UNIT–I: INTRODUCTION TO ARTIFICIAL INTELLIGENCE (10 periods)

Foundations of artificial intelligence, History of artificial intelligence, State of the art, Risks and benefits of AI, Intelligent agents – Agents and environments, The concept of rationality, Structure of agents.

1.1 The Foundations of Artificial Intelligence

Philosophy

Can formal rules be used to draw valid conclusions?

How does the mental mind arise from a physical brain?

Where does knowledge come from?

How does knowledge lead to action?

Aristotle (384-322 B.C.) was the first to formulate a precise set of laws governing the rational part of the mind. He developed an informal system of syllogisms for proper reasoning, which in principle allowed one to generate conclusions mechanically, given initial premises.

Much later, **Ramon Lull** (d. 13 15) had the idea that useful reasoning could actually be carried out by a mechanical artifact – "**concept wheels**".

Thomas Hobbes (1588-1679) proposed that reasoning was like numerical computation that "we add and subtract in our silent thoughts." The automation of computation itself was already well under way.

Around 1500, **Leonardo da Vinci** (1452-1519) designed but did not build a ***mechanical calculator***; recent reconstructions have shown the design to be functional.

The first known **calculating machine** was constructed around 1623 by the German scientist **Wilhelm Schickard** (1592-1635).

The Pascaline, built in 1642 by **Blaise Pascal** (1623-1662), is more famous. Pascal wrote that "***the arithmetical machine produces effects which appear nearer to thought than all the actions of animals***."

Gottfried Wilhelm Leibniz (1646-1716) built a mechanical device intended to carry out operations on concepts rather than numbers, but its scope was rather limited.

Now that we have the idea of a set of rules that can describe the formal, rational part of the mind, the next step is to consider **the mind as a physical system**.

Rene Descartes (1596-1650) gave the first clear discussion of the distinction between mind and matter and of the problems that arise.

One problem with a purely physical conception of the mind is that it seems to leave little **room for free will**: if the mind is governed entirely by physical laws, then it has no more free will than a rock "deciding" to fall toward the center of the earth.

Problems with Cartesian dualism

Alternative to dualism is **materialism,** which holds that the brain's operation according to the laws of physics *constitutes* the mind. <u>Free will is simply the way that the perception of available choices appears to the choice process"</u>.

“Given a physical mind that manipulates knowledge, the next problem is to establish the source of knowledge. The **empiricism** movement, starting with **Francis Bacon’s** (1561-1626) *Novum* is characterized by a dictum of **John Locke** (1632-1704): “***Nothing is in the understanding, which was not first in the senses.***“

David Hume‘s (171 1-1776) A *Treatise of Human Nature* (Hume, 1739) proposed what is now known as the principle of **induction:** that general rules are acquired by exposure to repeated associations between their elements.

Building on the work of **Ludwig Wittgenstein** (1889-1951) and **Bertrand Russell** (1872-1970), the famous Vienna Circle, led by **Rudolf Carnap** (1891-1970), developed the doctrine of **logical positivism.**

This doctrine holds that all knowledge can be characterized by logical theories connected, ultimately, to **observation sentences** that correspond to sensory inputs.

The **confirmation theory** of Carnap and Carl Hempel (1905-1997) attempted to understand how knowledge can be acquired from experience.

Carnap’s book *The Logical Structure of the World* (1928) defined an explicit computational procedure for extracting knowledge from elementary experiences. It was probably the first **theory of mind as a computational process**.

The final element in the philosophical picture of the mind is **the connection between knowledge and action**. This question is vital to AI, because intelligence requires action as well as reasoning. Moreover, only by understanding how actions are justified can we understand **how to build an agent** whose actions are justifiable (or rational).

Aristotle argued that actions are justified by a logical connection between goals and knowledge of the action’s outcome. Aristotle’s algorithm was implemented 2300 years later by Newel1 and Simon in their GPS program. Goal-based analysis is useful, but does not say what do when several actions will achieve the goal, or when no action will achieve it completely.

“Mathematics

What are the formal rules to draw valid conclusions?

What can be computed?

How do we reason with uncertain information?

Philosophers staked out most of the important ideas of AI, but the leap to a formal science required a level of **mathematical formalization** in three fundamental areas: **logic, computation, and probability.**

The idea of formal logic can be traced back to the philosophers of ancient Greece, but its mathematical development really began with the work of **George Boole** (1815-1864), who worked out the details of **propositional or Boolean logic**.

In 1879, **Gottlob Frege** (1848-1925) extended Boole’s logic to include objects and relations, creating the **first-order logic** that is used today as the most basic knowledge representation system.

Alfred Tarski (1902-1983) introduced a theory of reference that shows ***how to relate the objects in a logic to objects in the real world***.

