Introduction to Intelligent Technology

John Sum
Institute of Technology Management
National Chung Hsing University
Taichung 402, Taiwan

March 21, 2025

'I can only teach what I know and I can only demonstrate what I can do', JS.

Contents

List of Figures

List of Tables

1 Intelligent (AI) Technology

Let x bet the input of a function and y is its input. Assuming that $f(x)$ is a continuous function denoted as $f(x)$. In reality, noise must be existed. In this regard, the actual output of system is given by

$$y_k = f(x_k) + n_k \tag{1}$$

where (x_k, y_k) is the k^{th} sample and n_k is the noise corrupted.

Today, intelligent technology has been applied in almost every corner in the world. Every time a photo of friends has been uploaded to Facebook, Facebook will automatically square the faces and give name suggestions. The face recognition technology is an intelligent technology. While a LINE message is edited, the iPhone will automatically give next-word suggestions (predictive text). The technology behind this is an intelligent technology. A car can now drive from one place to another without human intervention. Clearly, auto-driving is yet another intelligent technology.

Many applications have applied intelligent technologies. Sometimes, we have even been unaware of being beneficial by intelligent technologies. For instance, news feed and friends recommendation in Facebook are developed based on an intelligent technology called recommendation system. Some network security systems have applied intelligent technology for intrusion detection. Table 1 lists a number of technologies that I have been using for my living and my work.

1.1 What is Intelligent Technology ?

Intelligent technology (or intelligent technologies), it usually refers to two types of technologies. For the first type (I) *it is a collection of technologies whose algorithm*[1] *designs are inspired by or copied from the ways how human thinks and the methods how a human being solves problems.* It could be a model of neuronal network with a number of neurons. The mathematical model of each neuron is modeled by the property of a biological neuron. It could also be a hypothetical (or psychological) model mimicking human stimuli-response behavior. For the second type (II) *it is a collection of technologies for solving problems used to be solved by human beings*, like chess playing and poker game playing.

1.1.1 Type (I) Intelligent Technology

The technology applied in the fuzzy logic-based auto-parking system is an example of the first type. Fuzzy logic is theory extended from the classical logic to tackle actions involving linguistic variables like 'turn more left', 'turn more

[1]An algorithm is essentially a step-by-step procedure (i.e. a program), an operations procedure, a process or a method for solving a problem. So, an algorithm design could be interpreted as a program design, an operations design, a procedural design or a methodology design.

Table 1: Use of Technologies. Note that not all of them are intelligent.

Technology/Software	Daily Live	Daily Work
Line	Personal communication	Official communication
WhatsApp	Personal communication	–
Siri	Voice-to-Text	–
Amazon Echo	Home automation	–
NCHU Webmail	–	Official communication
Gmail	Personal communication	Assignment submission
Yahoo! Mail	Communication	–
Gate Barrier	Plate recognition	Plate recognition
Navigation System	Drive direction	Drive direction
Google Map	Location search	–
Bus Schedule App	Bus schedule	Bus schedule
Hotspot	NB Internet connection	NB Internet connection
Spell Check	–	Editing
Word Suggestion	–	Editing
Spell Correction	–	Editing
Grammar Check	–	Editing
Google Search	–	Use of English
	–	Knowledge acquisition
Google Translate	–	Use of English
	–	Abstract generation
Wikipedia	–	Use of English
	–	Knowledge acquisition
ChatGPT	–	Knowledge acquisition
	–	Paraphrasing
	–	Research
	Chat	–
Google Gemini	–	Knowledge acquisition
	–	Paraphrasing
	–	Research
	Chat	–
TeamViewer	–	Research
Computer	–	Research
Matlab	–	Research
DevC	–	Teaching
Python	–	Teaching
	–	Research
MikTex	–	Editing
WinEdt	–	Editing
Texstudio	–	Editing
IPE	–	Diagram drawing

right', 'low speed' and 'high speed'. The technology applied in the optical character recognition (OCR) system is another example of the first type. Specifically, this technology being used in the 1990s and 2000s was neural network, in which the model is inspired by the neuronal network in human brain. Nowadays, the technology being used is deep neural network (equivalently, deep learning model).

For the first type of intelligent technologies, the technologies developed must have certain intelligent essence. The problems to be solved could be combinatorial optimization problems, financial prediction problems and system control problems. These problems could be solved by methods other than intelligent technologies. Intelligent technology is just an alternative method.

1.1.2 Type (II) Intelligent Technology

Driving a car is task to be done by a human driver. Today, many car manufacturers have equipped a car with an auto-driving system. Auto-driving system is a complicated system which applies various intelligent (and non-intelligent) technologies for object recognition, speed control, road event prediction and others. These intelligent technologies are examples of the second type. Google AlphaGO is a system designed to play GO game. To win a game, it is a problem to be solved by a human player. Thus, the technology behind AlphaGO is another example of the second type. Google translator applies an algorithm called Long-Short-Term-Memory (LSTM) model to learn to translate a sentence from one language to another. The LSTM is an intelligent technology of this type.

For the second type of intelligent technologies, the technologies developed for solving the problems used to be solved by human beings could have no intelligence nature. For example, to recognize an object in an image, the object has to be segmented in the first place. The object segmentation algorithm is basically an image processing algorithm which is not intelligent at all. That is to say, the technology developed for solving a problem used to be solved by an intelligent human could be non-intelligent.

1.2 Other References for Intelligent Technology

Apart from the above references, intelligent technology could be referred to (equivalently, interpreted as and perceived as) a *product* like Sony AiBo, a *service* like Amazon ECHO, a collection of technologies like neural machine translation and deep neural networks, a machine learning algorithm. The product and service are called the *intelligent product* and *intelligent service*.

Intelligent products and intelligent services could also be integrated and applied to develop other intelligent products and intelligent services. Moreover, the technologies developed in areas of *AI*, *machine learning* and *cognitive computing* are intelligent technologies.

1.3 Classes of Models in Intelligent Technologies

1.3.1 Foundational Classes

In view of the application scope, intelligent technology could be categorized in following foundational classes.

1. Technologies for audio/image/video signal processing applications. These technologies cover signal understanding and signal synthesization.

2. Robotic technology for building human-like robot, Figure 1.

3. Technologies for game playing. Deep Blue and AlphaGO are two examples.

4. Nature inspired technologies for solving optimization algorithms. Optimization problems include scheduling problems and production planning problems. Usually, genetic algorithms and simulated annealing are notable nature inspired approaches for solving these optimization problems. However, one should note that those technologies developed along these approaches are not intelligent.

5. Technologies for natural language processing (NLP) applications[2], like OpenAI ChatGPT4 and Google Gemini. The models developed could be categorized as general purpose large language models.

6. Technologies for reasoning, like OpenAI ChatGPT o1 and DeepSeek R1. The models developed could be categorized as reasoning models.

7. Technologies for procedure design and execution. The models developed could be categorized as agentic AI models. Once a user has stated a problem, the agentic AI model will generate the sequence of steps and then execute each step to get the final result. In simple words, an agentic AI model is an automated problem solver.

8. Technologies for knowledge generation. While preliminary works have been reported in the 2010s [3, 4, 5, 6], these technologies have not yet well developed and deployed.

1.3.2 General Purpose LLMs versus Reasoning Models

The differences between a general purpose large language model and a reasoning model can be found in the following URL from Microsoft Developer Community Blog.

`https://techcommunity.microsoft.com/blog/azuredevcommunityblog/how-reasoning-models-are-transforming-logical-ai-thinking/4373194`

[2]Text processing applications.

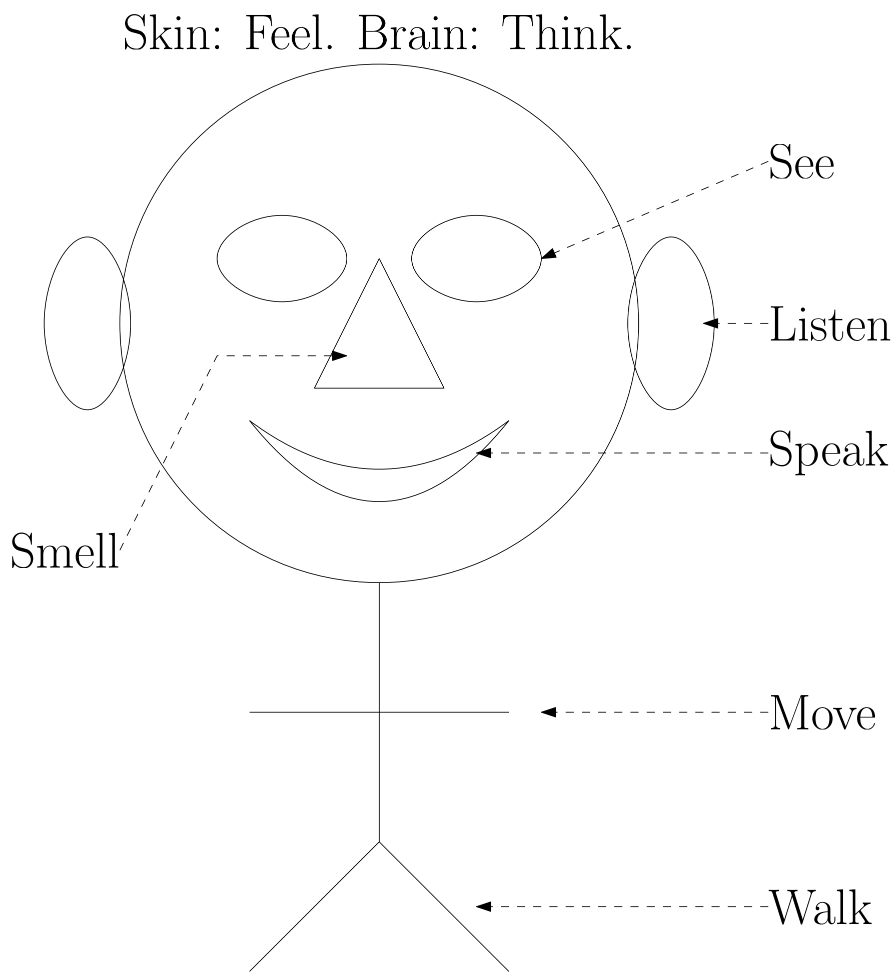


Figure 1: Towards a human-like robot. Robotic technologies are mainly applied in the mechanical systems controlling the body movements. Cameras, microphones, odor sensors and tactile sensors are the input devices for a robot to sense the environment. Loudspeaker and the locomotive mechanical systems for body movements are the output systems of a robot to react to the environment. A computer with a lot of AI programs is the brain of a robot. Building a human-like (equiv. humanoid) robot is already a difficult and huge project. Building a human-like robot with no Internet connection required is even difficult.

1.3.3 Agentic AI Model

From the usage point of view, an agentic AI model performs five tasks once it has received a request from the user.

T1. Understand the problem and identify its goals.

T2. Design the steps for solving the problem.

T3. For each step, generate computer program for it.

T4. Execute the computer programs following the sequence design in T2.

T5. Report the result to the user for validation.

Note that tasks T1 and T2 are reasoning tasks. They can readily be accomplished by a general purpose LLM and a reasoning model. Code generation is basically a task which can be accomplished by a general purpose LLM as well. Thus, the additional contribution of an agentic AI is on T4 – programs execution. Therefore, an agentic AI model has not much technological advancement as compared with a general purpose LLM together with a reasoning model.

One should further note that agentic AI model can only be applied in the problem involving data/information processing and the result is in digital form. If solving the problem requires physical materials and the result deployed is a physical product, agentic AI model can only complete the tasks delineated in T1, T2 and T3; and then give the user a suggestion on the procedure design for solving the problem.

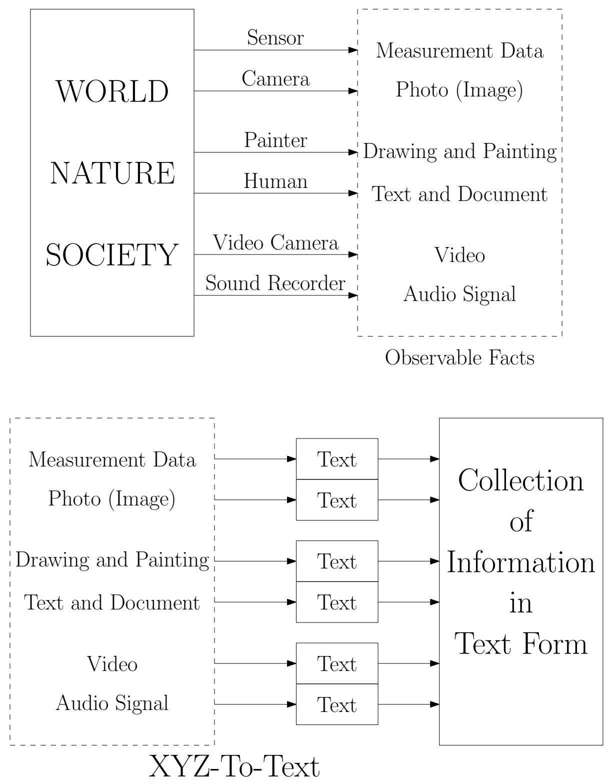
Nevertheless, the problem being solved by an agentic AI model cannot be a complex problem. For instance, compiling a survey report with presentation slides on AI is likely be automated by an agentic AI model. However, compiling an analysis report on AI technological forecast might not be accomplished by the current agentic AI models. By the same reason, compiling a master thesis or a doctoral dissertation by an agentic AI model is not possible.

1.4 Ingredients

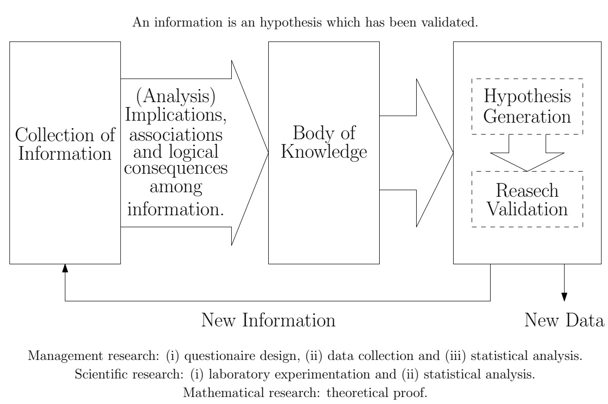
For an intelligent technology, it composes of three important ingredients: (1) the computational model of the intelligent system; (2) the learning rule governing the training of the computational model; and (3) the dataset for the training of the computational model. To assess the quality of a model after training, we need to have an assessment scheme.

1.5 Knowledge Generation

Knowledge generation is by far the most difficult problem to be solved. To generate a new knowledge, an intelligent system should be able to hypothesize and then validate if the hypothesis is correct, as shown in Figure 2, Figure 3 and Figure 4.



(a) Information extraction and collection.



(b) Knowledge generation.

Figure 2: Automatic knowledge generation.

