Chapter 2: Agents

2.1 Introduction to agents

An agent can be anything that perceives its environment through sensors and acts 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, and 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.

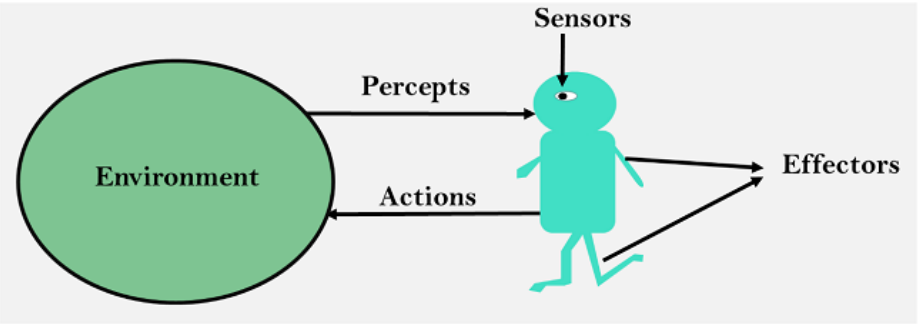
Hence the world around us is full of agents such as thermostat, cellphone, camera, and even we are also agents.

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 components of machines that convert 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.



What is an Ideal Rational Agent?

An ideal rational agent is the one, which is capable of doing expected actions to maximize its performance measure, on the basis of −

- Its percept sequence
- Its built-in knowledge base

Rationality of an agent depends on the following −

- The **performance measures**, which determine the degree of success.
- Agent's **Percept Sequence** till now.
- The agent's **prior knowledge about the environment**.
- The **actions** that the agent can carry out.

A rational agent always performs the right action, where the right action means the action that causes the agent to be most successful in the given percept sequence. The problem the agent solves is characterized by Performance Measure, Environment, Actuators, and Sensors (PEAS).

The Structure of Agents in Artificial Intelligence

- Agents in Artificial Intelligence follow this simple structural formula:
- Architecture + Agent Program = Agent

These are the terms most associated with agent structure:

- **Architecture:** This is the machinery or platform that executes the agent.
- **Agent Function:** The agent function maps a precept to the Action, represented by the following formula: f:P* - A
- **Agent Program:** The agent program is an implementation of the agent function. The agent program produces function f by executing on the physical architecture.

Many AI Agents use the PEAS model in their structure. PEAS is an acronym for Performance Measure, Environment, Actuators, and Sensors. For instance, take a vacuum cleaner.

- **Performance:** Cleanliness and efficiency
- **Environment:** Rug, hardwood floor, living room
- **Actuator:** Brushes, wheels, vacuum bag
- **Sensors:** Dirt detection sensor, bump sensor

2.2 Agents Performance

Certainly, here are the key points about agent performance in AI:

- **Task Completion:** Ability to achieve specific objectives or goals within a given environment.
- **Efficiency:** Accomplishing tasks in a timely manner while minimizing resource usage like computation and energy.
- **Adaptability:** Capacity to adjust strategies and behaviors in response to changing circumstances or environments.
- **Robustness:** Resilience to uncertainties and adversarial conditions, ensuring consistent performance.
- **Safety and Ethics:** Considerations for ensuring the agent's actions align with ethical principles and prioritize user safety.

2.3 Examples of Agents

- **Chatbots:** AI agents designed to converse with users, answer questions, provide recommendations, or assist with tasks. Examples include Apple's Siri, Google Assistant, and customer service chatbots on websites.
- **Autonomous Vehicles:** Self-driving cars employ AI agents to perceive their surroundings, make decisions, and navigate safely. Examples include Tesla's Autopilot, Waymo's autonomous driving technology, and Uber's self-driving vehicles.
- **Game Playing Agents:** AI agents capable of playing and mastering complex games. Examples include AlphaGo, developed by DeepMind,

which defeated world champions in the game of Go, and OpenAI's Dota 2 bot, which competes against professional players.