The next step was to determine the limits of what could be done with logic and computation.

The first nontrivial **algorithm** is thought to be Euclid's algorithm for computing greatest common denominators.

The study of algorithms as objects in themselves goes back to **al-Khowarazmi**, a Persian mathematician of the 9th century, whose writings also introduced Arabic numerals and algebra to Europe.

Boole and others discussed algorithms for logical deduction and by the late 19th century, efforts were under way to formalize general mathematical reasoning as logical deduction.

In 1900, **David Hilbert** (1862-1943) presented a list of **23 problems** that he correctly predicted would occupy mathematicians for the bulk of the century.

The final problem asks whether there is an algorithm for deciding the truth of any logical proposition involving the natural numbers – the famous **"decision problem"**. Essentially, Hilbert was asking whether there were fundamental limits to the power of effective proof procedures.

In 1930, **Kurt Godel** (1906-1978) showed that there exists an effective procedure to prove any true statement in the first-order logic of **Frege and Russell**, but that first-order logic could not capture the principle of mathematical induction needed to characterize the natural numbers.

In 1931, he showed that real limits do exist.

His **incompleteness theorem** showed that ***in any language expressive enough to describe the properties of the natural numbers, there are true statements that are undecidable in the sense that their truth cannot be established by any algorithm***.

This fundamental result can also be interpreted as showing that there are some functions on the integers that cannot be represented by an algorithm, that is, they cannot be computed.

This motivated **Alan Turing** (1912-1954) to try to characterize exactly which functions are capable of being computed. This notion is actually slightly problematic, because the notion of a computation or effective procedure really cannot be given a formal definition.

However, the **Church-Turing thesis**, which states that the Turing machine (Turing, 1936) is capable of computing any computable function, is generally accepted as providing a sufficient definition.

Turing also showed that there were some functions that no Turing machine can compute. For example, ***no machine can tell in general whether a given program will return an answer on a given input or run forever***.

Although **undecidability** and **noncomputability** are important to an understanding of computation, the notion of **intractability** has had a much greater impact.

A problem is called intractable if the time required to solve instances of the problem grows exponentially with the size of the instances.

Besides logic and computation, the third great contribution of mathematics to AI is the theory of **probability.** The Italian **Gerolamo Cardano** (1501-1576) first framed the idea of probability, describing it in terms of the possible outcomes of gambling events.

Probability quickly became an invaluable part of all the **quantitative sciences**, helping ***to deal with uncertain measurements and incomplete theories***.

Pierre Fermat (1601-1665), Blaise Pascal (1623-1662), James Bernoulli (1654-1705), F'ierre Laplace (1749-1827), and others advanced the theory and introduced new statistical methods.

Thomas Bayes (1702-1 761) proposed a rule for updating probabilities in the light of new evidence. Bayes' rule and the resulting field called **Bayesian analysis** form the basis of most modern approaches to uncertain reasoning in **AI** systems.

Economics (1776-present)

How should we make decisions so as to maximize payoff?

How should we do this when others may not go along?

How should we do this when the payoff may be far in the future?

The **science of economics** got its start in 1776, when Scottish philosopher **Adam Smith** (1723-1790) published *An Inquiry into the Nature and Causes of the Wealth of Nations.* While the ancient Greeks and others had made contributions to economic thought, Smith was the first to treat it as a science, using the idea that economies can be thought of as consisting of ***individual agents maximizing their own economic well-being***.

Most people think of economics as being about money, but economists will say that they are really studying ***how people make choices that lead to preferred outcomes or utility***.

Decision theory, which combines probability theory with utility theory, provides a formal and complete framework for decisions (economic or otherwise) made under uncertainty that is, in cases where probabilistic descriptions appropriately capture the decision-maker's environment.

This is suitable for **"large" economies** where each agent need pay no attention to the actions of other agents as individuals.

For **"small" economies**, the situation is much more like a game: the actions of one player can significantly affect the utility of another (either positively or negatively).

Von Neumann and Morgenstern's development of **game theory** included the surprising result that, for some games, a rational agent should act in a random fashion, or at least in a way that appears random to the adversaries.

For the most part, economists did not address the third question listed above, namely, how to make rational decisions when payoffs from actions are not immediate but instead result from several actions taken in *sequence.*

This topic was pursued in the field of **operations research**, which emerged in World War I1 from efforts in Britain to optimize radar installations, and later found civilian applications in complex management decisions.

The work of Richard Bellman (1957) formalized a class of sequential decision problems called **Markov decision processes**.

Work in economics and operations research has contributed much to our notion of rational agents.

For many years AI research developed along entirely separate paths. One reason was the apparent complexity of making rational decisions.

