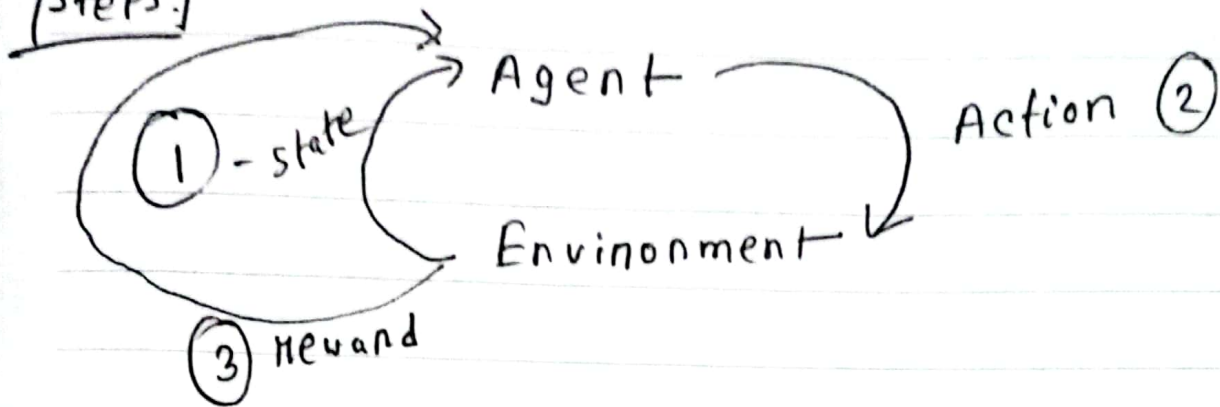


Reinforcement Learning:

Steps:



(4) Then go to next state

Def: is a ML technique that trains SW to make decisions to achieve the most optimal results

Ex: (Vacuum Cleaner)

→ Goal: To clean the floor efficiently

→ Actions: move forward, turn right/left, start/stop suction

→ reward: if picks up dirt → positive reward
bumping into wall → negative reward

* learns from reward and penalties

Markov Decision Process:

→ ~~markov~~ mathematically formulate reinforcement learning problem.

□ Markov property: current state completely characterises the state of the world.

→ defined by (S, A, R, P, γ)

S → set of possible states

A → " " Actions

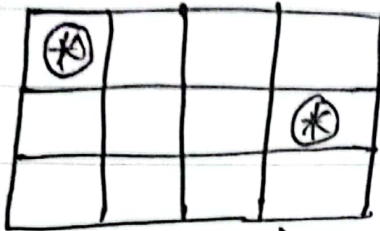
R → distribution of reward given (state, action) pain

P → transition probability. i.e: distribution over next state given (state, action) pain

γ → discount factor

* characteristic: Deterministic, but non-deterministic

It maximizes newands



(States)

actions = right \rightarrow

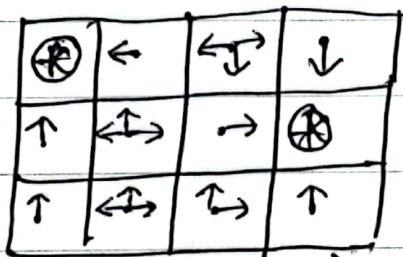
left \leftarrow

up \uparrow down \downarrow

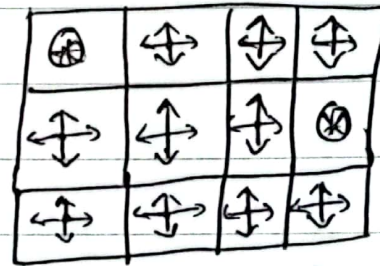
Objective:

any cell to
start cell in
min actions

for each transition,
newand, $r = -1$



(Optimal)
ans



(random)