- **Personalized Recommendations:** AI agents that analyze user preferences and behaviors to offer personalized recommendations. Examples include Netflix's recommendation system, Amazon's product recommendation engine, and Spotify's music recommendation algorithms.
- **Virtual Assistants:** AI agents that assist users with tasks such as scheduling appointments, setting reminders, and managing emails. Examples include Microsoft's Cortana, Amazon's Alexa, and Samsung's Bixby.
- **Medical Diagnosis:** AI agents that assist medical professionals in diagnosing diseases and interpreting medical images. Examples include IBM Watson for Oncology, which helps oncologists in treatment decisions, and Google's DeepMind Health, which analyzes medical scans for early detection of diseases.
- **Financial Trading:** AI agents used in algorithmic trading to analyze market data, identify patterns, and execute trades autonomously. Examples include high-frequency trading algorithms employed by financial institutions and hedge funds.
- **Emotion Recognition Systems:** AI agents that analyze facial expressions, voice tone, and other cues to recognize human emotions. They find applications in customer service, mental health monitoring, and human-computer interaction. Examples include Affectiva and Microsoft's Emotion API.
- **Supply Chain Optimization:** AI agents that optimize supply chain operations by predicting demand, optimizing inventory levels, and streamlining logistics. Examples include IBM's Watson Supply Chain, SAP's Integrated Business Planning, and Blue Yonder (formerly JDA Software).
- **Energy Management Systems:** AI agents that optimize energy consumption in buildings, factories, and transportation systems to improve efficiency and reduce costs. Examples include Siemens' Desigo CC, Schneider Electric's EcoStruxure, and C3.ai's Energy Management.
- **Content Generation:** AI agents that generate text, images, videos, and other multimedia content automatically. Examples include OpenAI's GPT

(Generative Pre-trained Transformer) models for text generation and DeepMind's WaveNet for speech synthesis.

- **Personalized Education:** AI agents that provide personalized learning experiences, adaptive tutoring, and educational content recommendation based on individual learning styles and progress. Examples include Duolingo for language learning, Khan Academy for personalized tutoring, and SMART Learning Suite for interactive classrooms.

2.4 Rationality and omniscience

1. **Rationality:** Rationality pertains to the ability to make logical decisions based on available information and goals. A rational agent selects actions that maximize its expected utility, given its knowledge and beliefs. Rationality is achievable within the constraints of available information and computational resources.
    - **Logic and Decision-Making:** Rationality involves making decisions and taking actions that are logically sound and coherent with one's goals or preferences.
    - **Optimization:** Rational agents aim to maximize their expected utility, selecting actions that lead to the most favorable outcomes given their knowledge and beliefs about the world.
    - **Goal-Directed Behavior:** Rationality implies acting in a manner that aligns with one's objectives, whether short-term or long-term, by considering the consequences of different courses of action.
    - **Consistency:** Rational agents exhibit consistent behavior over time, adhering to principles of coherence and avoiding contradictions in their decision-making processes.
    - **Bounded Rationality:** While ideal rationality assumes perfect knowledge and unlimited computational resources, in practice, agents operate under constraints such as limited information, time, and computational power. Bounded rationality acknowledges these limitations and focuses on making the best decisions given available resources.
    - **Adaptability:** Rationality encompasses the ability to adjust strategies and behaviors in response to changes in the environment or new information, aiming to optimize outcomes in dynamic and uncertain scenarios.
    - **Trade-offs:** Rational decision-making often involves evaluating trade-offs between competing objectives or constraints, seeking to strike a balance that maximizes overall utility or satisfaction.

- **Ethical Considerations:** Rational agents may incorporate ethical principles or moral values into their decision-making processes, ensuring that actions are not only logically consistent but also morally justifiable.