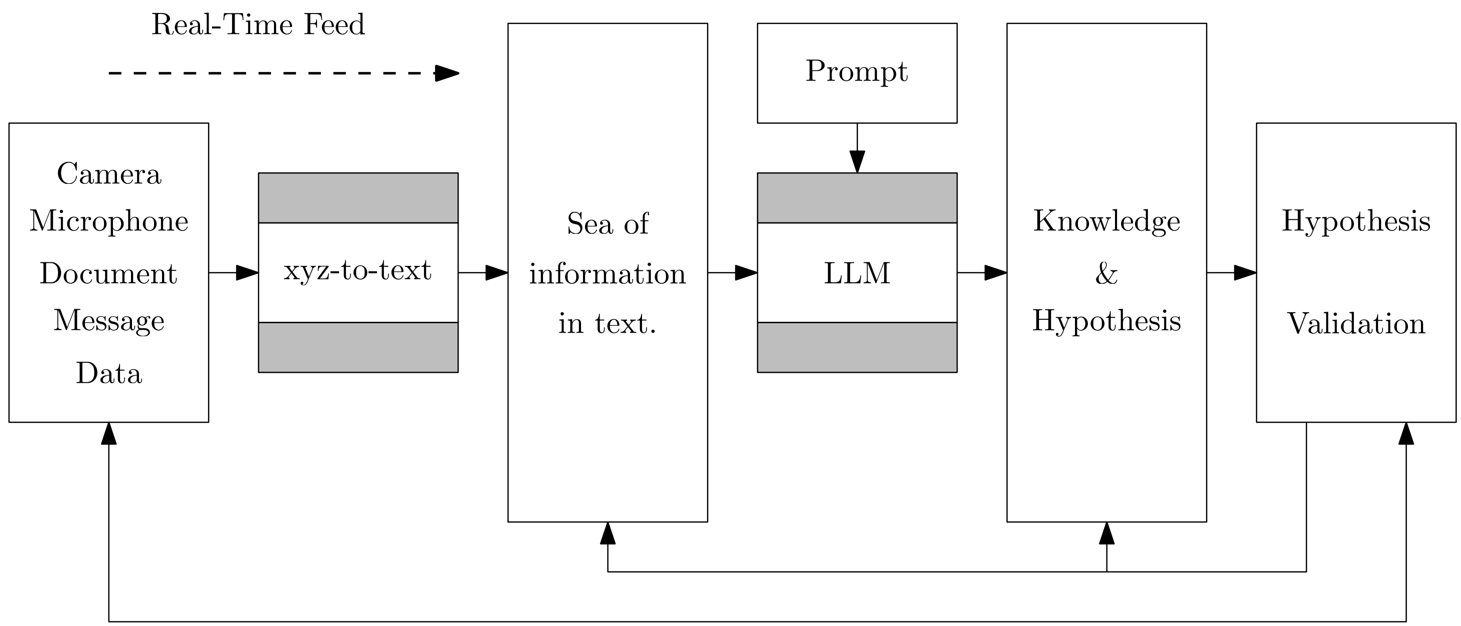


Figure 3: Human-in-a-Loop knowledge generation. Prompt entering and compiling the validation task are manually entered. Multi-modal data can be obtained real-time from all the apps connecting to the Internet. The sea of information could contain records of the activities of apps users on earth, with their time and place information.

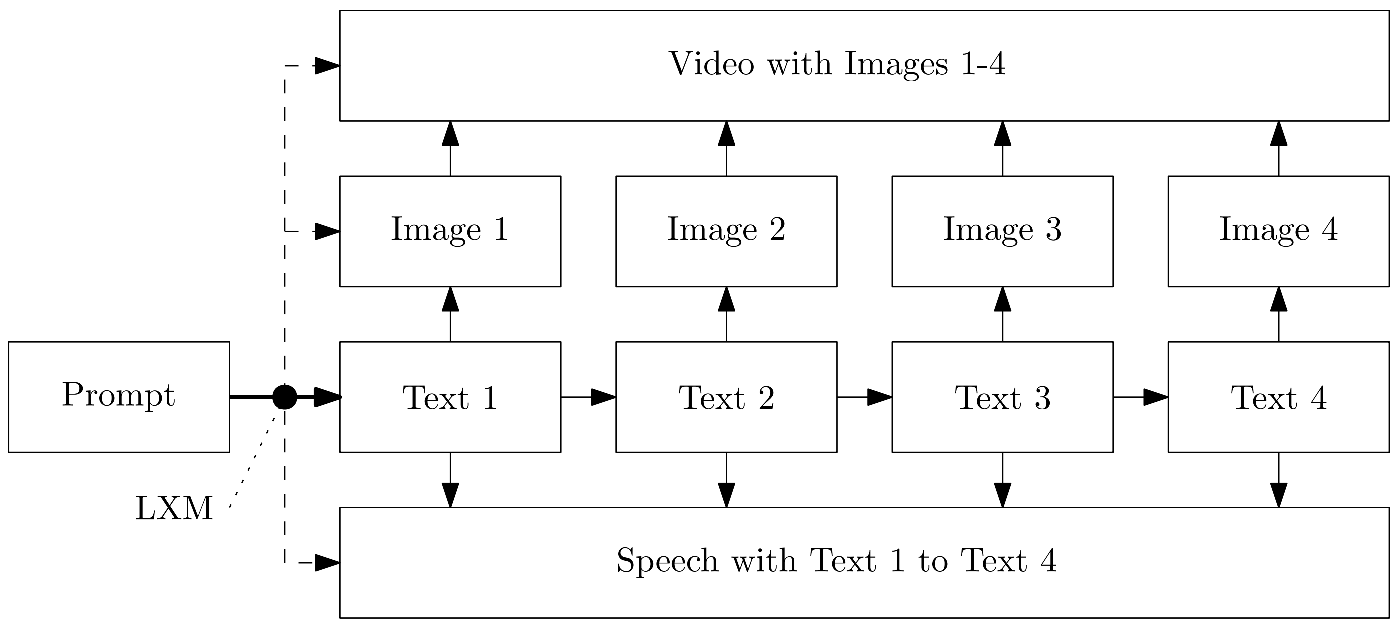


Figure 4: Multimodal sequence generation. The large model (LXM), the big black dot in between the prompt and Text 1, generates a sequence of text sentences in response to a prompt. By using text-to-image generator, a sequence of images could be generated. Follow by using a video generator with Images 1-4 as inputs, a video as shown on top embedding the information of the text sentences is hopefully generated. Similarly, a speech as shown on bottom embedding the information of the text sentences is generated. Clearly, the LXM should have embedded with *multiple semantic maps* for text sentences, images, videos and speeches. It might even be embedded with *higher order semantic maps* associated with all (resp. part) of these multimodal sources of information.

2 Intelligence & Learning

The very first concept to be clarified in intelligent technology or artificial intelligence is definitely the definition of intelligence. It is not an easy question. If you look up the information over the Internet[3], you will find dozen of definitions of intelligence which are defined by scholars from various fields, like in psychology [7]. But, how many of them you really understand. In regard to artificial intelligence, many scholars have also attempted for its definition or even its theory [8, 9] Nevertheless, as mentioned in [10], '*Despite a long history of research and debate, there is still no standard definition of intelligence. This has lead some to believe that intelligence may be approximately described, but cannot be fully defined.*'

2.1 Definition of Intelligence

In this section, I am going to define intelligence based upon my experience. For sure, you can disagree and give your own definition if you wish. First of all, one should admit that each of us is intelligent. Equivalently, human has intelligence. We have intelligence. A human being is intelligent because human can solve problem[4].

Axiom 1 (Human Intelligence) *Every person is intelligent. Every person has intelligence. Every person can solve at least one problem (i.e. survival) in his/her lifetime.*

Thus, we can have the following definition for human intelligence.

Definition 1 (Human Intelligence) *(1) Human intelligence is the capacity of a human in solving a problem. (2) Human intelligence is a noun attributed to the collective behavior of a sequence of purposeful actions a human takes in response to an external stimuli.*

Accordingly, we can have the following corollary regarding the intelligence of a living organism.

Corollary 1 (Non-Human Intelligence) *Living organism are intelligent as every living organism is able to find a way to survive.*

Now, we can give the definition of intelligence for human and living organism with an additional property for it. Intelligence must be observable or measurable.

Definition 2 (Intelligence) *(1) Intelligence is a noun attributed to the collective behavior of a sequence of purposeful actions a living organism takes in response to an external stimuli. (2) Intelligence must be observable.*

[3]`https://en.wikipedia.org/wiki/Intelligence` for instance.

[4]In the *Society of Mind* [11] which is authored by Marvin Minsky, intelligence merely means what people usually mean — the ability to solve hard problems.

A person's intelligence has to be observable from the person's reaction to the external stimuli. Therefore, a person cannot be claimed as intelligent driver if he/she claims that he/she can drive in his/her dream/imaginiation.

Regarding the problem and action as stated in Definition 1 and Definition 2, here are some examples for illustration.

Problem	Organism	Stimuli	Action(s)
Eat	Amoebae	Food	Move to the location of the food.
			Surround the food.
			Ingest the food.
Survive	Ant	Danger	Run away.
Survive	Baby	Hungry	Baby cries.
			(Mother comes.)
			(Put its head near a nipple.)
			Move its month to a nipple.
			Suck the breast milk.

Based on Definition 2, unicellular organism are intelligent. A virus is intelligent. A cancer cell is intelligent. Multicellular organisms, like plants and animals, are intelligent. Every animal is intelligent. Human is certainly intelligent. All living things are intelligent. For non-living things, like stone, I am not going to make any claim on their intelligence.

2.2 Definition of Artificial Intelligence

By the same token, a machine is intelligent if it can solve problem(s). Owing not to argue on the number and the nature of the problems, my **ground zero definition** on intelligent machine is stated below.

Definition 3 (Intelligent Machine) *A machine is intelligent if it can solve at least one problem.*

Based on the definitions for human intelligence and intelligence, machine intelligence and artificial intelligence could now be defined in similar manner.

Definition 4 (Machine Intelligence) *Machine intelligence could be referred to the capacity a machine which can solve a problem.*

Definition 5 (Artificial Intelligence) *Artificial intelligence is a noun which is attributed to the collective behavior of a sequence of purposeful actions a non-living organism takes in response to an external stimuli.*

Therefore, every intelligent machine has artificial intelligence or simply every intelligent machine has intelligence. On the other hand, every human has intelligence but no any human has artificial intelligence.

2.2.1 Note on learning

It should be noted that **learning is not, and should not be, a factor included in the definition of intelligence**. The reason could be explained by the following examples. A cancer cell is able to find a way to survive. A neuron is able to react to the signals received from its dendrites. One argument is that each cancer cell and neuron cell have been programmed to do so. Would there be any learning procedure designed in the program? It is hard to tell.

Here is another example. An auto-driving car is able to move in a city and park itself in a parking slot. The auto-driving system acts like a human driver controlling the car moving on a road and parking to a parking slot. This auto-driving system is commonly claimed as intelligent. However, the job it does is based on a program running in a computer installed inside the car. The program is able to sense the environment, recognize the objects around and then generate appropriate electrical signals to the mechanical system. Everything is programmed.

While learning is not, and should not be, a factor included in the definition of intelligence, it is clear that a person could solve more and more problems via learning.

2.2.2 Note on machine

A machine is not limited to a computer machine. It can be a mechanical computer. Like the Difference Engine[5] which was built by Charles Babbage in 1822, it was made of mechanical parts and it was applied to solve approximation problems. By Definition 3, Difference Engine is intelligent.

2.2.3 Note on problem difficulty

Here, difficulty of a problem is not included in the definitions of intelligence, machine intelligence, artificial intelligence and the definition of intelligent machines. One reason is that problem difficulty is not easily measured.

One might argue that problem difficulty could be measured by the complexity of the solution for solving such problem. The definition is in the same manner as the one defined in the area of computer algorithms[12]. While it seems reasonable, this definition has one important constraint. The complexity of a problem is defined as the complexity of the **best** algorithm for solving the problem. The best algorithm is the algorithm with the minimum number of steps for solving the problem. Thus, it comes to a question.

Let say, the best algorithm is designed by John. Peter does not know the best algorithm and designs another algorithm to solve the same problem. The algorithm requires more steps. Then, could we say that John is more intelligent than Peter? To me, I will not. I only say that both John and Peter are intelligent. So, *I will not say that a machine is intelligent if it can solve a hard problem or a machine is intelligent if it can solve hard problems.*

[5] `https://en.wikipedia.org/wiki/Difference_engine`.

2.3 Two Additional Properties of Intelligence

Two important properties have to be added to the notion of intelligence. *First, intelligence must be manifested from the interaction between a living organism and its environment.* Equivalently, it can only be attributed when a living organism has taken purposeful actions to an external stimuli. For human beings, intelligence can only be attributed only when a person has taken actions to solve a problem. No one should claim himself/herself being intelligent if he/she has not taken any action to solve any problem. Therefore, the following arguments are not valid under the above arguments.

- I can imagine how to solve a problem. So, I am intelligent.
- I can dream of controlling a fight jet. So, I am intelligent.

Imagination or dreaming of solving a problem does not actually take actions to solve the problem. Therefore, a person cannot claim himself/herself being intelligent solely based on his/her imagination or dreaming.

The second property of intelligence is that the end result of the actions taken to solve a problem must be assessable. From the end result, one can justify if the problem has been solved successfully. Let me take an example. A person has claimed that he/she has completed an assignment. The correctness (or quality) of the work presented in the assignment has to be assessed. Only when the problem has been successfully solved, the person who solved this problem is intelligent. The person has intelligence.

Let me take another example. Driving a car from my home to the school is a problem. Before starting the engine, I need to design the route to the school. Some people would consider that I am intelligent because I have designed the route. This comment is correct only if route design is the problem to be solved. However, I will say 'not yet'. It is because the actual problem is to drive a car to school. It has not yet been solved. Suppose that I have finally driven my car back to school. Only in this moment, I can claim that I am intelligent.

2.4 Turing Test

While we have intelligence quotient (IQ) to measure the intelligence capacity of a human being, there is no such measure for AI or machine intelligence. Only a test proposed by Alan Turing could be used to examine if a machine has intelligence.

Since Alan Turing has presented a number of conceptual computing machines for solving some decision and mathematical problems in 1930s [13, 14], he turned to think of using a computing machine to solve intelligent tasks [15]. In this regard, a very first question is how to assess a machine if it is intelligent. The answer is based on a test, the famous Turing Test, its procedure is presented in [15].