Herbert Simon, the pioneering AI researcher, won the **Nobel prize** in economics in 1978 for his early work showing that models based on **satisficing-making decisions** that are "good enough," rather than laboriously calculating an **optimal decision**. The earlier gives a better description of actual human behavior.

Neuroscience (1861-present)

How do brains process information?

Neuroscience is the study of the nervous system, particularly the brain. The exact way in which the brain enables thought is one of the great **mysteries of science.**

It has been appreciated for thousands of years that the brain is somehow involved in thought, because of the evidence that strong blows to the head can lead to **mental incapacitation**.

Brain is recognized as the **seat of consciousness**. Before then, candidate locations included the heart, the spleen, and the pineal gland.

Paul Broca's (1824-1880) study of ***aphasia*** (speech deficit) in brain-damaged patients in 1861 reinvigorated the field and persuaded the medical establishment of the existence of localized areas of the brain responsible for **specific cognitive functions**.

Despite these advances, we are still a long way from understanding how any of these cognitive processes actually work.

The truly amazing conclusion is that ***a collection of simple cells can lead to thought, action, and consciousness*** or, in other words, that ***brains cause minds*** (Searle, 1992).

The only real alternative theory is mysticism: that there is some mystical realm in which minds operate that is beyond physical science.

Brains and computers perform quite different tasks and have different properties.

Moore's Law predicts that the CPU's gate count will equal the brain's neuron count around 2020. Moore's Law says that the number of transistors per square inch doubles every 1 to 1.5 years. Human brain capacity doubles roughly every 2 to 4 million years.

E*ven though a computer is a million times faster in raw switching speed, the brain ends up being 100,000 times faster at what it does.*

Psychology (1879-present)

How do humans and animals think and act?

The origins of **scientific psychology** are usually traced to the work of the German physicist Hermann von **Helmholtz** (1821-1894) and his student Wilhelm **Wundt** (1 832-1920).

Helmholtz applied the scientific method to the study of human vision, and his *Handbook of Physiological Optics* is even now described as ***"the single most important treatise on the physics and physiology of human vision"*** (Nalwa, 1993, p.15).

In 1879, Wundt opened the first laboratory of **experimental psychology** at the University of Leipzig. Wundt insisted on carefully controlled experiments in which his workers would perform a perceptual or associative task while introspecting on their thought processes.

Biologists studying **animal behavior** on the other hand, lacked **introspective data** and developed an **objective methodology**.

Applying this viewpoint to humans, the **behaviorism** movement, led by **John Watson** (1878-1958), rejected *any* theory involving mental processes on the grounds that ***introspection could not provide reliable evidence***.

Behaviorists emphasize on studying only objective measures of the percepts given to an animal and its resulting actions (or *response)*.

The so-called mental constructs such as knowledge, beliefs, goals, and reasoning steps were dismissed as unscientific "**folk psychology**."

Behaviorism discovered a lot about rats and pigeons, but had less success at understanding humans.

The view of the **brain as an information-processing device**, which is a principal characteristic of **cognitive psychology,** can be traced back at least to the works of William James.

Cambridge's Applied Psychology Unit reestablished the **legitimacy of "mental" terms** as beliefs and goals, arguing that they are just as scientific as, say, using pressure and temperature to talk about gases, despite their being made of molecules that have neither.

Craik specified the **three key steps of a knowledge-based agent**:

1. the stimulus must be translated into an internal representation,
2. the representation is manipulated by cognitive processes to derive new internal representations, and
3. these are in turn retranslated back into action.

He clearly explained why this was a good design for an agent:

If the organism carries a "small-scale model" of external reality and of its own possible actions within its head, it is able to try out various alternatives, conclude which is the best of them, react to future situations before they arise, utilize the knowledge of past events in dealing with the present and future, and in every way to react in a much fuller, safer, and more competent manner to the emergencies which face it. (Craik, 1943)

Meanwhile, in the United States, the development of computer modeling led to the creation of the field of **cognitive science.** The field can be said to have started at a workshop in September 1956 at MIT. This is just two months after the confererence at which AI itself was "born."

At the workshop, George **Miller** presented *The Magic Number Seven,* Noam **Chomsky** presented *Three Models of Language,* and Allen **Newel1** and Herbert **Simon** presented *The Logic Theory Machine.*

These three influential papers showed <u>how computer models could be used to address the psychology of memory, language, and logical thinking</u>, respectively.

It is now a common view among psychologists that "***a cognitive theory should be like a computer program***" that is, it should describe a detailed information-processing mechanism whereby some cognitive function might be implemented.

<u>**Linguistics** (1957-present)</u>

How does language relate to thought?

In 1957, B. F. **Skinner** published ***Verbal Behavior***. This was a comprehensive, detailed account of the behaviorist approach to language learning, written by the foremost expert in the field.

