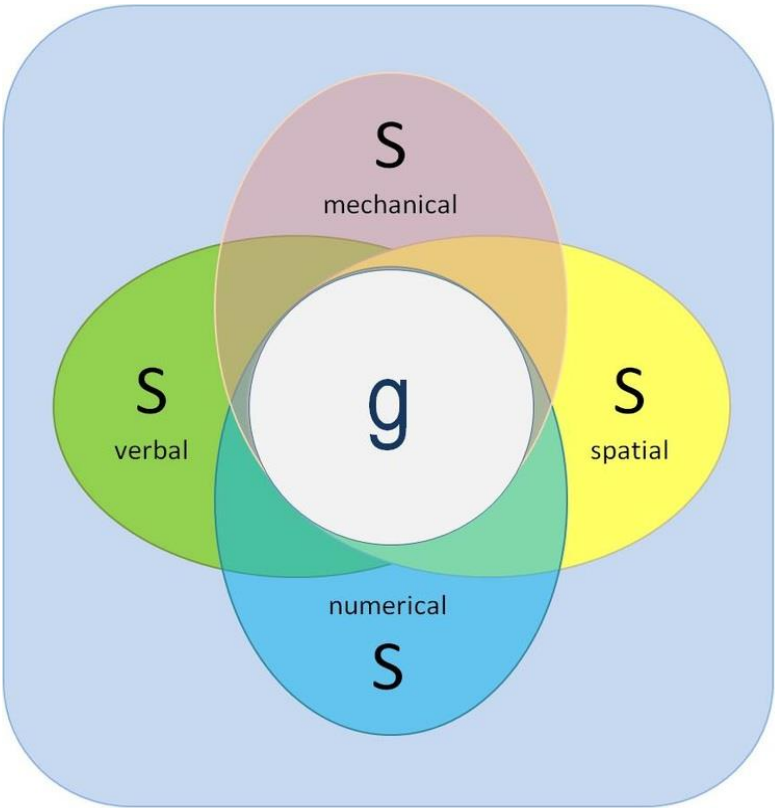
According to Galton, because reaction time and other physical attributes were evolutionarily advantageous, they would also provide a good indication of general mental ability. (SimplyPsychology, Charlotte Ruhl, 2020) Galton's study of intelligence in the laboratory setting and his theorization of the heritability of intelligence paved the way for decades of future research and debate in this field.

Like Galton, Charles Spearman also believed that intelligence was just one ability as opposed to it being the sum of multiple, different skills. Spearman was an English psychologist known for his work in statistics and factor analysis which he later used in his own model of intelligence. Factor analysis is a method used to study the correlation of related fields to find a common factor that is the cause for this correlation.

In his model, he proposed the idea of General Intelligence or 'g factor'. This was the 'common factor' underlying different fields of intelligence such as Numerical, Verbal, Spatial and Mechanical. Numerical intelligence is the ability to manipulate numbers in a variety of mathematical problems. When looking at AIs, this branch of intelligent behaviour is one which would arise naturally (if you can call it that) due to a computer's ability to complete complex calculations much faster than any single human. Moving on to the next category: Verbal intelligence is the ability to express yourself clearly using a wide vocabulary and to understand others in detail. Spatial intelligence, or spatial awareness, is being able to picture objects and data and manipulate them using your imagination. An example of this would be if I asked you to imagine a die with the 6 facing you.

Now, rotate it 2 faces to the right and then 2 faces up: what face are you looking at? Finally, Mechanical intelligence is a measure of fine motor skills and athleticism.

Spearman found that people who do well in one of these areas are also likely to perform higher in the others. In other words, there is a strong positive correlation across different fields of intelligence. (Richard Haier, 2022) From this, he concluded that there is one single 'g factor' representing a person's general intelligence, and another factor, 's', which represents their intelligence in one specific skill:



Credit: Samar Education

Today, almost all researchers who study intelligence use Spearman's theory of the g-factor. (Richard Haier, 2022) This model is commonly referred to as the 'Two-factor theory' of intelligence. Keep in mind that the model shown above is a simplified depiction of this theory as the four 's' factors can be subdivided into more specific categories.

Thurstone’s Primary Mental Abilities

Later on, Louis Leon Thurstone developed his own model of intelligence. Thurstone was an American psychologist, although his interests were initially rooted in mathematics and engineering. He got his doctorate in psychology in 1917 and in 1938 came up with his theory of Primary Mental Abilities.

Thurstone’s PMA theory of intelligence was based on Spearman’s work and the methods he used to develop his Two-factor Theory. Thurstone used the new advanced technique of factor analyses to Spearman’s factor analysis technique to understand the various factors that related to human intelligence. In his research, he conducted 60 mental abilities tests on a group of subjects and examined their test scores using the inter-correlation technique and matrix algebra. (Wendy Johnson, Thomas J. Bouchard Jr., 2005) The inter-correlation technique is a statistical tool used to analyse the correlation between variables such as test scores. At the end of his research, he arrived at a conclusion which contrasted Spearman’s work.

Thurstone came up with seven different mental abilities which he named the Primary Mental Abilities. He refused to accept Spearman’s theory as his findings didn’t seem to support it and leaned more towards his own PMA theory. In his work, he stated that everyone has a certain level of each of these seven abilities and he went further, saying that these abilities are not dependent on each other and that it is possible to measure them separately. (Louis Leon Thurstone – cited by Robert Sternberg, 2003) These are the seven PMAs:

Mental Ability	**Description**
Word Fluency	Ability to use words quickly and fluently in performing such tasks as rhyming, solving anagrams, and doing crossword puzzles.
Verbal Comprehension	Ability to understand the meaning of words, concepts, and ideas.
Numerical Ability	Ability to use numbers to quickly compute answers to problems.
Spatial Visualization	Ability to visualize and manipulate patterns and forms in space.

