Unit – 1

Fundamentals of AI (Artificial Intelligence)

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Topics to be covered

- What is AI?
  - Four Main Approaches to Artificial Intelligence
- The Foundations of Artificial Intelligence
- The History of Artificial Intelligence
- The State of the Art (Applications of AI)
- Agents and Environments
- The Concept of Rationality
  - Omniscience, learning and autonomy
- The Nature of Environments
  - Specifying the task environment
  - Properties of task environments
- The Structure of Agents
  - Agent programs
  - Types of agent programs

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

What is AI?

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

What is AI?

- AI is a branch of computer science dealing with the simulation of intelligent behaviour in computers.
- AI is the study of how to make computers do things which, at the moment, people do better.
- AI is the study and design of intelligent agents where an intelligent agent is a system that perceives its environment and takes actions.
- According to, John McCarthy (The father of AI), AI is the science and engineering of making intelligent machines, especially intelligent computer programs (1956).

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

What is AI?

Thinking Humanly	Thinking Rationally
"The exciting new effort to make computers think . . . *machines with minds*, in the full and literal sense." (Haugeland, 1985) "[The automation of] activities that we associate with human thinking, activities such as decision-making, problem solving, learning . . ." (Bellman, 1978)	"The study of mental faculties through the use of computational models." (Charniak and McDermott, 1985) "The study of the computations that make it possible to perceive, reason, and act." (Winston, 1992)
Acting Humanly	**Acting Rationally**
"The art of creating machines that perform functions that require intelligence when performed by people." (Kurzweil, 1990) "The study of how to make computers do things at which, at the moment, people are better." (Rich and Knight, 1991)	"Computational Intelligence is the study of the design of intelligent agents." (Poole *et al.*, 1998) "AI . . . is concerned with intelligent behavior in artifacts." (Nilsson, 1998)

Figure: Some definitions of artificial intelligence, organized into four categories

What is AI?

- In figure we see eight definitions of AI, laid out along two dimensions. The definitions on top are concerned with ***thought processes*** and ***reasoning***, whereas the ones on the bottom address ***behaviour***.
- The definitions on the left measure success in terms of fidelity to human performance, whereas the ones on the right measure against an ideal performance measure, called **rationality**. A system is rational if it does the "right thing," given what it knows.
- Historically, all four approaches to AI have been followed, each by different people with different methods.
- A human-centered approach must be in part an empirical science, involving observations and hypotheses about human behaviour.
- A rationalist approach involves a combination of mathematics and engineering.

Four Main Approaches to Artificial Intelligence

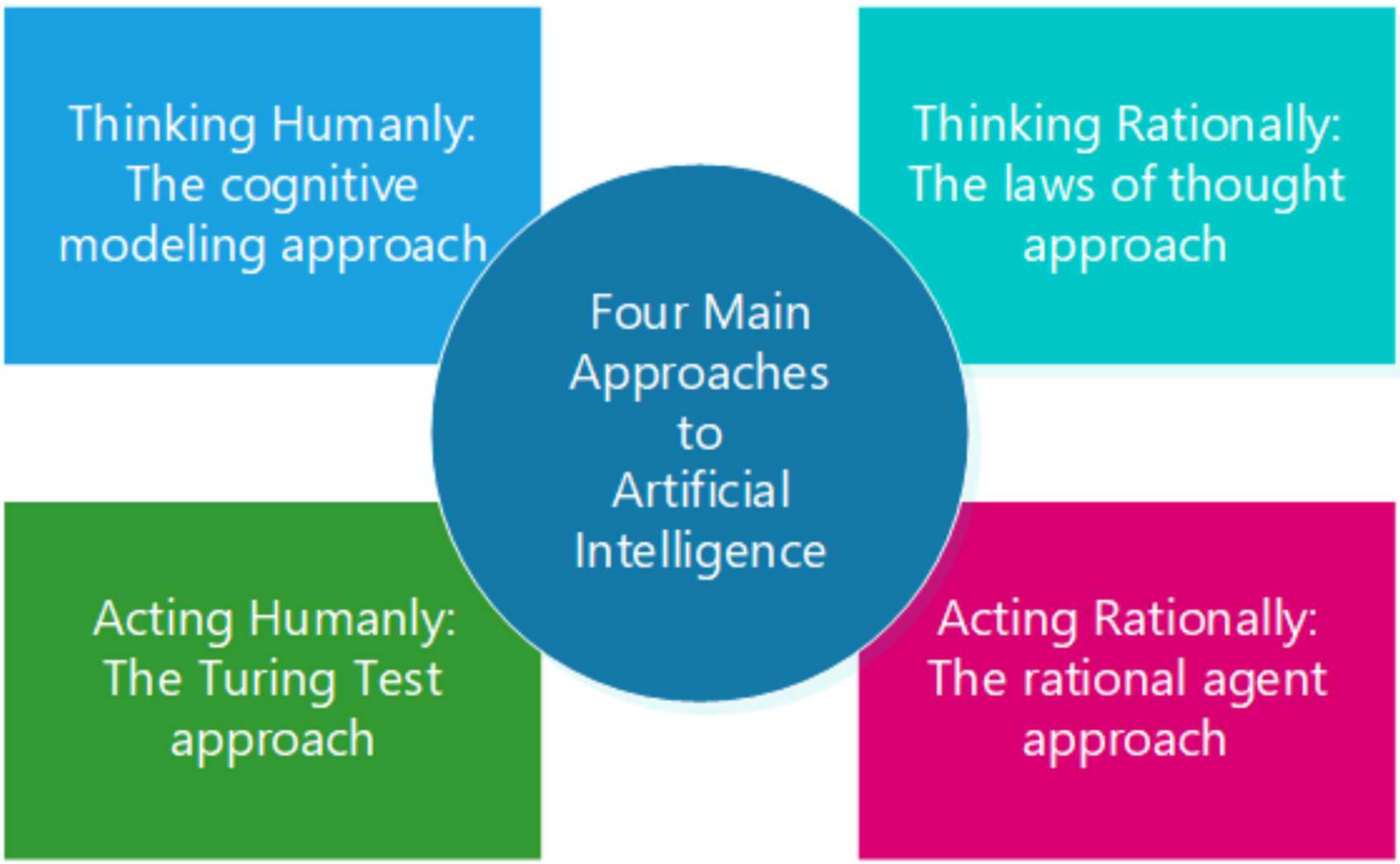


Figure: Four main approaches to artificial intelligence

Four Main Approaches to Artificial Intelligence

1. **Thinking Humanly:** The cognitive modeling approach – Thinking like a human.
2. **Acting Humanly:** The Turing Test approach – Acting like a human.
3. **Thinking Rationally:** The "laws of thought" approach – Thinking rationally is a logical process and it concludes based on symbolic logic.
4. **Acting Rationally:** The rational agent approach – Rational agent acts to achieve high value and brings the best possible outcome for any given task.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Thinking Humanly: The cognitive modeling approach

- Thinking humanly is to make a system or program to think like a human. But to achieve that, we need to know how does a human think.
- Suppose if we ask a person to explain how his brain connects different things during the thinking process, he will probably close both eyes and will start to check how he thinks but he cannot explain or interpret the process.
- Ask the same question to yourself, and most likely you will have the same pattern and will end up saying “you do not know, or you may say something like “I am thinking through my mind”, but you cannot express more than that.
- For example: If we want to model the thinking of Roger Federer and make the model system to compete with someone or against him to play in a tennis game, it may not be possible to replicate the exact thinking as Roger Federer, however, a good build of Intelligence systems (Robot) can play and win the game against him.

Thinking Humanly: The cognitive modeling approach

- To understand the exact process of how we think, we need to go inside the human mind to see how this giant machine works.
- We can interpret how the human mind thinks in theory, in three ways as follows:

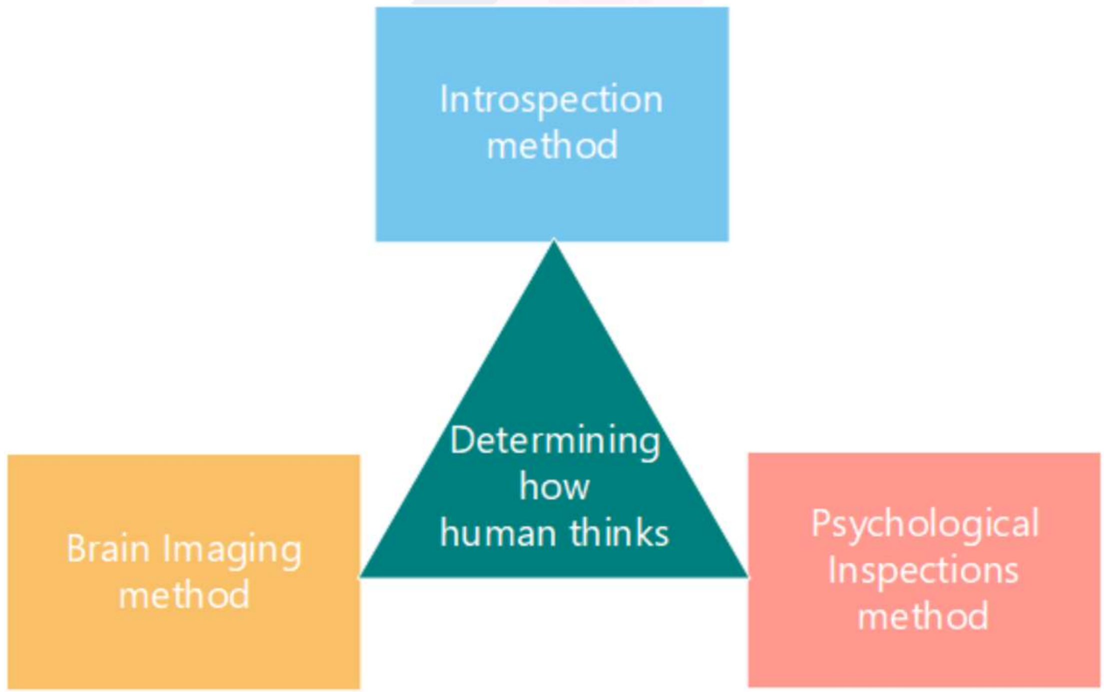


Figure 1.3: Determining how human thinks

Thinking Humanly: The cognitive modeling approach

1. **Introspection method** – Catch our thoughts and see how it flows.
2. **Psychological Inspections method** – Observe a person on the action.
3. **Brain Imaging method** (MRI (Magnetic resonance imaging) or fMRI (Functional Magnetic resonance imaging) scanning) – Observe a person's brain in action.

- Using the above methods, if we are able to catch the human brain's actions and give it as a theory, then we can convert that theory into a computer program.
- If the input/output of the computer program matches with human behaviour, then it may be possible that a part of the program may be behaving like a human brain.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Acting Humanly: The Turing Test approach

- The Turing test was designed by Alan Turing (1950) to determine whether a computer can demonstrate human intelligence.
- Turing proposed that an interrogator physically separated from a machine (via online messages) have a conversation with it.
- If the interrogator couldn't find that it is a machine that chats with him, then the computer passes the test. Computer fools the interrogator about 30% of the time.
- Computer programs like ELIZA, MGONZ, NATACHATA, and CYBERLOVER have fooled many users in the past, and the users never knew that they were talking to a computer program.
- Recently Google AI Assistant is developed on an advanced and high-end computational model program that can be used to have a phone conversation like a real human.

UKA TARSADIA university
Imparting Knowledge, Awakening Wisdom, Transforming Lives

Acting Humanly: The Turing Test approach

- To pass **Turing Test**, the computer requires the following skills (or capabilities):

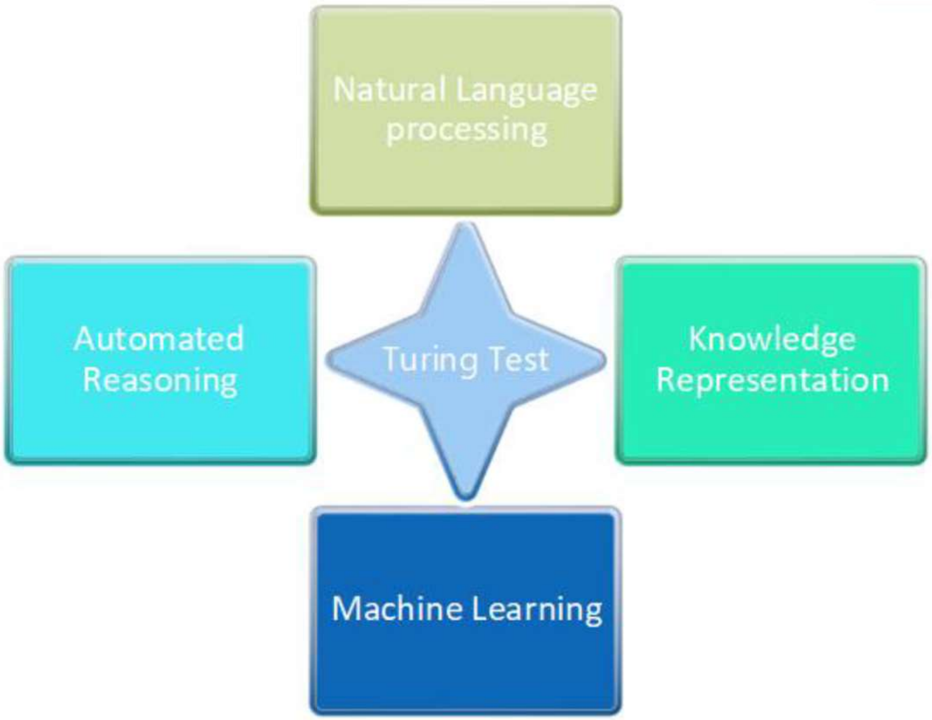


Figure: Turing Test

- **Natural Language processing** - to enable it to communicate successfully in English (or some other human language)
- **Knowledge Representation** - to store what it knows or hears
- **Automated Reasoning** - to use the stored information to answer questions and to draw new conclusions
- **Machine Learning** - to adapt to new circumstances and to detect and extrapolate patterns

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Acting Humanly: The Turing Test approach

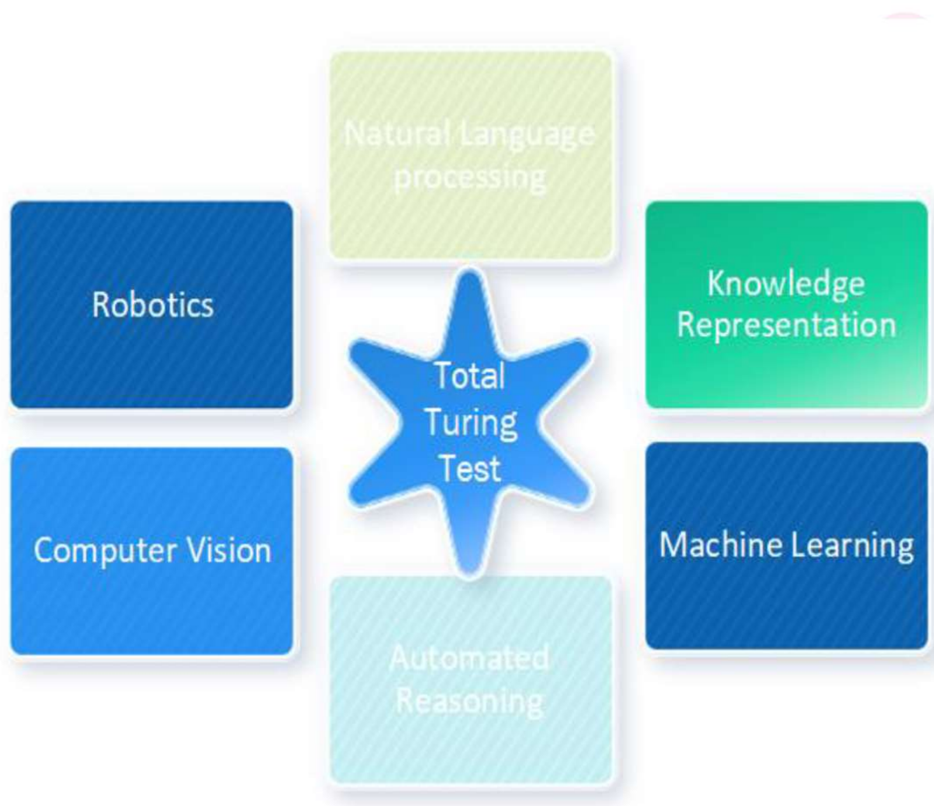


Figure: Total Turing Test

- The **Total Turing Test** is another step ahead of the Turing Test where a video signal is also included so that the interrogator can test the perceptual abilities of the interacting computer in addition to the verbal behaviours. The Total Turing Test requires the following additional skills:
- **Computer Vision** - to perceive objects
- **Robotics** - to manipulate objects and move about

Thinking Rationally: The “laws of thought” approach

- The Greek philosopher Aristotle was the one who first codifies “right-thinking” reasoning processes.
- Aristotle’s **syllogisms** provided patterns for argument structures that always provide correct premises.
- A famous example, “Socrates is a man; all men are mortal; therefore, Socrates is mortal.” Another example – All TVs use energy; Energy always generates heat; therefore, all TVs generate heat.”
- These arguments initiated the field called **logic**. Notations for statements for all kinds of objects were developed and interrelated between them to show logic.