A review of the book became as well known as the book itself, and served to almost kill off interest in behaviorism. The author of the review was **Noam Chomsky**, who had just published a book on his own theory, ***Syntactic Structures***.

Chomsky showed how the behaviorist theory did not address **the notion of creativity in language**. It did not explain how a child could understand and make up sentences that he or she had never heard before.

Chomsky's theory, based on **syntactic models** going back to the **Indian linguist Panini** (b.c.350), could explain this, and unlike previous theories, it was formal enough that it could in principle be programmed.

Modem linguistics and AI, then, were "born" at about the same time, and grew up together, intersecting in a hybrid field called **computational linguistics** or **natural language processing.**

The problem of understanding language soon turned out to be considerably more complex than it seemed in 1957.

Understanding language requires an understanding of the subject matter and context, not just an understanding of the structure of sentences.

This might seem obvious, but it was not widely appreciated until the 1960s. Much of the early work in **knowledge representation** (the study of how to put knowledge into a form that a computer can reason with) was tied to language and informed by research in Linguistics, which was connected in turn to decades of work on the philosophical analysis of language.

With the background material behind us, we are ready to cover the development of AI itself.

<u>Some of the important points are as follows:</u>

Different people think of AI differently. Two important questions to ask are: Are you concerned with thinking or behavior? Do you want to model humans or work from an ideal standard?

Intelligence is concerned mainly with **rational action.** Ideally, an **intelligent agent** takes the best possible action in a situation. We will study the problem of building agents that are intelligent in this sense.

Philosophers (going back to 400 B.c.) made AI conceivable by considering the ideas that the mind is in some ways like a machine, that it operates on knowledge encoded in some internal language, and that thought can be used to choose what actions to take.

Mathematicians provided the tools to manipulate statements of logical certainty as well as uncertain, probabilistic statements. They also set the groundwork for understanding computation and reasoning about algorithms.

Economists formalized the problem of making decisions that maximize the expected outcome to the decision-maker.

Psychologists adopted the idea that humans and animals can be considered information processing machines.

Linguists showed that language use fits into this model.

Computer engineers provided the artifacts that make AI applications possible. AI programs tend to be large, and they could not work without the great advances in speed and memory that the computer industry has provided.

Control theory deals with designing devices that act optimally on the basis of feedback from the environment. Initially, the mathematical tools of control theory were quite different from AI, but the fields are coming closer together.

1.2 The State of the Art

What can AI do today?

A concise answer is difficult, because there are so many activities in so many subfields.

Autonomous planning and scheduling

A hundred million miles from Earth, NASA's **Remote Agent Program** became the first on-board autonomous planning program. It control the scheduling of operations for a spacecraft

Remote Agent generated plans from high-level goals specified from the ground, and it monitored the operation of the spacecraft as the plans were executed-detecting, diagnosing, and recovering from problems as they occurred.

Game playing

IBM's Deep Blue became the first computer program to defeat the world champion in a chess match when it bested Garry Kasparov by a score of 3.5 to 2.5 in an exhibition match.

Kasparov said that he felt a "**new kind of intelligence**" across the board from him. ***Newsweek*** magazine described the match as "The brain's last stand." The value of IBM's stock increased by $18 billion.

Autonomous control

The ALVINN computer vision system was trained to steer a car to keep it following a lane. Video cameras that transmit road images to ALVINN, which then computes the best direction to steer, based on experience from previous training runs.

Diagnosis

Medical diagnosis programs based on probabilistic analysis have been able to perform at the level of an expert physician in several areas of medicine. Heckerman (1991) describes a case where a leading expert on lymph-node pathology scoffs at a program's diagnosis of an especially difficult case.

The creators of the program suggest he ask the computer for an explanation of the diagnosis. The machine points out the major factors influencing its decision and explains the subtle interaction of several of the symptoms in this case. Eventually, the expert agrees with the program.

Logistics Planning

During the Persian Gulf crisis of 1991, U.S. forces deployed a **Dynamic Analysis and Replanning Tool**, DART (Cross and Walker, 1994), to do automated logistics planning and scheduling for transportation. This involved up to 50,000 vehicles, cargo, and people at a time, and had to account for starting points, destinations, routes, and conflict resolution among all parameters.

The AI planning techniques allowed a plan to be generated in hours that would have taken weeks with older methods. The **Defense Advanced Research Project Agency** (DARPA) stated that this single application more than paid back DARPA's 30-year investment in AI.

Robotics

Many surgeons now use robot assistants in microsurgery. Computer vision techniques are used to create a three-dimensional model of a patient's internal anatomy and then uses robotic control to guide the insertion of a hip replacement prosthesis.