2. **Omniscience:** Omniscience refers to having complete or infinite knowledge. An omniscient entity possesses all possible knowledge about past, present, and future events. Omniscience is often considered an idealized concept, typically attributed to divine beings in religious contexts, and is not attainable within the confines of human or artificial intelligence.
    - **Complete Knowledge:** Omniscience entails possessing all possible knowledge about past, present, and future events, including all facts, truths, and possibilities.
    - **Infinite Awareness:** An omniscient entity is aware of every detail and aspect of existence, with a comprehensive understanding of all phenomena and their interrelationships.
    - **Unlimited Scope:** Omniscience encompasses knowledge of all domains, disciplines, and dimensions, transcending the boundaries of human comprehension and perception.
    - **Absolute Certainty:** Omniscience implies absolute certainty and infallibility in knowledge, devoid of any ambiguity, doubt, or error.
    - **Divine Attribution:** Omniscience is often attributed to divine or supernatural beings in religious and philosophical contexts, representing a state of ultimate wisdom and enlightenment.
    - **Temporal and Spatial Awareness:** Omniscience extends beyond temporal and spatial limitations, encompassing knowledge of events across all times and locations.
    - **Paradoxes and Limitations:** The concept of omniscience raises philosophical questions and paradoxes, such as the problem of free will and foreknowledge, highlighting potential contradictions between omniscience and human agency.
    - **Impossibility in Practice:** Achieving omniscience is considered impossible within the constraints of finite beings and limited cognitive capacities, remaining a theoretical ideal rather than an attainable reality.

In summary, while rationality involves making optimal decisions with available information, omniscience represents an unattainable state of knowing everything. Rationality is a practical concept applicable to decision-making processes, while omniscience remains a theoretical ideal.

2.5 Agent Environment and Types of Agent Environment

Agent Environment in AI

- An environment is everything in the world which surrounds the agent, but it is not a part of an agent itself. An environment can be described as a situation in which an agent is present.
- The environment is where the agent lives, operates and provides the agent with something to sense and act upon it. An environment is mostly said to be non-feministic.

Features of Environment

As per Russell and Norvig, an environment can have various features from the point of view of an agent:

1. Fully observable vs Partially Observable
2. Static vs Dynamic
3. Discrete vs Continuous
4. Deterministic vs Stochastic
5. Single-agent vs Multi-agent
6. Episodic vs sequential
7. Known vs Unknown
8. Accessible vs Inaccessible

1. Fully observable vs Partially Observable:

- If an agent sensor can sense or access the complete state of an environment at each point of time then it is **a fully observable** environment, else it is **partially observable**.
- A fully observable environment is easy as there is no need to maintain the internal state to keep track of the history of the world.

- An agent with no sensors in all environments then such an environment is called as **unobservable**.

2. Deterministic vs Stochastic:

- If an agent's current state and selected action can completely determine the next state of the environment, then such an environment is called a deterministic environment.
- A stochastic environment is random in nature and cannot be determined completely by an agent.
- In a deterministic, fully observable environment, an agent does not need to worry about uncertainty.

3. 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.
- However, in a Sequential environment, an agent requires memory of past actions to determine the next best actions.

4. Single-agent vs Multi-agent

- If only one agent is involved in an environment, and operating by itself then such an environment is called a single agent environment.
- However, if multiple agents are operating in an environment, then such an environment is called a multi-agent environment.
- The agent design problems in the multi-agent environment are different from single agent environments.

5. Static vs Dynamic:

- If the environment can change itself while an agent is deliberating then such an environment is called a dynamic environment else it is called a static environment.
- Static environments are easy to deal with because an agent does not need to continue looking at the world while deciding for an action.
- However, for a 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.

6. 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; else it is called continuous environment.
- A chess game comes under a 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.