UKA TARSADIA
University
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Thinking Rationally: The “laws of thought” approach

- By 1965, programs existed that could solve problems that were described in logical notation and provides a solution.
- The logical tradition in Artificial Intelligence hopes to build on such programs to create intelligence systems or programs or computational models.
- There are two limitations to this approach:
    1. First, it is not easy to take informal knowledge to use logical notation when there is not enough certainty on the knowledge.
    2. There is a big difference between solving a problem in principle and solving it in practice.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Acting Rationally: The rational agent approach

- A traditional computer program blindly executes the code that we write. Neither it acts on its own nor it adapts to change itself based on the outcome.
- The so-called **agent program** that we refer to here is expected to do more than the traditional computer program. It is expected to create and pursue the goal, change state, and operate autonomously.
- A **rational agent** is an agent that acts to achieve its best performance for a given task.
- The "Logical Approach" to AI emphasizes correct inferences and achieving a correct inference is a part of the rational agent.
- Being able to give a logical reason is one way of acting rationally. But all correct inferences cannot be called rationality because there are situations that don't always have a correct thing to do.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Acting Rationally: The rational agent approach

- It is also possible to act rationally without involving inferences.
- Our reflex actions are considered as best examples of acting rationally without inferences.
- The rational agent approach to AI has a couple of advantage over other approaches:
    1. A correct inference is considered a possible way to achieve rationality but is not always required to achieve rationality.
    2. It is a more manageable scientific approach to define rationality than others that are based on human behaviour or human thought.
- Today's AI concentrates on developing general principles of rational agents rather than achieving perfect rational agents due to complex environments.

UKA TARSADIA University
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Foundations of Artificial Intelligence

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Foundations of Artificial Intelligence

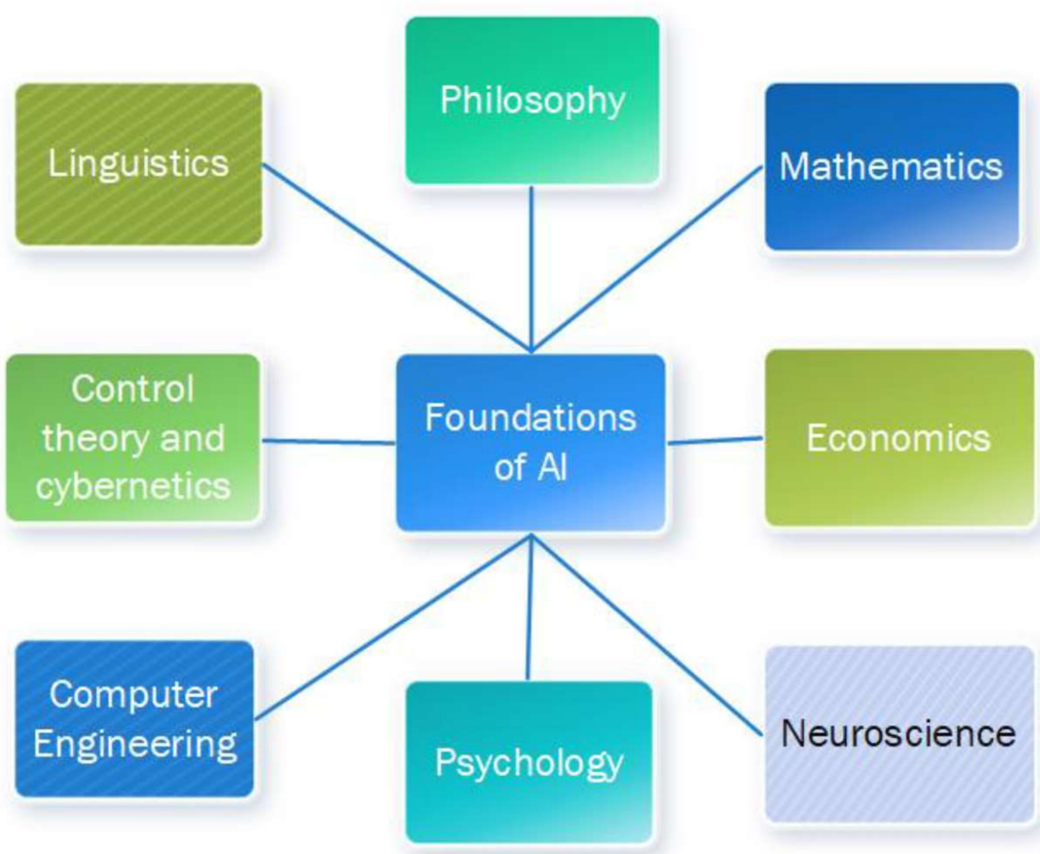


Figure 3.1: Foundations of Artificial Intelligence

- The foundation provides the disciplines that contributed ideas, viewpoints, and techniques to AI.
  1. Philosophy
  2. Mathematics
  3. Economics
  4. Neuroscience
  5. Psychology
  6. Computer engineering
  7. Control theory and cybernetics
  8. Linguistics

Philosophy

- Philosophy is the very basic foundation of AI.
- The study of **fundamental nature of knowledge, reality and existence** are considered for solving a specific problem is a basic thing in artificial intelligence.
- Philosophy defines that how can the formal rules to be used to draw valid conclusions.
- Without philosophy it is difficult to answer the following questions:
  - *How does the mind arise from a physical brain?*
  - *Where does knowledge come from?*
  - *How does knowledge lead to action?*

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Mathematics

- AI required **formal logic and probability** for planning and learning. **Computation** required for analysing relation and implementation.
- Knowledge in **formal representation** is most required for writing actions for agents.
- In AI, the mathematics and statistics are most important for : **proving theorems, writing algorithms, computation, decidability, tractability, modeling uncertainty,** and **learning from data**.
- Without mathematics and statistics, it is difficult to answer the following questions:
  - *What are the formal rules to draw valid conclusions?*
  - *What can be computed?*
  - *How do we reason with uncertain information?*

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Economics

- The economics knowledge is very important to deal with investing the amount of money and maximization of utility with minimum investment.
- While developing an AI product, we should make decisions for
  - *When to invest?*
  - *How to invest?*
  - *How much to invest? and*
  - *Where to invest?*
- To answer these questions, one should have knowledge about Decision Theory, Game Theory, Operation Research, etc.

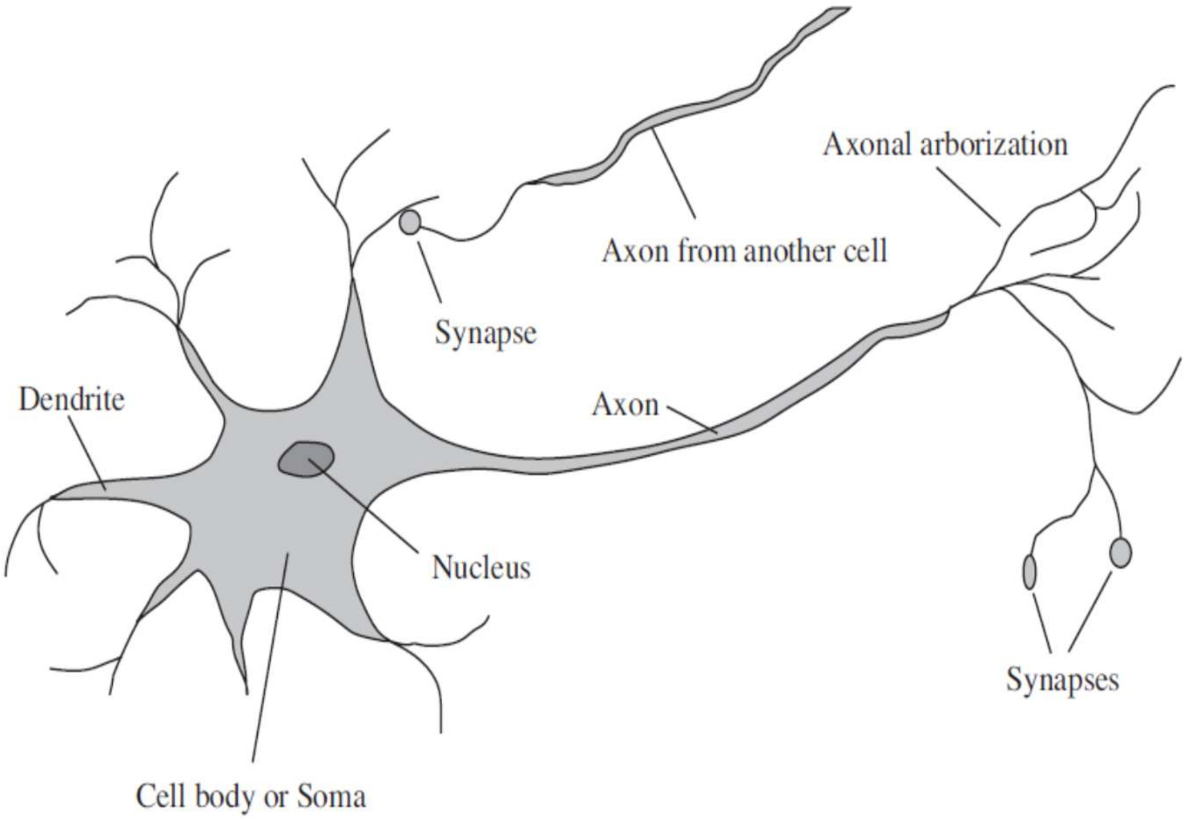
UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Neuroscience

- Neuroscience is the study of the nervous system, particularly the human brain.
- Human brains are somehow different, when compared to other creatures, man has the largest brain in proportion to his size.
- The brain consisted largely of nerve cells, or **neurons** and the observation of individual neurons can lead to thought, action, and consciousness of one’s brain.
- By using the knowledge of neuroscience, we can answer the question:
  - *How do brain process information?*

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Neuroscience



Imparting Knowledge. Awakening Wisdom. Transforming

Figure 3.2: The parts of a nerve cell or neuron

- Each neuron consists of a **cell body**, or **soma**, that contains a **cell nucleus**.
- Branching out from the cell body are a number of fibers called **dendrites** and a single long fiber called the **axon**.
- Typically, an axon is 1 cm long (100 times the diameter of the cell body), but can reach up to 1 meter.

