Subject : **ARTIFICIAL INTELLIGENCE**

Sub code : IS52

Credits : 2:0:1

Faculty : Kavya K S

Text Book:

1. Stuart Russel, Peter Norvig: Artificial Intelligence - A Modern Approach, 3rd Edition, Pearson Education, 2012.
2. Elaine Rich, Kevin Knight, Shivashankar B Nair: Artificial Intelligence, 3rd Edition, Tata McGraw Hill, 2011

References:

1. Peter Jackson, "Introduction to Expert Systems", 3rd Edition, Pearson Education, 2007.
2. Deepak Khemani "Artificial Intelligence", Tata Mc Graw Hill Education 2013.
3. http://nptel.ac.in

UNIT-I

Introduction: Definition of AI. Foundation of Artificial Intelligence. **Intelligent Agents:** Agents and Environments, Rationality, The Nature of Environments, The Structure of Agents.

Problem-solving by search: Problem-Solving Agents, Uninformed Search Strategies: Bidirectional Search. Informed Search Strategies: A* Search, Heuristic Functions.

UNIT II

Adversarial Search: Games, Optimal Decisions in Games, Alpha Beta Pruning, Imperfect Real-Time Decision. **Logical Agents:** Knowledge-Based Agents, The Wumpus World, Logic,

Propositional Logic: A very simple logic, Effective Propositional Model Checking, Agents Based on Propositional Logic. **First Order Logic:** Wumpus World representation, Knowledge Engineering in First-Order Logic.

UNIT III

Interference in First-order Logic: Propositional vs. First-Order Inference, Unification and Lifting, Forward chaining, Backward Chaining. Resolution. **Classical Planning:** Definition, Algorithms for Planning as State-Space Search, Planning Graphs, Other Planning Approaches.

UNIT IV

Knowledge Representation: ontological engineering, categories and objects, Events, mental objects and modal logic, Reasoning systems for categories. **Uncertainty**: Acting under Uncertainty, Inference using Full Joint Distributions, Independence.

UNIT V

Uncertainty: The Wumpus World Revisited, Learning from Examples: Forms of Learning. **Robotics:** Introduction, Hardware, Perception, Planning to Move, Planning Uncertain Movement, Moving, Robotic Software Architecture, Application Domains.

COURSE OUTCOMES:

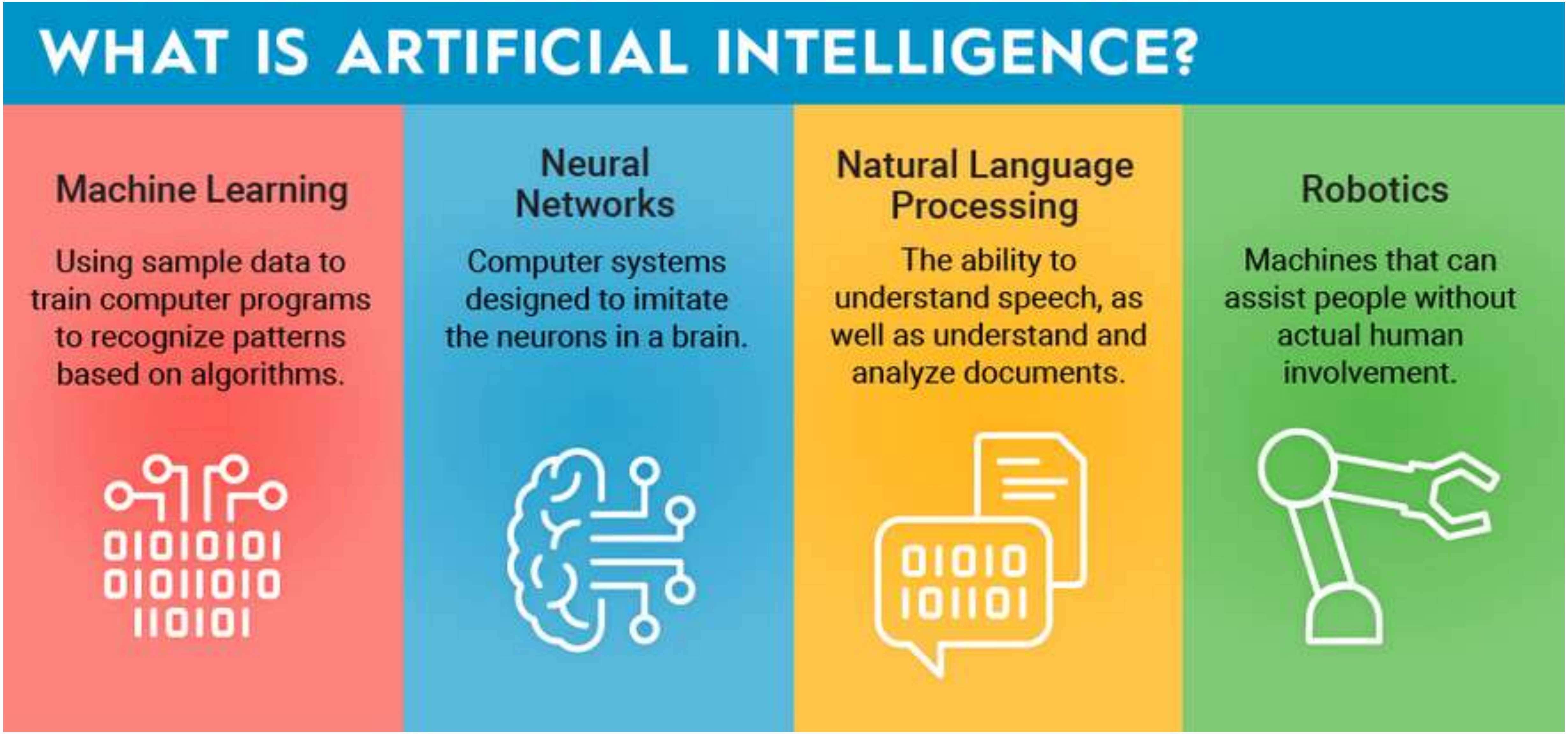
1. Identify the fundamental characteristics and challenging issues of Artificial Intelligence (AI) systems (PO-1,2,3,4,12, PSO-2)