Suppose a machine and a human are placed in two rooms. A human tester then asks a question and writes it on two pieces of papers. The question papers

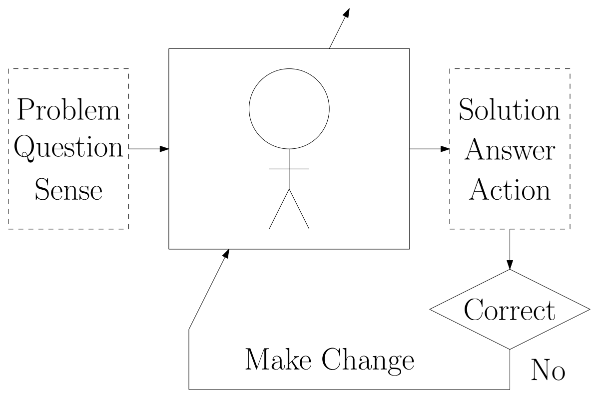


Figure 5: Schematic diagram for a human learning. Once a question is asked, a person will give an answer. If the answer is correct, there is no need to make change. If the answer is incorrect, the person will need to make some change in one's mind hoping to get correct answer in the future. The step of making change of one's mind is so-called a learning. In this example, the learning is called the error-correction learning. Reinforcement learning, as schematically shown in Figure 6, is another type of learning.

are then pass to both rooms. The machine (resp. human) responses to the question by putting the answer on a piece of paper. Finally, the human tester has to identify from their answers which room has a machine inside. If the human tester fails to identify, the machine is claimed to be intelligent.

By the Turing test, one can readily infer that Deep Blue, the machine defeated the world champion on chess game, and AlphaGo are not intelligent. It is because we can identify from the number of wins who is machine.

My definitions on intelligent machine (i.e. Definition 3), human intelligence (i.e. Definition 1) and intelligence (i.e. Definition 2) are inspired by the Turing Test. By using the same jargon, Turing Test is a bottom-line test. No matter in my definitions regarding intelligence and the Turing Test on an intelligent machine, the method how a human (resp. machine) is able to solve a problem or answer a question is not considered.

3 Informal Learning Theory

While some problem solving abilities are innate, many problem solving abilities are learned from our experience. These problem solving abilities are the end results attained after a number of learning processes. Figure 5 and Figure 6 show the schematic diagrams for two types of human learning. Once a question is asked, a person will give an answer. If the answer is correct, there is no need to make change. If the answer is incorrect, the person will need to make some change in one's mind hoping to get correct answer in the future.

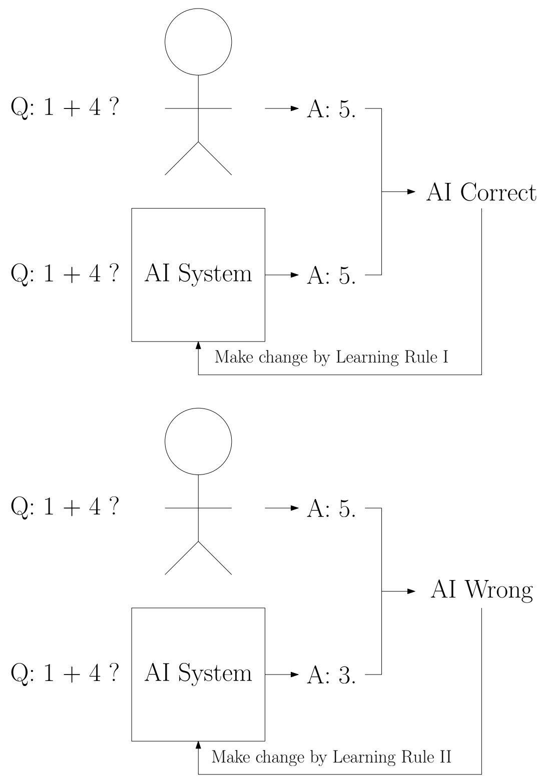


Figure 6: Reinforcement learning. **TOP:** If the AI system gives correct answer, it makes change by *Learning Rule I.* **BOTTOM:** If the AI system give incorrect answer, it makes change by *Learning Rule II.* As compared with the error-correction learning rule as shown in Figure 5, the AI system learns no matter the answer is correct or not. For the error-correction learning, the AI system only learns when the answer is incorrect.

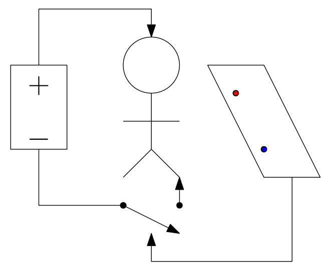


Figure 7: An exemplar reinforcement learning. A monkey is shown with either an apple or an orange. If the object is an apple, the monkey must press the red button. If the object is an orange, the monkey must press the blue button. If the monkey action is correct, it will get a food as a reward. If the action is wrong, it will get an electric shock as a penalty. In this example, the monkey is *reinforced* to learn to take the right action in accordance with the object shown.

3.1 Definition of Human Learning

No matter it is innate or problem driven, every learning process lets us acquire sufficient knowledge to solve a specific problem.

Definition 6 (Human Learning) *Learning is a process in which a person acquires sufficient knowledge to solve a problem by himself/herself. The knowledge includes (1) the factual events required and (2) the procedure for solving the problem.*

The factual events could be the data, the information and the casual relations among the events. The procedure[6] is the steps for solving the problem. If the factual events and the procedure are not available, acquisition of the factual events and the design of the procedure are yet another two problems to be solved. Therefore, we can have another two definitions for (two sub-tasks of) learning, as stated in the following definitions.

Definition 7 (Fact Acquisition) *Learning is a process in which a person acquires sufficient knowledge to acquire factual events for solving a problem by himself/herself.*

Definition 8 (Procedure Design) *Learning is a process in which a person acquires sufficient knowledge to design a procedure for solving a problem by himself/herself.*

Along with Definition 7, learning could also be defined as the process for us to get the regularities hidden in our world. The regularity could be a statistical model of certain events. For instance, the chance of getting '1' from a fair dice

[6]It can also be called the operation, the operational procedure, the solution, the method or the algorithm.

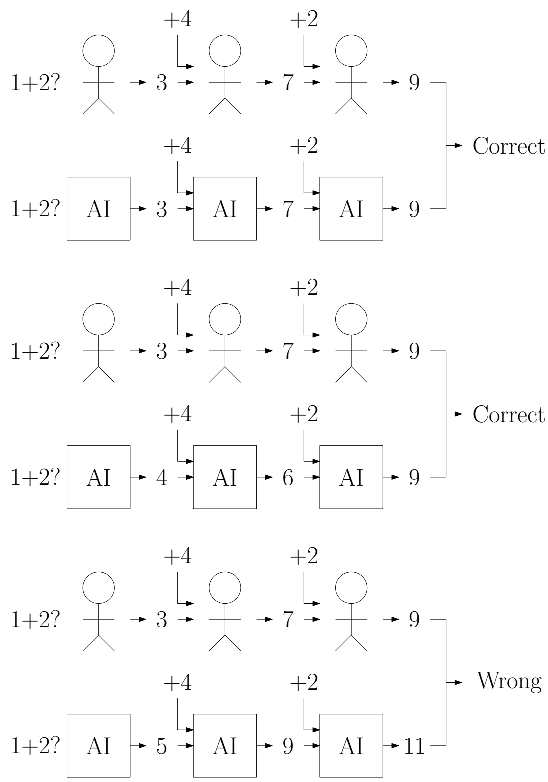


Figure 8: Reinforcement learning for solving a problem with multiple steps. For the AI system shown in the middle panel, it is clear that it is erroneous even though it gives correct answer. Therefore, simply applying *Learning Rule I* and *Learning Rule II* as shown in Figure 6 could be problematic. Specialized learning rule should be designed. One intuitive approach is design a learning rule which considers the correctness of all the intermediate steps. However, this intuitive approach is not applicable to the AI system which is developed for playing GO games. In an GO game, learning can only be conducted after the game has been over. Learning rule can only rely on the final 'WIN' or 'LOSE' result but the intermediate steps.

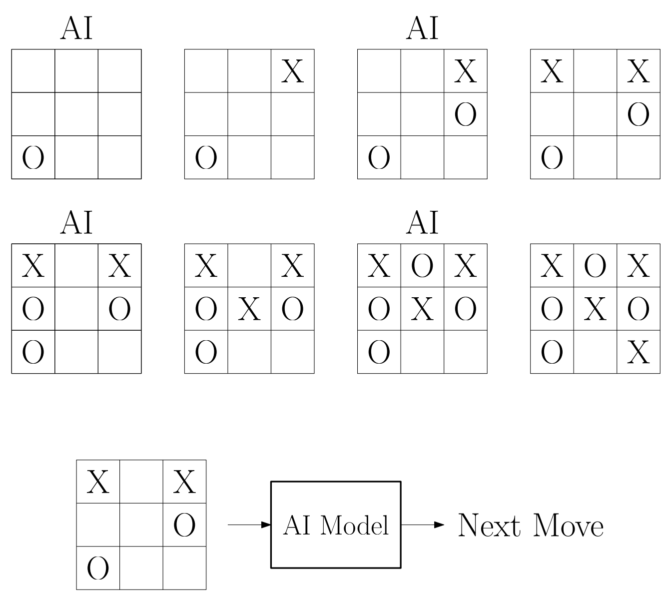


Figure 9: Reinforcement learning for solving a tic-tac-toe problem. The AI model takes a move every time after the opponent has taken a move. During the process of a game, the AI model can only determine what is the next move. The AI model has no way to learn. Only when the game is over and the result is known, the AI model makes change in accordance with the reward learning rule or the penalty learning rule. In this example, the AI model loses. So, it makes change in accordance with the penalty learning rule.

Table 2: A list of 60 stock prices in 60 consecutive trading days. What can you learn from this data? Any regularity you can find from them.

1-10	11-20	21-30	31-40	41-50	51-60
-0.1116	-1.1022	-1.8637	-2.0891	-1.6156	0.4193
-2.1471	-1.2414	-0.9226	-0.9674	-0.2519	-0.7084
-2.0689	-0.6808	-2.2141	-0.4475	-1.1924	-0.8022
-1.8095	-0.6871	-2.1135	0.1006	-0.1114	0.5877
-3.9443	-1.8649	-1.0068	0.5442	-1.7648	-1.8045
0.4384	-1.0301	0.5326	-0.9141	-2.4023	-0.3034
-0.6748	-1.1649	-1.7697	-2.4916	-2.4224	-0.1649
-1.7549	-0.3723	-0.6286	-1.7423	-0.5118	-1.2437
0.3703	0.0933	-1.2256	-2.0616	-1.1774	-0.7843
-2.7115	0.1093	0.1174	1.3505	-1.1961	-2.1658

is 1/6. It could be a casual model for a set of events. If the sky is getting darker and there are clouds in the sky, it is highly likely to have rain soon. So, Definition 7 could be re-stated in the following definition.

Definition 9 (Regularity Learning) *Learning is a process for a person to know the regularities of the world.*

The regularities being attained from learning are usually applied for solving problems. Learning process is always problem specific. We learn many things to solves many problems. Learning is process for us (or a machine) to be intelligent.

The simplest learning is to memorize all the factual events and the procedure. But, to me, it is not learning. The person who memorizes all the factual events and the procedure for him/her to solve the problem is no difference from a computer program. This person, be definition, is still intelligent. But, his/her intelligence would be questionable. Memorization should not be treated as learning. So now, what should be considered as learning?

3.2 Regularity Learning

Here, I have an idea. Let say, someone has designed (a procedure) a trading rule (based on the assumption that the stock price p of a listed company follows Gaussian distribution with stationary mean $\bar{p}$ and variance $\bar{S_p}$), *a stock can be sold when* $p \geq \bar{p}+2\times\sqrt{\bar{S_p}}$ *and the stock can be bought when* $p \geq \bar{p}-2\times\sqrt{\bar{S_p}}$. As an example, a list of 60 stock prices in 60 consecutive trading days is depicted in Table 2.

Trading the stock is now a problem to be solved. The procedure to solve the problem is very clear. But, you need to acquire the factual events, i.e. the **hidden regularity** of the change of the stock price. No way out, we need to learn from the historical stock price for the $\bar{p}$ and the $\bar{S_p}$. Clearly, it is the learning process defined in Definition 7. To do so, we need to develop a procedures for us to get $\bar{p}$ and $\bar{S_p}$. As $\bar{p}$ **and** $\bar{S_p}$ **are unknown**, the best that we can do is to estimate their values. The estimation is in fact the process of learning – learn from the historical and the future stock prices to estimate the values of $\bar{p}$ and $\bar{S_p}$.

3.2.1 Batch Mode Learning

We let $\hat{p}_N$ and $\hat{S}_N$ be the estimates of $\bar{p}$ and $\bar{S_p}$ based on N observable data $p_1, p_2, \cdots, p_N$.

$$\hat{p}_N = \frac{1}{N}\sum_{k=1}^{N} p_k, \quad \hat{S}_N = \frac{1}{N}\sum_{k=1}^{N}(p_k - \hat{p}_N)^2, \tag{2}$$

and $\bar{p} \approx \hat{p}_N$ and $\bar{S_p} \approx \hat{S}_N$. The goal of the learning process is to learn to known $\bar{p}$ and $\bar{S_p}$ as good as possible. Getting their values by using the formula in (2) is the bottom-line approach. We still can learn from it to get the values. However, it is not efficient. Especially when a new data coming in, say p_{N+1}, $\hat{p}_{N+1}$ and $\hat{S}_{N+1}$ will have to be evaluated based on the whole set of data $p_1, \cdots, p_{N+1}$. It will be time-consuming.

3.2.2 Sequential (Continuous) Learning

A philosophy behind learning is that we should not start from zero to learn something new. We should learn something new based on something already known. In this regard, another approach is to develop another learning algorithm. As it is known that $N\hat{p}_N = \sum_{k=1}^{N} p_k$ and $(N+1)\hat{p}_{N+1} = \sum_{k=1}^{N+1} p_k$, we can get that

$$(N+1)\hat{p}_{N+1} = \sum_{k=1}^{N+1} p_k = N\hat{p}_N + p_{N+1}.$$

Thus,

$$\begin{aligned}\hat{p}_{N+1} &= \frac{N}{N+1}\hat{p}_N + \frac{p_{N+1}}{N+1} \\ &= \hat{p}_N - \frac{1}{N+1}\left(\hat{p}_N - p_{N+1}\right).\end{aligned} \tag{3}$$

This is the way we learn from something new, i.e. p_{N+1}, based on something already known, i.e. $\hat{p}_N$, to get a better estimate on $\bar{p}$, a regularity of the stock price.

