

“The rotten tree-trunk, until the very moment when the storm-blast breaks it in two, has all the appearance of might it ever had.”

Isaac Asimov, *Foundation*



CMPUT 365

Introduction to RL

Reminder

You **should be enrolled in the private session** we created in Coursera for CMPUT 365.

I **cannot** use marks from the public repository for your course marks.

You **need to check, every time**, if you are in the private session and if you are submitting quizzes and assignments to the private section.

Some students who are enrolled in Coursera **haven't submitted any quizzes or assignments** in the private session, and that's all I can see.

The deadlines in the public session **do not align** with the deadlines in Coursera.

Please, interrupt me at any time!



Example: Value Function Computation

Consider the 7-state MDP on the side. It has four actions available: {up, down, left, and right}. Its dynamics are deterministic, except at the purple states, where the agent can go up with 40% chance, regardless of the action taken, and 60% chance one goes to the intended direction. The reward is +1 upon entering state s_6 , +2 upon entering the terminal state, and 0 otherwise. Let $\gamma = 0.8$. Consider the policy below:

s_1	s_2	s_3
s_4	s_5	
s_6	s_7	s_8

↓	↓	↓
↓	→	
→	0.5 ↑ 0.5 →	↑

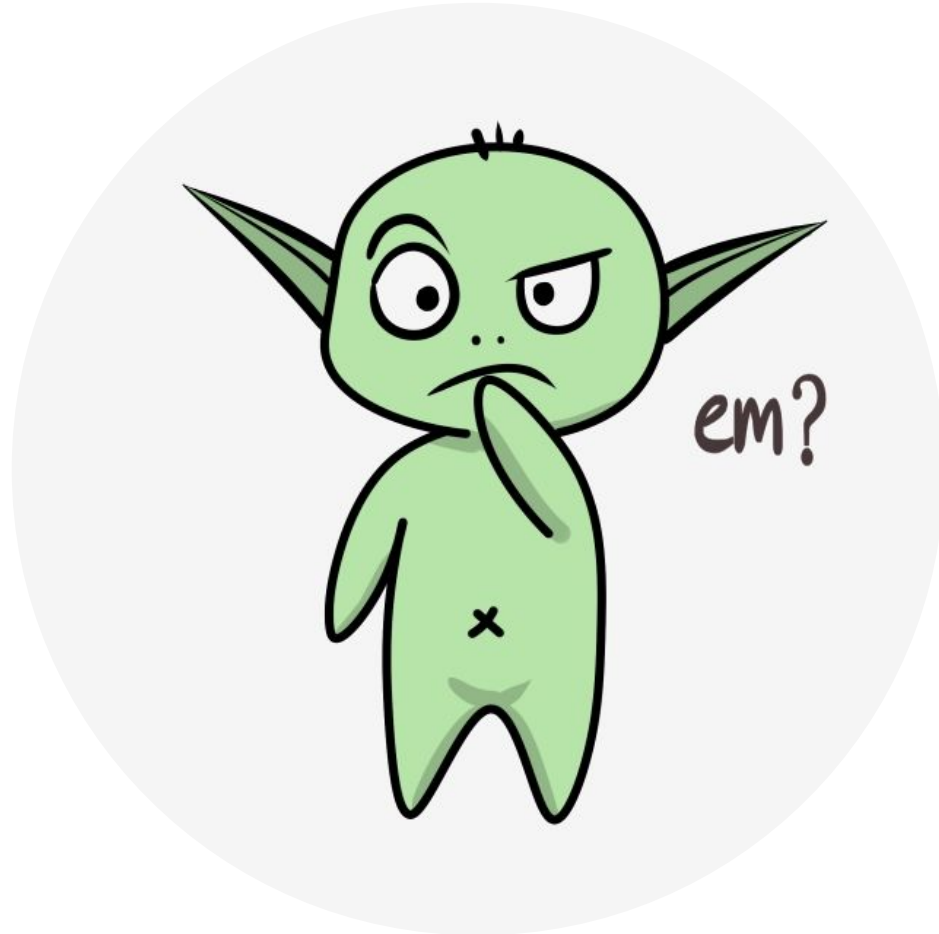
a) What's $v_{\pi}(s_4)$?

Recall

$$v_{\pi}(s) = \sum_a \pi(a|s) \sum_{s', r} p(s', r | s, a) [r + \gamma v_{\pi}(s')], \quad \text{for all } s \in \mathcal{S}$$

Solution: Value Function Computation

Solution: Value Function Computation



Next class

- What I plan to do:
 - Overview of Sutton & Barto's (2018) Chapter 4: Dynamic Programming
- What I recommend YOU to do for next class:
 - Read over Chapter 4, §4.1-§4.4 (pp. 73-84); §4.6-§4.7 (pp. 86-89).
 - Submit Practice Quiz for Fundamental of RL: Dynamic Programming (Week 4).
 - Start Programming Assignment for Fundamentals of RL: Dynamic Programming (Week 4).

Week	Date	Topic
5	Wed, Sep 25	Fundamentals of RL: Dynamic programming
5	Fri, Sep 27	Fundamentals of RL: Dynamic programming
	Mon, Sep 30	National Day for Truth and Reconciliation
5	Wed, Oct 2	Fundamentals of RL: Dynamic programming
	Fri, Oct 4	Midterm exam 1