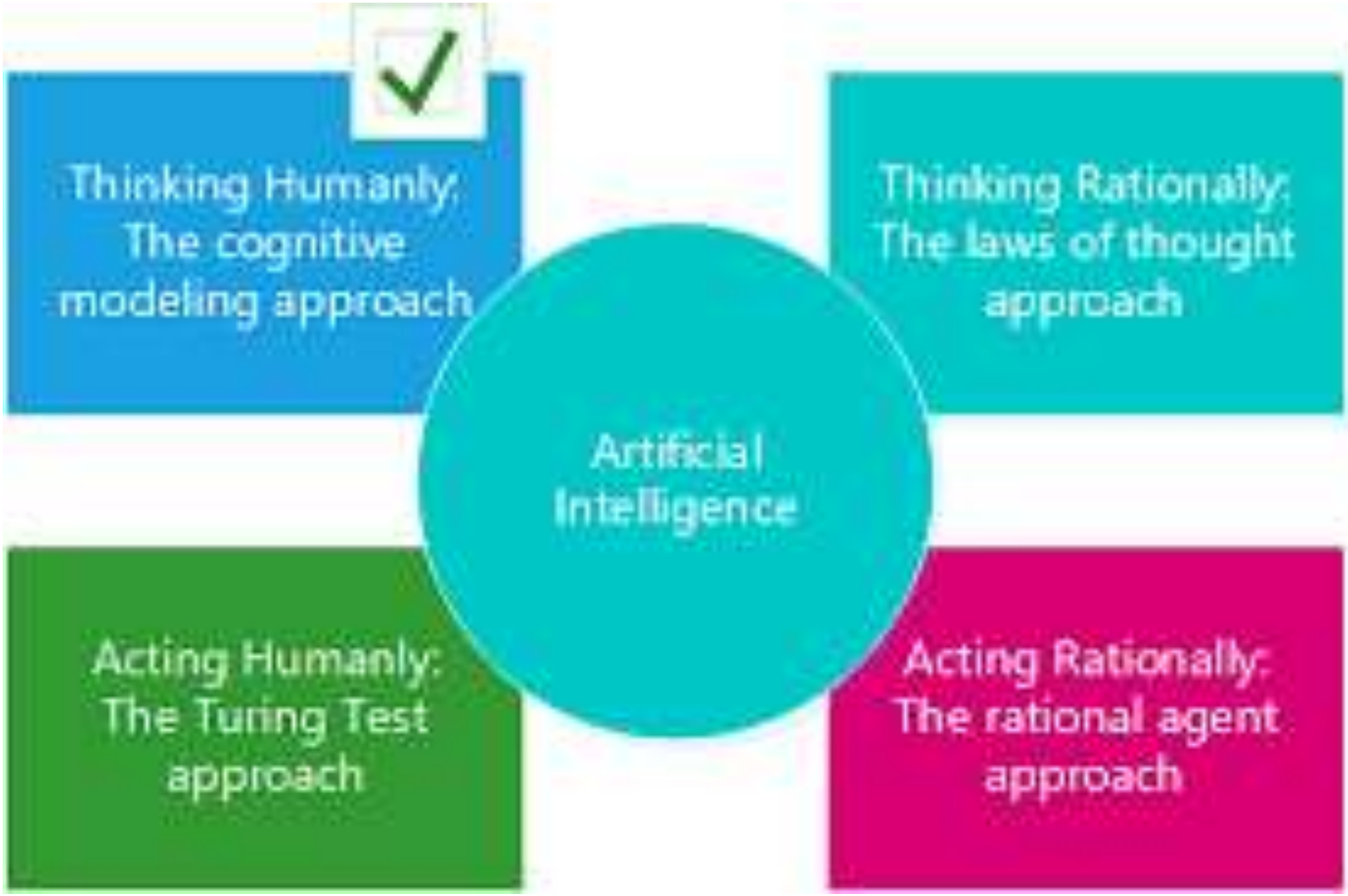
2. Apply various general purpose search algorithm as solutions for various problem-solving agents (PO -1,2,3,4,12, PSO-2)

3. Apply various symbolic knowledge representation to specify domains and reasoning tasks of a situated intelligent agent. (PO 4,5,9,12, PSO-2)

4. Apply algorithmic approach for planning and solving AI solutions that require problem solving, inference, perception, knowledge representation, and learning. (PO-1,4,5,6,7,PSO-2)

5. Extract conclusions on learning and quantify the uncertainty in the conclusions obtained from uncertain knowledge. (PO-5,9, PSO-2)