For the variances $\hat{S}_{N+1}$ and $\hat{S}_N$, we can go through similar step for an iteration equation as (3).

$$N\hat{S}_N = \sum_{k=1}^{N}\left(p_k - \hat{p}_N\right)^2, \quad (N+1)\hat{S}_{N+1} = \sum_{k=1}^{N+1}\left(p_k - \hat{p}_{N+1}\right)^2.$$

As

$$\begin{aligned}\sum_{k=1}^{N+1}\left(p_k-\hat{p}_{N+1}\right)^2 &= \sum_{k=1}^{N}\left(p_k-\hat{p}_{N+1}\right)^2+\left(p_{N+1}-\hat{p}_{N+1}\right)^2\\ &= \sum_{k=1}^{N}\left(p_k-\hat{p}_N+\hat{p}_N-\hat{p}_{N+1}\right)^2+\left(p_{N+1}-\hat{p}_{N+1}\right)^2\\ &= \sum_{k=1}^{N}\left(p_k-\hat{p}_N\right)^2+2\left(\hat{p}_N-\hat{p}_{N+1}\right)\sum_{k=1}^{N}\left(p_k-\hat{p}_N\right)\\ &\quad+\sum_{k=1}^{N}\left(\hat{p}_N-\hat{p}_{N+1}\right)^2+\left(p_{N+1}-\hat{p}_{N+1}\right)^2,\end{aligned}$$

we can get that

$$\hat{S}_{N+1}=\frac{\hat{S}_N+N\left(\hat{p}_N-\hat{p}_{N+1}\right)^2+\left(p_{N+1}-\hat{p}_{N+1}\right)^2}{N+1}. \tag{4}$$

The recursive equations (3) and (4) constitute the **learning algorithm** for getting the **hidden regularity of the stock price** is shown in Figure 10. This learning algorithm conforms to the learning process as stated in Definition 7. Once the model parameters have been obtained, the model can be used to generate new data, Figure 11.

In the above example, only S2.1 and S2.2 are the learning steps (or the learning algorithm). They learn from the incoming stock prices p_1, p_2 and so on for the mean and the variance of the stock prices. As the stock price is assumed to be following Gaussian distribution, the regularity of the stock prices will be known if the mean $\bar{p}$ and $\bar{S}$ are known. **In this example, learning is nothing else but simply doing estimation**.

3.2.3 Remarks

Here, three points should be noted. First, one factor leading to the success of the trading is based on the assumed model $p \sim \mathcal{N}(\bar{p}, \bar{S})$. Precisely, Gaussian distribution is a **family of models with the same mathematical definition** $\mathcal{N}(\bar{p}, \bar{S})$. $\mathcal{N}(0, 1)$ is a model. It is difference from $\mathcal{N}(1, 2)$. As $\bar{p}$ and $\bar{S}$ are the parameters of the family of Gaussian distributions, $\mathcal{N}(\bar{p}, \bar{S})$ is also called the parametric model for the family. Sometimes, an author might ignore the word 'parametric'.

Definition 10 (Parametric Model) [7] *A parametric model is referred to a family of mathematical model of the same form. A model is particular instance. It is a parametric model in which the parameters are defined to specific values.*

[7] With the parametric model estimated, one can apply the model to randomly generate unlimited number of data which fits to the distribution of the parametric model. Therefore, this parametric model is indeed a generative model.

S1 Initialize $\hat{p}_0 = 0$ and $\hat{S}_0 = 0$.

S2 Repeat the following steps whenever a new stock price p_k, for $k \geq 1$, has been input.

S2.1 Estimate the true but unknown mean $\bar{p}$ by $\hat{p}_k$, where

$$\hat{p}_k = \hat{p}_{k-1} - \frac{1}{k}\left(\hat{p}_{k-1} - p_k\right).$$

S2.2 Estimate the true but unknown variance $\bar{S}$ by $\hat{S}_k$, where

$$\hat{S}_k = \frac{\hat{S}_{k-1} + (k-1)\left(\hat{p}_{k-1} - \hat{p}_k\right)^2 + \left(p_k - \hat{p}_k\right)^2}{k}.$$

S2.3 Make the trading decision T_k, where

$$T_k = \begin{cases} \text{Buy} & \text{if } p_k \leq \hat{p}_k - 2\sqrt{\hat{S}_k}, \\ \text{Sell} & \text{if } p_k \geq \hat{p}_k + 2\sqrt{\hat{S}_k}, \\ \text{No action} & \text{Otherwise.} \end{cases}$$

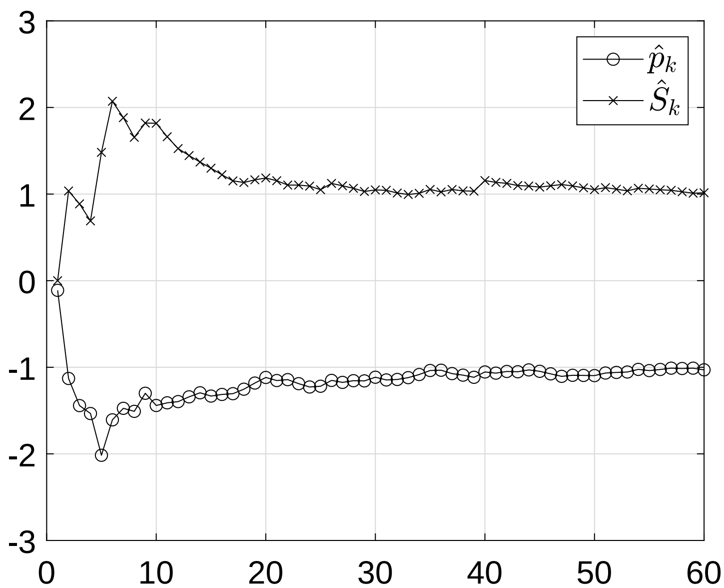


Figure 10: A continuous learning algorithm for getting the **hidden regularity of the stock price**, i.e. the **hidden generative model** with the values of $\bar{p}$ and S_p. The algorithm is developed based on (3) and (4) with an assumption that the model generating the stock prices is a Gaussian random number generator. Here, two analytical problems are aroused – (i) will $\hat{p}_k \to \bar{p}$ and (ii) $\hat{S}_k \to \bar{S}_p$ when $N \gg 1$ and $N \to \infty$.

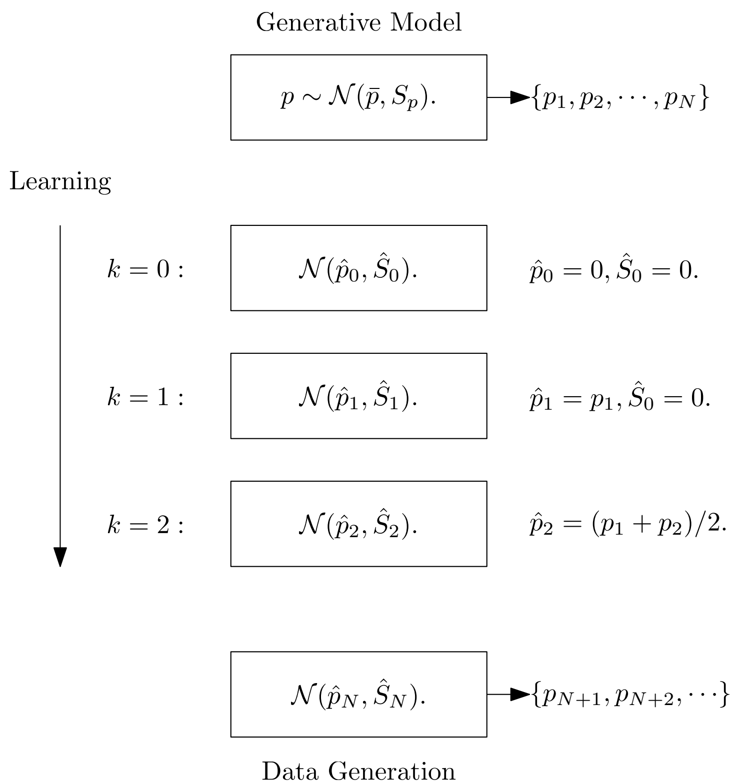


Figure 11: A generative model is able to generate new data after it has been trained. The new data is generated by $\mathcal{N}(\hat{p}_N, \hat{S}_N)$. Note that the assumption on the generative model is a key for the learning succeed.

Every learning algorithm must be developed based on an assumed parametric model. If there is something wrong with the assumed parametric model, the performance might be degraded. Therefore, for advanced learning theory, model selection is another big issue to be concerned.

Second, another factor leading to the success is the quality of learning, equivalently the quality of estimation. In other words, will $\hat{p}_k \to \bar{p}$ and $\hat{S}_k \to \bar{S}$? To answer this question, one would need to have good foundation on *Statistical Theory*. Without showing the detail proofs, it has been shown that $\hat{p}_k$ is a good estimate for $\bar{p}$. $\hat{S}_k$ is not a good estimate for $\bar{S}$.

Third, if the above algorithm is implemented by a computer program and the computer has been connected to the stock market, the computer can automatically trade the stocks for us and become an automated trading system. In accordance with the Definition 3, this computer can be claimed as an intelligent trading system even though you might not feel in that way.

3.3 Procedural Learning (Learn to Do)

In the above example, the decision rule has been designed – buy (resp. sell) the stock is $p_k \leq \bar{p} - 2\sqrt{\bar{S}}$ (resp. $p_k \leq \bar{p} + 2\sqrt{\bar{S}}$). However, there is a parameter in the decision rule. It is the value '2'. Why should it be '2', but not '1' or '3'? To answer this question, an investor will clearly tell you that this value can be changed via learning.

Recall that the purpose of learning is to let oneself to be more intelligent. However, the measure of intelligence is not definable. It is problem-specific and based on the *goal* of a person on the problem to be solved. Suppose the goal of the stock trading problem is to make the most profit. Then, we can decompose the stock trading problem into two sub-problems.

- P1 Estimate the values of $\bar{p}$ and $\bar{S}$.
- P2 Determine the factor α in the trading rule $p_k \leq \bar{p} - \alpha\sqrt{\bar{S}}$ (resp. $p_k \leq \bar{p} + \alpha\sqrt{\bar{S}}$).

For the second problem, a learning process will have to be developed so that the optimal value of α can be estimated.

3.4 Assessment

No matter which learning rule is applied, one problem is how to assess if learning has been completed. It brings out a question on the assessment scheme. In school, assessment on a student could be based on the scores on assignments, examinations, the quality of a project written report and performance in an oral presentation of the written report.

In AI, the assessment is usually not that comprehensive. For an AI system for object recognition, its assessment is simply based on its recognition rate, either by accuracy or error rate. For a large language model, its assessment will be more complicated. One simple way is to assess if given a sentence the

Table 3: Analogues between AI and human intelligence.

	AI	Human
Model	Computational models	Brain
Learning	Learning rules	??
Assessment	Accuracy	Examination score
Dataset	Training datasets	Textbooks Teachers Oneself
Cognitive stage[(a)]	Up to concrete operation	UP to formal operation

[(a)]With reference to Piaget's cognitive development stages [7].

LLM can generate the subsequent text which is identical to the text in the text database. Table 3 depicts an analogue between AI and human intelligence.

Suppose, there are two sentences in a text database : (1) *John Sum is a handsome guy and so he has many girlfriends.* (2) *John Sum is a guy and so he has a girlfriend right now.*

```
Input: John Sum is a handsome guy.
AI Output: He has many girlfriends.

Input: John Sum is a handsome guy.
AI Output: He has a girlfriend.

Input: John Sum is a handsome guy.
AI Output: He is an idiot.
```

The first output is clearly conformed to the sentence in the database. It should be assessed with the highest score. The second output should also be assessed with high score. However, the third output should be assessed with the lowest score. Now, the problem is how to define the assessment which conforms to this intuitive scoring. Its difficulty can be observed from this following response.

```
Input: John Sum is a handsome guy.
AI Output: He is a girlfriend.
```

This output is different from the second output by the word 'is' and 'has'. Based on the words in those sentences, they are similar. However, their semantic meanings are quite different. As a human, we can distinct their difference. How about the AI system?

4 Mathematical Learning Theory (*)

Learning involves the process of regularity learning and procedural learning. The theories of learning for these two different tasks have subtle difference.

Theory of regularity learning is normally referred to the so-called teacher learning. Theory of procedural learning is considered as goal-directed learning. While these two theories could be unified as a single learning theory, I am not going to do so in here. As an introductory text, I simply introduce in the following text the concepts regarding the theory of regularity learning.

Normally, an intelligent technology (but not all intelligent technologies) associated with a *parametric model* (equivalently, an hypothetical model) which generates the observations (resp. samples). As the parameters of the true model are unknown, it is inevitable to develop *an algorithm to update the parameters of a model (i.e. learning algorithm)* such that the parameters of the true model can be found eventually.

Here is a simple example. Let say, we have a set of N samples $\mathcal{D} = \{\mathbf{x}_k, y_k\}_{k=1}^N$, where $\mathbf{x}_k \in R^n$ and $y_k \in R$ for $k = 1, \cdots, N$. Assuming that this data set is generated by a linear regressor, i.e.

$$y_k = a + \mathbf{b}^T \mathbf{x}_k + \xi_k, \tag{5}$$

where ξ_k is a random noise. As a and $\mathbf{b}$ are unknown, we define the following model to learn from the samples the true parameters.

$$f(\mathbf{x}_k, \hat{a}, \hat{\mathbf{b}}) = \hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_k. \tag{6}$$

Given $\hat{a}$, $\hat{\mathbf{b}}$ and $\mathcal{D}$, the mean square error (MSE) between the parametric model (6) and the true model (5) is given by

$$\begin{aligned} E(\hat{a}, \hat{\mathbf{b}}) &= \frac{1}{N} \sum_{k=1}^{N} \left(y_k - f(\mathbf{x}_k, \hat{a}, \hat{\mathbf{b}}) \right)^2 \\ &= \frac{1}{N} \sum_{k=1}^{N} \left(y_k - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_k \right) \right)^2 . \end{aligned} \tag{7}$$

Taking derivative of (7) with respect to $\hat{a}$ and $\hat{\mathbf{b}}$, we get that

$$\frac{\partial E(\hat{a}, \hat{\mathbf{b}})}{\partial \hat{a}} = -\frac{2}{N} \sum_{k=1}^{N} \left(y_k - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_k \right) \right) \tag{8}$$

$$\frac{\partial E(\hat{a}, \hat{\mathbf{b}})}{\partial \hat{\mathbf{b}}}, = -\frac{2}{N} \sum_{k=1}^{N} \left(y_k - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_k \right) \right) \mathbf{x}_k. \tag{9}$$

Clearly, the true parameters could thus be estimated by setting the above equations to zeros. It works for N which is not large, say $N = 10^5$.

For large N, say $N = 10^{12}$, this method will not be feasible. An alternative approach is to design the search for true parameters by the following iterative equations.

$$\hat{a} \leftarrow \hat{a} + \mu \left(y_t - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_t \right) \right) \tag{10}$$

$$\hat{\mathbf{b}} \leftarrow \hat{\mathbf{b}} + \mu \left(y_t - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_t \right) \right) \mathbf{x}_t, \tag{11}$$

where μ is a small positive number called step size, $(\mathbf{x}_t, y_t)$ is a sample randomly picked from $\mathcal{D}$. The initial values of $\hat{a}$ and $\hat{\mathbf{b}}$ are small random numbers around zero. As a result, an algorithm to estimate the true parameters could be listed below.