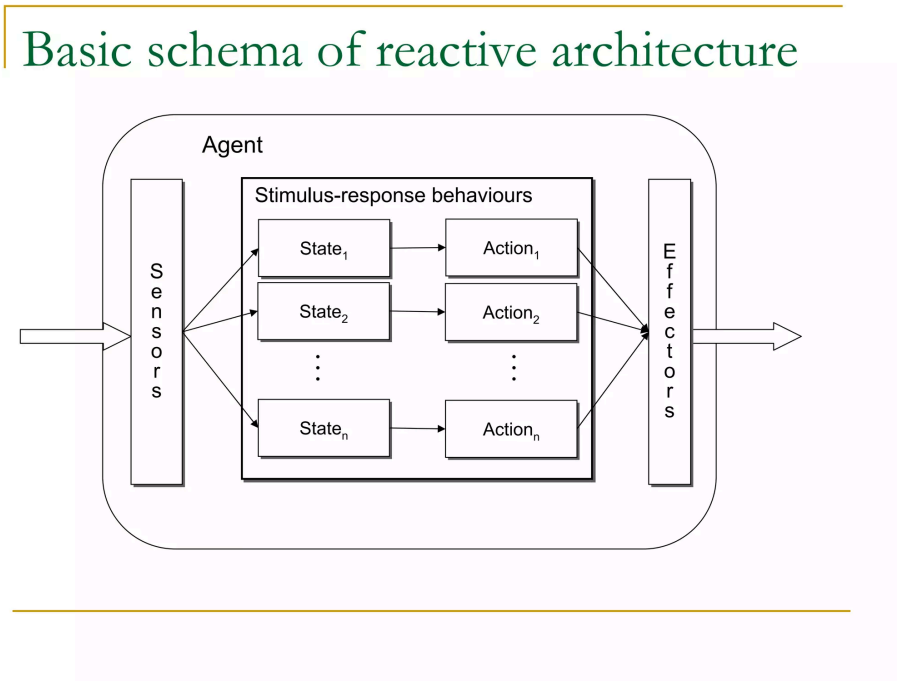
7. Known vs Unknown

- 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 an unknown environment, an agent needs to learn how it works in order to perform an action.
- It is quite possible for a known environment to be partially observable and an Unknown environment to be fully observable.

8. Accessible vs Inaccessible

- If an agent can obtain complete and accurate information about the state's environment, then such an environment is called an Accessible environment else it is called inaccessible.
- An empty room whose state can be defined by its temperature is an example of an accessible environment.
- Information about an event on earth is an example of Inaccessible environment.

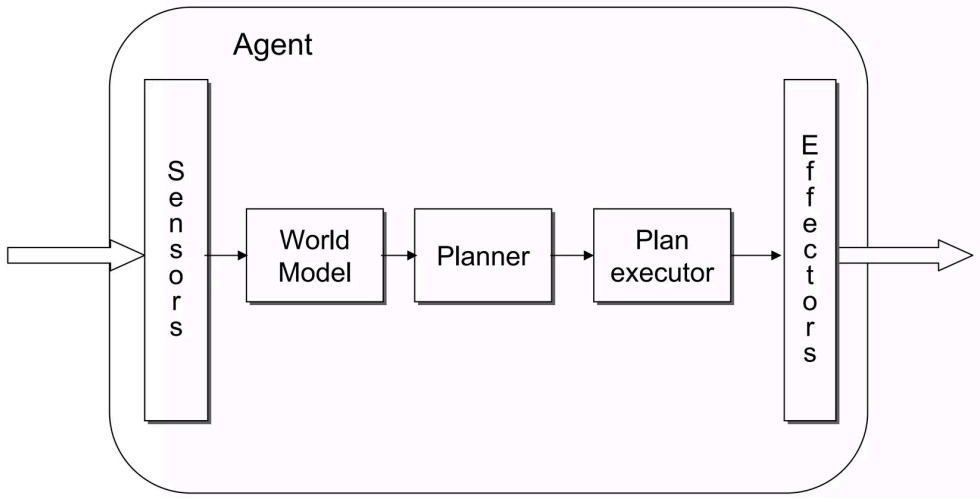
2.6 Agents Architecture in AI



Basic Schema Reactive Architecture of Agents:

- **Schema:** A reactive agent consists of a set of fixed action-response rules or schemas.
- **Reactivity:** It reacts to stimuli from the environment by executing predefined actions associated with specific conditions.
- **Example:** An agent programmed to move away from obstacles when it senses them.