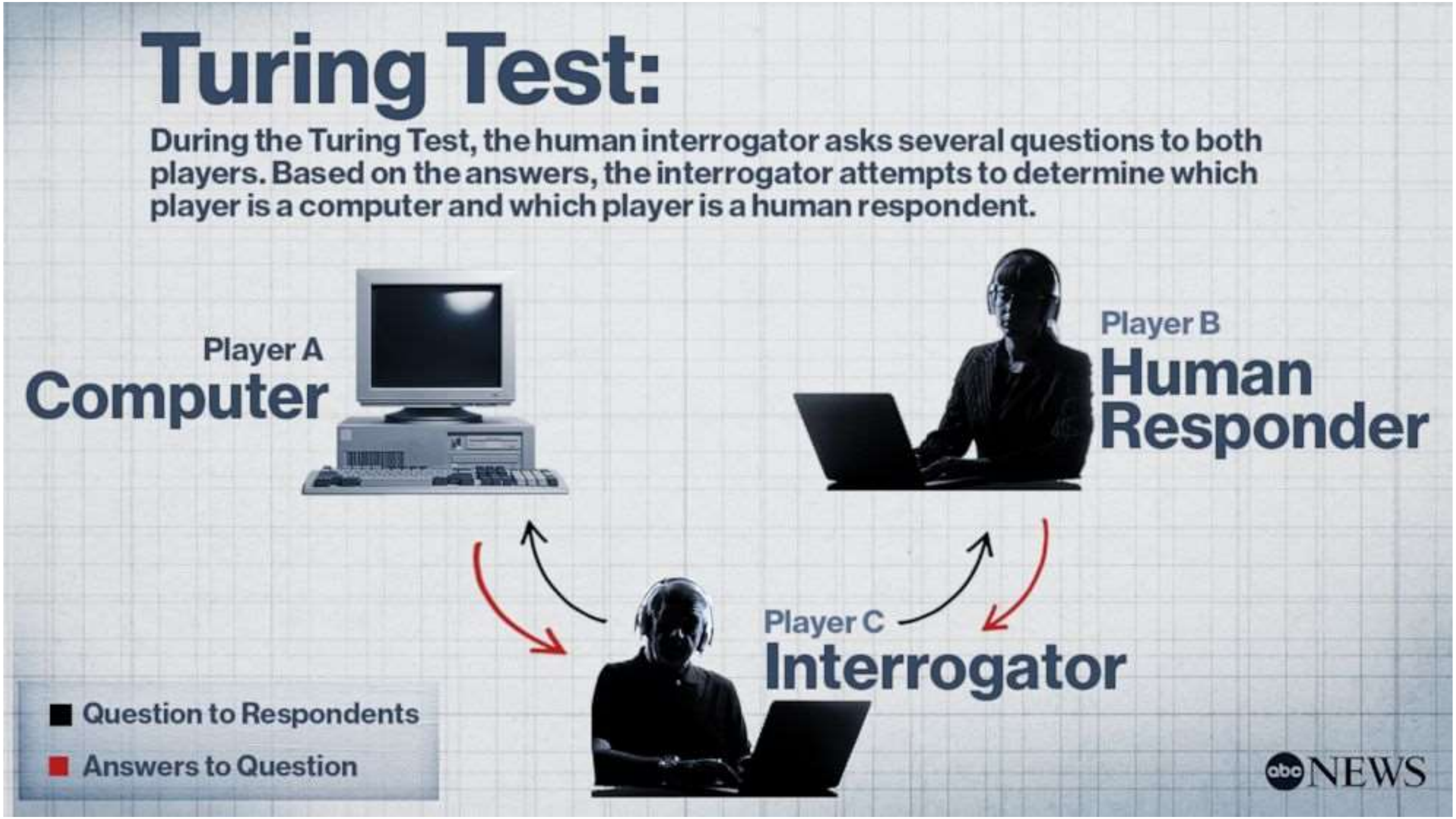

WHAT IS ARTIFICIAL INTELLIGENCE?
Machine Learning
Using sample data to train computer programs to recognize patterns based on algorithms.
01010101
01011010
110101
Neural Networks
Computer systems designed to imitate the neurons in a brain.
Natural Language Processing
The ability to understand speech, as well as understand and analyze documents.
01010
101101
Robotics
Machines that can assist people without actual human involvement.


Thinking Humanly: The cognitive modeling approach
Thinking Rationally: The laws of thought approach
Artificial Intelligence
Acting Humanly: The Turing Test approach
Acting Rationally: The rational agent approach

All four approaches to AI have been followed, each by different people with different methods.

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 1.1 Some definitions of artificial intelligence, organized into four categories.


Turing Test:
During the Turing Test, the human interrogator asks several questions to both players. Based on the answers, the interrogator attempts to determine which player is a computer and which player is a human respondent.
Player A
Computer
Player B
Human Responder
Player C
Interrogator
Question to Respondents
Answers to Question
abc NEWS

I Acting humanly: The Turing Test approach

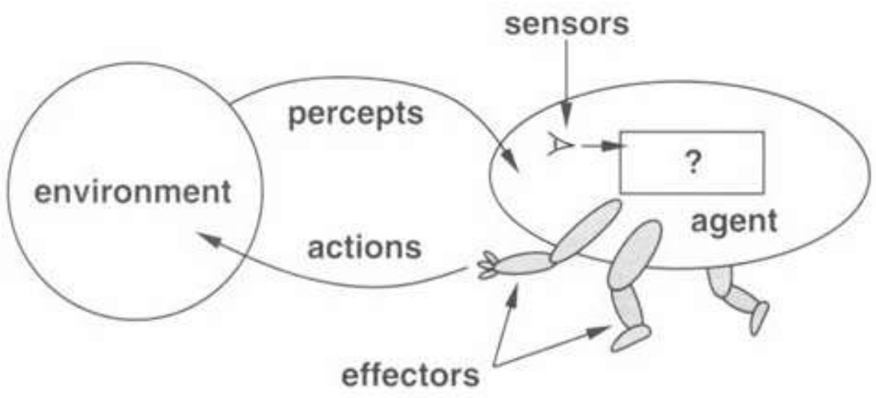
- The Turing Test, proposed by Alan Turing (1950), was designed to provide a satisfactory operational definition of intelligence. A computer passes the test or a human interrogator, after posing some written questions, cannot tell whether the written responses came from a person or from a computer.

The computer would need to possess the following capabilities:

- **natural language processing** to enable it to communicate successfully in English
- **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

Total Turing Test: Includes a video signal so that the interrogator can test the subject's perceptual abilities, as well as the opportunity for the interrogator to - pass physical objects "through the hatch." To pass the total Turing Test, the computer will need

Total Turing Test



- Computer Vision (or other senses)
- Computer manipulation
- Robot
  - A machine able to extract information from its environment and use knowledge about its world to move safely in a meaningful manner
  - Physically embodied Intelligent Agent
    - As opposed to a "brain-in-a-box"

II Thinking humanly: The cognitive modeling approach

To determine how humans think, one needs to get *inside* the actual workings of human minds.

Following are the 3 ways to do this:

1. Through introspection—trying to catch our own thoughts as they go by;
2. Through psychological experiments—observing a person in action;
3. Through brain imaging—observing the brain in action.

Example, Allen Newell and Herbert Simon, who developed GPS, the “General Problem Solver”.

III Thinking rationally: The "laws of thought" approach

- The Greek philosopher Aristotle was one of the first to attempt to codify "right thinking," that is, irrefutable reasoning processes. His **syllogisms** provided patterns for argument structures that always yielded correct conclusions when given correct premises.

- For example, "Socrates is a man; all men are mortal; therefore, Socrates is mortal."

- These laws of thought were supposed to govern the operation of the mind; their study initiated the field called **logic**.

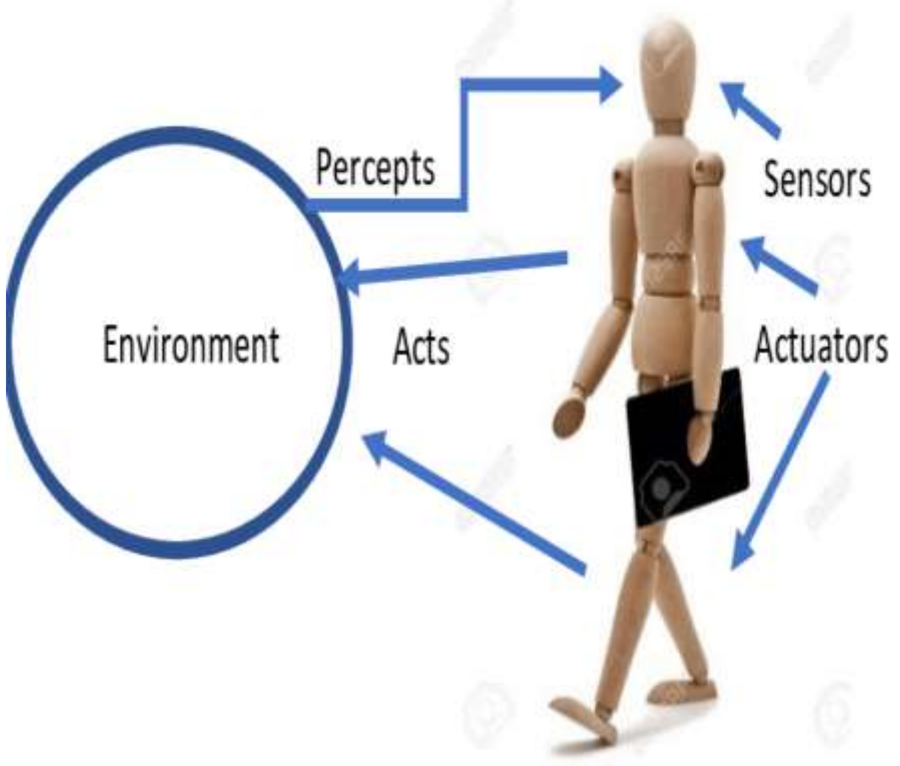
- The so-called **logicist** tradition within artificial intelligence hopes to build on such programs to create intelligent systems.