S1 Initialize $\hat{a}$ and $\mathbf{b}$ to small random numbers around zero. Set $\mu = 0.01$.

S2 Repeat the following steps until the square error is smaller than 0.0001.

  S2.1 Pick a sample randomly from $\mathcal{D}$ and set it to be $(\mathbf{x}_t, y_t)$.

  S2.2 $\hat{a} \leftarrow \hat{a} + \mu \left(y_t - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_t \right) \right)$.

  S2.3 $\hat{\mathbf{b}} \leftarrow \hat{\mathbf{b}} + \mu \left(y_t - \left(\hat{a} + \hat{\mathbf{b}}^T \mathbf{x}_t \right) \right) \mathbf{x}_t$.

The above procedure is called the *learning algorithm* for the model (6) to learn from the data set $\mathcal{D}$ the behavior of the true model (5).

Thus, the *mathematical model*, the *cost function* to measure the fitness of the model and the *learning algorithm* are the essential components for an intelligent technology. To understand the working principle and the limitation of an intelligent technology, one would need to understand these three components. One note to add, the *cost function* is also called the *learning objective function*, *learning objective* or *objective function*. By showing its value after each round of learning, i.e. the steps S2.1, S2.2 and S2.3, one can check the progress of learning.

5 Related Issues

5.1 Learning Machines

The term *learning machine* has already appeared in 1959 in an article authored by Friedberg on specialized learning machine [16] and in 1965 by Nils J. Nilsson on a general introduction on the learning machines in that era [17]. In that period of time, a learning machine was referred to a digital computing machine like IBM 704 or a specialized designed machine like Perceptron Mark I that is able to implement a learning algorithm. Today, the term *learning machine* has rarely been linked to an intelligent technology or a hardware with intelligent technology inside.

5.2 AI and Machine Learning

Artificial intelligence and machine learning are two closely related areas (resp. collections of technologies). Roughly speaking, one can consider both area are the same. Strictly speaking, both areas have subtle difference.

AI refers to a collection of technologies (like tools, methods and systems) which are applied in solving problems that require human intelligence. Exemplar problems include chess playing and gambling.

Machine learning refers to a collection of technologies which make a computing machine learn to solve a problem. The learning rule can be arbitrary and not human oriented. The problem can be any problem.

5.3 Smart Home/City/Material

While the Chinese translation of the 'smart' in smart home, smart city and smart material is the same as the 'intelligent' in intelligent technology, one should not confuse that the actual meanings of both of them are very difference. Strictly speaking, smart homes, smart cities and smart materials are not intelligent, while intelligent technology could be applied in making part of a home (resp. city) smart. A smart home is normally referred to a home with fully automated control of the home appliances. For instance, the A/C could be set to be automatically on at 18:00 every day. All the lights in the living room will be off if the sensors sense no any conversation at home for more than 15 minutes. Amazon ECHO and Google HOME are two intelligent systems that can used for making a home smart. However, the services delivered by Amazon ECHO and Google HOME are a lot more than making a home smart.

Regarding smart material, the goal is even far different from intelligent. The ultimate goal of smart material is to synthesize new materials for special applications. The materials include the material for making lighter cloth for athletes and soldiers, the harder and stronger material for making fighting jets, the material coating on a fighting jet to make it invisible under any radar system. So, smart material is not an intelligent technology. It has nothing related to intelligence. One should be confused.

5.4 Intelligent Technology : What is It ?

Based on the definitions of *human intelligence* and *intelligence*, i.e. Definition 1 and Definition 2, intelligent technology could simply be defined as following.

Definition 11 *Any technology that can be applied to solve a problem used to be solved by a living organism is an intelligent technology.*

It is a ground zero definition. Once a technology is able to be applied to solve a problem, it is intelligent. So, almost all technologies are intelligent technologies based on the above definition.

Pretty clear, not every one accepts this ground zero definition. So, after all, would there be a better definition for intelligent technology? Here, I give one.

Definition 12 *An intelligent technology is a technology that is able to solve a problem used to be solved by human beings.*

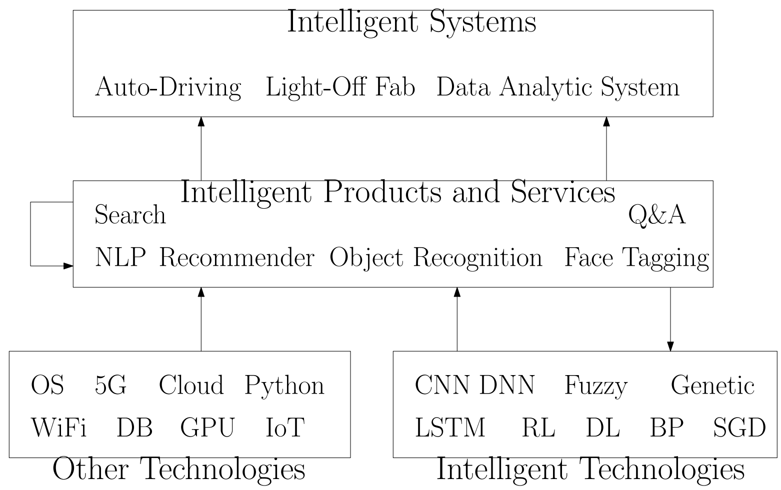


Figure 12: Intelligent technologies, intelligent products/services and intelligent systems. Genetic, RL, DL, BP and SGD inside the intelligent technologies block are different learning rules. CNN, DNN, fuzzy models are generic AI models.

This technology could have no any inspiration from human behaviors or biological neural structure. On the other hand, intelligent technology could be a technology its *model* is inspired by *human behaviors* or *biological neuronal structure*. It is a technology its *learning algorithm* is inspired by *human genetic evolution*. These intelligent technologies are applied to a wide range of problems including engineering problems, management problems and others. The relations amongst intelligent technologies, intelligent products/services and intelligent systems are shown in Figure 12. In the bottom level, there are (generic) intelligence technologies, including those models that you can find in AI/ML textbooks, and non-intelligent technologies, including computer technologies and communication technologies. These two types of technologies could thus be applied to develop intelligent products and services.

5.5 iService using both iTech and non-iTech

Let us have an example. iPhone is an intelligent system with two intelligent services FaceID and Siri. FaceID is a built-in security system to authenticate the user. It uses a face recognition software to capture the 3D face features of the user and use them as the key to unlock the iPhone. The face recognition software is developed based on intelligent technologies together with image processing technologies (non-intelligent technologies) for face recognition.

Siri is another built-in system for converting speech to text. Once the Siri is on, user could speak out a speech and then Siri will convert the speech to a text message. Imagine that your speech is a command like 'phone call Mary please'. It is clear that Siri can get this text message as well. If this message conforms

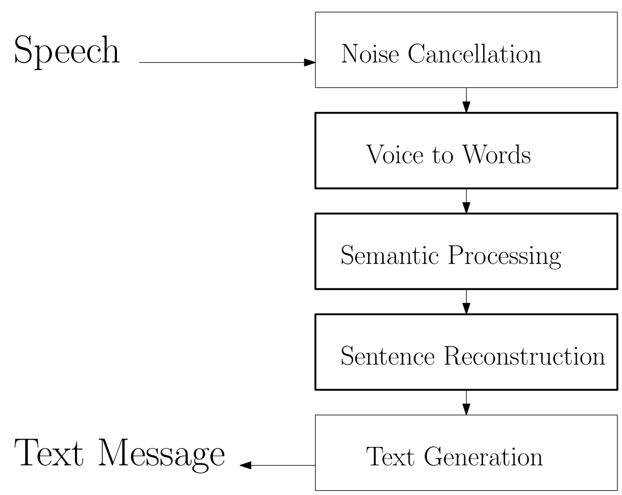


Figure 13: Schematic diagram for the speech-to-text service.

to the format of a voice command and the name Mary is listed on the phone book, the voice command module in the iPhone will act on behalf of the user to make the phone call to Mary.

Figure 13 shows the schematic diagram of the technologies behind the speech-to-text service. A user speaks a speech which is then sensed by the built-in microphone and converted to a series of electrical signal. As background noise exists in the environment, the electrical signal consists of both the speech signal and the background noise. So, the electrical signal generated by the microphone will pass to a filter for noise cancellation and get a clean speech signal. In this first step, the technology applied is not intelligent. It is a simple signal processing technique.

For the second step, the clean signal is then passed to a voice-to-word module. The work to be done in this step is complicated.

- Word segment identification – To identify which part of the signal is likely to be a word.
- Word segmentation – Find and cut the signal into segments. Each segment corresponds to an unknown word.
- Word recognition – For each segmented signal, find out the corresponding word.
- Word concatenation – Combine the words to form a sentence.

After the second step has completed, a word strings will be get. For example, the sentence could read like below.

After the second step: **I an are hand some man**.

Clearly, this sentence seems not quite correct. So, this text message will then pass to another module for further processing. In this step, the technology for word recognition is an intelligent technology.

In the third and forth step, the string of texts will then be passed for semantic processing. This step is even complicated. So, I am not going to tell the detail. The result is that the module tries to find the appropriate semantic meaning of the text and makes correction. Finally, a new string of texts.

After the forth step: **I am a handsome man**.

Here, the technologies for semantic processing and sentence reconstruction are intelligent. In technical terms, they are natural language processing (NLP) or language understanding technology.

In the final step, the reconstructed text message is sent to the APP, like LINE and WhatsAPP, for display. Clearly, the technology for this step is not an intelligent technology.

5.6 Intelligent Services and Intelligence Infrastructure

In Section 5.5, we have mentioned two intelligent services, Siri and FaceID. As a matter of fact, various tech giants have already released a number of intelligent services on their cloud platforms. Here are some examples[8].

- IBM Cloud – Watson Speech to Text, Watson Text to Speech, Watson Language Translator, Watson Visual Recognizer, IBM Watson Services for CoreML, etc.
- Amazon AWS – Amazon Lex (voice-to-text), Amazon Polly (text-to-voice), Amazon Rekonigtion for image analysis, Amazon Machine Learning, etc.
- Microsoft Azure – AI Services like Azure Cognitive Services and Azure Machine Learning; AI Tools and Framework; and AI Infrastructure, etc.
- Google Cloud – Cloud Vision API, Cloud Intelligence API, Natural Language API, Cloud Translation API, Speech-to-Text API, Text-to-Speech API, Tensor Processing Unit (TPU), Google Bard, other Cloud ML services.
- OpenAI – ChatGPT.

These cloud platforms delivering intelligence services would serve as the intelligence infrastructures for the development of higher level of intelligence application systems, as shown in Figure 14. To ensure that the intelligence services are deliverable, information infrastructure has to be accessible in 24/7 manner. So that, user device[9] is able to connect to the information infrastructure via Internet at anytime and anywhere.

[8] `https://www.eweek.com/artificial-intelligence/aiaas-companies/`.

[9] In recent years, user device is technically called an edge device. A limitation of this edge device is that its computational power is not sufficient to support all computational processes

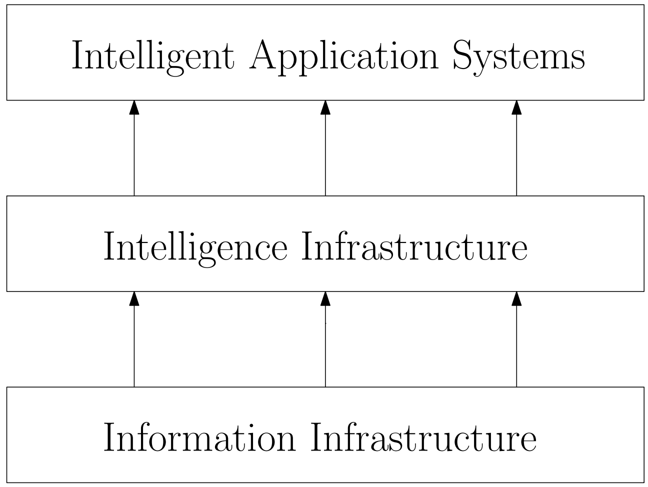


Figure 14: Intelligence infrastructure advocated by Michael I. Jordan [1].

If we consider that the intelligence infrastructure includes the (1) intelligent technologies, the (2) hardware and (3) software technologies specially developed to support the intelligent technologies, an overall picture on intelligent service delivery and intelligent service development could be shown in Figure 15. Each block on the left hand side corresponds to a collection of technologies. Each block on the right hand side corresponds to a collection of development tools for technology development. One should be noted that some development tools might have applied intelligent technologies.

Today, we have a lot more emerging technologies available in the *information infrastructure*. They include the personal area network (PAN), high speed wireless communication, Internet of Things (IoT), Internet of Vehicles (IoV), global positioning systems (GPS), mobile devices (smart phones, pads, watches and wearable devices), virtual reality (VR) headsets, augmented reality (AR) headsets like Microsoft Hololen, 5G communication technologies, cloud platforms and others.

The services delivered on top of this information infrastructure, like Google Map and Facebook, would definitely facilitate the development of intelligent services to be added to the *intelligence infrastructure*. Intelligent systems development could be even faster then ever.

in an AI application. One AI application mentioned is the Siri. An iPhone must have the Internet access and then the iPhone is able to connect to the corresponding server in the Apple Cloud and let the server to complete the NLP processes and return the result to the iPhone. In other words, the iPhone is just a terminal device connecting to a powerful server for the Siri service.

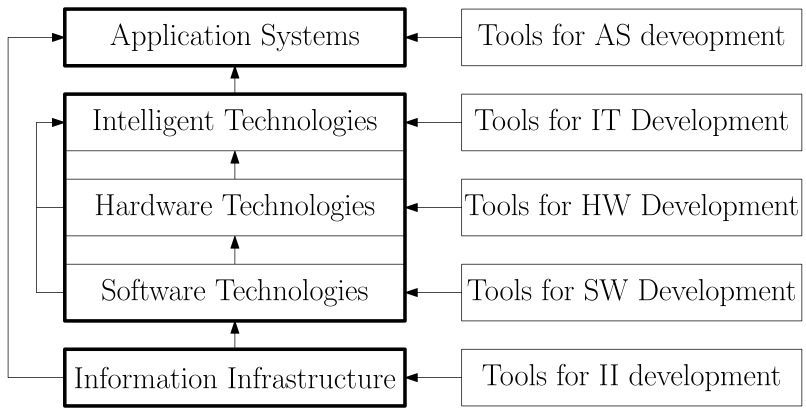


Figure 15: Intelligence infrastructure includes the (1) intelligent technologies, the (2) hardware and (3) software technologies specially developed to support the intelligent technologies. One should be noted that some development tools might have applied intelligent technologies.

5.7 Real World Applications

With the aforementioned technologies and other advanced technologies, a number of real world applications have been on the move. In a blog, Aayushi Johari has introduced 10 real world AI applications[10] in the areas of (1) marketing, (2) banking, (3) finance, (4) agriculture, (5) health care, (6) gaming, (7) space exploration, (8) autonomous vehicles, (9) chatbots and (10) artificial creativity.