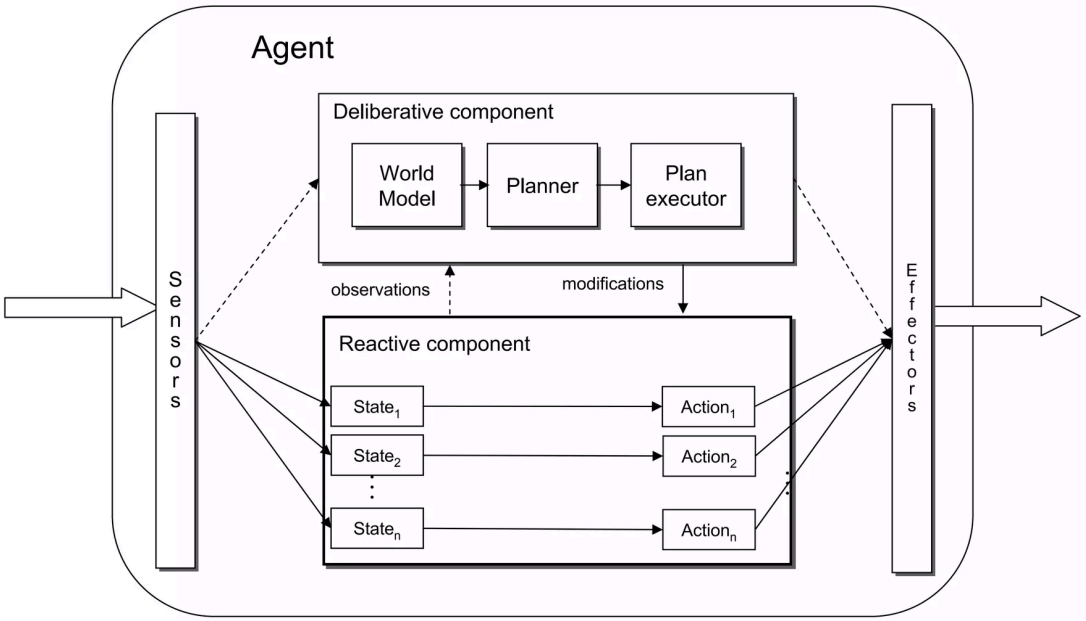
Basic deliberative architecture



Basic Deliberative Architecture:

- **Deliberation:** Focuses on decision-making through reasoning and planning.
- **Knowledge Representation:** It involves representing the world, goals, and actions in a formal way that allows for logical inference.
- **Example:** A deliberative agent might use a search algorithm to plan a route from point A to point B considering various factors like traffic conditions, road closures, and time constraints.

Hybrid agent architecture



Hybrid Agent Architecture:

- **Combination:** Integrates elements of both reactive and deliberative architectures to leverage their respective strengths.
- **Flexibility:** Offers the ability to react quickly to immediate stimuli while also being capable of complex planning and decision-making.
- **Example:** A robot navigating a dynamic environment might use reactive components for obstacle avoidance but switch to a deliberative approach for long-term path planning considering goals and constraints.

2.7 PEAS in AI

- We know that there are different types of agents in AI.
- PEAS System is used to categorize similar agents together.
- The PEAS system delivers the performance measure with respect to the environment, actuators, and sensors of the respective agent.
- Most of the highest performing agents are Rational Agents.
- **Rational Agent:** The rational agent considers all possibilities and chooses to perform a highly efficient action. For example, it chooses the shortest path with low cost for high efficiency. **PEAS** stands for a *Performance measure, Environment, Actuator, Sensor*.
    1. **Performance Measure:** Performance measure is the unit to define the success of an agent. Performance varies with agents based on their different precepts.
    2. **Environment**: Environment is the surrounding of an agent at every instant. It keeps changing with time if the agent is set in motion. There are 5 major types of environments:
        - Fully Observable & Partially Observable
        - Episodic & Sequential
        - Static & Dynamic
        - Discrete & Continuous
        - Deterministic & Stochastic
    3. **Actuator**: An actuator is a part of the agent that delivers the output of action to the environment.
    4. **Sensor**: Sensors are the receptive parts of an agent that takes in the input for the agent.

