Assignment 1 – Defining & Solving RL Environments

Part 1: Defining RL Environments

1. Describing deterministic and stochastic environments.
Deterministic Environment:

1. Definition: The deterministic environment follows a strict rule-based approach where the drone's actions lead to predictable outcomes and there is no randomness which means every action results in its intended effect.

2. States:

- The environment is represented as a 6×6 grid (default).
- The drone has a (x, y) position and a binary flag (0 or 1) indicating whether it has picked up the package.
- The grid contains no-fly zones, a warehouse (pickup location), and a customer destination (drop-off point).

3. Actions:
The drone can take the following six actions:

1. UP – Move one cell up.
2. DOWN – Move one cell down.
3. LEFT – Move one cell left.
4. RIGHT – Move one cell right.
5. PICKUP – Collect the package at the warehouse.
6. DROPOFF – Deliver the package at the customer's location.

4. Rewards:

- Moving: -1 (penalty for taking steps).
- Entering a no-fly zone: -20 (severe penalty).
- Picking up a package: +10.
- Successfully delivering the package: +20 (episode ends if not in multi-delivery mode).

5. Main Objective:
The goal is to deliver the package efficiently while minimizing movement penalties and avoiding no-fly zones. Since the environment is fully predictable, the agent can plan its route perfectly.

Stochastic Environment:

1. Definition: The stochastic environment introduces random disturbances, meaning actions may not always produce the intended movement. There is a 10% probability that the drone's movement deviates to an adjacent cell in a random direction.

2. States:

- The grid structure and state definitions are identical to the deterministic environment.
- The drone's position may unexpectedly shift due to stochastic effects.

3. Actions:

- The available six actions (UP, DOWN, LEFT, RIGHT, PICKUP, DROPOFF) are the same as in the deterministic environment.
- However, when moving (UP, DOWN, LEFT, RIGHT), there is a 10% chance that the drone will move off-course in a random adjacent direction.

4. Rewards:

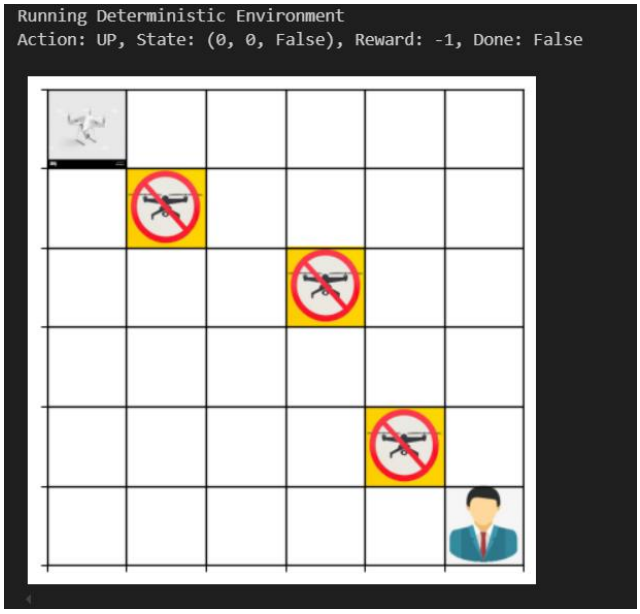
- The reward system remains the same:
  - -1 for each step,
  - -20 for entering no-fly zones,
  - +10 for picking up a package,
  - +20 for successful delivery.
- Due to randomness, additional penalties may be incurred if the drone is forced into a longer path or an obstacle.

5. Main Objective:

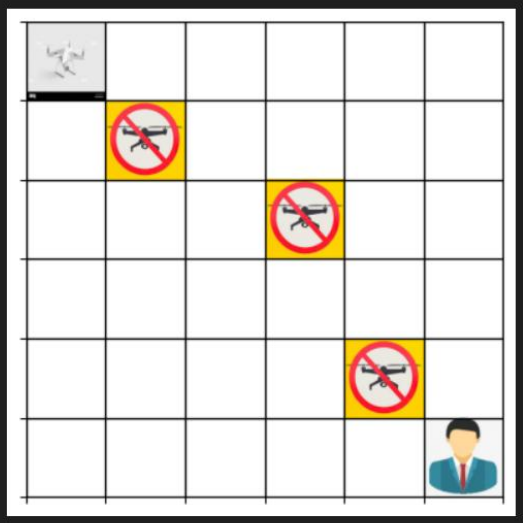
- The goal remains efficient package delivery, but the agent must now adapt to uncertainty.
- The drone must account for potential deviations and choose more robust paths to minimize the risk of failure.
- Strategic planning becomes crucial to compensate for random errors in movement.