Perceptual Speed	Ability to grasp perceptual details quickly and accurately and to determine similarities and differences between stimuli.
Memory	Ability to recall information such as lists or words, mathematical formulas, and definitions.
Inductive Reasoning	Ability to derive general rules and principles from presented information.

The difference in the conclusions drawn from the two theories can be due to the different research methods used by the two psychologists. For instance, Spearman's sample of subjects consisted of randomly chosen individuals and it mostly was made up of children. Thurstone, on the other hand, conducted his research exclusively on school-students. They also used a different choice of tests. Spearman's tests were very different from each other because if the tests were identical or if they were to share an overwhelming batch of similarities, the 's' factor could have overlapped. This would've caused additional correlations among variables which then would have led to incorrect results. However, the tests that were used by Thurstone were almost identical. (StudiousGuy, 2020)

It is interesting to point out that when measuring the g-factor of one individual, it is virtually irrelevant what battery of tests you pick out. If you were to pick 10 randomly selected tests and applied them to estimate one's g-factor (this is essentially what an IQ test does) and then repeated this with a completely separate battery of equally random tests, the scores obtained by the individual will have a correlation coefficient over 0.9. (Richard Haier, 2022)

Even though Thurstone stated that the seven PMAs were independent of each other, he never managed to prove this. When he attempted to prove it, he instead rediscovered the presence of the g-factor. In his final interpretation of his theory, he proposed that the g-factor 'coexists' with the PMAs. However, we cannot confidently say that his theory isn't valid to some extent. The PMA theory explains that people possess various unique mental abilities and it has been shown that when a student is completing a task, at least one of Thurstone's PMAs are involved. (StudiousGuy, 2020)

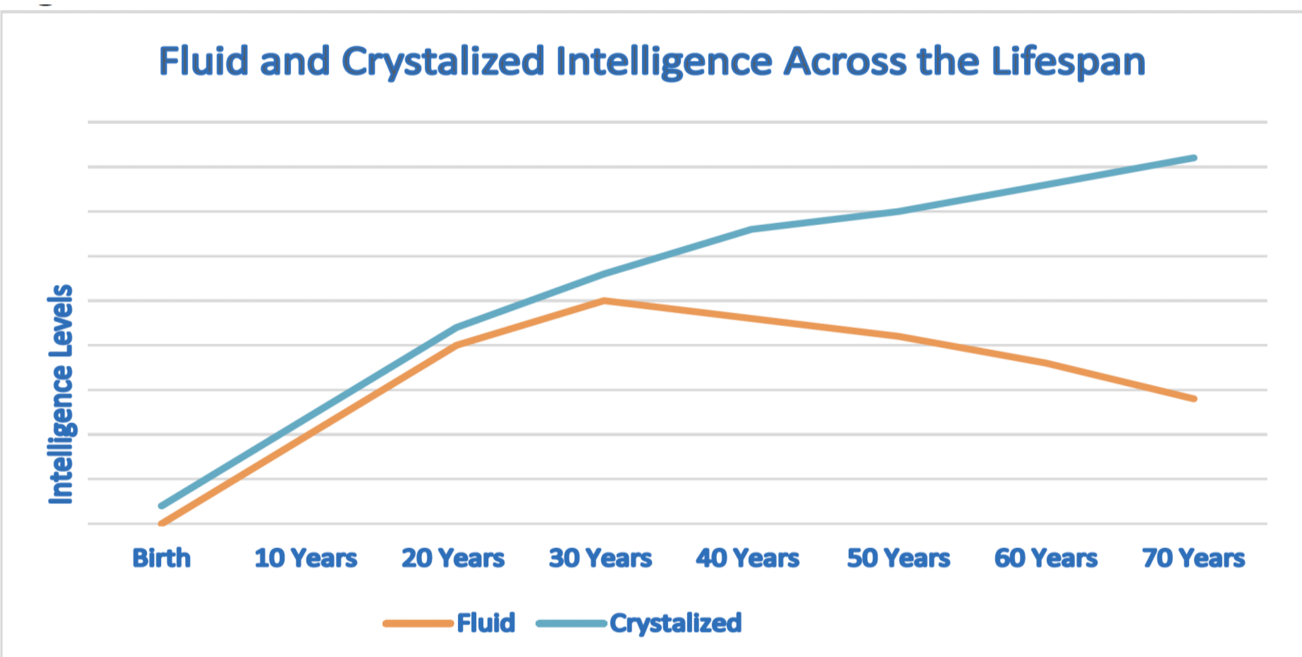
The Cattell-Horn Theory – Fluid and Crystallised Intelligence

Raymond Cattell was a British-American psychologist; the first to suggest the concepts of Fluid and Crystallised intelligence in 1963. This was developed further by John Horn, a cognitive psychologist and scholar who played an important role in developing theories of multiple intelligence. (Wikipedia, 2021)

Cattell's theory was based on Thurstone's PMA theory and the research he conducted in the 1930s. He stated that Fluid Intelligence (Gf) included inductive and deductive reasoning abilities, saying they were influenced by biological, neurological and environmental factors. He also believed that Crystallised Intelligence (Gc) consisted of previously acquired knowledge and previously learned abilities/skills. (Encyclopaedia of Special Education by Elaine Fletcher-Janzen and Cecil R. Raynolds, 2014)

Fluid Intelligence, as the name suggests, is continuous. It is the ability to solve problems in unique situations without any prior knowledge, but rather through the use of logical and abstract thinking. As one grows older, Gf increases however later starts to decrease typically in the late 20s. Crystallised Intelligence is the knowledge which you accumulate over your life. This includes facts learned in school, specific motor skills or muscle memory. (Raymond Cattell, 1963) With age, Gc increases as you, hopefully, learn new skills.

