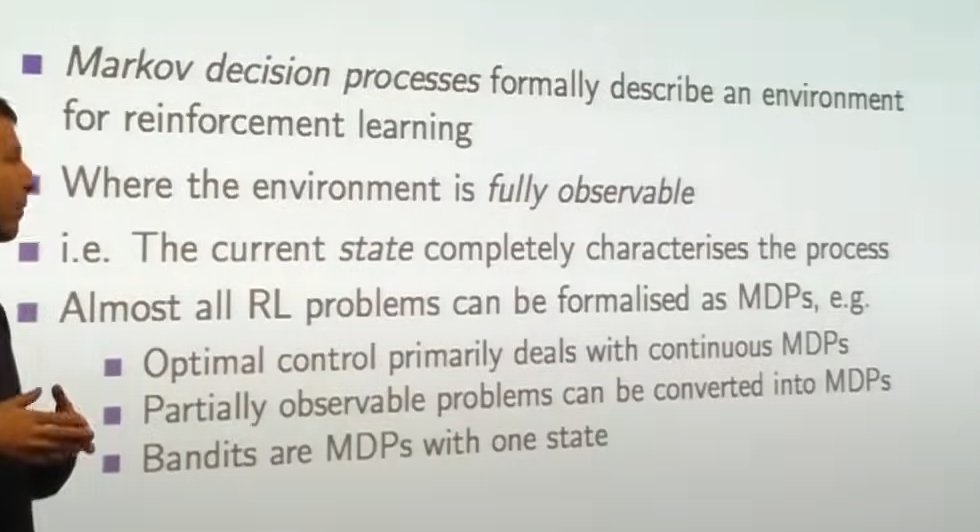
## horizontal line

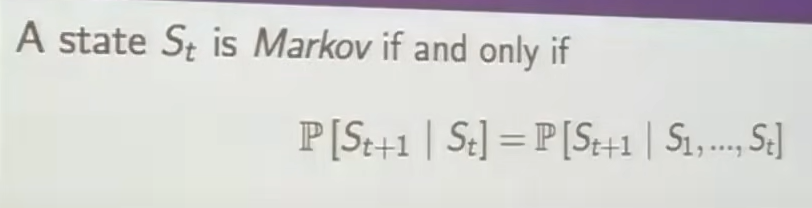
Markov Decision Process

18.10.2025

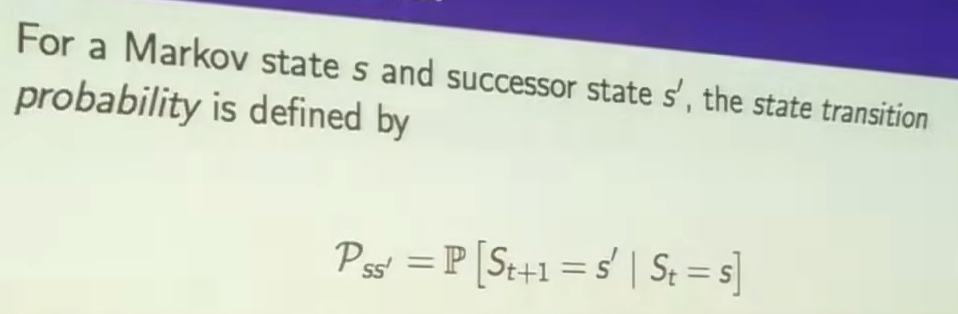


Markov Property :