2. Provide visualizations of your environments.

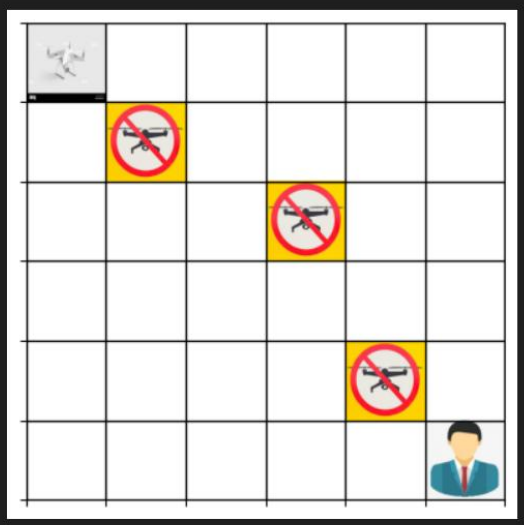
Deterministic Environment:



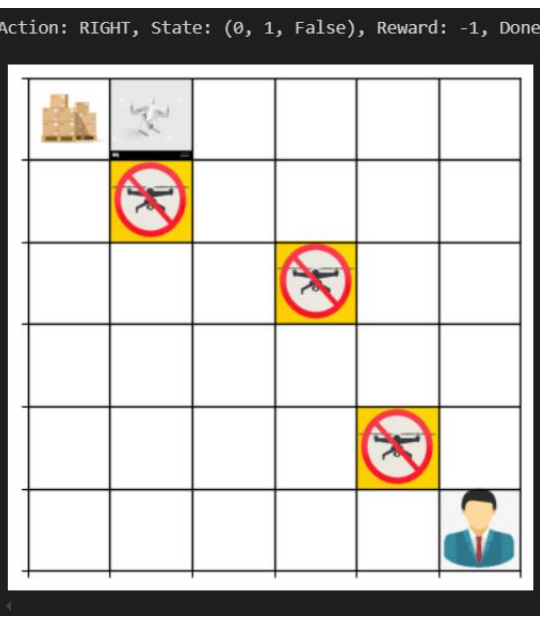
Action: LEFT, State: (0, 0, False), Reward: -1, Done: False



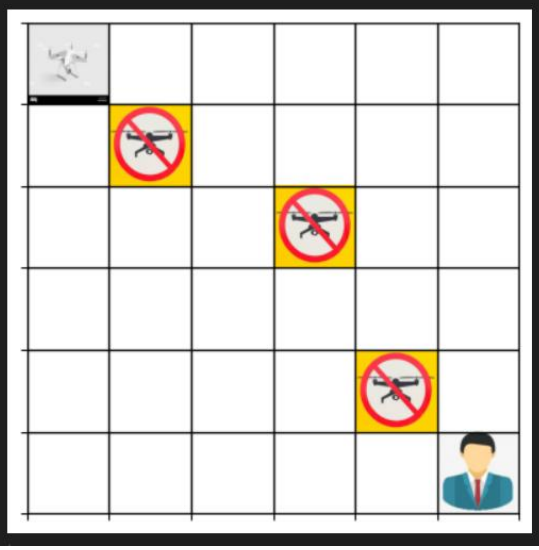
Action: LEFT, State: (0, 0, False), Reward: -1, Done: False



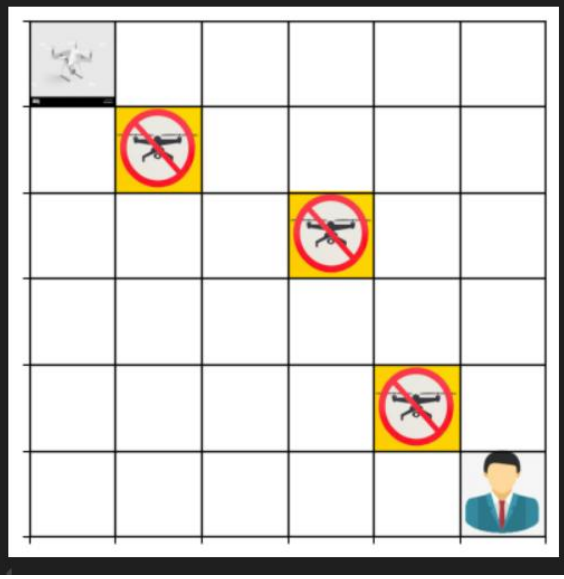
Action: RIGHT, State: (0, 1, False), Reward: -1, Done: False