The graph below shows how Gf and Gc change over the course of a normal life:



Credit: Simply Psychology

Over a long period of time, it is entirely possible for Gf to become Gc. In other words, methods and different processes can become second nature. So, it could be said that overall, the total intelligence of one individual roughly stays the same as Gf slowly moves into the area of Gc. IQ tests prove this as the scores obtained on an IQ test at an early age are highly correlated with scores in later life.

Emotional Intelligence

We've talked about certain types of intelligence which are mostly based on logical thinking and life-experiences. However, in everyday life we don't only need to be able to think logically: it's also vitally important to understand each other and build deductions based on information about how other people are feeling.

Emotional Intelligence is the ability to understand your own and other people's emotions and to differentiate between emotions. It is also important for using information about someone's emotions to help modulate your thinking and behaviour in a way that is appropriate for that specific situation. (Peter Salovey and John D. Mayer, 1990)

The four main parts to overall Emotional Intelligence (EQ) are **Self-Awareness** (the ability to focus on yourself and how your actions, thoughts, and emotions relate to your internal standards), **Self-Management** (being able to regulate your behaviour and emotions and keep them in check), **Social-Awareness** (this is commonly known as Empathy) and **Relationship-Management** (how well you communicate with and understand others). (Travis Bradberry and Jean Greaves, 2006)

Credit: Travis Bradberry and Jean Greaves

For a computer to be considered as intelligent as a human, it would have to possess EQ with all four of its attributes. However, in order to have Self-Awareness and Self-Management, the computer would also have to be conscious, and whether or not it's possible for a computer to have consciousness is another question altogether. So, for the purposes of this study, we will define EQ by the other two qualities which are about others rather than one's self. Therefore, when it comes to EQ, a computer has to have Social-Awareness and Relationship-Management in order to function (emotionally) like a human.

Revised Intelligence Model

So now, based on the models mentioned before, we can build a fairly accurate picture of a revised model of intelligence which can then be applied to an AI. This is going to be a checklist of sorts; we'll need to pick out key points from the intelligence models (only points which hold ground in all the models) and build a summary of criteria which need to be met for something to satisfy the definition of intelligence.

First of all, Spearman's model: the g-factor, as we discussed, is the general intelligence of any individual and it's the 'common factor' underlying specific (S) fields of intelligence. Namely Verbal, Mechanical, Numerical and Spatial. So, the first four points in the revised model will have to be these abilities. It is also useful that the g-factor can be estimated using a standard IQ test. The IQ score of an individual is an indicator of one's g-factor **relative to the rest of the population**. (Richard Haier, 2022) This means that on the IQ Bell-Curve, there is actually no absolute 0.

Next is Thurstone's PMA model. Now, this model and Spearman's Two-Factor theory have a lot in common: Verbal Intelligence in Spearman's model can be split into Word-Fluency and Verbal Comprehension from the PMA model. This gives us a more detailed look into this branch of intelligence. Similarly, Numerical Intelligence and Spatial Visualisation are also mentioned in both theories. The three PMAs which are only seen in Thurstone's model are Perceptual Speed, Memory and Inductive Reasoning. So, these will be the next addition to our model.

The Cattell-Horn theory can also be summarised by parts of previous models. This is not surprising as it was based on Thurstone's work. Fluid Intelligence has been defined as the ability to solve problems in unique situations using logical and abstract thinking. It is made up of Deductive and Inductive reasoning skills. We've already mentioned Inductive reasoning, however, Deductive reasoning is just as important. Crystallised intelligence can be said to be a part of memory as it involves recalling previously acquired knowledge. In fact, it is part of Long-term memory. Therefore, we can split Memory up into Long-term and Short-Term memory and this will now include Gc.

Finally, Emotional Intelligence: as we've said previously, we'll need to use a 'simplified model' of EQ where Self-Awareness and Self-Management have been removed. With that, we're left with Social-Awareness and Relationship Management.

On the next page is a summary of the final, revised model:

Attribute	***Definition***
Word Fluency	The ability to use words quickly and fluently in performing such tasks as rhyming, solving anagrams, and doing crossword puzzles.
Verbal Comprehension	The ability to understand the meaning of words, concepts, and ideas.
Numerical Ability	The ability to use numbers to quickly compute answers to problems.
Mechanical Intelligence	A measure of fine-motor skills and athleticism.
Spatial Visualisation	The ability to visualize and manipulate patterns and forms in space.
Perceptual Speed	The ability to grasp perceptual details quickly and accurately and to determine similarities and differences between stimuli.
Long-term Memory	The ability to recall specific information and knowledge after a long period of time (for example a year).
Short-term Memory	The ability to hold detailed information about something for a short time (for example a few minutes)
Inductive Reasoning	The ability to derive general rules and principles from presented information.
Deductive Reasoning	The ability to arrive at a logical conclusion from multiple points of information.
Social Awareness	The practice of Empathy.
Relationship Management	The ability to communicate with and understand others.

Part II: Applying the Revised Model to AI

Overview

In this part, we'll analyse current and possible future AI using the model from Part I. The aim here is to find out whether or not it's possible for a computer to possess the 12 attributes listed in the model and thus be considered intelligent in the same way as a human.

