

Lecture 2: Markov Decision Processes

└ Extensions to MDPs

└ Average Reward MDPs

Ergodic MDP

(no exam)

Definition

An MDP is ergodic if the Markov chain induced by any policy is ergodic.

For any policy π , an ergodic MDP has an *average reward per time-step* ρ^π that is independent of start state.

$$\rho^\pi = \lim_{T \rightarrow \infty} \frac{1}{T} \mathbb{E} \left[\sum_{t=1}^T R_t \right]$$