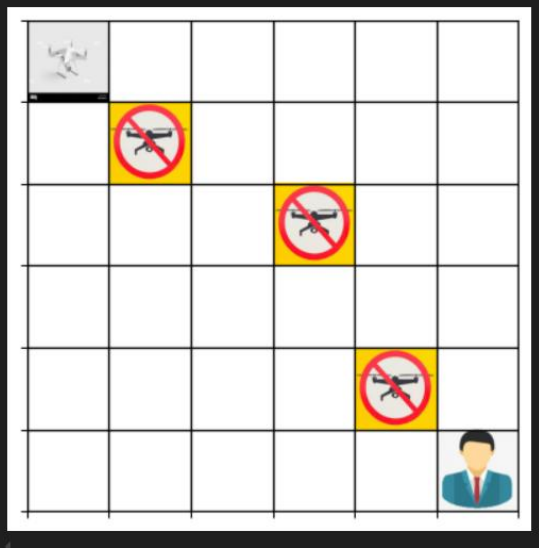
Action: LEFT, State: (0, 0, False), Reward: -1, Done: False



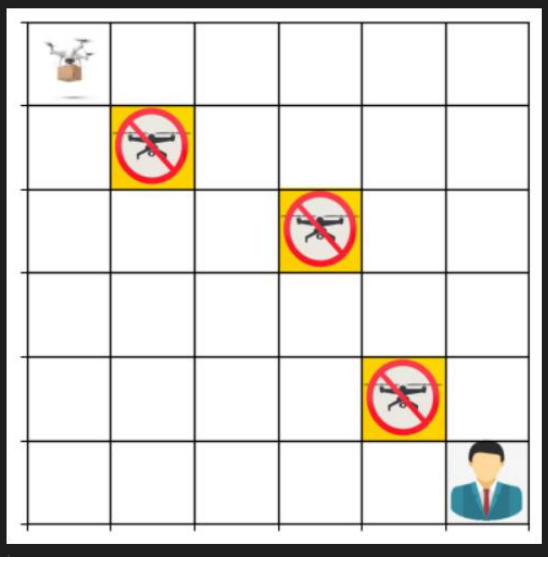
Action: UP, State: (0, 0, False), Reward: -1, Done: False



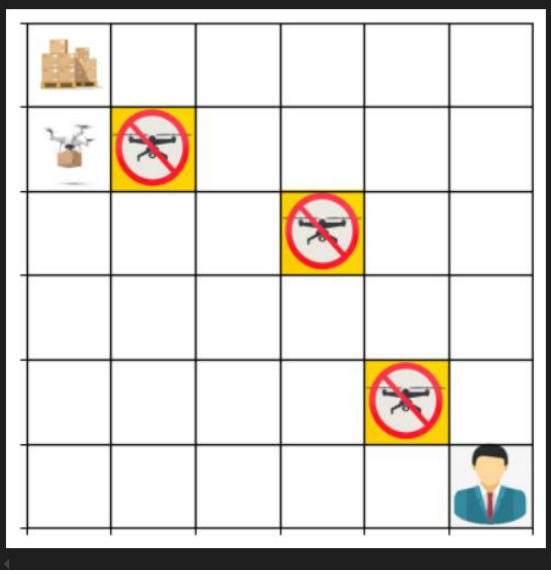
Action: UP, State: (0, 0, False), Reward: -1, Done: False