Two obstacles with the approach of defining AI as building rationally thinking agents:

1. When knowledge is not 100% certain, it is not easy to take informal knowledge and state it in the formal terms.
2. There is a big difference between solving a problem “in principle” and solving it in practice. Even problems with just a few hundred facts can exhaust the computational resources of any computer unless it has some guidance as to which reasoning steps to try first.

IV Acting rationally: The rational agent approach:

- An agent is just something that acts.
- All computer programs do something, but computer agents are expected to do more: operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, and create and pursue goals.
- A **rational agent** is one that acts so as to achieve the best outcome or, when there is uncertainty, the best-expected outcome.



- In the **“laws of thought”** approach to AI, the emphasis was on correct inferences.
- Making correct inferences is sometimes *part* of being a rational agent because one way to act rationally is to reason logically to the conclusion that a given action will achieve one’s goals and then act on that conclusion.

The rational-agent approach has **two advantages** over the other approaches.

- First, it is more general than the “laws of thought” approach because correct inference is just one of several possible mechanisms for achieving rationality.
- Second, it is more amenable (acceptance) to scientific development than are approaches based on human behavior or human thought.

1.2 The foundations of AI

1. Philosophy
2. Mathematics & Statistics
3. Economics
4. Neuroscience
5. Psychology
6. Computer science & engineering
7. Control Theory & 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 be used to draw valid conclusions.
- With out philosophy it is difficult to answer the following questions.
- How does the mind arise from a physical brain?
- Where does the knowledge come from?
- How does the knowledge lead to action?

Mathematics & Statistics

- AI required Formal Logic and Probability for planning and learning.
- Computation required for analyzing relation, and implementation.
- Knowledge in Formal Representation are most required for writing actions for agents.
- In AI, the Mathematics & Statistics are most important for
  - Proving theorems,
  - Writing algorithms,
  - computation,
  - decidability,
  - tractability,
  - modeling uncertainty,
  - learning from data.
- What are the formal rules to draw valid conclusions?
- What can be computed?
- How do we reason with uncertain information?

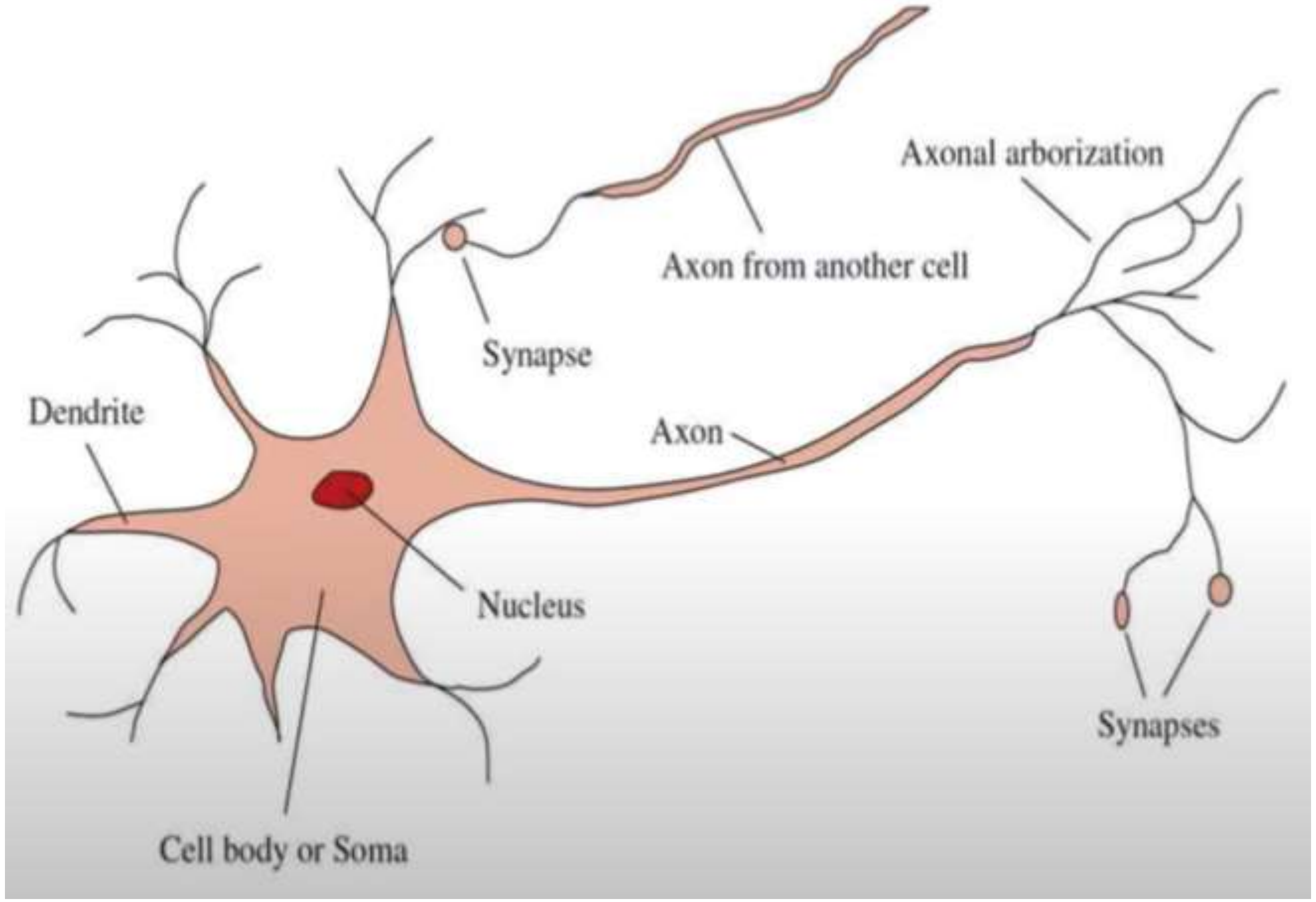
Economics

- Deals with investing the amount of money and Maximization of utility with minimal 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, and etc.

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.*
- How do brains process information?

Nerve cell or Neurons:



- **Dendrites**–A branch-like structure that **functions by receiving messages** from other neurons and allow the transmission of messages to the cell body.
- **Axon**–Axon is a tube-like structure that functions by **carrying an electrical impulse** from the cell body to the axon terminals for passing the impulse to another neuron.
- **Synapse**– This structure functions by permitting the entry of a neuron to move an **electrical or chemical signal** from one neuron to another neuron.

Psychology/Cognitive Science

- The scientific method to the study of human vision.
- Problem solving skills,
- how do people behave,
- perceive,
- Process cognitive information, and
- represent knowledge.
- How do humans and animals think and act?