Except that, AI had been successfully applied in automated mail-sorting machines in the US Postal Office. The key technology being applied is a neural network for optical character recognition which was developed by Yann LeCun and his collaborators in AT&T Bell Lab [11]. AI programs have been developed and applied in the US legal system [18].

AI programs have now been applied in scientific researches. Before 2000, AI program had been developed to read articles and then generate a summary for the articles [19]. Today, this technology has been even advanced. AI programs have been developed to read thousands of articles and generate a monograph [20]. These technologies could help a researcher to spend more time on the solution of a research rather than reading research articles. Even more, some AI programs are able to make hypotheses from the articles read [3]. Researchers could thus select from the set of hypotheses a few hypotheses for laboratory researches.

[10] https://www.edureka.co/blog/artificial-intelligence-applications/

[11] http://yann.lecun.com/ex/research/index.html.

5.8 Workflow Management

In Table 1, a number of technologies have been listed. Those technologies have been used for my daily living, teaching and research. Some technologies might have applied intelligent technologies and some of them might not have. For further explanation, let me consider the case of business adminstration. Table 4 lists a number of technologies that might facilitate the works in business administration.

Table 4: Technologies for Business Administration.

Technology/Software	Business administration
Line	Unofficial group communication
WhatsApp	Unofficial group communication
Siri	Voice-to-Text
Cell phone	Official and personal communication
Webmail (Official)	Official communication
Gmail or Yahoo!Mail	Personal communication
MS WORD spell check	Reporting & documentation
MS WORD word suggestion	Reporting & documentation
MS WORD spell correction	Reporting & documentation
MS WORD grammar check	Reporting & documentation
MS PowerPoint	Presentation slide preparation
Google Search	Use of English for documentation
Google Translate	Use of English for documentation
ChatGPT or Google Bard	Paraphrasing
Google Meet	Virtual meeting
Database management system	Data management
Information system	Adminstration workflow(*) management
	Document management
Computer	Work
Projector	Reporting
Network communication	Information infrastructure

5.8.1 Project budget request

Imagine that a team in a department would like to request for a budget for a specialized project, see Figure 16. The team will first need to file a budget request form and pass it to get approval from its department manager. Once the department manager has signed to approve the request, the form will then be passed to the financial department for budget availability. If it is available, the request form will be passed to the chief financial officer (CFO) for the final approval. The request form is then passed back to the team confirming the budget approval. A budget approval form is then passed to the financial

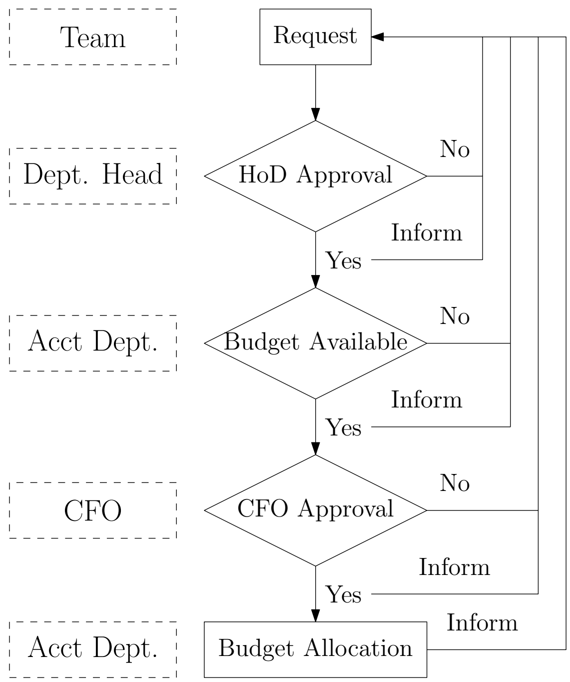


Figure 16: The workflow of a project budget request process. In each stage of decision, the corresponding party has to inform the department team for the decision so that the team is able to know the progress of its request.

department to reserve the budget for the team.

The above flow of the budget request from a department team to the CFO is so-called the workflow for budget request. Clearly, this workflow can definitely be facilitated by the use of an information system. The flow of the budget request form is made electronically and each approval step can be done online.

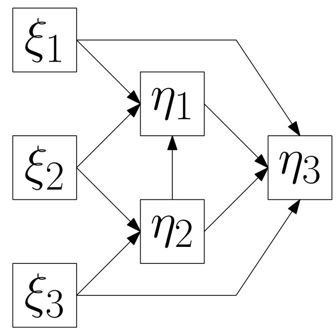
5.8.2 Business operations

Clearly, many workflow management systems have been developed for business operations like *order placement and fulfillment* and *manufacturing*. A key benefit of these systems is simply to facilitate the workers to report their job status online. The managers can thus monitor the job completion status via the workflow management systems and made appropriate decisions if necessary.

5.8.3 No Intelligent Technology

From the above explanation, it should be clear that a workflow management system should have not applied any intelligent technology. It is just like an online reporting system. Happen to be, this reporting system is able to help a

Model Complexity: 12 parameters.



$$\eta_1 = \gamma_{11}\xi_1 + \gamma_{12}\xi_2 + \beta_{12}\eta_2 + \alpha_1.$$
$$\eta_2 = \gamma_{22}\xi_2 + \gamma_{23}\xi_3 + \alpha_2.$$
$$\eta_3 = \gamma_{31}\xi_1 + \gamma_{33}\xi_3 + \beta_{31}\eta_1 + \beta_{32}\eta_2 + \alpha_3.$$

Figure 17: A conceptual (hypothetical) model for the purchasing intention of a home video game console. This model is not complex as it consists of 12 parameters only.

manager to monitor the progress of a job.

6 Model & Learning Complexities

A model complexity is referred to the amount of memory space required for the use of an AI model and the computational time required for the AI model in response to a request. Technically, we call them the memory complexity and the computational complexity. They reveal the amounts of memory resource and computational resource for the usage of an AI system. Reducing the amounts of memory and computational complexities is always a challenge to the AI system developers.

6.1 Computational Model and Learning Rule

Complexity of a computational model and the complexity of the associated learning rule are two important factor for the development of an AI system. If the model has large number of parameters and the learning rule is complex, it might take months for an AI system to learn. Nevertheless, a user might need to wait minutes for the AI system to generate results.

Today, many AI models are very complex. Thus, computational resources for using and training such models are huge. To understand the complexity of a model, an easy way is to count on its number of parameters.

Table 5: Number of parameters of deep neural networks for object recognition.

Model (Year)	No. of Parameters
LeNet[a] (1998)	60,000
AlexNet (2012)	60,000,000
GoogLeNet (2014)	7,000,000
VGG16 (2014)	138,000,000
VGG19 (2014)	144,000,000
ResNet18 (2015)	11.700,000
ResNet50 (2015)	25,600,000
ResNet101 (2015)	44,600,000

[a] LeNet is a model for hand-written character recognition `http://yann.lecun.com/index.html`.

6.2 Conceptual Model for Purchasing Intention

Figure 17 shows a conceptual model for the study of the purchasing intention of home video game consoles. Here, η_3 is the factor for the purchasing intention. ξ_1, ξ_2, ξ_3, η_1 and η_2 are the factors governing the purchasing intention for a home video game console[12].

6.3 Computational Models for Object Recognition and NLP

This conceptual model is a simple model as it consists of 12 parameters. For those systems in applications to object recognition, their numbers of parameters range from 1.2 million to 2440 million as depicted in Table 5[13]. Just on the memory space for storing the parameters, the size is already huge. Not to mention about the time spent on training such models, their time spent are in term of days and even months. Using Nivida GPUs and cloud GPUs for conducting the training processes are thus inevitable. As been aware by some scholars, there is a growing influence of the technology firms in AI research[14]. For large language models, like ChatGPT and Google Gemini, their model complexities are way more than your imagination, as depicted in Table 6.

[12] The conceptual model is extracted from the master thesis of one of my master students who investigated the factors governing the purchasing intension of a home video game console. The thesis can be downloaded from the link `john.digi-pack.io/papers/wt_thesis_09.doc`.

[13] `https://paperswithcode.com/sota/image-classification-on-imagenet`.

[14] Ahmed, N., Wahed, M., and Thompson, N. C. (2023). The growing influence of industry in AI research. *Science*, 379(6635), 884-886.

Table 6: Number of parameters in a large language model or a large reasoning model (Feb 2025).

Large Language Model	Number of Parameters	Memory Space[a]
Apple Siri (2019)[b]	133 Kilo[b]	497KB[b]
OpenAI ChatGPT 1	0.117 Billion	0.468GB[b]
OpenAI ChatGPT 2	1.5 Billion	6GB
OpenAI ChatGPT 3	175 Billion	700GB
OpenAI ChatGPT 4	1760 Billion	$\approx$ 7000GB
OpenAI ChatGPT 4o	200 Billion	800GB
OpenAI GhatGPT 4o-mini	8 Billion	32GB
Claude	52 Billion	208GB
Claude 3.5	175 Billion	700GB
Goolge Bard	137 Billion	548GB
Google Gemini Pro	500 Billion	1000GB
Google Gemini Ultra	1000 Billion	4000GB
BLOOM	176 Billion	704GB
Mistral 7B LLM	7 Billion	28GB
Meta LLaMA-7B	7 Billion	28GB
Meta LLaMA-13B	13 Billion	52GB
Meta LLaMA-33B	33 Billion	132GB
Meta LLaMA-65B	65 Billion	260GB
Meta LLaMA2-7B	7 Billion	28GB
Meta LLaMA2-13B	13 Billion	52GB
Meta LLaMA2-70B	70 Billion	280GB
Baidu Ernie	260 Billion	$\approx$ 1000GB
Tencent Hunyuan	100 Billion	400GB
BAAI WuDao 2.0	1750 Billion	7000GB
Huawei Pangu-α	$\approx$ 200 Billion	$\approx$ 400GB
Huawei Pangu Bot	0.35, 2.6 Billion	1.4GB, 10.4GB
Huawei Pangu-Σ	1085 Billion	$\approx$ 4300GB
DeepSeek V3	671 Billion	2700GB
DeepSeek R1	671 Billion	2700GB

[a] Assume that each parameter is encoded in single precision floating point format. Thus, each parameter requires four bytes memory space for storage.

[b] Zhao, S. *et al.* (2019, July). Raise to Speak: An accurate, low-power detector for activating voice assistants on smartwatches. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2736-2744).

[*] A list of large language models can be found in `https://en.wikipedia.org/wiki/List_of_large_language_models`.

7 Graphical Processing Unit (GPU)

One important hardware which accelerates the advancement of intelligent technology is the graphical processing unit (GPU). GPU is a special-designed processor used to be applying in handle intensive mathematical calculations in real-time video processing. Imagine that you are playing an on-line game, in which the background images have to be processed in real-time. As a general purpose CPU, like Intel CPU, normally takes much longer time to render the animation, the player will feel uncomfortable lag on the animation and thus quit the game. With a video card in which a GPU is installed, rendering animation would become a piece of cake. Uncomfortable lag is obsolete.

7.1 Mathematics Co-Processor

GPU is a processor designed particularly to handle specific mathematical calculations for image rendering and video processing. Its design is much simpler than a general purpose CPU which is designed to handle everything, like keyboard input, panel output, logical operations and arithmetic operations and others. Thus, the processing power of a GPU in mathematical calculations could be more than thousand time faster than a general purpose CPU.

As a GPU is particularly designed to handle mathematical calculations, it has then been applied to handle complicated and time-consuming learning algorithm. In the previous example about learning algorithm, the hypothetical model is simply a linear regressor. While the sample size N could be very large, say $N = 10^{12}$, the computational complexity (i.e number of multiplications) per step is just in the order of $\mathcal{O}(n)$, where n is the size of the parametric vector $\mathbf{b}$.

7.2 Use of GPU in AlexNet

However, for some specific models like deep neural network models, the per-step computational complexity could be in the order of $\mathcal{O}(n^3)$ and n could be larger than 10^6. In this regard, the learning process would take weeks to complete in a computer with general purpose CPU only. Hence, in the 2010s, researchers in AI/ML started to map the learning process to GPU and demonstrated that the processing time for a learning could be reduced to just a few days and even a few hours[15]. Subsequently, many research groups followed and purchased GPUs to accelerate their researches on the development and the applications of the machine learning algorithms.

7.3 A Driving Force in AI

Therefore, GPU has to be worth mentioned with intelligent technology as it is a major driving force for the advancement of intelligent technology. In the early days, Nvidia is the major GPU designer and chip maker. Today, many firms have been involved in design and/or making GPU. Intel, AMD and Apple are

[15] Search from the Google for the information about ImageNet competition and AlexNET.

three other players in the market. The A-series system-on-chip (SoC) processor by Apple has already embedded with multiple CPUs and multiple GPUs in it. Intelligent service developed from non-in-house Apple developers for such on-chip GPUs has yet to be explored.

7.4 XPU

Today (2025), many GPUs have largely been applied in many large language models development, reasoning models development and agentic AI models development. Here, one should be noted that many special-designed processing units have been developed to handle some specialized computational task for an AI model. These special-designed processing units aim to provide a faster computational power than the general purpose GPUs.

8 Intelligent Technology Management

Intelligent technology management (equivalently, management of intelligent technology) is a management process a firm has to cater. One should realize that all intelligent technology-based application systems are developed to support the business processes in a firm, Figure 18.

8.1 Definition of Intelligent Technology Management

Intelligent technology management covers at least the following three tasks.

(1) The management of the usage of intelligent technologies with access right granted for work (resp. living).

(2) The management of the development of a new intelligent technology if needed but it is not yet available.

(3) The management of the access right (resp. licensing) of the new intelligent technology developed.

An example in (1) is the use of ChatGPT to generate a survey report on a topic in *research method*. Another example is the use of ChatGPT to assist someone to learn new knowledge. In the above examples, ChatGPT is able to shorten the time for a user to complete his/her work.

In (2) and (3), the intelligent technology could be treated as a product (resp. service). Let me called it an intelligent technology product (resp. service). For instance, ChatGPT is an intelligent technology service for text generation. Self-driving car is an intelligent technology product for a driver and his/her family for transportation.

As long as an intelligent technology is a product (resp. service), a technology firm should either set a selling price for the intelligent technology product or setting a licensing fee for its access right. Revenue from these selling price or licensing fee is important for sustainable development of intelligent technology.