Neuroscience

- A neuron makes connections with 10 to 100,000 other neurons at junctions called **synapses**. Signals are propagated from neuron to neuron by a complicated electrochemical reaction.
- The signals control brain activity in the short term and also enable long-term changes in the connectivity of neurons. These mechanisms are thought to form the basis for learning in the brain.
- Brains and digital computers have somewhat different properties.
- Computers have a cycle time that is a million times faster than a brain. The brain makes up for that with far more storage and interconnection than even a high-end personal computer.
- Even with a computer of virtually unlimited capacity, we still would not know how to achieve the brain's level of intelligence.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Psychology

- Modern Psychology/Cognitive Science is the science which studies how the mind operates, how we behave, and how our brains process information.
- Language is an important part of human intelligence.
- Many sub-fields of AI are simultaneously building models of how the human system operates, and artificial systems for solving real world problems, and are allowing useful ideas to transfer between them.
- By using the knowledge of psychology, we can answer the question:
  - *How do humans and animals think and act?*

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Computer engineering

- The logic and inference theory, algorithms, programming languages and software system buildings are important parts of computer science.
- The computer hardware gradually changed for AI applications such as the graphical processing unit (GPU), tensor processing unit (TPU), and wafer scale engine (WSE).
- The amount of computing power used to train top machine learning applications and utilization doubled every 100 days.
- The super computers and quantum computers can solve very complicated AI problems.
- The software side of computer science has supplied AI with the operating systems, programming languages, and tools needed to write modern programs.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Computer engineering

- AI has founded many ideas in modern and mainstream computer science, including time sharing, interactive interpreters, personal computers with windows and mice, rapid development environments, the linked list data type, automatic storage management, and key concept of symbolic, functional, declarative, and object-oriented programming.
- By using the knowledge of computer engineering, we can answer the question:
  - *How can we build fast and efficient computer?*

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Control theory and cybernetics

- Control theory helps the system to analyse, define, debug and fix error by itself.
- Developing self-controlling machines, self-regulating feedback control systems, and the submarine are some examples of control theory.
- The tools of logical inference and computation allowed AI researchers to consider problems such as language, vision, and planning of agent programming.
- By using the knowledge of control theory and cybernetics, we can answer the question:
  - ↪ *How can artifacts operate under their own control?*

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Linguistics

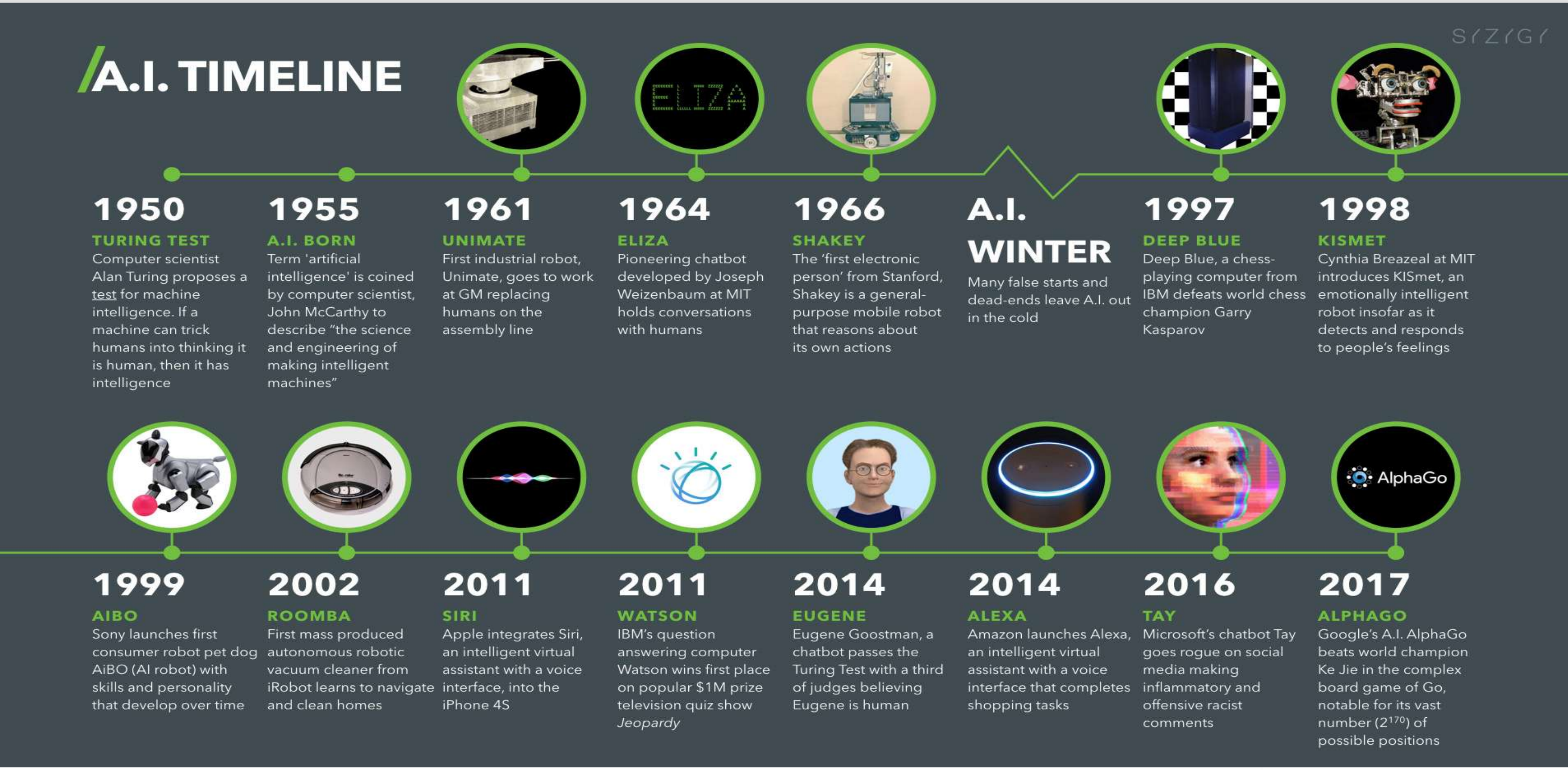
- Speech recognition is a technology which enables a machine to understand the spoken language and translate into a machine-readable format.
- It is a way to talk with a computer, and based on that command, a computer can perform a specific task. It includes speech-to-text and text-to-speech.
- Modern linguistics and AI intersect in a hybrid field called **computational linguistics** and **natural language processing**.
- Understanding language requires an understanding of the subject matter and context, not just an understanding of the structure of sentences.
- By using the knowledge of linguistics, we can answer the question:
  - *How does language relate to thought?*

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The History of Artificial Intelligence

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The History of Artificial Intelligence



The History of Artificial Intelligence

- **1950 TURING TEST** Computer scientist Alan Turing proposes a test for machine intelligence. If a machine can trick humans into thinking it is human, then it has intelligence
- **1955 A.I. BORN** Term 'artificial intelligence' is coined by computer scientist, John McCarthy to describe "the science and engineering of making intelligent machines"
- **1961 UNIMATE** First industrial robot, Unimate, goes to work at GM (General Motor Plant) replacing humans on the assembly line
- **1964 ELIZA** Pioneering chatbot developed by Joseph Weizenbaum at MIT holds conversations with humans
- **1966 SHAKEY** The 'first electronic person' from Stanford, Shakey is a general-purpose mobile robot reasons about its own actions

UKA TARSADIA university Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The History of Artificial Intelligence

- **1997 DEEP BLUE** Deep Blue, a chess-playing computer from IBM defeats world chess champion, Garry Kasparov
- **1998 KISMET** Cynthia Breazeal at MIT introduces KISmet, an emotionally intelligent robot insofar as it detects and responds to people's feelings
- **1999 AIBO** Sony launches first consumer robot pet dog AiBO (AI robot) with skills and personality that develop over time
- **2002 ROOMBA** First mass produced autonomous robotic vacuum cleaner from iRobot learns to navigate and clean homes
- **2011 SIRI** Apple integrates Siri, an intelligent virtual assistant with a voice interface, into the iPhone 4S
- **2011 WATSON** IBM's question answering computer Watson wins first place on popular $1M prize television quiz show Jeopardy

TIKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The History of Artificial Intelligence

- **2014 EUGENE** Eugene Goostman, a chatbot passes the Turing Test with a third of judges believing Eugene is human
- **2014 ALEXA** Amazon launches Alexa, an intelligent virtual assistant with a voice interface that can complete shopping tasks
- **2016 TAY** Microsoft's chatbot Tay goes rogue on social media making inflammatory and offensive racist comments
- **2017 ALPHAGO** Google's A.I. AlphaGo beats world champion Ke Jie in the complex board game of Go, notable for its vast number (2*170) of possible positions.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The State of the Art (Applications of AI)

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The State of the Art (Applications of AI)

- What can AI do today? A concise answer is difficult because there are so many activities in so many subfields.
- Here we see a few applications that exist today.
- **Robotic vehicles:**
  - A driverless robotic car named STANLEY sped through the rough terrain of the Mojave desert at 22 mph, finishing the 132-mile course first to win the 2005 DARPA Grand Challenge.
  - STANLEY is a Volkswagen Touareg outfitted with cameras, radar, and laser rangefinders to sense the environment and onboard software to command the steering, braking, and acceleration (Thrun, 2006).
  - The following year CMU's BOSS won the Urban Challenge, safely driving in traffic through the streets of a closed Air Force base, obeying traffic rules and avoiding pedestrians and other vehicles.