Computer Science & Engineering

- Logic and inference theory, algorithms, programming languages, and system building are important parts of Computer science.
- Computer hardware gradually changed for AI applications, such as the graphics 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 the utilization doubled every 100 days.
- **The Super computers and quantum computers** can solve very complicated AI problems
- the software side of computer science, supplied the operating systems, programming languages, and tools needed to write modern programs

Computer Science & 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 concepts of symbolic, functional, declarative, and object-oriented programming.
- How can we build fast and efficient computer?

Control theory

- Control theory helps the system to analyze, define, debug and fix errors by itself.
- Developing self-controlling machine, self-regulating feedback control systems and the submarine are some examples of control theory
- Calculus and matrix algebra, and the tools of control theory, provide themselves to systems that are describable by **fixed sets of continuous variables**, are foundation of AI

Control theory …

- Knowledge representation, grammars, computational linguistics or natural language processing (NLP) are significant to developing AI applications.
- The tools of logical inference and computation provide the language, vision, and symbolic planning of agent programming.
- How can artifacts operate under their own control?

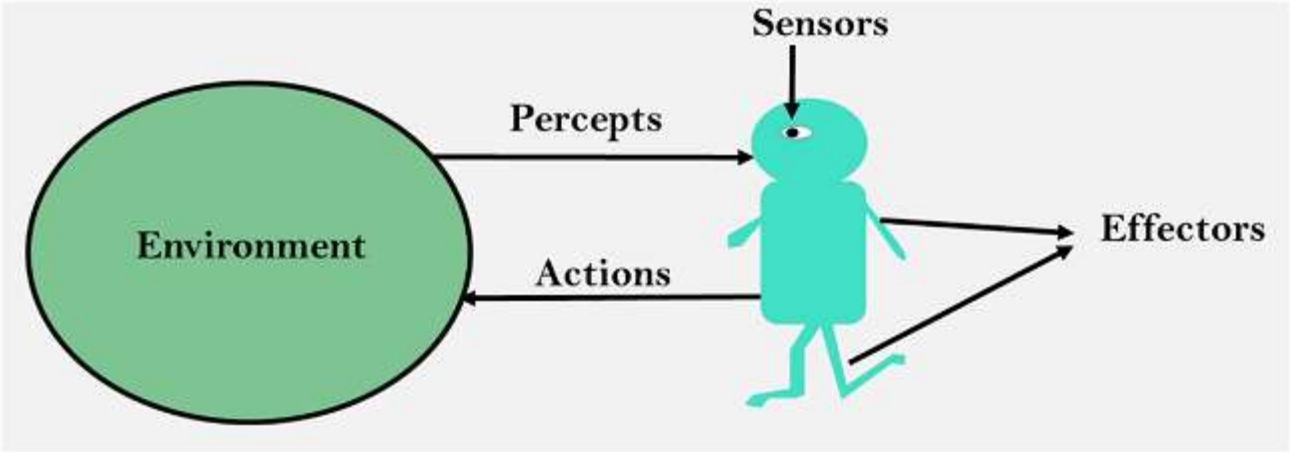
Linguistics

- 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 on the basis of that command, a computer can perform a specific task.
- It includes **Speech to text, Text to speech.**
- How does language relate to thought?

Intelligent Agents: Agents and Environment

- An agent can be anything that perceive its 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.

- **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.

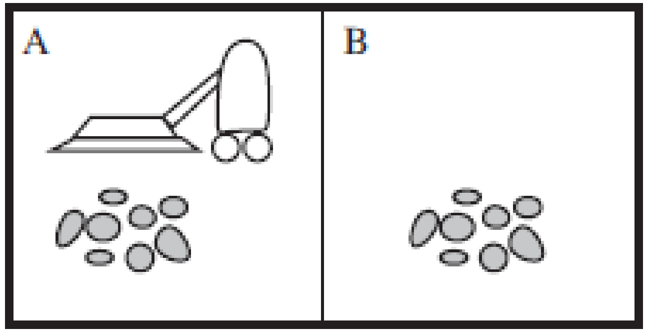
Rationality

What is rational at any given time depends on four things:

- The performance measure that defines the criterion of success.
- The agent’s prior knowledge of the environment.
- The actions that the agent can perform.
- The agent’s percept sequence to date.

A very simple example—the **vacuum-cleaner** world: This particular world has just 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.

- One very simple agent function is the following: if the current square is dirty, then suck up the dirt; otherwise, move to the other square.
- A partial tabulation of this agent function is shown in Figure 2.3 and an agent program that implements it



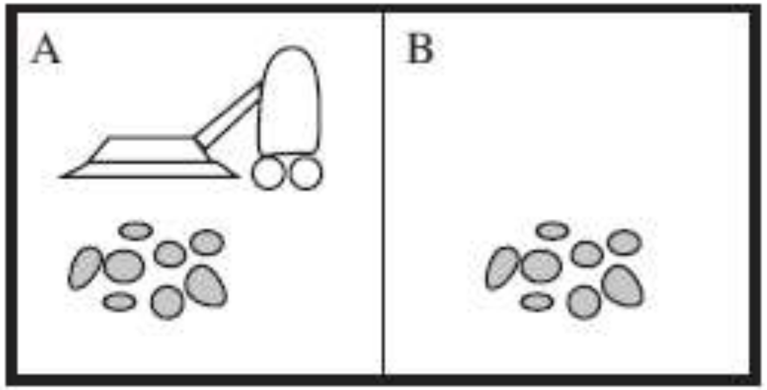


Figure 2.2 A vacuum-cleaner world with just two locations.

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*
$\vdots$	$\vdots$
$[A, Clean], [A, Clean], [A, Clean]$	*Right*
$[A, Clean], [A, Clean], [A, Dirty]$	*Suck*
$\vdots$	$\vdots$

Figure 2.3 Partial tabulation of a simple agent function for the vacuum-cleaner world shown in Figure 2.2.

Consider the simple vacuum-cleaner agent that cleans a square if it is dirty and moves to the other square if not.

Q. Is this a rational agent?

A. First, we need to say what the performance measure is, what is known about the environment, and what sensors and actuators the agent has.

Let us assume the following:

- The performance measure awards one point for each clean square at each time step
- Known – geography of environment
- Unknown - dirt distribution and the initial location of the agent are not.

- Clean squares stay clean and sucking cleans the current square. The Left and Right actions move the agent left and right except when this would take the agent outside the environment, in which case the agent remains where it is.
- The only available actions are Left , Right, and Suck.
- The agent correctly perceives its location and whether that location contains dirt.
- We claim that under these circumstances the agent is indeed rational; its expected performance is at least as high as any other agent’s

- For example, once all the dirt is cleaned up, the agent will oscillate needlessly back and forth; if the performance measure includes a penalty of one point for each movement left or right, the agent will fare poorly.
- A better agent for this case would do nothing once it is sure that all the squares are clean. If clean squares can become dirty again, the agent should occasionally check and re-clean them if needed. If the geography of the environment is unknown, the agent will need to explore it rather than stick to squares A and B.