Agent	Performance Measure	Environment	Actuators	Sensors
Human Agent	Subjective: Satisfaction, Efficiency	Varied (Indoor, Outdoor, Social)	Limbs (Hands, Legs), Voice	Eyes (Vision), Ears (Hearing), Skin (Touch)
Robotic Agent	Objective: Task Completion, Energy Efficiency	Varied (Factory Floor, Household, Space)	Motors, Grippers, Tools	Cameras, Sensors (Temperature, Proximity, etc.)
Taxi Driving Agent	Objective: Passenger Satisfaction, Profitability	Urban Environment, Traffic, Weather	Steering Wheel, Accelerator, Brake, Navigation System	Cameras, GPS, Lidar, Ultrasonic Sensors
8-Queens Problem	Objective: Find Solution, Minimize Conflicts	Abstract, 8×8 Chessboard with Queens	Placing Queens on the Board	Board State Representation, Conflict Detection

Agent	Performance Measure	Environment	Actuator	Sensor
Hospital Management System	Patient's health, Admission process, Payment	Hospital, Doctors, Patients	Prescription, Diagnosis, Scan report	Symptoms, Patient's response
Automated Car Drive	The comfortable trip, Safety, Maximum Distance	Roads, Traffic, Vehicles	Steering wheel, Accelerator, Brake, Mirror	Camera, GPS, Odometer
Subject Tutoring	Maximize scores, Improvement is students	Classroom, Desk, Chair, Board, Staff, Students	Smart displays, Corrections	Eyes, Ears, Notebooks

Part-picking robot	Percentage of parts in correct bins	Conveyor belt with parts; bins	Jointed arms and hand	Camera, joint angle sensors
Satellite image analysis system	Correct image categorization	Downlink from orbiting satellite	Display categorization of scene	Color pixel arrays

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

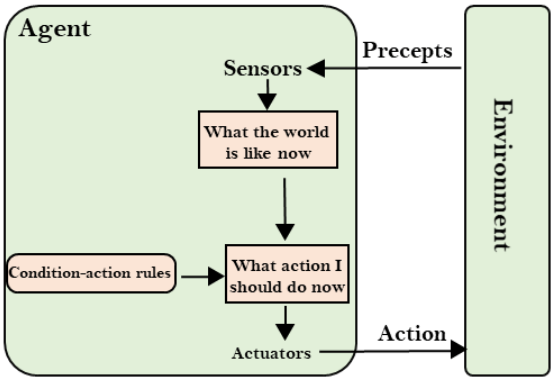
2.8 Types of AI Agents

Agents can be grouped into five classes based on their degree of perceived intelligence and capability. All these agents can improve their performance and generate better action over time. These are given below:

- Simple Reflex Agent
- Model-based reflex agent
- Goal-based agents
- Utility-based agent
- Learning agent

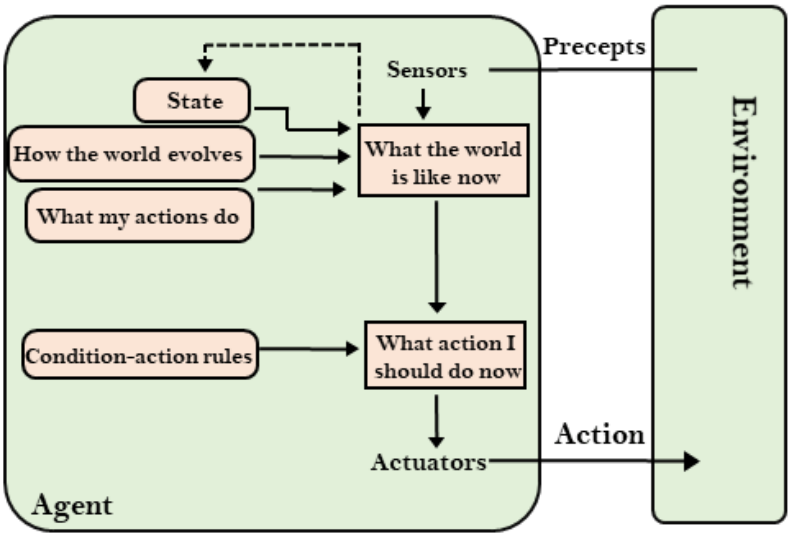
1. Simple Reflex agent:

- The Simple reflex agents are the simplest agents. These agents take decisions on the basis of the current percepts and ignore the rest of the percept history.
- These agents only succeed in the fully observable environment.
- The Simple reflex agent does not consider any part of perceived history during their decision and action process.
- The Simple reflex agent works on Condition-action rule, which means it maps the current state to action. Such as a Room Cleaner agent, it works only if there is dirt in the room.
- Problems for the simple reflex agent design approach:
    - They have very limited intelligence
    - They do not have knowledge of non-perceptual parts of the current state
    - Mostly too big to generate and to store.
    - Not adaptive to changes in the environment.



2. Model-based reflex agent

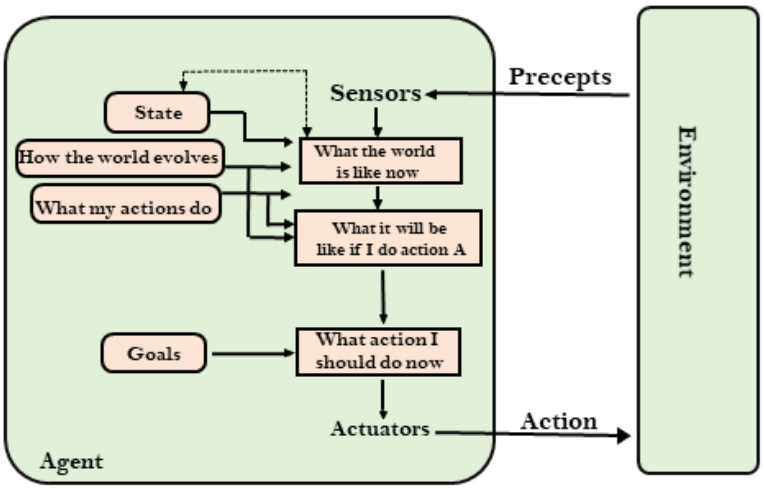
- The Model-based agent can work in a partially observable environment, and track the situation.
- A model-based agent has two important factors:
    a. **Model:** It is knowledge about "how things happen in the world," so it is called a Model-based agent.
    b. **Internal State:** It is a representation of the current state based on percept history.
- These agents have the model, "which is knowledge of the world" and based on the model they perform actions.
- Updating the agent state requires information about:
    a. How the world evolves
    b. How the agent's action affects the world.



3. Goal-based agents

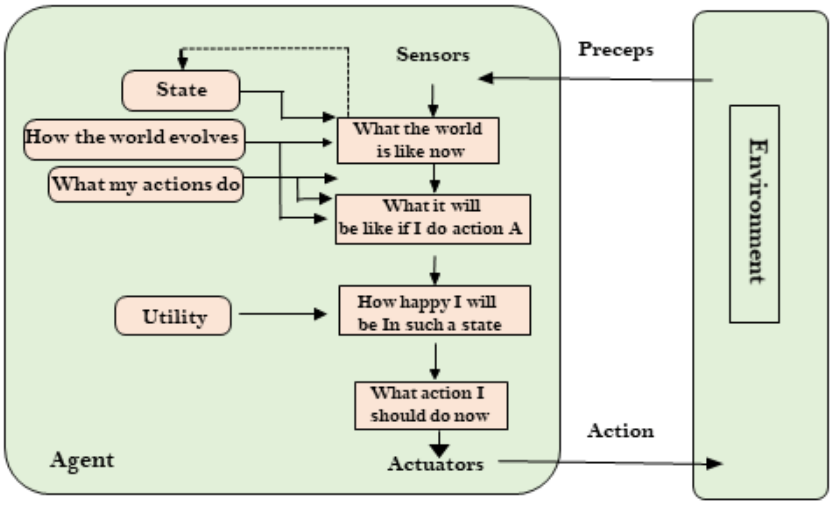
- The knowledge of the current state environment is not always sufficient to decide for an agent what to do.