Verbal Intelligence – Word Fluency and Verbal Comprehension

Verbal Intelligence (as we saw in the model) can be split into Word Fluency and Verbal Comprehension. These intelligences are vital as they are the way you understand others without using body language or facial expressions. This would be a part of EQ. They also play a big role in how you comprehend ideas and concepts and how well you find connections across multiple ideas.

Today, AIs can already 'understand' sentences and are able to respond to them appropriately. For example, Amazon's Alexa understands instructions in many different ways. This means that you could tell her to do something using two unique sentences and she would understand that while the words you chose were different, you actually meant the same thing. Example: "set up a 10-minute timer" can also be said as "start a timer for 10 minutes".

However, AIs don't fully comprehend language the same way as we do. Many AIs which seem like they understand language and ones that score higher than humans on a common set of comprehension tasks usually don't notice when the words in a sentence are jumbled up. The problem here is in the way that natural-language processing (NLP) systems are trained to understand words and comprehend sentences. (Will Douglas Heaven, MIT Technology Review)

Researchers at Adobe Research and Auburn University realised this when they told NLP systems to give logical explanations for their claims. An example would be to explain why it judged two different sentences to have the same meaning. When they used this approach, they found that changing the order of words in the sentence generally didn't change the AIs explanation in any way. (Will Douglas Heaven, MIT Technology Review)

Natural Language Processing is a specific branch of computer science which deals with a computer's ability to understand language in a similar way we humans do. NLP combines computational linguistics (a rule-based model of language) with statistical, machine learning, and deep learning models. These revolutionary techniques allow AIs to process human language and to 'understand' what it means. (IBM, 2020)

The research team tested multiple state-of-the-art NLP systems by feeding them normal sentences at first, and then the same sentences but with mixed-up words. The AIs were convinced that both versions were saying/asking the same thing. For example, they thought that "Does marijuana cause cancer?" and "Does cancer cause marijuana?" were the exact same question. (Will Douglas Heaven, MIT Technology Review)

Based on this, we can see that modern AIs are not fully equipped when it comes to Verbal Comprehension. It is clear that, even though they understand instructions when we're talking about AIs such as Alexa or Siri, they don't actually understand the meaning of those instructions. This can be translated to Word Fluency as a computer would need to understand language in order to have the same speaking ability as a human.

Numerical Ability

This, as we've said before, arises 'naturally' in computers. Any sophisticated modern machine needs to be able to carry out millions and sometimes billions of calculations per second. On average, a modern processor has a clock speed of 3.5GHz to 4GHz. (HP, 2018)

Complex mathematics (such as calculus) is made up of two main types of operations: numerical and symbolic. Numerical operations only involve numbers and there are no unknowns. Symbolic operations deal with unknown variables which need to be represented as symbols or letters such as in algebra which involves symbols like 'x' and 'y'. (Peter Dunn, MIT, 2010)

Numerical operations can be split up into addition, subtraction, multiplication, and division. In order to use a wider range of numerical values without exhausting memory and processing resources, computers use the floating-point system. In this system, large integers such as 25000000 are replaced with floating-point values such as 2.5 x 10^7. This usually only approximates the result, however, with a few minor tweaks an answer could be reached which has negligible error compared to the true value. (Peter Dunn, MIT, 2010)

Symbolic operations are much more difficult to manage in a computer system. The first problem is to represent symbols using only binary code. This is done by coding for 'x' with one number and for 'y' with a different one. The computer interprets the numbers as codes and not as numerical values. More complex expressions can be represented by the decomposition (breaking down) of expressions into simpler parts which are connected by pointers in the computer's memory. This allows the representation and the processing of expressions such as 'x + 2y'. For example, the differentiation of a complex expression (such as SinX + Cos(X+½)) can be broken down into operations that differentiate simpler expressions (SinX then Cos(X+½) separately). Even though these procedures can prove to be quite long, today's microprocessors are performing billions of operations per second which cuts down the calculation time immensely. (Peter Dunn, MIT, 2010)

Mechanical Intelligence

When it comes to humans, this is extremely important. We need to be able to navigate our movement on a daily basis, and while it may appear easy and simple at first glance, this is actually a really complex skill and requires a surprising amount of brain-power.

With modern technology, we are extremely close to achieving human-level athleticism in robots. In fact, one might say we have already done it. The most famous example is Boston Dynamics' Atlas. Atlas is a bipedal humanoid robot which can already complete multiple complex athletic movements such as running, jumping over obstacles, balancing at an angle, balancing on one leg and it can even do the backflip. (Boston Dynamics, 2021)

Credit: Boston Dynamics

This obviously shows an incredible feat of engineering as Atlas demonstrates human-level physical ability. However, Mechanical Intelligence is more of a necessity for humans than it is for AI. A computer can behave intelligently even though it doesn't have the ability to move. Most of what we consider intelligence (11/12 attributes) is a product of mental ability rather than physical. For us humans, being able to control our bodies is obviously one of the most important things in everyday life. For an AI, as long as it demonstrates all the 'mental' abilities that a human has, the ability to move can be ignored.

Spatial Visualisation

With this skill, it is very important to realise that we are only talking about the ability to **manipulate** objects and this does not include necessarily recognising them.

In order to visualise and manipulate objects in space, first we need to have a way of passing the information about the object into the AI's 'brain'. We need it to be able to picture the object in a similar way you and I do. One way to do this is to pass the information about the object's physical dimensions and its shape into the AI's memory. We can do this using 3D Scanning.

3D Scanning is the process of analysing 3-dimensional objects and collecting data in order to construct 3D models. There are two main types of 3D Scanning: Contact Scanning, where the object being analysed needs to be in contact with the Scanner, and Non-Contact, where contact between the Scanner and the object is not necessary. (Wikipedia) For an AI to perceive objects similar to how we do (mainly with sight), it would have to use Non-Contact Scanning (NCS).