Language understanding and problem solving

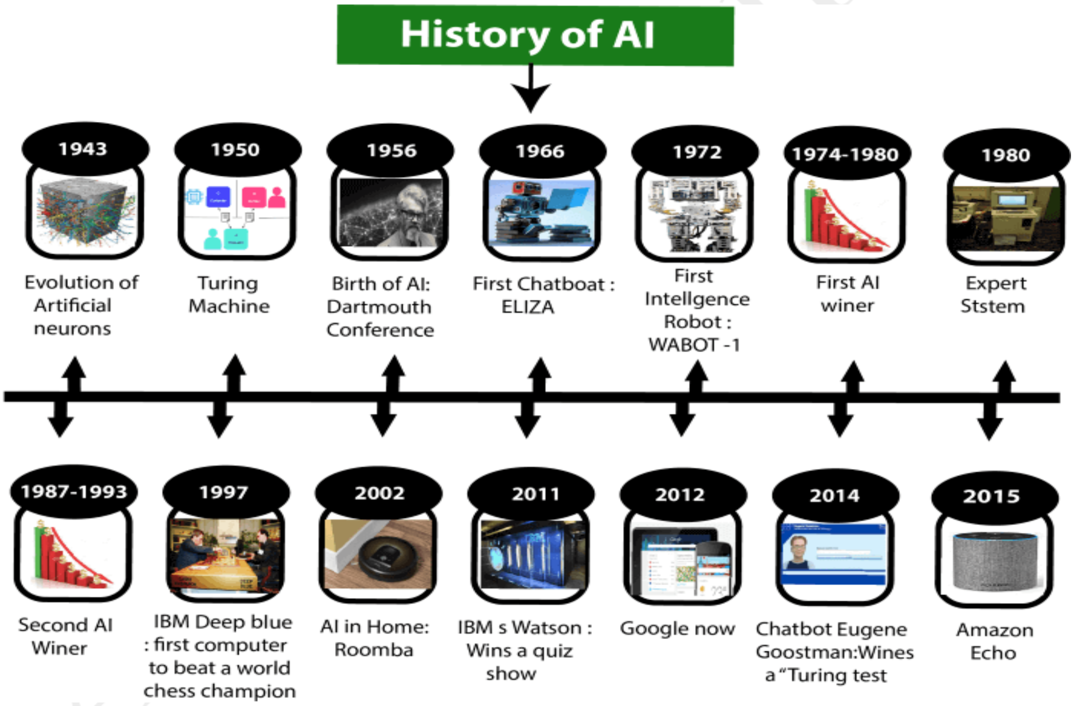
PROVERB is a computer program that solves crossword puzzles better than most humans, using constraints on possible word fillers, a large database of past puzzles, and a variety of information sources including dictionaries and online databases such as a list of movies and the actors that appear in them.

1.3 The history of AI

The history of AI has had cycles of success, misplaced optimism, and resulting cutbacks in enthusiasm and funding. There have also been cycles of introducing new creative approaches and systematically refining the best ones.

History of Artificial Intelligence

Artificial Intelligence is not a new word and not a new technology for researchers. This technology is much older than you would imagine. Even there are the myths of Mechanical men in Ancient Greek and Egyptian Myths. Following are some milestones in the history of AI which defines the journey from the AI generation to till date development.



Maturation of Artificial Intelligence (1943-1952)

- Year 1943: The first work which is now recognized as AI was done by Warren McCulloch and Walter pits in 1943. They proposed a model of artificial neurons.
- Year 1949: Donald Hebb demonstrated an updating rule for modifying the connection strength between neurons. His rule is now called Hebbian learning.
- Year 1950: The Alan Turing who was an English mathematician and pioneered Machine learning in 1950. Alan Turing publishes "Computing Machinery and Intelligence" in

which he proposed a test. The test can check the machine's ability to exhibit intelligent behavior equivalent to human intelligence, called a Turing test.

The birth of Artificial Intelligence (1952-1956)

- Year 1955: An Allen Newell and Herbert A. Simon created the "first artificial intelligence program"Which was named as "Logic Theorist". This program had proved 38 of 52 Mathematics theorems, and find new and more elegant proofs for some theorems.
- Year 1956: The word "Artificial Intelligence" first adopted by American Computer scientist John McCarthy at the Dartmouth Conference. For the first time, AI coined as an academic field.

At that time high-level computer languages such as FORTRAN, LISP, or COBOL were invented. And the enthusiasm for AI was very high at that time.

The golden years-Early enthusiasm (1956-1974)

- Year 1966: The researchers emphasized developing algorithms which can solve mathematical problems. Joseph Weizenbaum created the first chatbot in 1966, which was named as ELIZA.
- Year 1972: The first intelligent humanoid robot was built in Japan which was named as WABOT-1.