- The agent needs to know its goal which describes desirable situations.
- Goal-based agents expand the capabilities of the model-based agent by having the "goal" information.
- They choose an action, so that they can achieve the goal.
- These agents may have to consider a long sequence of possible actions before deciding whether the goal is achieved or not. Such considerations of different scenarios are called searching and planning, which makes an agent proactive.



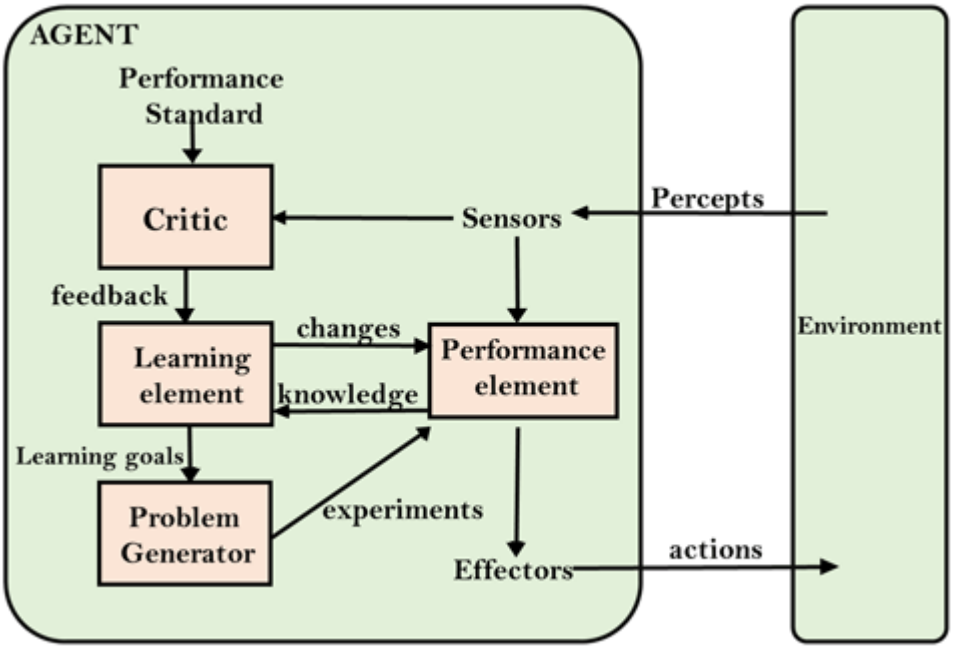
4. Utility-based agents

- 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.
- The utility function maps each state to a real number to check how efficiently each action achieves the goals.



5. Learning Agents

- A learning agent in AI is the type of agent which can learn from its past experiences, or it has learning capabilities.
- It starts with basic knowledge and then is able to act and adapt automatically through learning.
- A learning agent has mainly four conceptual components, which are:
    a. **Learning element:** It is responsible for making improvements by learning from environment
    b. **Critic:** Learning element takes feedback from the critic which describes how well the agent is doing with respect to a fixed performance standard.
    c. **Performance element:** It is responsible for selecting external action
    d. **Problem generator:** This component 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.


AGENT
Performance Standard
Critic
Sensors
Percepts
feedback
changes
Learning element
Performance element
knowledge
Environment
Learning goals
Problem Generator
experiments
Effectors
actions