The two main methods of NCS we will talk about are Time-of-Flight NCS and Triangulation NCS. Time-of-Flight NCS works by using a laser telemeter to shine laser light at the object and measure the round-trip it took for the light to travel to a specific point on the object and back. This time is then used to calculate the distance between a particular point on the object and the telemeter. (Wikipedia) The distance is found using the following equation:

$$d = \frac{ct}{2}$$

Where ***d*** is the distance between the telemeter and the point observed, ***c*** is the speed of light and ***t*** is the time taken for the laser to return.

Once a sufficient number of points have been mapped, they are put together to construct a 3D model of the object.

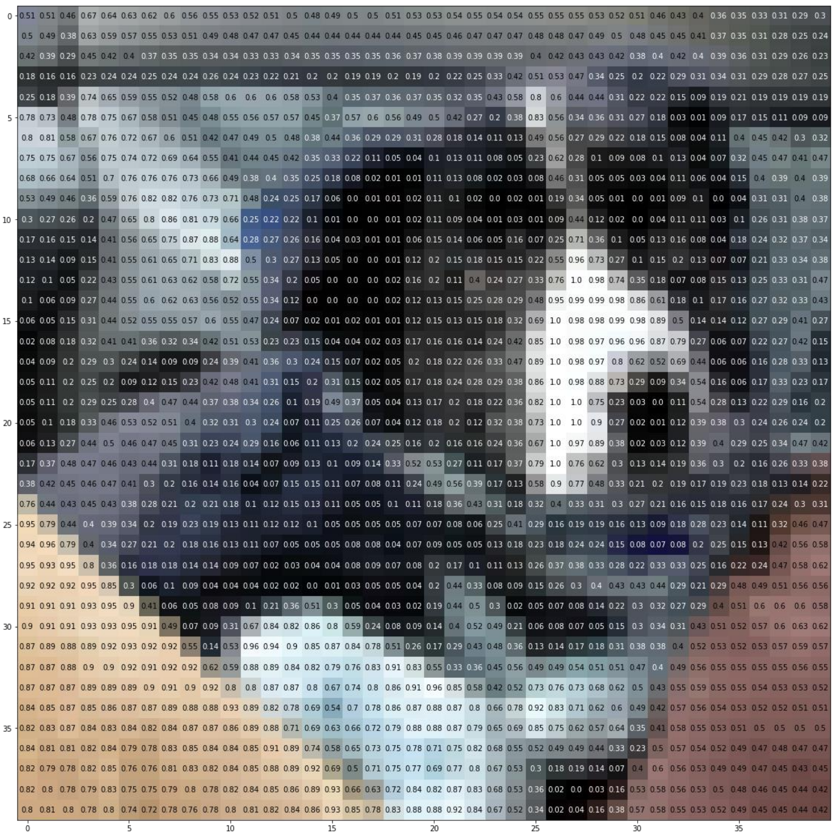
Triangulation is similar to ToF in that it too shines a laser at the object and uses information about the travel of light. In this method, the laser is shined at the object and a camera is used to capture the reflected light. Then, we look for the location of the laser dot on the image. Depending on how the laser reflects from the surface of the object, the dot will appear at different places in the camera's field of view and thus on the image. This is called 'Triangulation' because the object, camera and laser emitter form a triangle. (Wikipedia)

Looking at these two methods, it is obvious that ToF NCS is much more practical as it requires less equipment. The AI would need a built-in laser (or a different way to emit light) and some kind of device to detect the light reflected from the object (for example an LDR or Photodiode).

Once the 3D model of the object is fully captured, the data can be passed into the AI's memory and it can interpret and manipulate it.

However, if we want the AI to recognise objects and images, we need to use a different technique. A digital image is made up of pixels; each pixel has a discrete numerical representation for its grey-level (intensity). Computers process images using these values: in order to recognise an image, it has to process the patterns in the data that it is given.

An example of how a computer processes an image:



Credit: Great Learning

Recognising an image involves the use of neural networks. Neural networks in AI work similar to the neural networks found in the human brain: their job is to take the information they were given (the pixels' intensity), analyse them by comparing them to a set of pre-labelled images, and reach a conclusion about the contents of the image. The pre-labelled images are the 'reference' that the computer has to compare images to which it's trying to recognise. The more images there are in the training set, the more accurate the final conclusion will be. Comparing the images results in percentage certainties about the contents of the image: if we feed the image above to an AI which has the job of categorising pictures, it will come up with multiple possibilities as to what that picture is of. (Great Learning Team, 2022) So, it may say something similar to the following:

Dog – 87%

Car – 10%

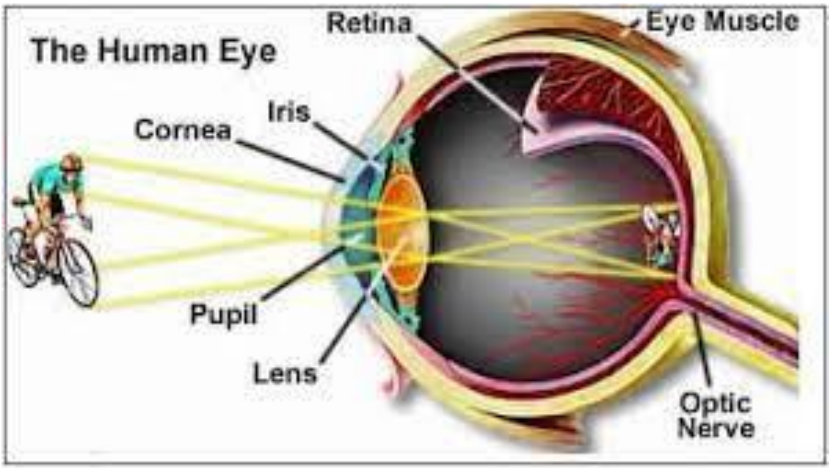
Bed – 1.2%

Microphone – 1.8%

Then, it will choose the object with the highest percentage of confidence; in this case, a dog. Furthermore, images which were correctly identified could be added to the reference set to further improve future image identification. If we want to apply this to 3D objects, all we need to do is first take a picture of the object and then repeat the process.