Omniscience, learning, and autonomy:

- An omniscient agent knows the actual outcome of its actions and can act accordingly; but omniscience is impossible in reality.
- This example shows that rationality is not the same as perfection. Rationality maximizes expected performance, while perfection maximizes actual performance.
- **Learning:** Our definition requires a rational agent 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 do to compensate for partial or incorrect prior knowledge.

Specifying the task environment:

- The rationality of the simple vacuum-cleaner agent, we had to specify the performance measure, the environment, and the agent's actuators and sensors. We group all these under the heading of the **task environment**. We call this as PEAS (**P**erformance, **E**nvironment, **A**ctuators, **S**ensors).
- The vacuum world was a simple example; let us consider a more complex problem: an automated taxi driver.
- Figure 2.4 summarizes the PEAS description for the taxi's task environment.

PEAS environment for Automatic Taxi

AGENT TYPE	PERFORMANCE	ENVIRONMENT	ACTUATORS	SENSORS
Taxi Driver	Safe, fast, legal, comfortable trip, maximise profits	Roads, other traffic, pedestrians, customers	Steering, accelerator, brake, signal, horn, display	Camera, sonar, speedometer, GPS, odometer, accelerator, engine sensors, keyboard

AGENT TYPE	PERFORMANCE	ENVIRONMENT	ACTUATORS	SENSORS
Medical diagnosis system	Healthy patient, reduced costs, accuracy	Hospital, clinic, patients, staff, doctors	Display, MRI, Scan, ECG	Temp sensor, pressure
Satellite image analysis system	Clarity, object detection, coverage	Space, orbiting satellite, weather	Controller	Camera
Part-picking robot				
Refinery controller				
Interactive English tutor				

Properties of task environments

- Fully observable vs. partially observable
- Single agent vs. multi-agent
- Deterministic vs Stochastic
- Episodic vs. sequential
- Static vs Dynamic
- Discrete vs Continuous
- Known vs Unknown

Task Environment	Observable	Agents	Deterministic	Episodic	Static	Discrete
Crossword Puzzle	Fully	Single	Yes	No	Static	Yes
Chess with a clock	Fully	Multi	Yes	Yes	No	Yes
Poker	Partially	Multi	No	Yes	Yes	Yes
Backgammon						
Taxi driving	Partially		No			
Medical Diagnosis						
Image analysis						
Part-picking robot						
Refinery controller						
Interactive English tutor						

The Structure of Agents:

- The job of AI is to design an **agent program** that implements the agent function—the mapping from perceptions to actions.
- We assume this program will run on some sort of computing device with physical sensors and actuators—called the **architecture**:

agent = architecture + program

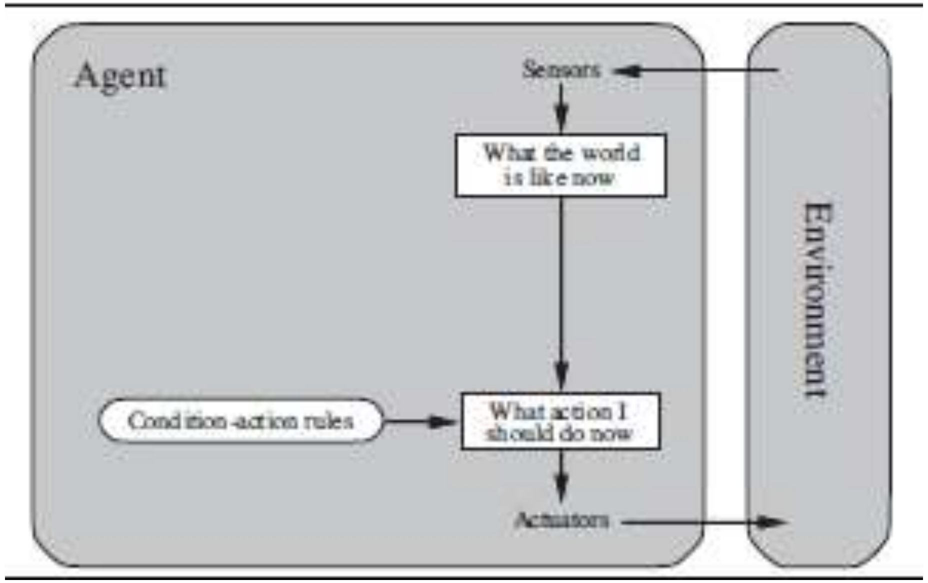
- **Agent programs:**The agent program take the current percept as input from the sensors and return an action to the actuators. The difference between the agent program, which takes the current percept as input, and the agent function, which takes the entire percept history.

- **Simple reflex agents:** These agents select actions on the basis of the *current* percept, ignoring the rest of the percept history. For example, the vacuum agent, because its decision is based only on the current location and on whether that location contains dirt.

```
function REFLEX-VACUUM-AGENT([location,status]) returns an action
  if status = Dirty then return Suck
  else if location = A then return Right
  else if location = B then return Left
```

```
function SIMPLE-REFLEX-AGENT(percept) returns an action
  persistent: rules, a set of condition–action rules

  state ← INTERPRET-INPUT(percept)
  rule ← RULE-MATCH(state, rules)
  action ← rule.ACTION
  return action
```



- The reflex agent will work only if the correct decision can be made on the basis of only the current percept—that is, only if the environment is **fully observable**.
- Suppose that a simple reflex vacuum agent is deprived of its location sensor and has only a dirt sensor. Such an agent has just two possible percepts: [Dirty] and [Clean]. It can Suck in response to [Dirty];
- What should it do in response to [Clean]? Moving Left fails (forever) if it happens to start in a square A, and moving Right fails (forever) if it happens to start in square B. Infinite loops are often unavoidable for simple reflex agents operating in partially observable environments.

Model-based reflex agents:

- 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.
- 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.

- When the agent turns the steering wheel clockwise, the car turns to the right, or after driving for five minutes north on the freeway, one is usually about five miles north of where one was five minutes ago.
- 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**.

