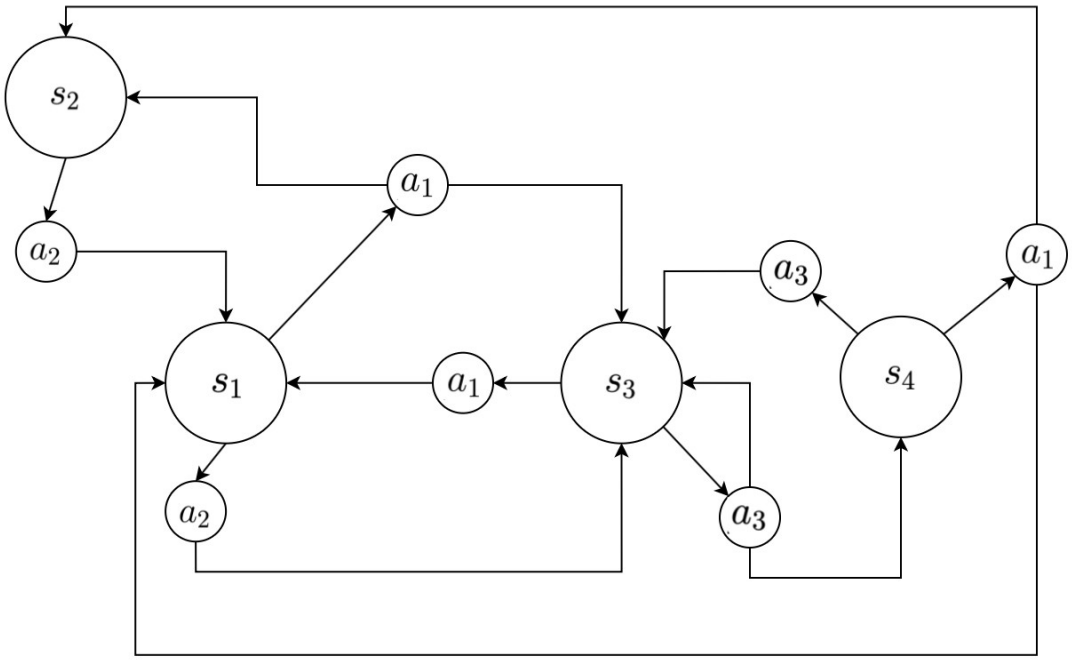
Task

Below is a diagram of a Markov decision process.



For all states, the probability distributions of the choice of available actions are known $\pi(a \vee s)$, probability distributions of transitions to the next states $P^a_{ss'}$ and mathematical expectations $R^a_{ss'}$, awards received.

$$\pi(a_1|s_1) = 0.4,\ \ \pi(a_2|s_1) = 0.6,$$

$$P_{s_1 s_2}^{a_1} = 0.3,\ \ P_{s_1 s_3}^{a_1} = 0.7,$$

$$R_{s_1 s_2}^{a_1} = 2.0,\ \ R_{s_1 s_3}^{a_1} = 3.0,$$

$$P_{s_1 s_3}^{a_2} = 1,$$

$$R_{s_1 s_3}^{a_2} = 1.0.$$

$$\pi(a_2|s_2) = 1.0,$$

$$P_{s_2 s_1}^{a_2} = 1.0,$$

$$R_{s_2 s_1}^{a_2} = 3.0.$$

$$\pi(a_1|s_3) = 0.5,\ \ \pi(a_3|s_3) = 0.5,$$

$$P_{s_3 s_1}^{a_1} = 1.0,$$

$$R_{s_3 s_1}^{a_1} = -3.0,$$

$$P_{s_3 s_3}^{a_3} = 0.2,\ \ P_{s_3 s_4}^{a_3} = 0.8,$$

$$R_{s_3 s_3}^{a_3} = 1.0,\ \ R_{s_3 s_4}^{a_3} = 6.0.$$

$$\pi(a_1|s_4) = 0.5,\ \ \pi(a_3|s_4) = 0.5,$$

Write down the Bellman equations and get the values of the state value if the discount factor $\gamma = 0.8$.

Bellman equation:

$$v_\pi(s) = \sum_{a \in A(s)} \pi(a|s) \sum_{s' \in S} \mathcal{P}^a_{ss'} \left(\mathcal{R}^a_{ss'} + \gamma v_\pi \left(s' \right) \right)$$