



$P(s_0) \equiv$ Initial state distribution

$$E_{a_t \sim \pi(a_t)} \sum_{s_{t+1} \in \mathcal{S}} P(s_{t+1} | s_t, a_t)$$

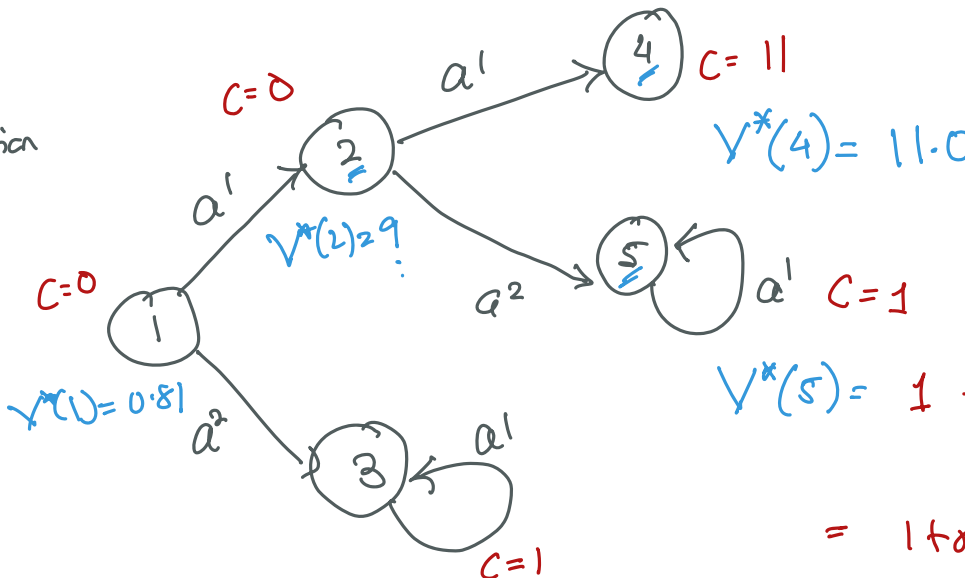
$P(a_0 | s_0) \equiv$ Comes from π

$P(s_1 | s_0, a_0) \equiv$ TRANSITION FN DEFINED MDP

$P(a_1 | s_1) \equiv$ Comes from π

What is the optimal action at state 1?

What is the optimal value at state 1?



Discount $\gamma = 0.9$

$$V^*(5) = 1 + \gamma \cdot 1 + \gamma^2 \cdot 1 + \gamma^3 \cdot 1 + \dots \infty$$

$$= 1 + \gamma + \gamma^2 + \dots \infty$$

$$= \frac{1}{1-\gamma}$$

$$= 10.0$$

$$V^*(2) = \min_{a_1, a_2} \left\{ \begin{aligned} & \overset{=0}{c(2, a_1)} + \gamma V^*(4) \\ & \overset{=0}{c(2, a_2)} + \gamma V^*(5) \end{aligned} \right. = \gamma V^*(5) = 0.9 \times 10 = 9.$$