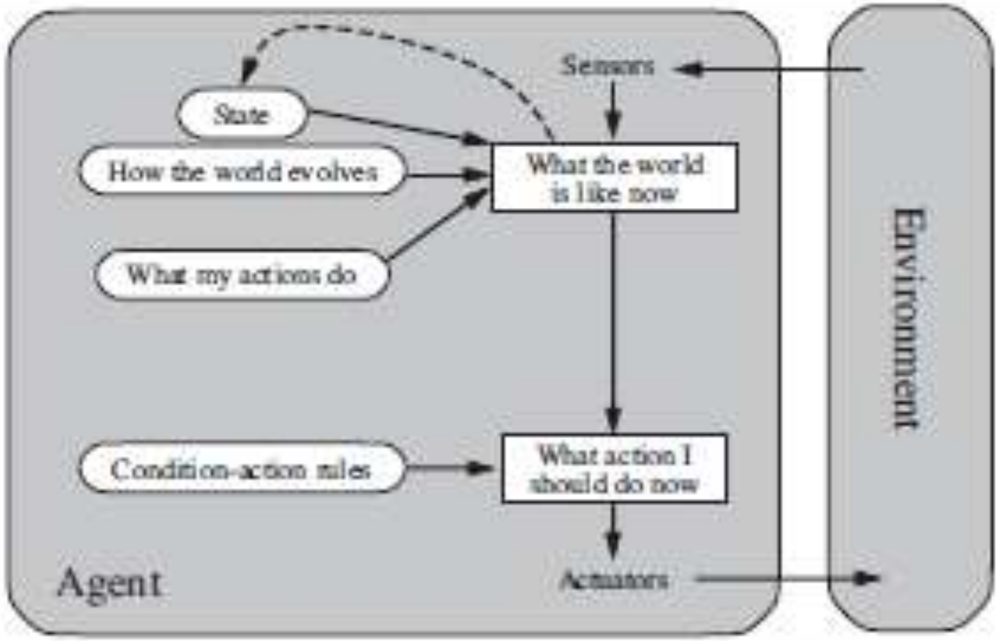


Figure 2.11 A model-based reflex agent.

function MODEL-BASED-REFLEX-AGENT(*percept*) **returns** an action
  persistent: *state*, the agent's current conception of the world state
    model, a description of how the next state depends on current state and action
    rules, a set of condition–action rules
    action, the most recent action, initially none

  state ← UPDATE-STATE(*state*, *action*, *percept*, *model*)
  rule ← RULE-MATCH(*state*, *rules*)
  action ← *rule*.ACTION
  return *action*

Figure 2.12 A model-based reflex agent. It keeps track of the current state of the world, using an internal model. It then chooses an action in the same way as the reflex agent.

- **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.
- Sometimes goal-based action selection is straightforward or sometimes tricky (when the agent has to consider long sequences of twists and turns in order to find a way to achieve the goal)

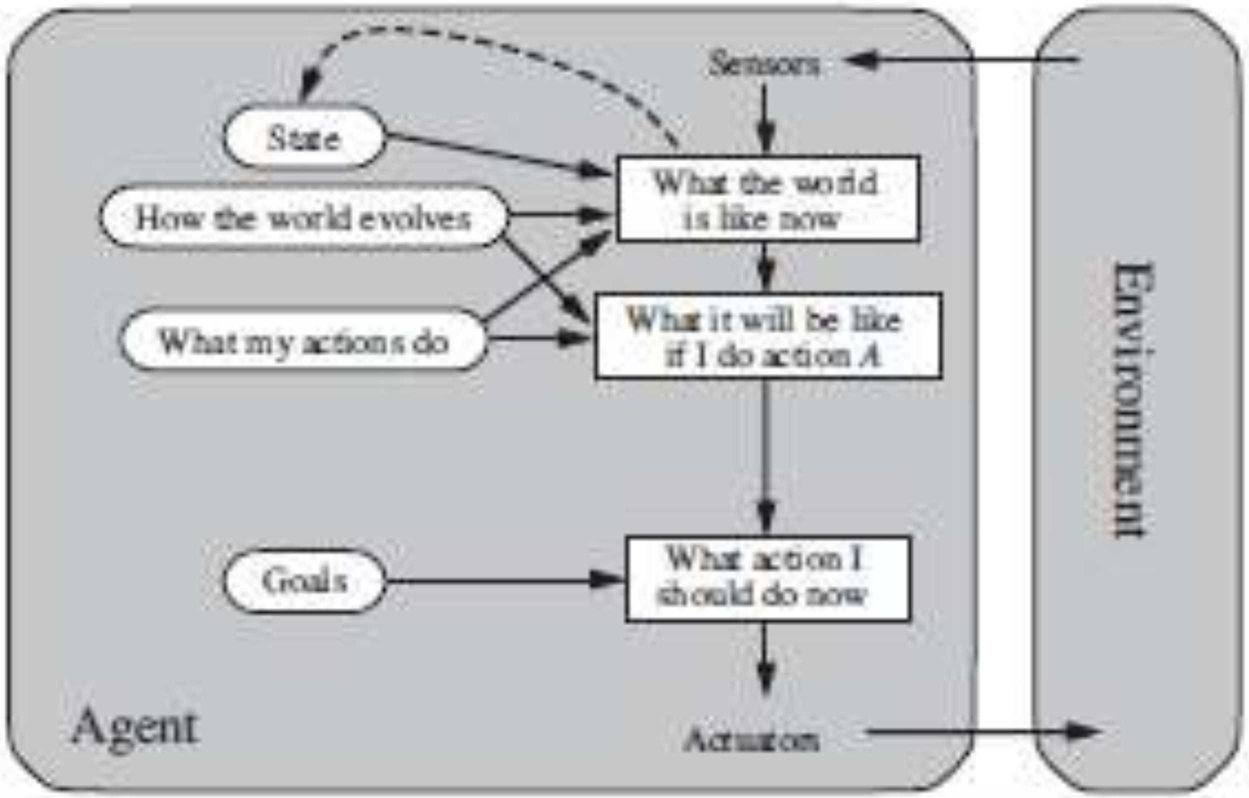


Figure 2.13 A model-based, goal-based agent. It keeps track of the world state as well as a set of goals it is trying to achieve, and chooses an action that will (eventually) lead to the achievement of its goals.

- **Utility-based agents:** Goals alone are not enough to generate high-quality behavior 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.
- More general performance measure should allow a comparison of different world states according to exactly how happy they would make the agent. Because “happy” does not sound very scientific, economists and computer scientists use the term **utility** instead.