The first AI winter (1974-1980)

- The duration between years 1974 to 1980 was the first AI winter duration. AI winter refers to the time period where computer scientist dealt with a severe shortage of funding from government for AI researches.
- During AI winters, an interest of publicity on artificial intelligence was decreased.

A boom of AI (1980-1987)

- Year 1980: After AI winter duration, AI came back with "Expert System". Expert systems were programmed that emulate the decision-making ability of a human expert.
- In the Year 1980, the first national conference of the American Association of Artificial Intelligence was held at Stanford University.

The second AI winter (1987-1993)

- The duration between the years 1987 to 1993 was the second AI Winter duration.
- Again Investors and government stopped in funding for AI research as due to high cost but not efficient result. The expert system such as XCON was very cost effective.

The emergence of intelligent agents (1993-2011)

- Year 1997: In the year 1997, IBM Deep Blue beats world chess champion, Gary Kasparov, and became the first computer to beat a world chess champion.

- Year 2002: for the first time, AI entered the home in the form of Roomba, a vacuum cleaner.
- Year 2006: AI came in the Business world till the year 2006. Companies like Facebook, Twitter, and Netflix also started using AI.

Deep learning, big data and artificial general intelligence (2011-present)

- Year 2011: In the year 2011, IBM's Watson won jeopardy, a quiz show, where it had to solve the complex questions as well as riddles. Watson had proved that it could understand natural language and can solve tricky questions quickly.
- Year 2012: Google has launched an Android app feature "Google now", which was able to provide information to the user as a prediction.
- Year 2014: In the year 2014, Chatbot "Eugene Goostman" won a competition in the infamous "Turing test."
- Year 2018: The "Project Debater" from IBM debated on complex topics with two master debaters and also performed extremely well.
- Google has demonstrated an AI program "Duplex" which was a virtual assistant and which had taken hairdresser appointment on call, and lady on other side didn't notice that she was talking with the machine.
- Now AI has developed to a remarkable level. The concept of Deep learning, big data, and data science are now trending like a boom. Nowadays companies like Google, Facebook, IBM, and Amazon are working with AI and creating amazing devices. The future of Artificial Intelligence is inspiring and will come with high intelligence.

1.4 Benefits of Artificial Intelligence:

AI machines use machine learning algorithms to mimic the cognitive abilities of human beings and solve a simple or complex problem.

1. Increase work efficiency

- AI-powered machines are great at doing a particular repetitive task with amazing efficiency. The simple reason is that they remove human errors from their tasks to achieve accurate results every time they do that specific task.
- Moreover, such machines can work 24X7, unlike humans. Thus, they eliminate the need to deploy two sets of employees working in day and night shifts to work on important tasks. For example, AI-powered chat assistants can answer customer queries and provide support to visitors every minute of the day and boost the sales of a company.

2. Work with high accuracy

- Scientists are working to teach artificial intelligence powered machines to solve complex equations and perform critical tasks on their own so that the results obtained have higher accuracy as compared to their human counterparts.
- Their high accuracy has made these machines indispensable to work in the medical field particularly, owing to the criticality of the tasks. Robots are getting better at diagnosing serious conditions in the human body and performing delicate surgeries to minimize the risk of human lives.

3. Reduce cost of training and operation

- AI uses machine learning algorithms like Deep Learning and neural networks to learn new things like humans do. This way they eliminate the need to write new code every time we need them to learn new things.
- There is significant Research and Development going on in the world to develop AI machines that optimize their machine learning abilities so that they learn much faster about new processes. This way the cost of training robots would become much lesser than that of humans. Moreover, machines already reduce the cost of operations with their high efficiency and accuracy of doing work. For example, machines don't take breaks and can perform the same mundane task again and again without any downtime or change in results.

4. Improve Processes

- The best part about AI-powered machines being deployed for work is that they let us gather humongous amounts of data related to their work. Such data can be processed to gather deep insights into the processes with quantitative analysis so that we can optimize them even further.
- Machine learning abilities of AI machines are increasing further and further to do even the analysis by themselves.

1.5 Risks of Artificial Intelligence

Although hailed as a boon for humanity by tech pundits, artificial intelligence is feared by a lot of scientists and regular citizens alike. This fear has made it to the silver screen several times in the form of movies depicting dystopian futures created by AI machines that took over the planet. The most notable of these is the Matrix and the Terminator.

1. AI is Unsustainable

- Intelligent machines have characteristically high computing powers contributed by an array of several processers. These computer chips have rare earth materials like Selenium as a major constituent. Besides, the batteries of such machines run on Lithium, again a rare element in earth's crust. The increased mining of these materialsis irreversibly damaging our environment at a rapid pace. Moreover, they consume huge amounts of power to function, that is putting severe pressure on our power plants and again harming the environment.