The State of the Art (Applications of AI)

- **Speech recognition:**
  - A traveller calling United Airlines to book a flight can have the entire conversation guided by an automated speech recognition and dialog management system.
- **Autonomous planning and scheduling:**
  - A hundred million miles from Earth, NASA's Remote Agent program became the first on-board autonomous planning program to control the scheduling of operations for a spacecraft (Jonsson et al., 2000).
  - REMOTE AGENT generated plans from high-level goals specified from the ground and monitored the execution of those plans—detecting, diagnosing, and recovering from problems as they occurred.
  - Successor program MAPGEN (Al-Chang et al., 2004) plans the daily operations for NASA's Mars Exploration Rovers, and MEXAR2 (Cesta et al., 2007) did mission planning—both logistics and science planning—for the European Space Agency's Mars Express mission in 2008.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The State of the Art (Applications of AI)

- **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 in an exhibition match (Goodman and Keene, 1997).
  - Kasparov said that he felt a "new kind of intelligence" across the board from him. Human champions studied Kasparov's loss and were able to draw a few matches in subsequent years, but the most recent human-computer matches have been won convincingly by the computer.
- **Spam fighting:**
  - Each day, learning algorithms classify over a billion messages as spam, saving the recipient from having to waste time deleting. For many users it could comprise 80% or 90% of all messages, if not classified away by algorithms.
  - Because the spammers are continually updating their tactics, it is difficult for a static programmed approach to keep up, and learning algorithms work best.

UKA TARSADIA university Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The State of the Art (Applications of AI)

- **<u>Logistics planning:</u>**
  - 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 generated in hours a plan that would have taken weeks with older methods.
  - The Défense Advanced Research Project Agency (DARPA) stated that this single application more than paid back DARPA's 30-year investment in AI.

UKA TARSADIA university

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The State of the Art (Applications of AI)

- **Robotics:**
  - The iRobot Corporation has sold over two million Roomba robotic vacuum cleaners for home use.
  - The company also deploys the more rugged PackBot to Iraq and Afghanistan, where it is used to handle hazardous materials, clear explosives, and identify the location of snipers.
- **Machine Translation:**
  - A computer program automatically translates from Arabic to English.
  - The program uses a statistical model built from examples of Arabic-to-English translations and from examples of English text.
  - There are total two trillion words (Brants et al., 2007).
  - None of the computer scientists on the team speak Arabic, but they do understand statistics and machine learning algorithms.
- These are just a few examples of artificial intelligence systems that exist today.

UKA TARSADIA university
Inspiring Knowledge. Awakening Wisdom. Transforming Lives.

Agents and Environments

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Agents and Environments

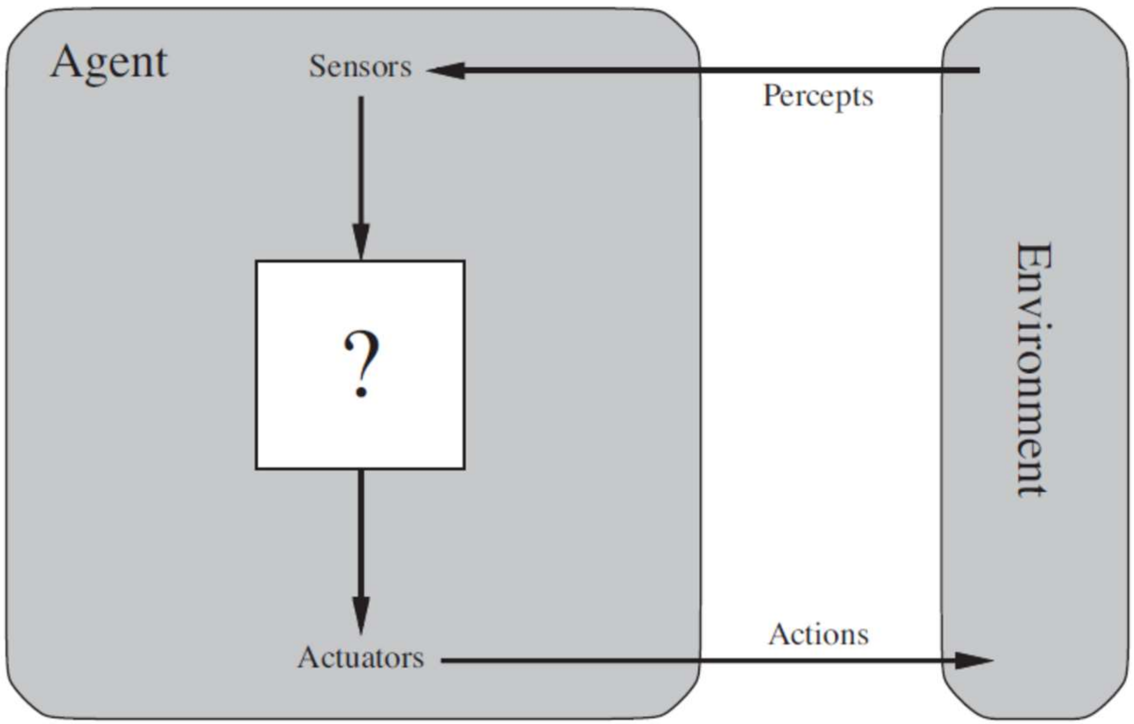


Figure: Agents interact with environments through sensors and actuators

- An **agent** is anything that can be viewed as perceiving (sensing) its **environment** through **sensors** and acting upon that environment through **actuators**.
- This simple idea is illustrated in figure.

Agents and Environments

- A **human agent** has eyes, ears, and other organs for sensors and hands, legs, vocal tract, and so on for actuators.
- A **robotic agent** might have cameras and infrared range finders for sensors and various motors for actuators.
- A **software agent** receives keystrokes, file contents, and network packets as sensory inputs and acts on the environment by displaying on the screen, writing files, and sending network packets.
- We use the term **percept** to refer to the agent's perceptual inputs at any given instant. An agent's **percept sequence** is the complete history of everything the agent has ever perceived.
- In general, an agent's choice of action at any given instant can depend on the entire percept sequence observed to date, but not on anything it hasn't perceived.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Agents and Environments

- An agent's behaviour is described by the **agent function** that maps any given percept sequence to an action. We can imagine tabulating the agent function that describes any given agent; for most agents, this would be a very large table.
- We can construct this table by trying out all possible percept sequences and recording which actions the agent does in response.
- The table is an external characterization of the agent. Internally, the agent function for an artificial agent will be implemented by an **agent program**.
- It is important to keep these two ideas distinct. The agent function is an abstract mathematical description; the agent program is a concrete implementation, running within some physical system.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Agents and Environments – Vacuum-Cleaner World

- To illustrate these ideas, we use a very simple example - the vacuum-cleaner world shown in Figure.
- It has two locations: squares A and B.
- The vacuum agent perceives which square it is in and whether there is dirt in the square.
- It can choose to move left, move right, suck up the dirt, or do nothing.

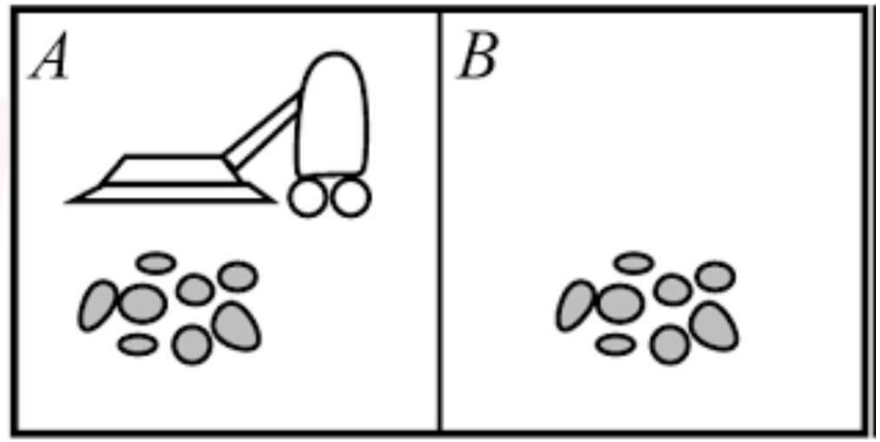


Figure: A vacuum-cleaner world with just two locations

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Agents and Environments – Vacuum-Cleaner World

- One very simple agent function is the following: if the current square is dirty, then suck; otherwise, move to the other square. A partial tabulation of this agent function is shown below.