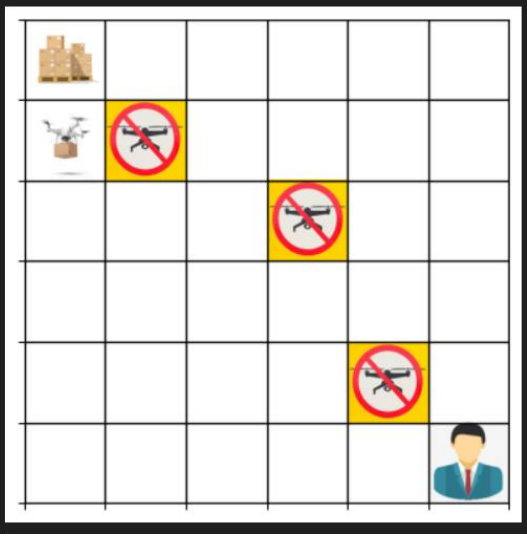
Action: PICKUP, State: (0, 0, True), Reward: 9, Done: False



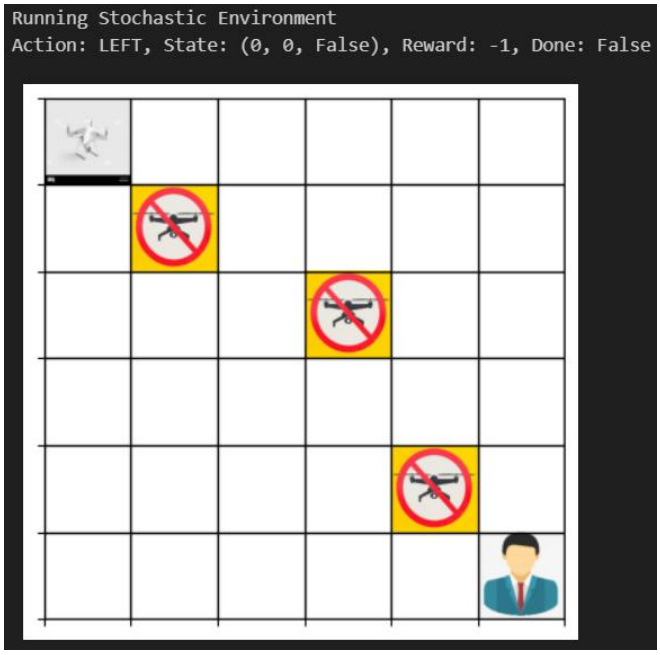
Action: DOWN, State: (1, 0, True), Reward: -1, Done: False