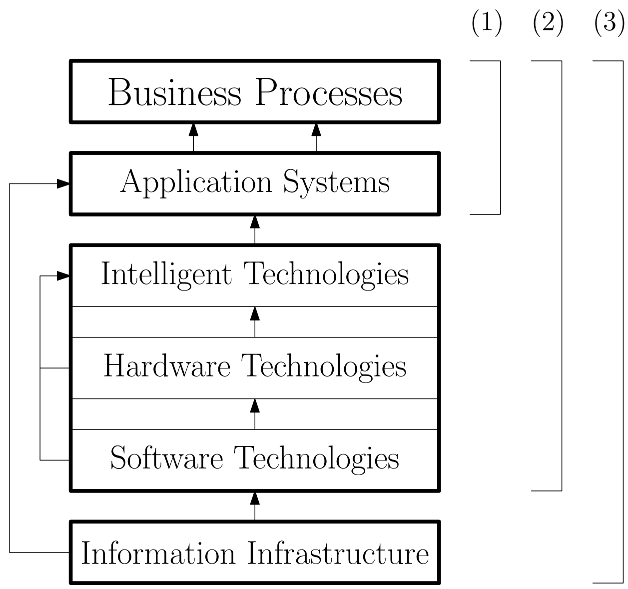


Figure 18: *Intelligence infrastructure* (including intelligent technologies, hardware and software technologies) are developed to support the development of the *application systems* to support the business processes in a firm. In view of *intelligent technology management*, three scopes have to be clearly identified. (1) ITM only concerns on the usages of the existing intelligent application systems. For (2) and (3), ITM covers the management of the (intelligent technologies) for the development of new application systems and/or new intelligent technologies for use.

8.2 Versus the Definition of Technology Management

This definition is similar to that of the definition for *technology management* and *management*. *Technology management* is a management process covering (1) the management of the usage of technologies with access right granted for work, (2) the management of the development of a new technology and (3) the management of the access right of the new technology.

8.3 Versus the Definition of Management

Management is a management process covering (1) the management of the usage of resources with access right granted for work, (2) the management of the development of a new resource and (3) the management of the access right of the new resource.

It should be noted that the new resource could be a new product or an intellectual property. Either one of them could be selling for profit. Profit is clearly a financial resource of a firm. By the same principle, human resource is yet another resource of a firm. Human resource could be trained to be more capable in solving problems for a firm.

8.4 Versus the Definition of Operation Management

In a firm, an operation could be referred to a marketing process, a production process, a project management, new product/service development, human resource management, a procurement process, a financial investment or administration.

Thus, operation management could be defined as (1) the management of the usage of the resources with access granted for the design of an operation, (2) the management of the usage of the resources with access granted for completion of an operation and (3) the management of the access right of the new operation design.

8.5 Versus the Definition of XYZ Management

From the definition of management, either intelligent technology or technology is a resource of a firm. That is to say, XYZ refers to a resource of a firm. XYZ can be referred to financial resource, information, human resource, strategy, operation design and product design.

Therefore, by the same principle as for the definition of management, financial management, information management, human resource management and strategic management could be defined in similar manner if one considers finance as a resource, human worker as a resource, information as a resource or strategy as a resource of a firm.

Replicate the definition of intelligent technology management, we can come up with the definition of XYZ management. XYZ management covers at least the following three tasks.

(1) The management of the usage of XYZ with access right granted for work (resp. living).

(2) The management of the development of a new XYZ if needed but it is not yet available.

(3) The management of the access right (resp. licensing) of the new XYZ developed or acquired.

8.6 KEY: Operation Analysis and Design

Similar to any other XYZ management, the key in the ITM is on the operations analysis and design. Only if a person has a clear picture on the steps of an operation together with the job descriptions of the people involved, he/she is able to determine in which step and what intelligent technology should be applied.

Clearly, it is not a simple problem. As the year around 1990 when computer and network technologies were emerged in the market, the uses of computer and network technologies in an enterprise became a crucial problem to many enterprises. Eventually, two new management disciplines were evolved in the 1990s – *technology management* and *business process re-engineering*.

9 Artificial General Intelligence (AGI)

Today, technology firms like Google [21] and OpenAI [2] have laid frameworks for the development of an AGI system. Figure 19 shows the framework laid by OpenAI.

9.1 General Applications of AI

One should note that their definitions and frameworks on AGI, to me, are basically not AGI. Their definitions target on the use of AI in our living and works with no any creativity and innovation. Those routine administrative works like generating a survey report, generating presentation slides, collecting data from the Internet and conducting statistical analysis are possibly be completed by an AI system. Conducting a market survey in which the data can be collected from the Internet, this type of survey report with presentation slides can then be done by the current AI systems.

9.2 Not Able to Generate Knowledge

With reference to my idea as shown in Figure 2, an AGI system should be able to generate new knowledge. The current AI systems are way lagged behind my goal. Current AI systems can best be applied in solving problems as stated by the users and then report the results. The problem stated by the user can be formulated in multimodal form which includes text, image, voice and video. The result in response can also be in multimodal form.

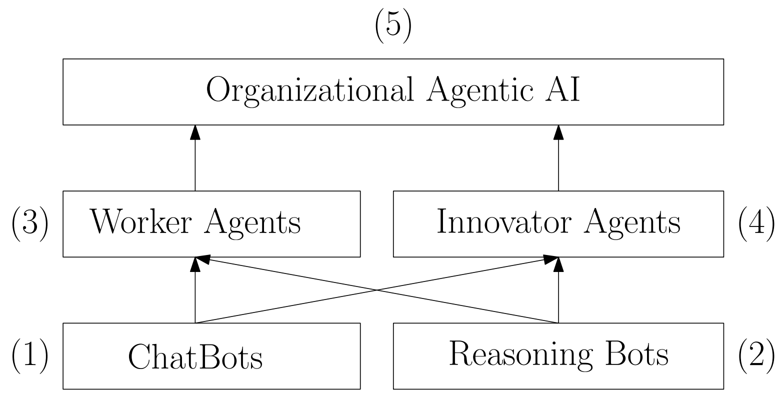


Figure 19: Framework of AGI laid by OpenAI in 2024 [2]. It is a give-stage development framework for making AI systems to be applicable to in every level in an organization. Integration of AI systems and business process re-engineering are clearly two crucial and yet difficult tasks for the success of an organizational level AGI. If it succeeds, an electronic organization can be realized to replace its physical counterpart. One should be noted that this framework is not about the level of intelligence.

9.3 No Yet Up to the Formal Operational Stage

As those AI systems are trained by existing information available, the responses of those systems are in essence based upon existing knowledge. Those systems are unable to discover any problem (resp. generate any hypothesis) and then design the procedure to solve the problem (resp. validate the hypothesis). They are yet to generate new knowledge. They have not yet achieved the formal operational stage in the Piaget cognitive development model [7].

References

[1] M. I. Jordan, "Artificial Intelligence – The Revolution Hasn't Happened Yet," *Harvard Data Science Review*, vol. 1, no. 1, July 2019. [Online]. Available: https://hdsr.mitpress.mit.edu/pub/wot7mkc1

[2] T. Duenas and D. Ruiz, "The path to superintelligence: A critical analysis of OpenAI's five levels of AI progression," *ResearchGate, 2024b. doi*, vol. 10, 2024.

[3] S. Spangler, A. D. Wilkins, B. J. Bachman, M. Nagarajan, T. Dayaram, P. Haas, S. Regenbogen, C. R. Pickering, A. Comer, J. N. Myers *et al.*, "Automated hypothesis generation based on mining scientific literature," in *Proceedings of the 20th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2014, pp. 1877–1886.

[4] Y. Gil, D. Garijo, V. Ratnakar, R. Mayani, R. Adusumilli, H. Boyce, and P. Mallick, "Automated hypothesis testing with large scientific data repositories," in *Proceedings of the Fourth Annual Conference on Advances in Cognitive Systems (ACS)*, vol. 2, 2016, p. 4.

[5] J. Sybrandt, M. Shtutman, and I. Safro, "Moliere: Automatic biomedical hypothesis generation system," in *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2017, pp. 1633–1642.

[6] S. J. Wilson, A. D. Wilkins, M. V. Holt, B. K. Choi, D. Konecki, C.-H. Lin, A. Koire, Y. Chen, S.-Y. Kim, Y. Wang *et al.*, "Automated literature mining and hypothesis generation through a network of medical subject headings," *BioRxiv*, p. 403667, 2018.

[7] J. Piaget, *The Psychology of Intelligence*. Routledge and Kegan Paul, 1950.

[8] M. Minsky, "Steps toward artificial intelligence," *Proceedings of the IRE*, vol. 49, no. 1, pp. 8–30, 1961.

[9] J. S. Albus, "Outline for a theory of intelligence," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 3, pp. 473–509, 1991.

[10] S. Legg, M. Hutter *et al.*, "A collection of definitions of intelligence," *Frontiers in Artificial Intelligence and applications*, vol. 157, p. 17, 2007.

[11] M. Minsky, *Society of Mind*. Simon and Schuster, 1988.

[12] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. W. H. Freeman and Company, 1979.

[13] A. M. Turing, "On computable numbers, with an application to the entscheidungsproblem," *Proceedings of the London Mathematical Society*, vol. s2-42, no. 1, pp. 230–265, 1937.

[14] ——, "On computable numbers, with an application to the entscheidungsproblem. A correction," *Proceedings of the London Mathematical Society*, vol. s2-43, no. 1, pp. 544–546, 1938.

[15] ——, "Computing machinery and intelligence," *Mind*, vol. 59, no. 236, pp. 433 – 460, 1950.

[16] R. M. Friedberg, "A learning machine: Part I," *IBM Journal of Research and Development*, vol. 2, no. 1, pp. 2–13, 1958.

[17] N. J. Nilsson, *Learning Machines: Foundations of Trainable Pattern-Classifying Systems.* McGraw-Hill, 1965.

[18] B. Alarie, A. Niblett, and A. H. Yoon, "How artificial intelligence will affect the practice of law," *University of Toronto Law Journal*, vol. 68, no. supplement 1, pp. 106–124, 2018.

[19] G. Salton, A. Singhal, M. Mitra, and C. Buckley, "Automatic text structuring and summarization," *Information Processing & Management*, vol. 33, no. 2, pp. 193–207, 1997.

[20] B. Writer, *Lithium-Ion Batteries : A Machine-Generated Summary of Current Research.* Springer, 2019.

[21] M. R. Morris, J. Sohl-Dickstein, N. Fiedel, T. Warkentin, A. Dafoe, A. Faust, C. Farabet, and S. Legg, "Levels of AGI: Operationalizing progress on the path to AGI," *arXiv preprint arXiv:2311.02462*, 2023.

[22] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115–133, 1943.

[23] M. Minsky and S. Papert, *Perceptrons : An introduction to computational geometry.* MIT Press, 1969.

[24] B. Widrow, "Generalization and information storage in networks of Adaline neurons," in *Self-Organizing Systems.* Spartan Books, 1962, pp. 435–461.

[25] F. Rosenblatt, "The Perceptron: a probabilistic model for information storage and organization in the brain." *Psychological Review*, vol. 65, no. 6, p. 386, 1958.

[26] ——, *Principles of Neurodynamics: Perceptions and the theory of brain mechanisms.* Spartan, 1962.

Appendix

A McCulloch-Pitts Neuron Model

McCulloch-Pitts model is the first mathematical model proposed by W. McCulloch and W. Pitts in 1943 [22]. This model abstracts the *all-or-none* property of a neuron – *If the stimulus feeding to a neuron is larger enough, the neuron fires.*

A.1 Mathematical Model

Consider a neuron with n inputs and let $x_1, \cdots, x_n$ be the inputs. $x_i \in \{0, 1\}$ for $i = 1, \cdots, n$. Let $y = f(\mathbf{x})$ be the neuron output. The output is defined as follows :

$$f(\mathbf{x}) = h\left(\sum_{i=1}^{n} w_i x_i - b\right), \tag{12}$$

where

$$h(u) = \begin{cases} 1 & \text{if } u > 0, \\ 0 & \text{if } u \leq 0. \end{cases} \tag{13}$$

A.2 Example

Figure 20 shows a model with two inputs. In the figure, w_1 and w_2 are called the synaptic weights. They act like scaling factors controlling the effects of the inputs to the neuron. b is called the threshold (or bias). If the weighted sum of the inputs is larger than the threshold b, the neuron fires (equivalently, $f(\mathbf{x}) = 1$. The $h(\cdot)$ in the neuron is a step function as defined in (13). One should be noted that the inputs are all binary variables which are non-negatives. If the value of w_1 (resp. w_2) is positive, the effect of x_1 (resp. x_2) to the neuron is *excitatory*. If the value of w_1 (resp. w_2) is negative, the effect of x_1 (resp. x_2) to the neuron is *inhibitory*.

A.3 Logical Operations

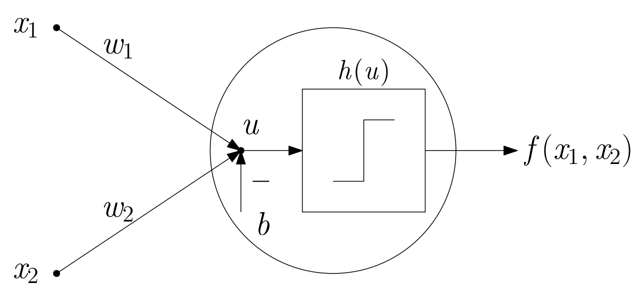
For a single two-input McCulloch-Pitts neuron with specified values for w_1, w_2 and b, one can use the neuron to perform some logical operations. Figure 21 shows the geometries of three logical operations and their truth tables.

A.3.1 $w_1 = w_2 = 1$, $b = 1.5$ (AND)

For $w_1 = w_2 = 1$, $b = 1.5$, the neuronal model is given by

$$f(x_1, x_2) = h(x_1 + x_2 - 1.5). \tag{14}$$

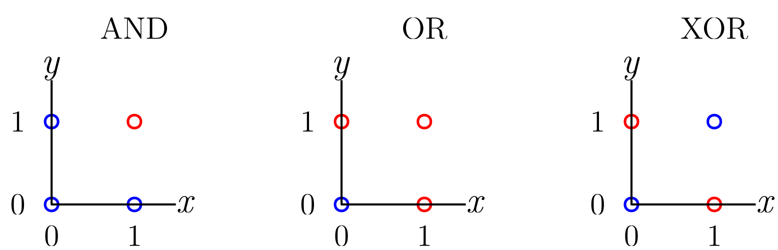
As $x_1, x_2 \in \{0, 1\}$, $f(x_1, x_2) = 1$ if and only if $x_1 = x_2 = 1$. The neuron as defined by (14) performs logical AND. It acts as an AND gate.



$$f(x_1, x_2) = h(\underbrace{w_1 x_1 + w_2 x_2 - b}_{u}).$$

$$h(u) = \begin{cases} 1 & \text{if } u > 0 \\ 0 & \text{if } u \leq 0. \end{cases}$$

Figure 20: A McCulloch-Pitts model of a neuron with two inputs. Here, w_1 and w_2 are the synaptic weights; b is called the bias and $h(\cdot)$ is a step function. If the value of w_1 (resp. w_2) is positive, the effect of x_1 (resp. x_2) to the neuron is excitatory. If the value of w_1 (resp. w_2) is negative, the effect of x_1 (resp. x_2) to the neuron is inhibitory.