2. Lesser Jobs

- There is no doubt that machines do routine and repeatable tasks much better than humans. Many businesses would prefer machines instead of humans to increase their profitability, thus reducing the jobs that are available for the human workforce.

3. A threat to Humanity(?)

- Elon Musk is considered to be one of the smartest person working on AI in present times. He has also stated publicly that AI is the biggest threat to human civilization in the future. This means that the dystopian future that sci-fi movies show is not impossible. Also, Stephen Hawking has always shown his disagreement with the advancement in thefield of AI.

The biggest risk associated with AI is that machines would gain sentience and turn against humans in case they go rogue.

1.6 Intelligent Agents

Agents in Artificial Intelligence

An AI system can be defined as the study of the rational agent and its environment. The agents sense the environment through sensors and act on their environment through actuators. An AI agent can have mental properties such as knowledge, belief, intention, etc.

What is an Agent?

An agent can be anything that perceiveits environment through sensors and act upon that environment through actuators. An Agent runs in the cycle of perceiving, thinking, and acting. An agent can be:

- Human-Agent: A human agent has eyes, ears, and other organs which work for sensors and hand, legs, vocal tract work for actuators.
- Robotic Agent: A robotic agent can have cameras, infrared range finder, NLP for sensors and various motors for actuators.
- Software Agent: Software agent can have keystrokes, file contents as sensory input and act on those inputs and display output on the screen.

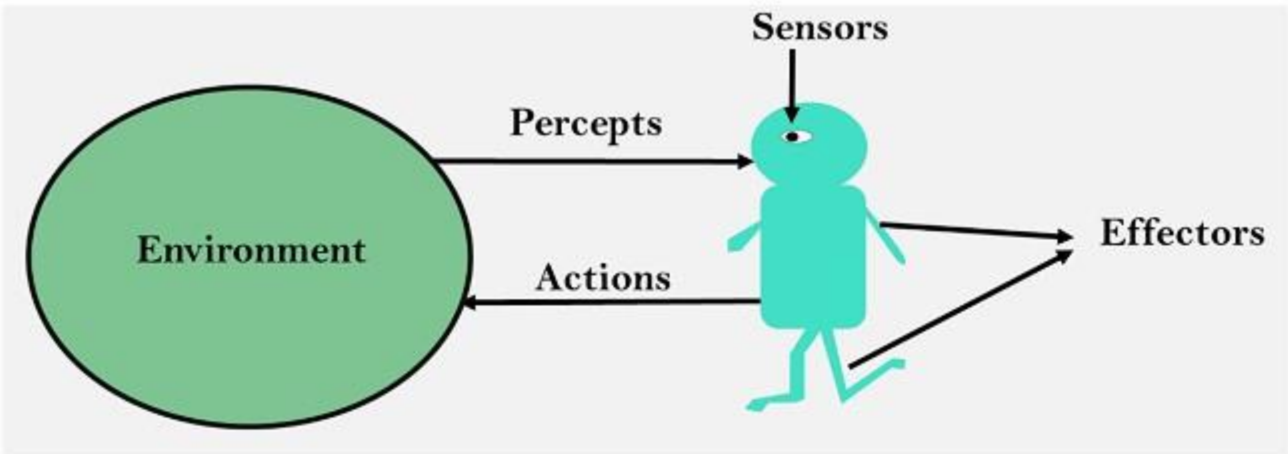
Hence the world around us is full of agents such as thermostat, cellphone, camera, and even we are also agents.

Before moving forward, we should first know about sensors, effectors, and actuators.

Sensor: Sensor is a device which detects the change in the environment and sends the information to other electronic devices. An agent observes its environment through sensors.

Actuators: Actuators are the component of machines that converts energy into motion. The actuators are only responsible for moving and controlling a system. An actuator can be an electric motor, gears, rails, etc.

Effectors: Effectors are the devices which affect the environment. Effectors can be legs, wheels, arms, fingers, wings, fins, and display screen.



Intelligent Agents:

An intelligent agent is an autonomous entity which act upon an environment using sensors and actuators for achieving goals. An intelligent agent may learn from the environment to achieve their goals. A thermostat is an example of an intelligent agent.

Following are the main four rules for an AI agent:

- Rule 1: An AI agent must have the ability to perceive the environment.
- Rule 2: The observation must be used to make decisions.
- Rule 3: Decision should result in an action.
- Rule 4: The action taken by an AI agent must be a rational action.

