RL makes decisions based on learning from past experiances

It is a feedback based technique where the agent learns how to behave by performing actions and checking their results

Reward desired behaviour and punish bad behaviour

Hit and trail practise

Direct interaction between agent and environment

Solve difficult problems using control optimization and decision making can be solved using RL

Agent interacts with environment, observes the state or environment, selects actions and recieves rewards or penalties based on the actions. Over time, agent learns to take actions that maximize the award.

Key Concepts in RL

Agent : Agent takes actions

Environment : World in which agent exists & operates

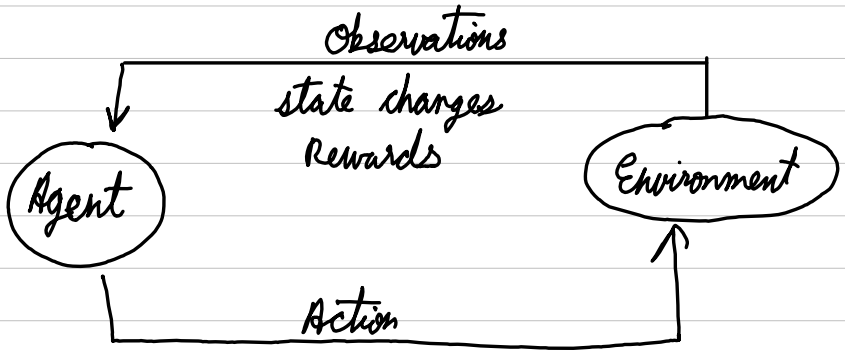
Actions : A move that the agent can make

Action space A : Set of possible actions an agent can make in the environment
Action space can be discrete or continuouse

Observations : Input from the environment

State : situation which the agent percieves

Reward : Feedback that measures the success or failure of the agents action



Policy : Which action to do in which state (π)

Value function : How good the situation is.

Reward

Agent wants to maximize the reward

At any time t, the agent has a total reward

$$R_t = \sum_{i=t}^{\infty} r_i$$

R_t: Total reward at time t

r_i: reward returned at time i

In order to give more importance to immediate rewards, a γ factor is added

$$R_t = \sum_{i=t}^{\infty} \gamma^i r_i$$

γ: discounting factor

$(0 \leq \gamma \leq 1)$

Q function

Q function captures the expected total future reward an agent in state "s" can recieve by executing a certain action "a"

$$Q(S_t, a_t) = \mathbb{E}(R_t \mid S_t, a_t)$$

- Q: Q function
- S_t: state
- a_t: action
- $\mathbb{E}$: Expected
- R_t: total reward

Q function takes in a state and a possible action and tells how much total reward will be obtained.

Choose the action that maximizes future reward

Advantages →
1. Versatility
2. Adaptability
3. Efficiency

Disadvantages →
1. Need a simulatable environment
2. Reward design difficult

Application →
1. Robotics
2. Game player (chess, go)
3. Petrollum refinery parameter controller