AND

x	y	z
0	0	0
0	1	0
1	0	0
1	1	1

OR

x	y	z
0	0	0
0	1	1
1	0	1
1	1	1

XOR

x	y	z
0	0	0
0	1	1
1	0	1
1	1	0

Figure 21: Geometrical interpretation of three logical operations, namely logical AND (left), logical OR (middle) and XOR (right), and their truth tables.

A.3.2 $w_1 = w_2 = 1$, $b = 0.5$ (OR)

For $w_1 = w_2 = 1$, $b = 0.5$, the neuronal model is given by

$$f(x_1, x_2) = h(x_1 + x_2 - 0.5). \tag{15}$$

As $x_1, x_2 \in \{0, 1\}$, $f(x_1, x_2) = 0$ if and only if $x_1 = x_2 = 0$. The neuron as defined by (15) performs logical OR. It acts as an OR gate.

A.3.3 $w_1 = w_2 = -1$, $b = -1.5$ (NAND)

For $w_1 = w_2 = -1$, $b = -1.5$, the neuronal model is given by

$$f(x_1, x_2) = h(-x_1 - x_2 + 1.5). \tag{16}$$

As $x_1, x_2 \in \{0, 1\}$, $f(x_1, x_2) = 0$ if and only if $x_1 = x_2 = 1$. The neuron as defined by (16) performs logical NAND. It acts as an NAND gate.

A.3.4 $w_1 = w_2 = -1$, $b = -0.5$ (NOR)

For $w_1 = w_2 = -1$, $b = -0.5$, the neuronal model is given by

$$f(x_1, x_2) = h(-x_1 - x_2 + 0.5). \tag{17}$$

As $x_1, x_2 \in \{0, 1\}$, $f(x_1, x_2) = 1$ if and only if $x_1 = x_2 = 0$. The neuron as defined by (17) performs logical NOR. It acts as an NOR gate.

A.4 XOR Operation

For the above logical operations, their successes rely on proper designs of their decision boundaries given by

$$w_1 x_1 + w_2 x_2 - b = 0. \tag{18}$$

For each of the above logical operations, only one decision boundary is needed. As highlighted in [23], a single two-input McCulloch-Pitts neuron is unable to perform XOR operation. To do so, three two-input McCulloch-Pitts neurons are needed.

Figure 22 shows the network of three neurons which performs the XOR operation. Two neurons are needed in the (so-called) hidden layer. The outputs of the hidden neurons feed their output to the output neuron. The neurons in the hidden layer are defined to perform logical OR and logical NAND. Let $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ be the outputs of the hidden neurons. By (14), (15) and (16), we get that

$$\begin{aligned} f_1(x_1, x_2) &= h(x_1 + x_2 - 0.5), \\ f_2(x_1, x_2) &= h(-x_1 - x_2 + 1.5), \\ f(x_1, x_2) &= h(x_1 + x_2 - 1.5). \end{aligned}$$

That is to say, with the settings of $w_{11} = w_{12} = 1$, $b_1 = 0.5$, $w_{21} = w_{22} = -1$, $b_2 = -1.5$, $\alpha_1 = \alpha_2 = 1$, $\beta = 1.5$ for the three two-input McCulloch-Pitts neurons as shown in Figure 22, XOR can be implemented.

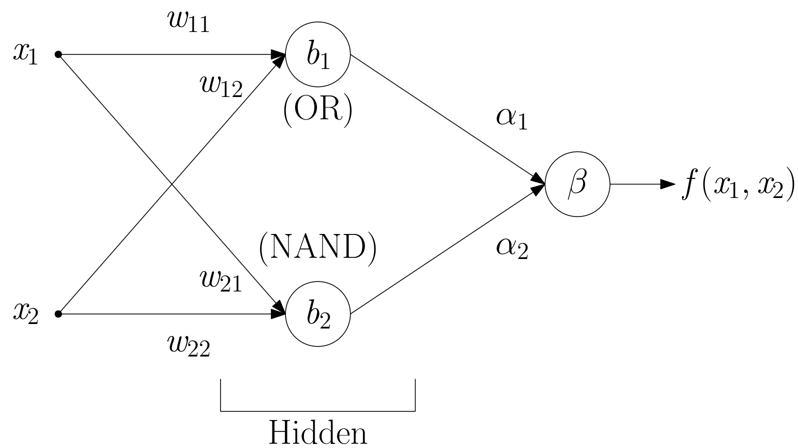


Figure 22: A network of three two-input McCulloch-Pitts neurons performs XOR operation. The neurons in the hidden layer are defined to perform logical OR and logical NAND. The parameters of the model are set to be $w_{11} = w_{12} = 1$, $b_1 = 0.5$, $w_{21} = w_{22} = -1$, $b_2 = -1.5$, $\alpha_1 = \alpha_2 = 1$, $\beta = 1.5$

A.5 Network of McCulloch-Pitts Neurons

To go beyond, one can claim that all multiple-input-multiple-output binary system can be implemented by a network of two-input McCulloch-Pitts neurons. In view of the processing in each neuron, these networks are basically computational models. Given an input $\mathbf{x}$, the network simply computes the outputs in accordance with the computations of the neurons in the network. A network of two-input McCulloch-Pitts neurons is essentially a *computational model*. Precisely, it is a *multiple-binary-input-multiple-binary-output computational model*[16].

A.5.1 Decision Network

To play Tic-Tac-Toe, one needs to block the opponent to fill up a line. If a line has already filled up with two opponent symbols, we should fill in the reminding un-filled cell with our symbol. To make this decision, Figure 23 shows a network of McCulloch-Pitts neurons for this decision making – *Should a symbol be put on the cell corresponding to x_1, x_2 or x_3?*

A.5.2 3-Input-4-Output Neuronal Network

To accomplish this, an AI model with three inputs and four outputs can be designed. In it, there are even types of neurons. Some of them are single-input-

[16]Note that this model is a special class of models. For the input (resp. output) value is not limited to binary, the model is simply called multiple-input-multiple-output (MIMO) model (equi. system).

single-output neurons (a-type and b-type). Some of them are two-input-single-output neurons (f-type and g-type) and some of them are three-input-single-output neurons (c-type, d-type and e-type). Their mathematical models are given as follows :

$$\begin{aligned}
f_a(x_i) &= h(-x_i - 0.5), && (19)\\
f_b(x_i) &= h(x_i - 0.5), && (20)\\
f_c(y_1, y_2, y_3) &= h(y_1 + y_2 + y_3 - 1.5), && (21)\\
f_d(y_1, y_2, y_3) &= h(-y_1 - y_2 - y_3 + 2.5), && (22)\\
f_e(y_1, y_2, y_3) &= h(y_1 + y_2 + y_3 - 1.5), && (23)\\
f_f(f_c, f_d) &= h(f_c + f_d - 1.5), && (24)\\
f_g(f_f, z_i) &= h(f_e - z_i - 0.5), && (25)
\end{aligned}$$

for $i = 1, 2, 3$. The outputs are defined as follows :

$$o_1 = f_g(f_f, z_1), \;\; o_2 = f_g(f_f, z_2), \;\; o_3 = f_g(f_f, z_3), \;\; o_4 = f_e(y_1, y_2, y_3). \quad (26)$$

If $o_i = 1$, fill in the cell x_i with a symbol. If $o_4 = 1$, the game is over.

While there are one-input-one-output neurons and two-input-one-output neurons in this network, we can replace them by using three-input-one-output neurons. The idea is straight forward. To implement an one-input-one-output neuron, we can set the weights of two inputs to zeros as shown in Figure 24. This idea can be extended to N-input-one-output neuron. The network as shown in Figure 23 can be implemented as a multilayered network as shown in Figure 25. With proper design on the values for the weights and biases, this network is able to (but not limited to) replicate the functionalities of the network as shown in Figure 23.

A.5.3 9-Input-10-Output Neuronal Network

Note that there are eight lines to be diagnosed – three rows, three columns and two diagonals. Decision on the next move in a tic-tac-toe game can be solved by a (bigger) network consisting of nine inputs and ten outputs. This big network is basically a consolidation of eight of the above network of McCulloch-Pitts neurons. Each network makes decision on the next possible move for a line.

A.6 Digital Computer and M-P Networks

Note that a computer is essentially constructed by a network of AND, OR, NAND, NOR and XOR logic gates to perform both logical and arithmetics operations. As a network of two-input McCulloch-Pitts neurons can perform the operations as the logic gates, a digital computer can thus be implemented by these two-input McCulloch-Pitts neurons. In this regard, a connection between computer and brain was established. *A human brain can do more than a digital computer.*

x_1	x_2	x_3

$$x_1, x_2, x_3 \in \{-1, 0, 1\}. \quad o_1, o_2, o_3, o_4 \in \{0, 1\}.$$

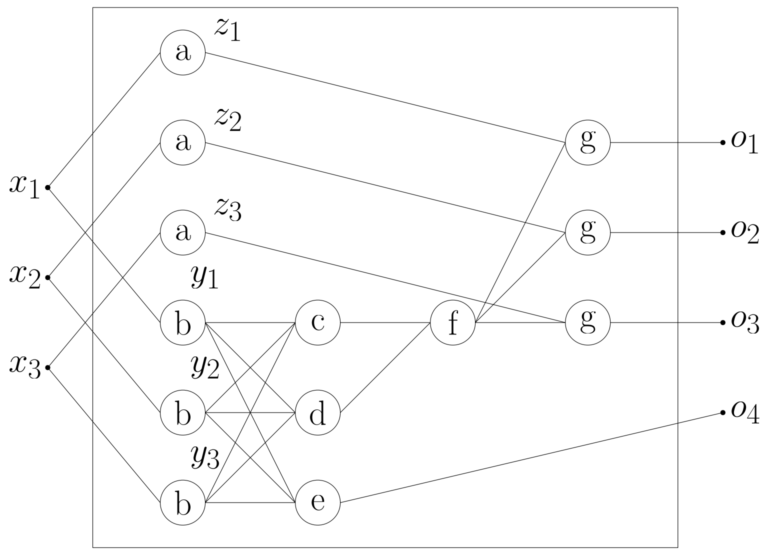
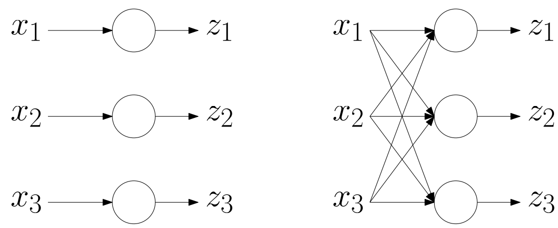


Figure 23: A network of McCulloch-Pitts Neurons could make decision for a step in a Tic-Tac-Toe game – Should a symbol be put on the cell corresponding to x_1, x_2 or x_3? In this network, there are seven different types of neurons. Some of them are single-input-single-output neurons (a-type and b-type). Some of them are two-input-single-output neurons (f-type and g-type) and some of them are three-input-single-output neurons (c-type, d-type and e-type). Here, if $o_4 = 1$, it means *game over*. Note that there are eight lines to be diagnosed – three rows, three columns and two diagonals. Decision on the next move in a tic-tac-toe game can be solved by a network consisting of eight of this network of McCulloch-Pitts Neurons.



$$
\begin{aligned}
z_1 &= h(-x_1 + 0x_2 + 0x_3 - 0.5) \\
z_2 &= h(0x_1 - x_2 + 0x_3 - 0.5) \\
z_3 &= h(0x_1 + 0x_2 - x_3 - 0.5)
\end{aligned}
$$

Figure 24: Implementation of an one-input-one-output neuron by three-input-one-output neuron. For the redundant weights, we simply set them to be zeros.

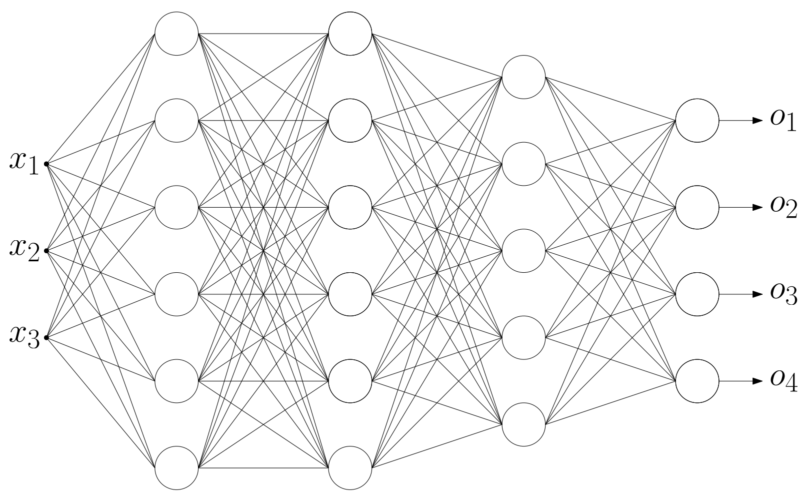


Figure 25: A 3-input-4-output multilayered network of all N-input-one-output neurons. Note that this network is also called a computational model. With proper design on the values for the weights and biases, this network is able to (but not limited to) replicate the functionalities of the network as shown in Figure 23.

A.6.1 Network Complexity

It is clear that a McCulloch-Pitts neuron can sometimes outperform a logic gate. An obvious example is on the number of inputs. A logic gate can only accept two inputs, while a McCulloch-Pitts neuron can accept more than two inputs. For the logical operation with three inputs and its output '1' if and only if all three inputs are '1', two AND logic gates are needed for this operation. Using McCulloch-Pitts neuron, we need only one. The network complexity could be reduced.

A.6.2 Beyond Digital Computations

Moreover, McCulloch-Pitts neuron accepts scalar inputs instead of binary. This neuron can be designed to solve problems with scalar inputs. Therefore, a network of McCulloch-Pitts neurons can be designed to solve 2-class classification problems – object recognition problems in which only two classes of objects are to be recognized. Along this line of thought, multiple networks of McCulloch-Pitts neurons can thus be applied to general object recognition problems with multiple classes of objects to be recognized. Furthermore, the model of McCulloch-Pitts neuron was applied in signal processing [24].

A.7 M-P Network as a Computational Model

It is no doubt that a network of McCulloch-Pitts neurons is essentially a computational model. As long as all the neuronal models have been defined, the operations of the network are defined accordingly. Each neuron simply performs a computation and gives results. The computational models developed along this line are called *Perceptrons* which are developed and advocated by Frank Rosenblatt in the 1950s to 1960s [25, 26].

In the example delineated in Figure 23, all parameters in the network are pre-defined by me. One question is then aroused. *What if the parameters are not given, is it possible to develop a learning algorithm for this model to get these parameters?* The answer is clearly YES. The learning rule associated with *Perceptrons* were later named as *Perceptron learning rule* in [23].