# Reinforcement Learning Workshop
# Day 2 – Student Activities

## Topics

- Markov Decision Processes (MDPs)

---

## 1 Activity 1: Computing Returns by Hand

Consider the following trajectory:

$$(s_0, a_0, r_1 = 1),\ (s_1, a_1, r_2 = 2),\ (s_2, a_2, r_3 = -3),\ (s_3 \text{ terminal})$$

1. Compute the total return with $\gamma = 1$
2. Compute the total return with $\gamma = 0.9$

**Question:** How does the choice of $\gamma$ change the importance of future rewards?

—

## 2 Activity 3: Evaluating a Policy

A MDP has state space $\mathscr{S} = 1, 2$ and action space , , . All actions are available in all states. The transition probability and reward matrices for each state are:

$$P^a = \begin{bmatrix} 0.2 & 0.8 \\ 0.7 & 0.3 \end{bmatrix} \quad R^a = \begin{bmatrix} 10 & 7 \\ 12 & 15 \end{bmatrix}$$

$$P^b = \begin{bmatrix} 0.4 & 0.6 \\ 0.1 & 0.9 \end{bmatrix} \quad R^b = \begin{bmatrix} 5 & 11 \\ 14 & 7 \end{bmatrix}$$

$$P^c = \begin{bmatrix} 0.8 & 0.2 \\ 0.2 & 0.8 \end{bmatrix} \quad R^c = \begin{bmatrix} 14 & 3 \\ 2 & 12 \end{bmatrix}$$

Consider the deterministic policy $\pi(1) = c$, $\pi(2) = b$

(a) Write the Bellman equations for the state value function $V^\pi(s)$ for this policy.

(b) Write the equations in matrix form $AV^\pi = b$

(c) Solve linear system to find $V^\pi$ (either by hand or using a computer).

## 3 Optional Coding Activity

Go to the GitHub repository for the course (`https://github.com/lfmartins/rl_cimpa_2026` and click on the “Day 1 Activities in Colab” link.

—