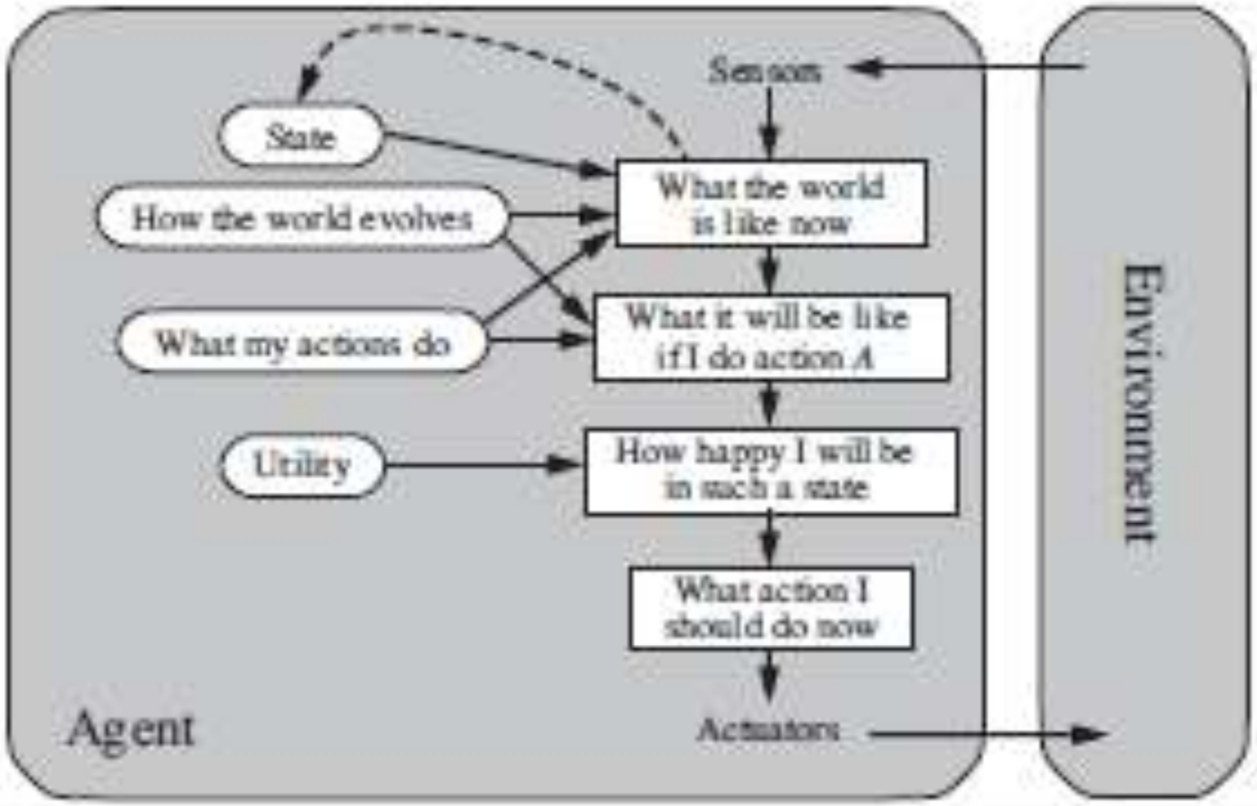


Figure 2.14 A model-based, utility-based agent. It uses a model of the world, along with a utility function that measures its preferences among states of the world. Then it chooses the action that leads to the best expected utility, where expected utility is computed by averaging over all possible outcome states, weighted by the probability of the outcome.

Learning agents:

A learning agent can be divided into four conceptual components:

- **learning element:** responsible for making improvement
- **performance element: r**esponsible for selecting external actions.
- **Critic:** The learning element uses feedback from the **critic** on how the agent is doing and determines how the performance element should be modified to do better in the future.
- **problem generator:** responsible for suggesting actions that will lead to new and informative experiences.

- Ex: Automated Taxi: The performance element consists of whatever collection of knowledge and procedures the taxi has for selecting its driving actions.
- The taxi goes out on the road and drives, using this performance element
- The critic observes the world and passes information along to the learning element. For example, after the taxi makes a quick left turn across three lanes of traffic, the critic observes and learning element is able to formulate a rule saying this was a bad action, and the performance element is modified by installation of the new rule
- The problem generator might identify certain areas of behavior in need of improvement and suggest experiments

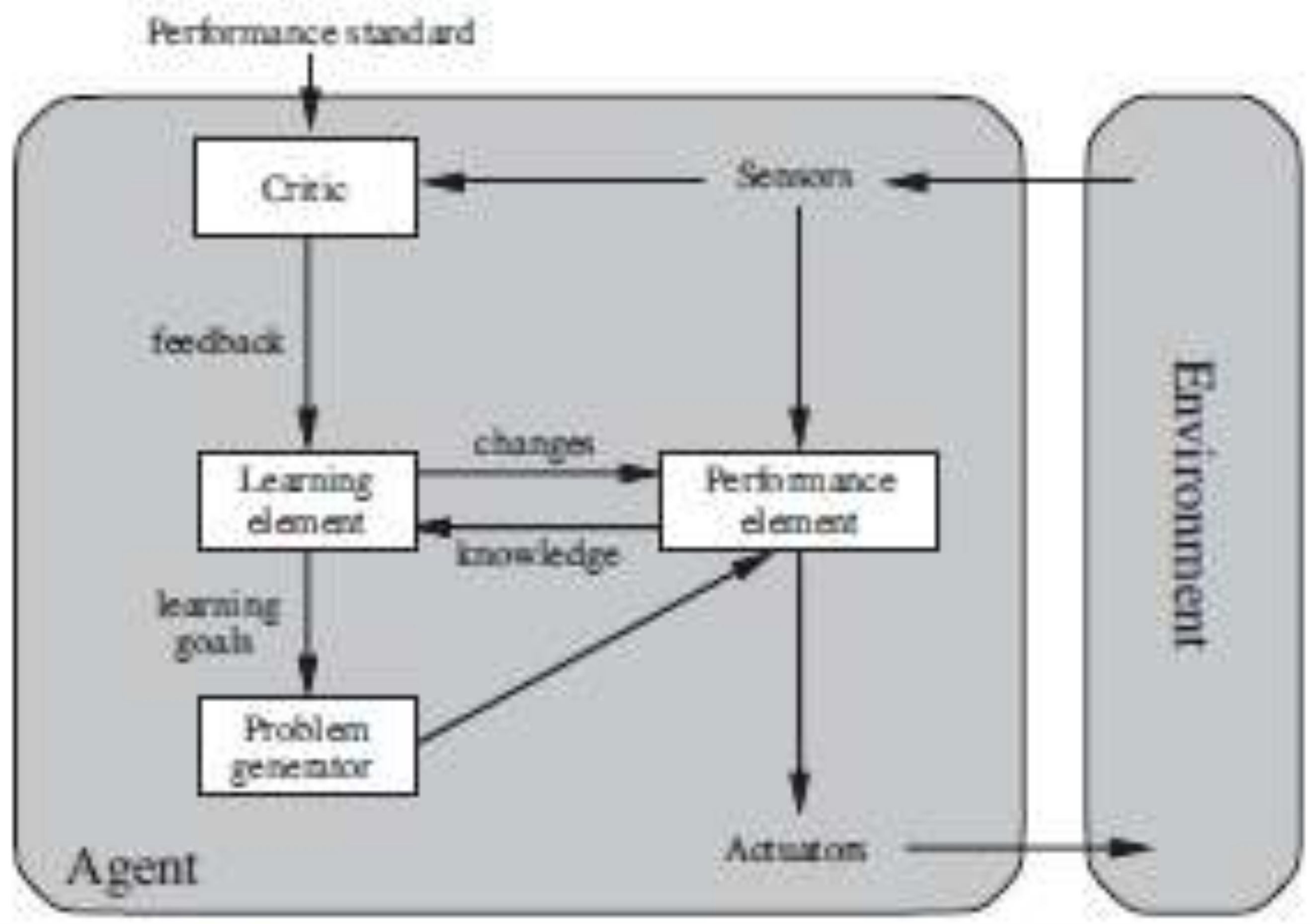


Figure 2.15 A general learning agent.