Percept sequence	Action
[A, *Clean*]	*Right*
[A, *Dirty*]	*Suck*
[B, *Clean*]	*Left*
[B, *Dirty*]	*Suck*
[A, *Clean*], [A, *Clean*]	*Right*
[A, *Clean*], [A, *Dirty*]	*Suck*
⋮	⋮
[A, *Clean*], [A, *Clean*], [A, *Clean*]	*Right*
[A, *Clean*], [A, *Clean*], [A, *Dirty*]	*Suck*
⋮	⋮

Figure: Partial tabulation of a simple agent function for the vacuum-cleaner world

The Concept of Rationality

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Concept of Rationality

- A **rational agent** is one that does the right thing - means in the table for the agent function is filled out correctly.
- Doing the right thing is better than doing the wrong thing. Right action is the one that will cause the agent to be most successful. Therefore, we will need to some way to measure success.
- **Performance measure:** A performance measure embodies the criterion for success of an agent's behaviour. When an agent is plunked down in an environment, it generates a sequence of actions according to the percepts it receives. This sequence of actions causes the environment to go through sequence of states. If the sequence is desirable, then the agent has performed well. This notion of desirability is captured by a **performance measure** that evaluates any given sequence of environment states.

The Concept of Rationality

- There is not one fixed measure suitable for all agents.
- We could ask the agent for the subjective opinion of how happy it is with its own performance, but some agents would be unable to answer, and other would delude themselves.
- **Example: vacuum - cleaner agent :** We might propose to measure performance by the amount of dirt cleaned up in a single eight-hour shift. A rational agent can maximize this performance measure by cleaning up the dirt, then dumping it all on the floor, then cleaning it up again, and so on.
- *As a general rule, it is better to design performance measures according to what one actually wants in the environment, rather than according to how one thinks the agent should behave.*

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Concept of Rationality

- What is rational at any given time depends on four things:
  1. The performance measure that defines the criterion of success.
  2. The agent’s prior knowledge of the environment.
  3. The actions that the agent can perform.
  4. The agent's percept sequence to date.
- This leads to a **definition of a rational agent**:
- *For each possible percept sequence, a rational agent should select an action that is expected to maximize its performance measure, given the evidence provided by the percept sequence and whatever built-in knowledge the agent has.*

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Omniscience, learning and autonomy

- An **omniscient agent** knows the actual outcome of its actions and can act accordingly; but omniscience is impossible in reality.
- Doing actions in order to modify future percepts - sometimes called **information gathering** - is an important part of rationality.
- Definition of a rational agent requires not only to gather information but also to **learn** as much as possible from what it perceives.
- To the extent that an agent relies on the prior knowledge of its designer rather than on its own percepts, we say that the agent lacks **autonomy**.
- A rational agent should be **autonomous** - it should learn what it can to compensate for partial or incorrect prior knowledge.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Nature of Environments

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Specifying the task environment

- Task environments are the "problems" to which rational agents are the "solutions."
- Task environments specify the performance measure, the environment, and the agent's actuators and sensors needed to make an agent.
- We call this the **PEAS description:**
  - **P**erformance measure
  - **E**nvironment
  - **A**ctuators
  - **S**ensors
- In designing an agent, the first step must always be to specify the task environment as fully as possible.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Examples of Agent types and its PEAS description

Agent Type	Performance Measure	Environment	Actuators	Sensors
Taxi driver	Safe, fast, legal, comfortable trip, maximize profits	Roads, other traffic, pedestrians, customers	Steering, accelerator, brake, signal, horn, display	Cameras, sonar, speedometer, GPS, odometer, accelerometer, engine sensors, keyboard

PEAS description of the task environment for an automated taxi

university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Examples of Agent types and its PEAS description

Agent Type	Performance Measure	Environment	Actuators	Sensors
Medical diagnosis system	Healthy patient, reduced costs	Patient, hospital, staff	Display of questions, tests, diagnoses, treatments, referrals	Keyboard entry of symptoms, findings, patient's answers
Satellite image analysis system	Correct image categorization	Downlink from orbiting satellite	Display of scene categorization	Color pixel arrays
Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arm and hand	Camera, joint angle sensors
Refinery controller	Purity, yield, safety	Refinery, operators	Valves, pumps, heaters, displays	Temperature, pressure, chemical sensors
Interactive English tutor	Student's score on test	Set of students, testing agency	Display of exercises, suggestions, corrections	Keyboard entry

Properties of task environments

- A number of dimensions along which task environments can be categorized:
  1. Fully observable vs. partially observable
  2. Single agent vs. multiagent
  3. Deterministic vs. stochastic
  4. Episodic vs. sequential
  5. Static vs. dynamic
  6. Discrete vs. continuous
  7. Known vs. unknown
- These dimensions help determine agent design as well as the applicability of each of the principle family of techniques for agent implementation.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Properties of task environments

- **<u>Fully observable vs. partially observable:</u>**
  - If an agent's sensors give it access to the complete state of the environment at each point in time, then we say that the task environment is **fully observable**.
  - Fully observable environments are convenient because the agent need not maintain any internal state to keep track of the world.
  - An environment might be **partially observable** because of noisy and inaccurate sensors or because parts of the state are simply missing from the sensor data - for example, a vacuum agent with only a local dirt sensor cannot tell whether there is dirt in other squares, and an automated taxi cannot see what other drivers are thinking.
  - If the agent has no sensors at all then the environment is **unobservable**.

Properties of task environments

- **Single agent vs. multiagent:**
  - If only one agent is involved in an environment, and operating by itself then such an environment is called **single agent** environment.
  - However, if multiple agents are operating in an environment, then such an environment is called a **multiagent** environment.
  - For example, an agent solving a crossword puzzle by itself is clearly in a single-agent environment, whereas an agent playing chess is in a two-agent environment.
  - The agent design problems in the multiagent environment are different from single agent environment. A multiagent environment can be **competitive** or **cooperative**.
  - A game of chess is an example of competitive environment and taxi-driving is an example of cooperative environment.

UKA TARSADIA University
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Properties of task environments

- **Deterministic vs. stochastic:**
  - If an agent's current state and selected action can completely determine the next state of the environment, then such environment is called a **deterministic** environment. For example, vacuum cleaner agent.
  - A **stochastic** environment is random in nature and cannot be determined completely by an agent. For example, taxi-driving.
  - In a deterministic, fully observable environment, agent does not need to worry about uncertainty.
  - A **nondeterministic** environment is one in which actions are characterized by their possible outcomes, but no probabilities are attached to them.
  - Nondeterministic environment descriptions are usually associated with performance measures that require the agent to succeed for all possible outcomes of its actions.

Properties of task environments

- **Episodic vs. sequential:**
  - In an **episodic** environment, there is a series of one-shot actions, and only the current percept is required for the action. For example, classification tasks.
  - The agent's experience is divided into atomic episodes. In each episode the agent receives a percept and then performs a single action. The next episode does not depend on the actions taken in previous episodes.
  - However, in **sequential** environment, an agent requires memory of past actions to determine the next best actions. For example, Chess and taxi driving.
  - In sequential environment, the current decision could affect all future decisions.
  - Episodic environments are much simpler than sequential environments because the agent does not need to think ahead.

UKA TARSADIA university
Imparting Knowledge Awakening Wisdom Transforming Lives

Properties of task environments

- **Static vs. dynamic:**
  - If the environment can change while an agent is deliberating, then we say the environment is **dynamic** for that agent; otherwise, it is **static**.
  - Static environments are easy to deal because an agent does not need to continue looking at the world while deciding for an action.
  - However, for dynamic environment, agents need to keep looking at the world at each action.
  - Taxi driving is an example of a dynamic environment whereas Crossword puzzles are an example of a static environment.
  - If the environment itself does not change with the passage of time but the agent's performance score does, then we say the environment is **semi-dynamic**. Game of chess played with clock is an example of it.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Properties of task environments

- **Discrete vs. continuous:**
  - If in an environment there are a finite number of percepts and actions that can be performed within it, then such an environment is called a **discrete** environment; otherwise, it is called **continuous** environment.
  - The discrete/continuous distinction applies to the state of the environment, to the way time is handled, and to the percepts and actions of the agent.
  - A chess game comes under discrete environment as there is a finite number of moves that can be performed. A self-driving car is an example of a continuous environment.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Properties of task environments

- **Known vs. unknown:**
  - Known and unknown are not actually a feature of an environment, but it is an agent's state of knowledge to perform an action.
  - In a **known** environment, the results for all actions are known to the agent. While in **unknown** environment, agent needs to learn how it works to perform an action.
  - It is quite possible that a known environment to be partially observable and an unknown environment to be fully observable.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Examples of task environments and their characteristics