Rational Agent:

- A rational agent is an agent which has clear preference, models uncertainty, and acts in a way to maximize its performance measure with all possible actions.
- A rational agent is said to perform the right things. AI is about creating rational agents to use for game theory and decision theory for various real-world scenarios.
- For an AI agent, the rational action is most important because in AI reinforcement learning algorithm, for each best possible action, agent gets the positive reward and for each wrong action, an agent gets a negative reward.

Note: Rational agents in AI are very similar to intelligent agents.

1.7 Rationality:

The rationality of an agent is measured by its performance measure. Rationality can be judged on the basis of following points:

- Performance measure which defines the success criterion.
- Agent prior knowledge of its environment.
- Best possible actions that an agent can perform.
- The sequence of percepts.

Note: Rationality differs from Omniscience because an Omniscient agent knows the actual outcome of its action and act accordingly, which is not possible in reality."

1.8 Structure of an AI Agent

- The task of AI is to design an agent program which implements the agent function. The structure of an intelligent agent is a combination of architecture and agent program. It can be viewed as:

1. Agent = Architecture + Agent program

Following are the main three terms involved in the structure of an AI agent:

Architecture: Architecture is machinery that an AI agent executes on.

Agent Function: Agent function is used to map a percept to an action.

1. $f:P^* \rightarrow A$

Agent program: Agent program is an implementation of agent function. An agent program executes on the physical architecture to produce function f.

PEAS Representation

PEAS is a type of model on which an AI agent works upon. When we define an AI agent or rational agent, then we can group its properties under PEAS representation model. It is made up of four words:

P: Performance measure

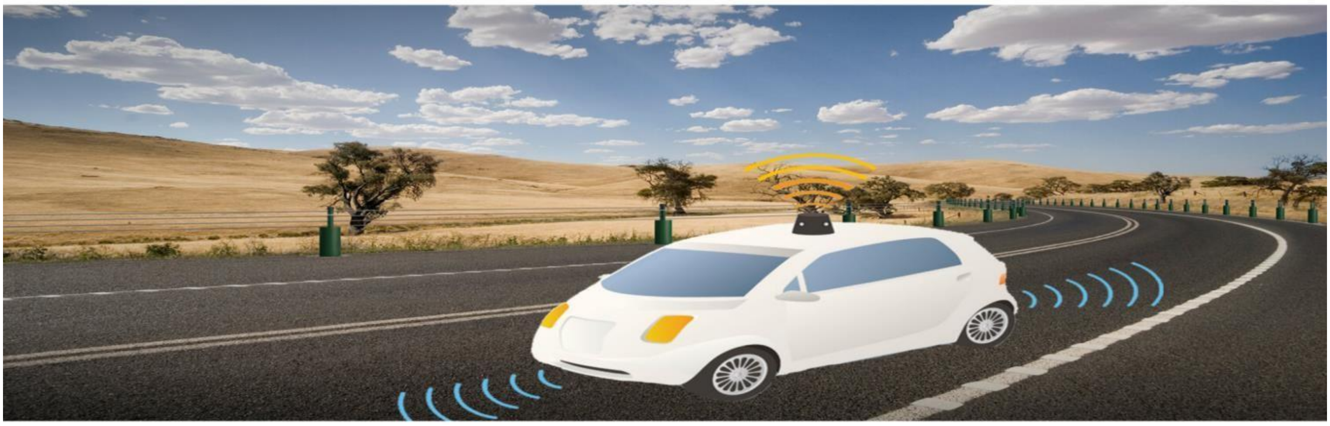
E: Environment

A: Actuators

S: Sensors

Here performance measure is the objective for the success of an agent's behavior.

PEAS for self-driving cars:



Let's suppose a self-driving car then PEAS representation will be:

- Performance: Safety, time, legal drive, comfort
- Environment: Roads, other vehicles, road signs, pedestrian
- Actuators: Steering, accelerator, brake, signal, horn
- Sensors: Camera, GPS, speedometer, odometer, accelerometer, sonar.

Example of Agents with their PEAS representation

Agent	Performance measure	Environment	Actuators	Sensors
1. Medical Diagnose	○ Healthy patient	○ Patient ○ Hospital	○ Tests ○ Treatments	Keyboard (Entry of symptoms)

	○ Minimized cost	○ Staff		
2. Vacuum Cleaner	○ Cleanness ○ Efficiency ○ Battery life ○ Security	○ Room ○ Table ○ Wood floor ○ Carpet ○ Various obstacles	○ Wheels ○ Brushes ○ Vacuum Extractor	○ Camera ○ Dirt detection sensor ○ Cliff sensor ○ Bump Sensor ○ Infrared Wall Sensor
3. Part - picking Robot	○ Percentage of parts in correct bins.	○ Conveyor belt with parts, ○ Bins	○ Jointed Arms ○ Hand	○ Camera ○ Joint angle sensors.