Perceptual Speed

The ability to perceive details at high speeds can be implemented with relative ease. The human eye can see at approximately 60 FPS. (Ann Marie Griff, O.D. and Jennifer Larson, 2020) This allows for detailed perception of images and motion, and in turn for a relatively high Perceptual Speed. The Brain processes images using the light which passes through the cornea reflected from an observed object. This light then travels until it hits the Lens in your eye which focuses it on the Retina located at the back of your eye. Photoreceptor cells in the Retina convert the light into electrical signals; Rods and Cones then pick up on the motion. Finally, the Optic nerve carries the electrical signals to your brain which converts the signals into images. (Ann Marie Griff, O.D. and Jennifer Larson, 2020)



Credit: Light Physics

Digital cameras work by replicating this process and then storing the image in the same way as presented in the previous section. In particular, High-speed cameras can be used to capture motion and images at a much higher frame rate than that of the human eye. The fastest High-speed camera in the world currently has a frame rate of about 1 trillion FPS while typical, commercially available High-speed cameras have a frame rate of about 250 FPS. (Emily Velasco, Caltech) A higher frame rate means that smaller differences in motion can be detected. These frames would then be stored in the AI's memory in order to analyse them and determine the difference between them. In fact, in order to have a high Perceptual Speed, all the AI would need is the process of recognising images as described in the previous section and a High-speed camera.

Since with Perceptual Speed you're always scanning for difference between exactly 2 pieces of stimuli, it could be calculated using the following equation:

$$S = \frac{1}{F} + T_1 + T_2$$

S (Perceptual Speed) is equal to **1/F** (the time between the two frames) plus **T1** (the time taken to analyse Frame 1) plus **T2** (the time taken to analyse Frame 2). The time between two frames is equal to 1 second divided by the Frame Rate (**F**). A more accurate way of determining the Perceptual Speed would be to consider an arbitrary number of frames and then to take the average of the total time it took to find the difference between pairs of frames. This would mean we'd have to consider the time between each pair of frames which would be equal to the time between two frames multiplied by **n-1** where **n** is the number of frames considered. For example, if we were to find the time between pairs of frames in a sample of 10 frames, we would have to use **1/F** as the time between two frames, and then multiply this by 9 (10 – 1). Then, we'd need to add to this the sum of the times it took to analyse each frame. This would give us the total time needed to analyse the sample:

$$T_t = \frac{(n-1)}{F} + T_1 + T_2 + \cdots + T_n$$

$$T_t = \frac{(n-1)}{F} + \sum_{m=1}^{n} T_m$$

The capital sigma seen in line 2 is the mathematical representation of a sum of an arbitrary number (**n**) of objects from some starting index (here, **m = 1**); in this case, the 'object' is time. The Perceptual Speed would be this total time (**Tt**) divided by the total number of pairs of frames; this is equal to **n-1**:

$$S = \frac{T_t}{n-1}$$

The reason why the number of pairs of frames is equal to **n-1** would be best explained by the following representation:

F1 p F2 p F3 p F4 p F5 p F6 p F7 p F8 p F9 p F10

I have placed a red 'p' between each two frames to represent pairs of frames. If you count the number of p's, you'll find that it is exactly one less than the number of frames (9). This will hold for any number of frames since each frame creates a brand-new pair.

Long-Term and Short-Term Memory (LTM and STM)

This is another area in which computers tower over us. A computer can hold information indefinitely while we are oftentimes known to be forgetful. When it comes to LTM, computers have a few resources available for substitution. Modern computers which are commercially available usually use HDDs (Hard Disk Drives) and/or SSDs (Solid-State Drives) to store data in the long term. (Bob Reeves, 2015)

SSDs would be more suitable for use in an AI since they're more portable than an HDD and they have no moving parts. This means that they're more durable and more reliable since there are fewer individual points of failure. HDDs use a rotating magnetic disk and a read-write head to manipulate the data. In simple terms, the RW head 'floats' over the rapidly-rotating disk; it detects data on the disk written in binary when it reads and writes in binary on the disk. When in 'write mode', an electrical current is passed through the RW head which alters the electrical field on the surface of the disk. This is interpreted as a 0 or a 1. (Bob Reeves, 2015)

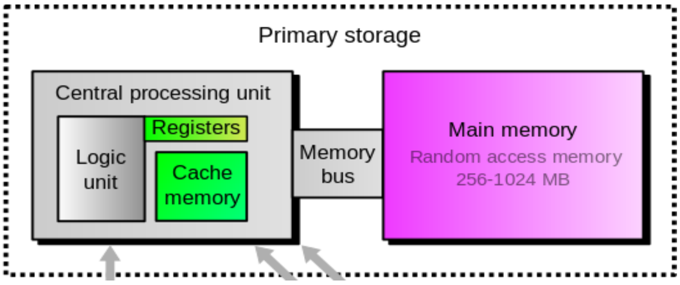
SSDs use a grid of electrical cells to store and change data. These grids are separated into sections called 'pages'. Data is ultimately stored on these pages which are grouped together in 'blocks'. (Bob Reeves, 2015)

There are also a number of solutions to STM. In computers, STM can be defined in a similar way as in humans: we can define it as the collection of data which is currently in use or which has been used with a relatively high frequency compared to other pieces of data in a short space of time. In fact, we do have a human equivalent for the first half of that definition: Working Memory (WM).

