🕐 Round 1: Completed   #classroom   Weight: 1.0

:rowd 👹

nent 1
orcement Learning
xi
nanth

By 👹 **Alcrowd** & 🏛 **IIT Madras**   👁 1534   👤 78   👥 0   🚀 342   ❤ 3   Follow

Overview   Leaderboard   Notebooks   Discussion   Insights   Resources   Submissions   Rules

🕵️ Introduction

In this problem we have a a taxi driver, who serves three cities A, B and C. On a regular workday, the taxi driver can find a new ride by choosing one of the following **actions**:

1.   Cruise the streets looking for a passenger.

2.   Go to the nearest taxi stand and wait in line.

3.   Wait for a call from the dispatcher (this is not possible in town B because of poor reception).

For a given town and a given action, there is a probability that the next trip will go to each of the towns A, B and C and a corresponding reward in monetary units associated with each such trip. This reward represents the income from the trip after all necessary expenses have been deducted. Please refer to the table below for the rewards and transition probabilities.

Town i	Actions k	Probabilities p_{ij}^k j = A B C	Rewards r_{ij}^k j = A B C
A	1 2 3	A B C 1/2 1/4 1/4 1/16 3/4 3/16 1/4 1/8 5/8	A B C 10 4 8 8 2 4 4 6 4
B	1 2	A B C 1/2 0 1/2 1/16 7/8 1/16	A B C 14 0 18 8 16 8
C	1 2 3	A B C 1/4 1/4 1/2 1/8 3/4 1/8 3/4 1/16 3/16	A B C 10 2 8 6 4 2 4 0 8

- p^k_{ij} is the probability of getting a ride to town **j**, by choosing an action **k** while the driver was in town **i**.

NOTEBOOKS   See all

1 ❤   **[Starter Notebook] RL - Taxi Problem**   💬 2
By ashivani 5 months ago

- r^k_{ij} is the immediate reward of getting a ride to town **j**, by choosing an action **k** while the driver was in town **i**.

Instructions

1. Implement the DP algorithm. Find the pseudocode below. Let **S** be the state space and **A** the action space.

Initialize $J_N(s) = 0, \forall s \in S$;
For $i = N-1, \cdots, 0$

- $J_i(s) = \max_{a \in \mathcal{A}} \mathbb{E}_{s' \in \mathcal{S}}[r(s, a, s') + J_{i+1}(s')], \forall s \in \mathcal{S}$;
- $\pi_i(s) = \operatorname{argmax}_{a \in \mathcal{A}} \mathbb{E}_{s' \in \mathcal{S}}[r(s, a, s') + J_{i+1}(s')], \forall s \in \mathcal{S}$.

---

2. Tabulate the optimal policy and optimal value for each state in each round for **N** = 10.

3. Consider a policy that always forces the driver to go to the nearest taxi stand, irrespective ofthe state. Is it optimal? Justify your answer.

You will be writing your solutions & making a submission through a notebook. You can follow the instructions in the starter notebook.

💾 Dataset

📁 FILES

Under the `Resources` section you will find data files that contains parameters for the environment for this problem.

🚀 Submission

Submissions will be made through a notebook following the instructions in the starter notebook.

📱 Contact

- RL TAs