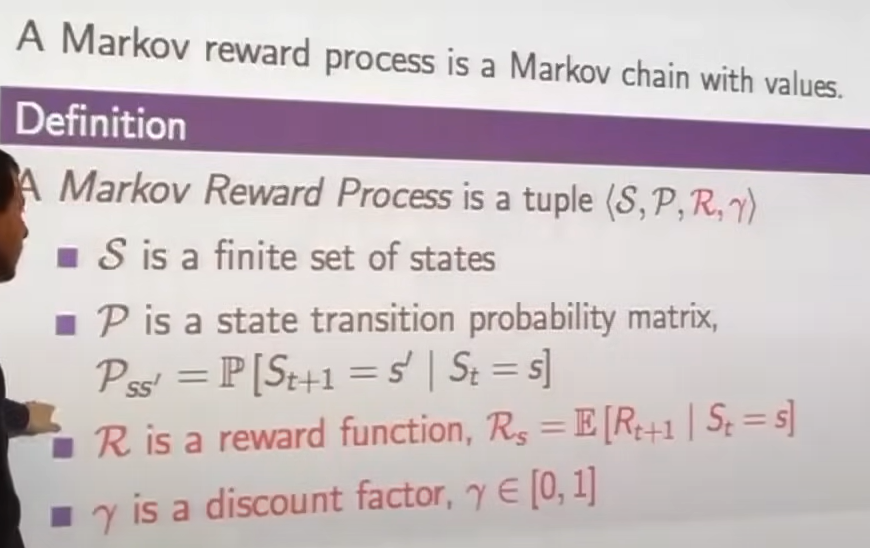
“The Future is independent of the past given the present.”

1. The State captures all relevant information from history.
2. Once the state is known , the history may be thrown away.
3. The state is a sufficient statistic of the future.  
   

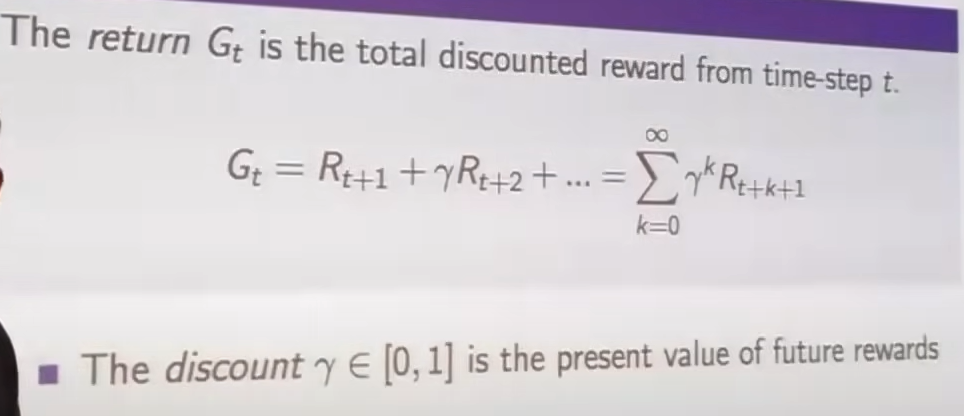
State Transition Matrix , P defines transition probabilities from all states s to all successor states s’.



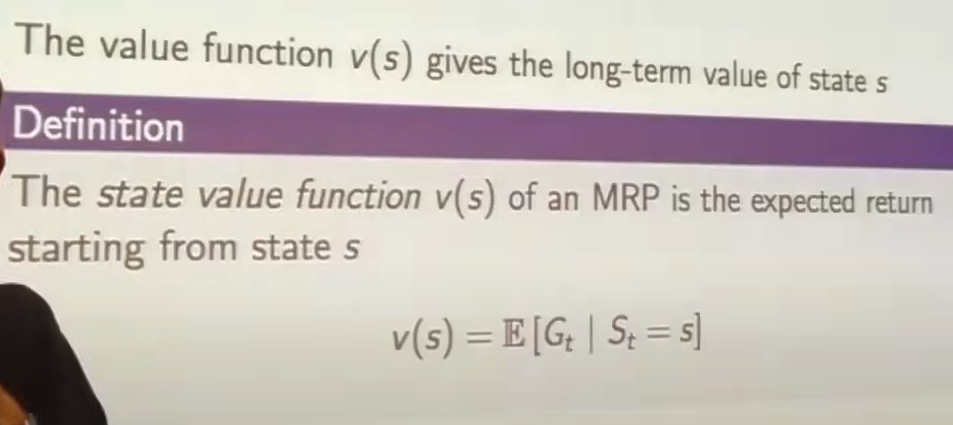
Markov Process

****

Return Gt , The reason to use discounted factor :

1. Mathematically convenient 
2. Avoids infinite returns
3. Uncertainty about the future

Value Function



The value function can be decomposed into two parts :