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword puzzle	Fully	Single	Deterministic	Sequential	Static	Discrete
Chess with a clock	Fully	Multi	Deterministic	Sequential	Semi	Discrete
Poker	Partially	Multi	Stochastic	Sequential	Static	Discrete
Backgammon	Fully	Multi	Stochastic	Sequential	Static	Discrete
Taxi driving	Partially	Multi	Stochastic	Sequential	Dynamic	Continuous
Medical diagnosis	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Image analysis	Fully	Single	Deterministic	Episodic	Semi	Continuous
Part-picking robot	Partially	Single	Stochastic	Episodic	Dynamic	Continuous
Refinery controller	Partially	Single	Stochastic	Sequential	Dynamic	Continuous
Interactive English tutor	Partially	Multi	Stochastic	Sequential	Dynamic	Discrete

The Structure of Agents

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

The Structure of Agents

- The job of AI is to design the **agent program** that implements the agent function - mapping percents to actions.
- This will run on some sort of computing device with physical sensors and actuators. We call this the **architecture**.

agent = architecture + program

- **<u>Agent programs:</u>** The agent programs take the current precept as input from the sensors and return an action to the actuators.
  - Notice the difference between the agent program, which takes the current percept as input and the agent function, which takes the entire percept history.
  - The agent program takes just the current percept as input because nothing more is available from the environment; if the agent's actions need to depend on the entire percept sequence, the agent will have to remember the percepts.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Types of agent programs

- There are four basic kinds of agent programs that embody the principles underlying almost all intelligent systems:
  1. Simple reflex agents
  2. Model-based reflex agents
  3. Goal-based agents
  4. Utility-based agents
- Each kind of agent program combines particular components in particular ways to generate actions.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Simple reflex agents

- The simplest kind of agent is the **simple reflex agent**.
- These agents select actions on the basis of the current percept, ignoring the rest of the percept history.
- Imagine yourself as the driver of the automated taxi. If the car in front brakes and its brake lights come on, then you should notice this and initiate braking.
- In other words, some processing is done on the visual input to establish the condition, we call “The car in front is braking.” Then, this triggers some established connection in the agent program to the action “initiate braking.”
- We call such a connection a **condition - action rule**, written as

__if__ car-in-front-is-braking __then__ initiate-braking.

UKA TARSADIA
university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Simple reflex agents

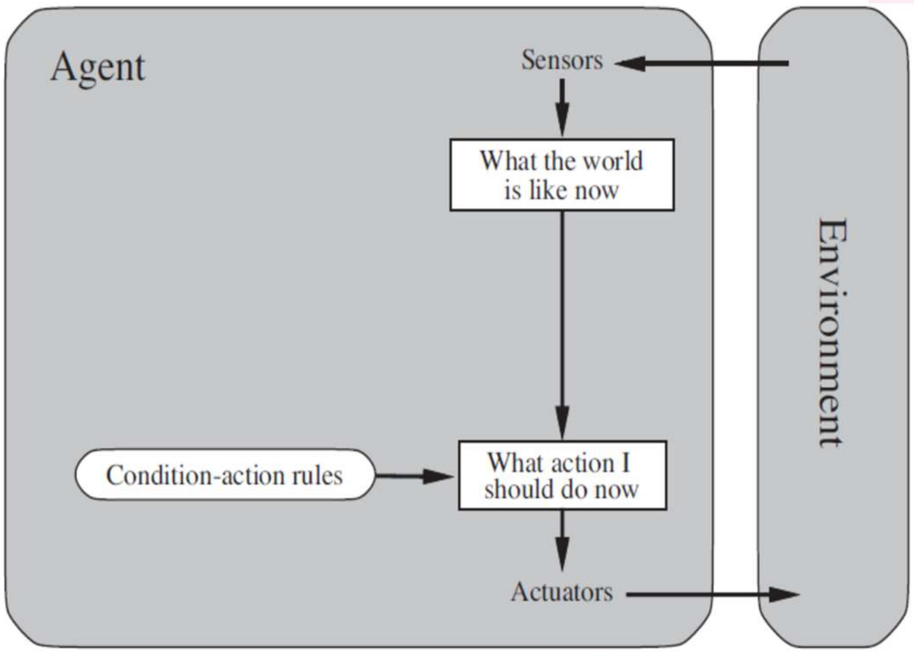


Figure: Schematic diagram of a simple reflex agent

- A more general and flexible approach is first to build a general-purpose interpreter for condition - action rules and then to create rule sets for specific task environments.
- Figure gives the structure of this general program in schematic form, showing how the condition - action rules allow the agent to make the connection from percept to action.
- Simple reflex agents have the admirable property of being simple, but they turn out to be of limited intelligence.

Model-based reflex agents

- Model-based reflex agents maintain **internal state** to track aspects of the world that are not evident in the current percept.
- The most effective way to handle partial observability is for the agent to keep track of the part of the world it can't see now. That is, the agent should maintain some sort of internal state that depends on the percept history and thereby reflects at least some of the unobserved aspects of the current state.
- For the braking problem, the internal state is not too extensive - just the previous frame from the camera, allowing the agent to detect when two red lights at the edge of the vehicle go on or off simultaneously.
- For other driving tasks such as changing lanes, the agent needs to keep track of where the other cars are if it can't see them all at once.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Model-based reflex agents

- Updating this internal state information requires two kinds of knowledge to be encoded in the agent program.
- First, we need some information about how the world evolves independently of the agent.
- Second, we need some information about how the agent's own actions affect the world.
- This knowledge about "how the world works" - whether implemented in simple Boolean circuits or in complete scientific theories - is called a **model** of the world.
- An agent that uses such a model is called a **model-based agent**.

UKA TARSADIA University
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Model-based reflex agents

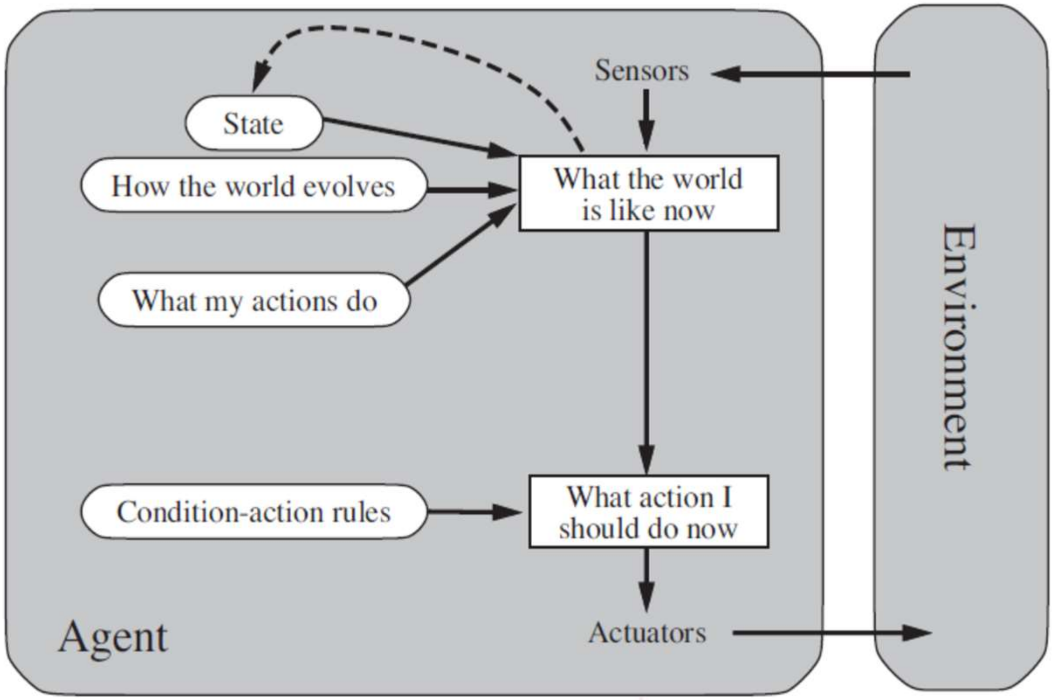


Figure: A model-based reflex agent

- Figure shows the structure of the model-based reflex agent with internal state, showing how the current percept is combined with the old internal state to generate the updated description of the current state, based on the agent's model of how the world works.