STM and WM are often used interchangeably since they both describe a type of 'temporary storage'. WM allows data to be actively manipulated. For example, when solving an arithmetic problem in your head, you are constantly changing the data and thus using your WM. (Dennis Norris, University of Cambridge, 2017)

The most common solutions to STM in computers (although they are never considered to be a solution to this exactly) are collectively called Primary Memory. Also known as Main Memory, this includes Registers and Cache in the CPU (Central Processing Unit) and RAM (Random Access Memory) which is accessed by the CPU via the Memory Bus. Primary memory is volatile, meaning that information is lost once the computer powers off. (Bob Reeves, 2015) They are placed close to the CPU to allow for fast information retrieval, simulating STM and WM.

Their job is to hold information which is currently being used by the CPU. Cache memory in particular holds information which has been frequently accessed in a short space of time and not necessarily used continuously.



Credit: Wikiwand

Inductive and Deductive Reasoning

The difference between Inductive and Deductive reasoning is very subtle. At first glance, they may even be mistaken to be the same. However, they are two separate skills. When you use inductive reasoning, you are **extracting general rules** from a specific set of information. (Louis Thurstone, 1938) However, deductive reasoning is the process of **drawing precise conclusions** from a set of information.

Inductive reasoning can be modelled as a series of 3 main steps: making an observation, recognising a pattern from that observation, and drawing a general conclusion. (Pritha Bhandari, 2022) An example would be sitting in a hospital and watching doctors go by (observation), realising that all doctors who were observed wore white coats (pattern recognition) and stating that all doctors must wear white coats (general conclusion). In order for an AI to have Inductive Reasoning skills, it would need to be able to carry out the 3 main steps on its own.

Observation: Observing a specific stimulus can be said to be the same as extensively collecting data about it. For instance, when you're observing the doctors walk by in the hospital, you're likely not only looking at their coats but mostly everything about them (the colour of their hair, their height, gender etc...). This can be applied to computers since they are excellent at analysing large quantities of data. And, this isn't exclusive to visual information; it can be applied to any format.

Pattern Recognition: This is something which we have perfected over the course of evolution. It is the reason why you sometimes find yourself discovering a face in a pillow on your bed. Pattern recognition is widely used in Computer Science; in fact, it's used in image recognition software which we have talked about in the section on Spatial Visualisation. The same process can be applied to the set of data form the Observation phase (especially when considering visual data). With this, the AI could identify points in the data which have changed persistently or which have stayed the same and determine whether or not the change is significant enough to be considered a pattern. Again, with the use of artificial neural networks (ANN) as discussed previously.

General Conclusion: Once the AI has discovered the presence of a pattern, it can highlight the variables involved in it, reaching a conclusion. This could again be achieved using ANNs and reference sets. For example, after analysing multiple pictures of the doctors, it would find that the images had a high number and concentration of white pixels. Once this pattern is discovered, the items in the reference sets which contained a similar distribution of white pixels would help determine the contents of the images and thus help reach a conclusion: doctors wear white coats.

Deductive reasoning is a few steps more complicated. It can be modelled as a series of logical steps just like Inductive reasoning, however, instead of 3 steps, it's made up of 5: recognising an existing theory, formulating a hypothesis, collecting data to prove/disprove the hypothesis, analysing the data, reaching a conclusion. (Pritha Bhandari, 2022)

The steps are similar to those in Inductive reasoning but there are a few differences. First of all, Deductive reasoning requires the presence of an already existing theory (such as "All doctors wear white coats"). Then, before collecting data (observing), a brand-new hypothesis is required ("Only some doctors wear white coats"). The last 3 steps are the exact same as in Inductive reasoning; however, we draw a specific conclusion instead of a general rule. So, the only extra skill required in Deductive reasoning is the process of creating an original hypothesis.

This requires a certain level of creativity since the AI will need to come up with something new. One of the ways to exercise creativity is to look at the opposite of what you're studying. A friend and colleague of Randall Davis (Professor of computer science at MIT) was working on a software which was to discover new avenues in the study of elementary number theory. One of the main areas the program focused on was the study of prime numbers. Unexpectedly, the program started looking at the opposite of prime numbers: numbers which have an especially high number of devisors. He was obviously extremely intrigued by this so he started researching this area and found only one other example of someone looking at what were called maximally divisible numbers: the Indian mathematician Srinivasa Ramanujan. Therefore, this could easily be categorised as a creative process carried out by the program. (Jason M. Rubin, MIT Technology Review, 2011)

So, looking at this example, it is entirely possible for a computer to 'think' creatively; if not now, sometime in the future. Therefore, it is also possible for them to possess Deductive reasoning skills.

Social Awareness and Relationship Management

Finally, we have arrived to what could be said to indicate true human intelligence: Emotional Intelligence. Broadly speaking, both Social Awareness (SA) and Relationship Management (RM) describe the ability to recognise people's behaviour and emotions and to respond to them appropriately. The response here is the easy part since once the AI detects an emotion or social cue, it can be taught how to react using the same method as described when we talked about image recognition (with the use of ANNs and reference sets). This is especially important in SA.