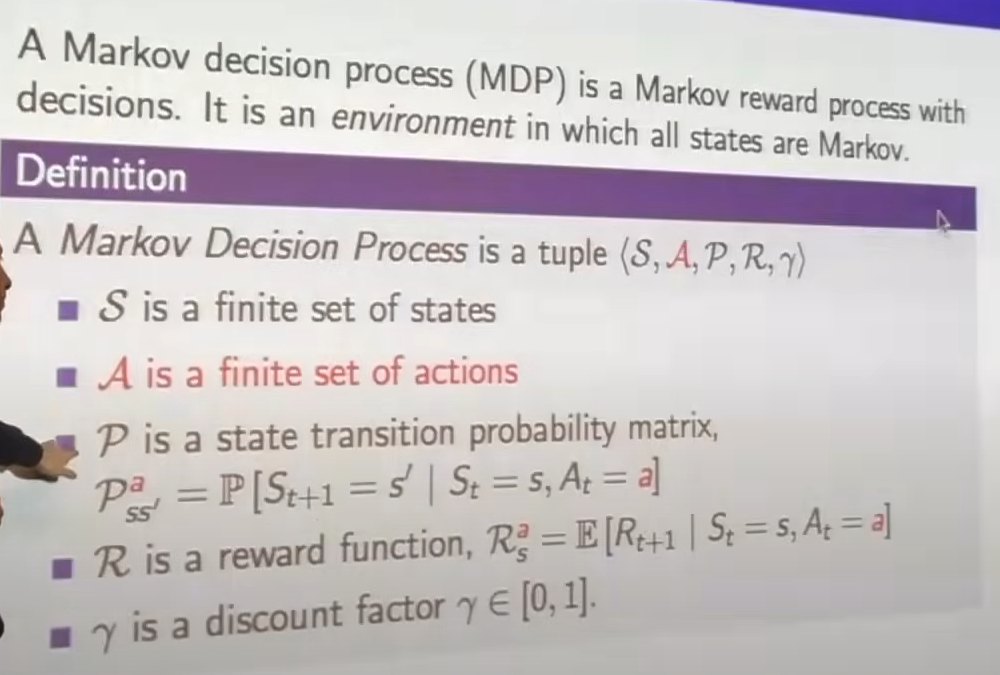
1. Immediate reward R(t+1)
2. Discounted value of successor state.



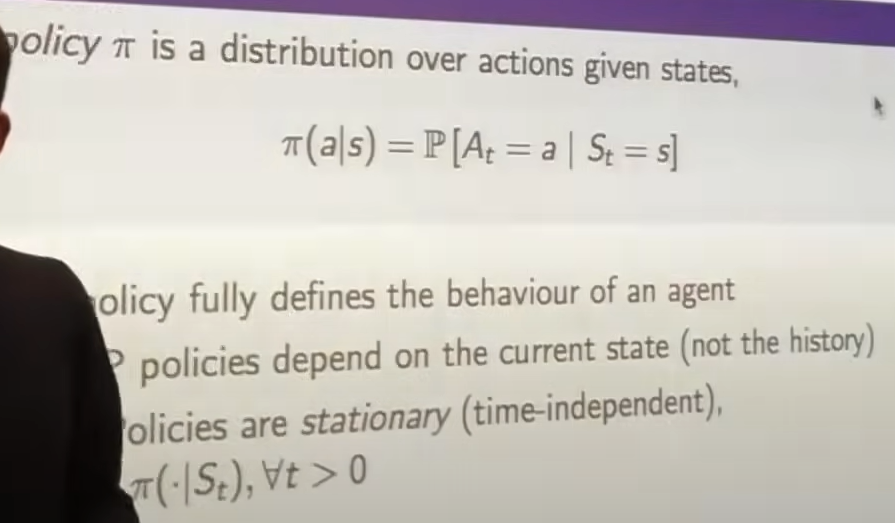
# Bellman Equation :

The Bellman equation is a fundamental formula in reinforcement learning that helps an agent estimate how good a state or action is by combining immediate rewards with expected future rewards. It’s used to find optimal strategies (policies) for decision-making.

# Markov Decision Process



Policy



State Value Function Vs Action Value Function



# Optimal Value Function

# 

Optimal Policy