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Goal-based agents

- Knowing something about the current state of the environment is not always enough to decide what to do.
- For example, at a road junction, the taxi can turn left, turn right, or go straight on. The correct decision depends on where the taxi is trying to get to.
- In other words, as well as a current state description, the agent needs some sort of **goal** information that describes situations that are desirable - for example, being at the passenger's destination.
- The agent program can combine this with the model (the same information as was used in the model-based reflex agent) to choose actions that achieve the goal.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Goal-based agents

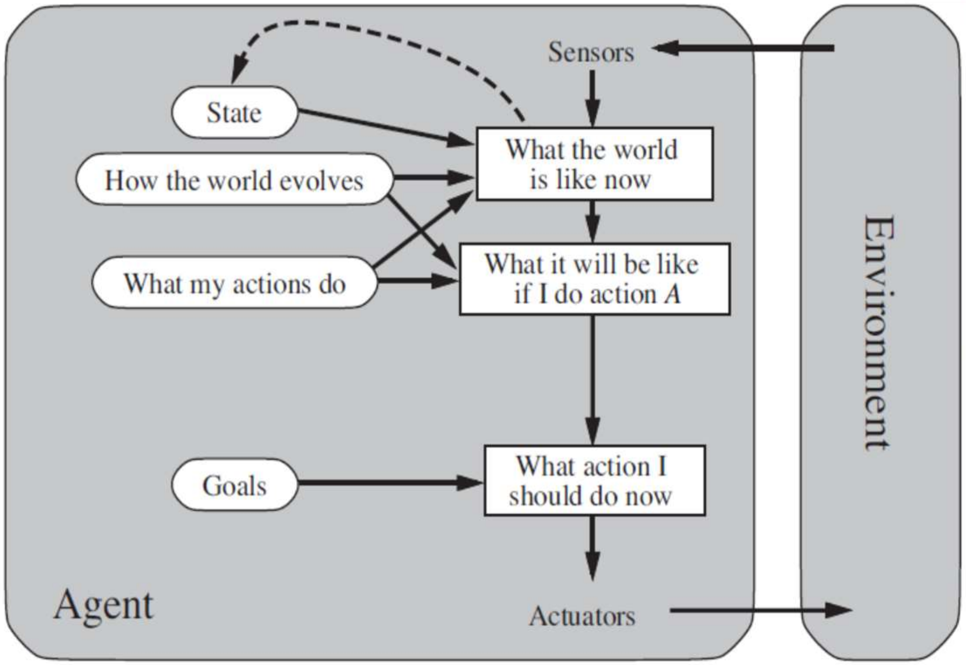


Figure: A model-based, goal-based agent

- Fig. shows goal-based agent's structure.
- It keeps track of the world state as well as a set of goals it is trying to achieve and chooses an action that will lead to the achievement of its goals.
- **Search** and **planning** are the subfields of AI devoted to finding action sequences that achieve the agent's goals.
- Although the goal-based agent appears less efficient, it is more flexible because the knowledge that supports its decisions is represented explicitly and can be modified.

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Utility-based agents

- Goals alone are not enough to generate high-quality behaviour in most environments.
- For example, many action sequences will get the taxi to its destination (thereby achieving the goal) but some are quicker, safer, more reliable, or cheaper than others.
- These agents are similar to the goal-based agent but provide an extra component of utility measurement which makes them different by providing a measure of success at a given state.
- Utility-based agent act based not only goals but also the best way to achieve the goal.
- The Utility-based agent is useful when there are multiple possible alternatives, and an agent has to choose in order to perform the best action.

UKA TARSADIA university
Imparting Knowledge, Awakening Wisdom, Transforming Lives

Utility-based agents

- The utility function maps each state to a real number to check how efficiently each action achieves the goals.
- A complete specification of the utility function allows rational decisions in two kinds of cases where goals are inadequate.
- First, when there are conflicting goals, only some of which can be achieved (for example, speed and safety), the utility function specifies the appropriate trade-off.
- Second, when there are several goals that the agent can aim for, none of which can be achieved with certainty, utility provides a way in which the likelihood of success can be weighed against the importance of the goals.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Utility-based agents

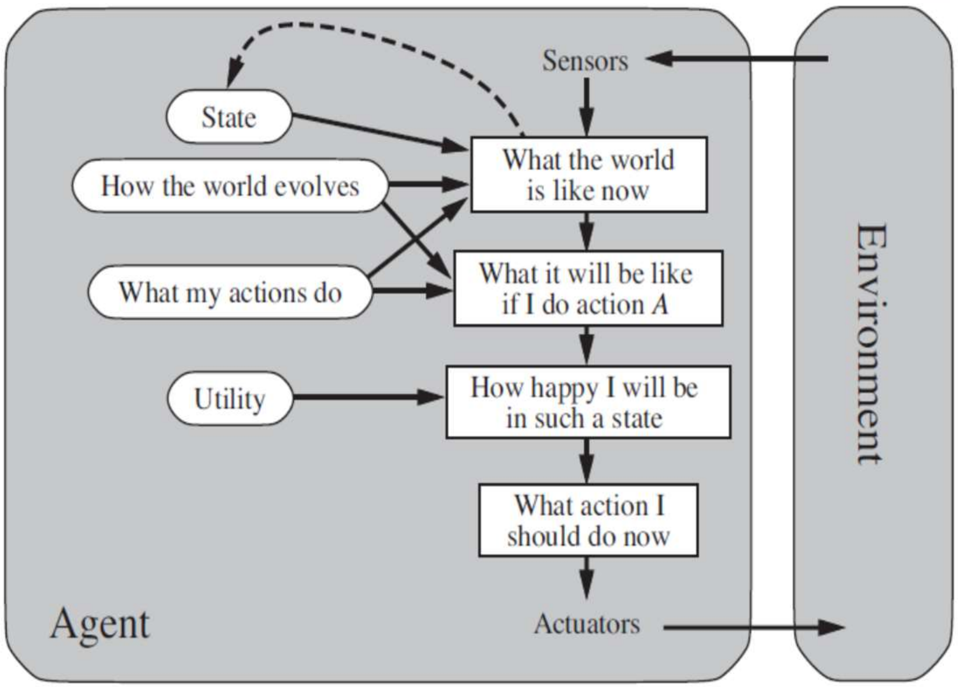


Figure: A model-based, utility-based agent

- Fig. shows utility-based agent's structure.
- It uses a model of the world, along with a utility function that measures its preferences among states of the world.
- Then it chooses the action that leads to the best expected utility, where expected utility is computed by averaging over all possible outcome states, weighted by the probability of the outcome.

Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Learning Agents

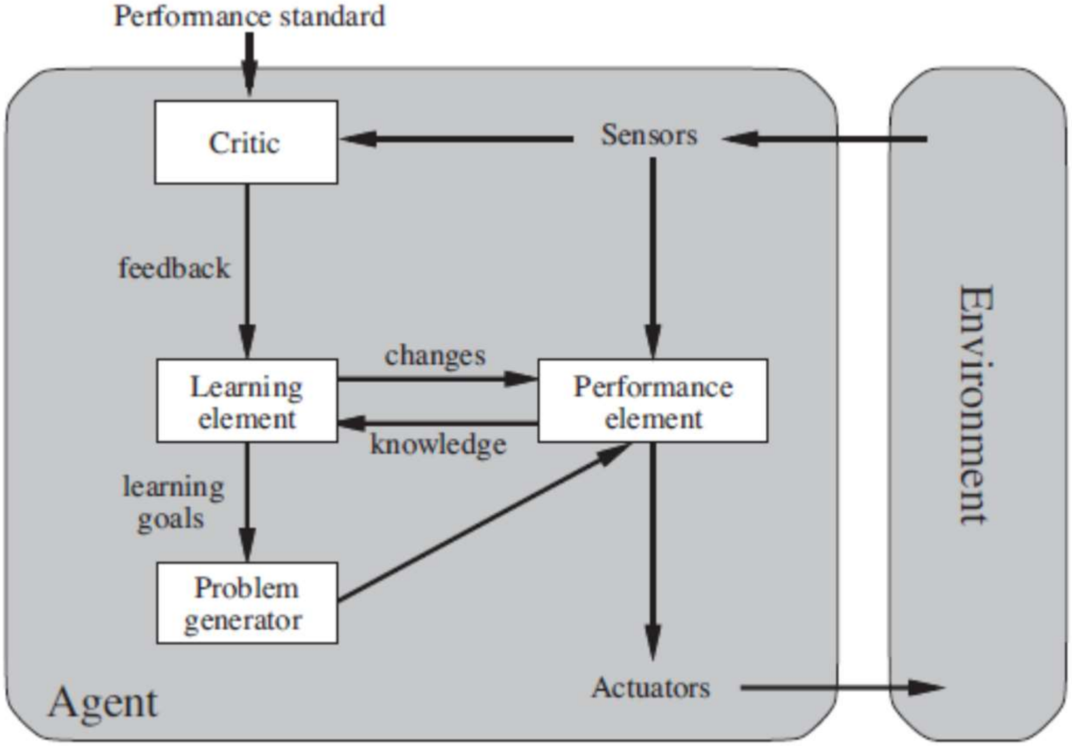


Figure: A general learning agent

- A learning agent in AI is the type of agent which can learn from its past experiences, or it has learning capabilities. It starts to act with basic knowledge and then able to act and adapt automatically through learning.
- A learning agent can be divided into four conceptual elements as shown in figure, which are: **learning element, performance element, critic** and **problem generator**.

Learning Agents

1. **Learning element**: which is responsible for making improvements to the performance elements.
2. **Performance element:** which is responsible for selecting external actions, which is what we have previously considered to be the entire agent: it takes in percepts and decides on actions.
3. **Critic:** which provides feedback to the learning element on how well the performance element is doing, as well as information on how the performance element should be modified to do better in the future.
4. **Problem generator**: which is responsible for suggesting actions that will lead to new and informative experiences.

- Hence, learning agents are able to learn, analyze performance, and look for new ways to improve the performance.

UKA TARSADIA university
Imparting Knowledge. Awakening Wisdom. Transforming Lives.

Thank You

UKA TARSADIA university

Imparting Knowledge. Awakening Wisdom. Transforming Lives.