So, the main question becomes whether an AI is capable of recognising subtle behavioural cues in body-language and emotions. Well, if we think about it, to recognise an emotion, the AI can treat it like an image and apply the same method as with image recognition. For example, if it wants to detect how someone is feeling, it could take a picture of them and compare it to a set of reference images. A database such as this already exists (with the name: 'ImageNet') and is being used to train facial recognition software. (Alexa Hegarty, University of Cambridge and Alexandra Albert, UCL) From here, it's not a big leap to suggest that this technology could be used to **imitate** empathy by teaching the AI how to respond to different emotions. It is important to remember that we can only imitate Emotional Intelligence with our current AI and Machine Learning technology since we cannot make a computer actually understand what emotions mean. This is the same with Verbal Intelligence which we have talked about at the beginning of this part.

Now to RM. This is about being able to communicate with others through non-verbal means. This includes facial expressions and body-language.

In 2017, Researchers at Carnegie Mellon University's Robotics Institute have created an AI which is capable of understanding body poses and movements of multiple people in real time. This, for the first time ever, included hand-gestures. Since recognising the movements and positions of multiple people in a scene by simply tracking their movement has proven to be ineffective, the team came up with an inventive solution. Rather than keeping track of each individual's movements as a whole, the AI identifies body-parts first and then links them to the people in the scene. (Byron Spice, Carnegie Mellon University) This was done with the help of the Panoptic Studio which is a two-story dome equipped with 500 cameras. The data collected was then analysed using datasets of different body-parts.

On the next page is an image of the Panoptic Studio.

Credit: Carnegie Mellon University

However, this wasn't so simple for hand gestures. A camera is highly unlikely to see all parts of the hand in detail and simultaneously as they can move much more rapidly than the rest of the body and can perform finer movements. Furthermore, sufficiently large reference sets do not exist of hand images which have also been annotated with labels of parts and positions. This is where the Panoptic Studio came to be of the most use in the project. A single shot in the studio gave the team 500 shots of hand positions from different angles. Using this technique, they were able to build a large set which could then be used to analyse the images fed to the AI. (Byron Spice, Carnegie Mellon University)

When we combine these two technologies, we can see that it is entirely possible for an AI to possess SA and RM as long as we can consider the display of these abilities just as valuable as understanding the presented data.

Conclusion

Now, we can sum up everything we've covered. As we've seen for almost all of the attributes, AIs are not necessarily capable of understanding what they're doing. For this reason, they can only **simulate** intelligent behaviour. While there is a big difference between appearing as intelligent and being intelligent, if we allow ourselves to consider the fact that in essence both of these states function the same way, we can see how an AI could be considered to have similar abilities as a human.

First of all, Verbal Intelligence (Verbal Comprehension and Word Fluency). This, perhaps, is the only area where we saw that AIs don't reach quite the same level as us. Human language is incredibly complicated as it comes with a surprisingly large barrage of minor details which can influence the meaning of a sentence.

This is not to say that AIs which are close to human level linguistic ability don't exist. An example would be Sophia: a humanoid robot which can carry a fluid conversation and can also recognise sarcasm. However, as we saw, NLP systems are not perfect as they often have trouble recognising that some very similar sentences have different meanings. Such as "Pour the tea in the mug" and "Pour the mug in the tea". While it's obvious that only one of these sentences makes sense, NLP systems would have trouble realising this.

On the other hand, a computer's mathematical ability easily surpasses that of any human. I think we can safely say that you and I would have a hard time calculating 12x23x34x45x56x67 in our heads even though it is only 'simple' multiplication. However, for a computer this would be a walk in the park. This isn't only true for maths; as we've seen, computers happen to be a lot better than us in more than one area. Long and Short-Term Memory and Working Memory also operate on a higher standard in computers. This is for the simple reason that they need to: if a computer were to 'forget' something, most likely there will be at least one process sometime in the future which will be impossible to carry out.

There also seem to be some areas in which future AIs have the potential of reaching full competence given slightly improved technology. These areas are Spatial Visualisation, Inductive/Deductive Reasoning, Social Awareness and Relationship Management. With the exception of Inductive/Deductive Reasoning, the main problem is the size of the equipment required to rise to these abilities: the Panoptic Studio and most laser telemeters are obviously too large and expensive for use in an AI. However, perhaps as technology continues to bulldoze forward through the problems it faces, we'll find a way to incorporate these machines as well.

The problem then with Inductive/Deductive Reasoning is the fact that AIs would need to be able to come up with their own ideas and theories in the process of their 'thinking'. Algorithms which allow for such a skill do not exist yet, or at least they aren't as sophisticated so that they could be used in serious projects.

Lastly, Mechanical Intelligence and Perceptual Speed. Both of these areas near human-level in computers (more specifically, robots when talking about Mechanical Intelligence). However, physical movement isn't as important if we're talking about intelligent behaviour. So, while Boston Dynamics have done an incredible job with Atlas and all their other products, the place for them lies in use for help around sites which aren't safe for people to investigate in person (such as rescuing someone from burning/broken-down building).

Finally, to sum up, here's a table of the 12 attributes from the revised model with respective AI Capabilities:

Attribute	***Current AI Capability***
Word Fluency	Not fully capable
Verbal Comprehension	Not fully capable
Numerical Ability	Fully capable
Mechanical Intelligence	Fully capable (not necessary)
Spatial Visualisation	Fully capable (Needs technological improvement)
Perceptual Speed	Fully capable
Long-term Memory	Fully capable
Short-term Memory	Fully capable
Inductive Reasoning	Fully capable (Needs technological improvement)
Deductive Reasoning	Fully capable (Needs technological improvement)
Social Awareness	Fully capable (Needs technological improvement)
Relationship Management	Fully capable (Needs technological improvement)