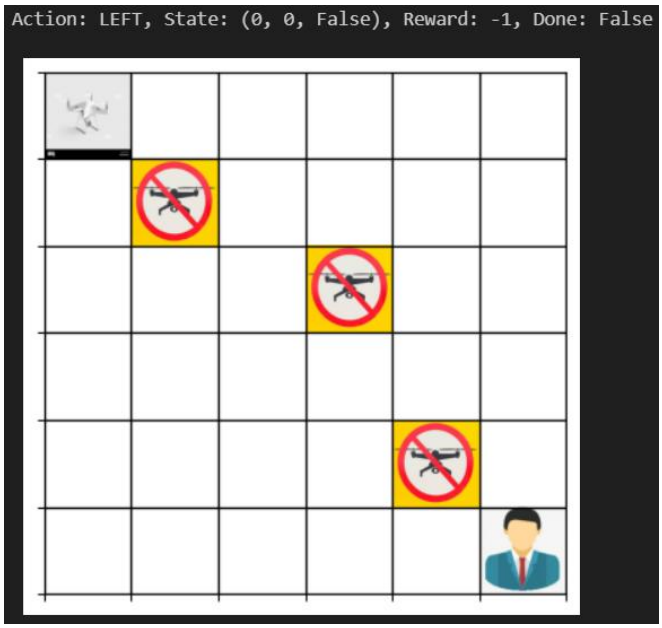


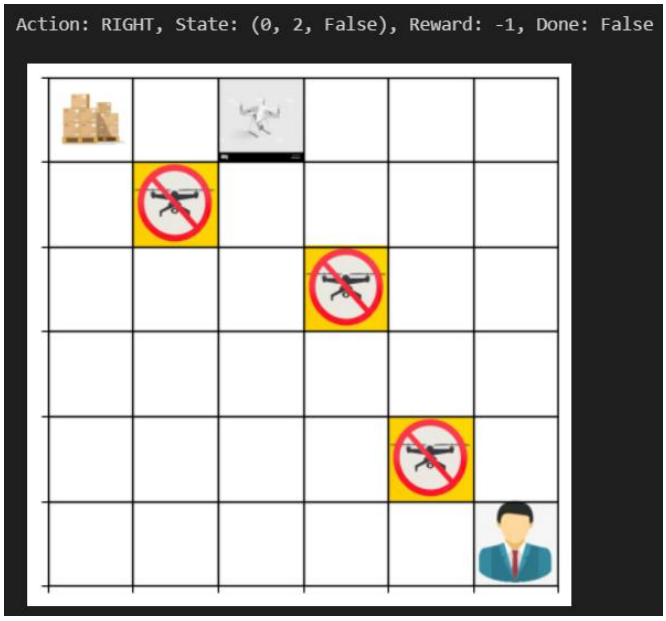
Action: DROPOFF, State: (1, 0, True), Reward: -1, Done: False



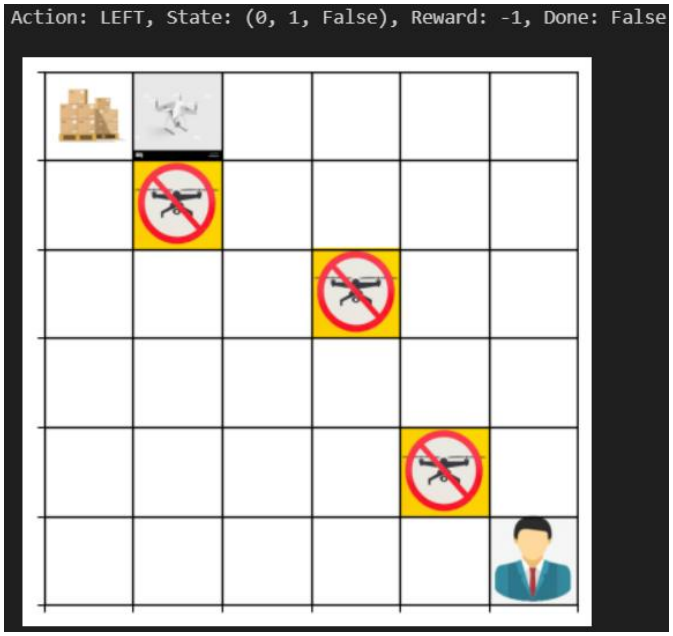
Stochastic Environment:



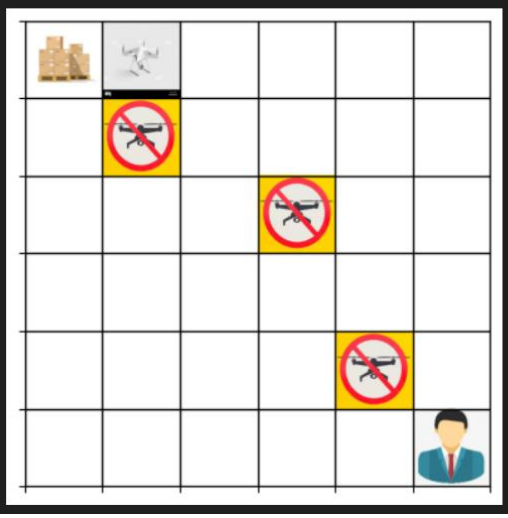




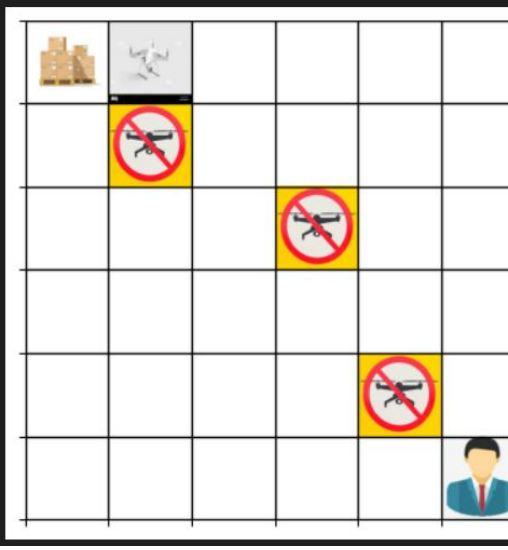
Here we can observe that the drone moved right by 2 boxes, this is because of the action RIGHT, the drone went right 90% but due to stochastic wind conditions of 10% it went right one more box. This proves the stochastic logic.



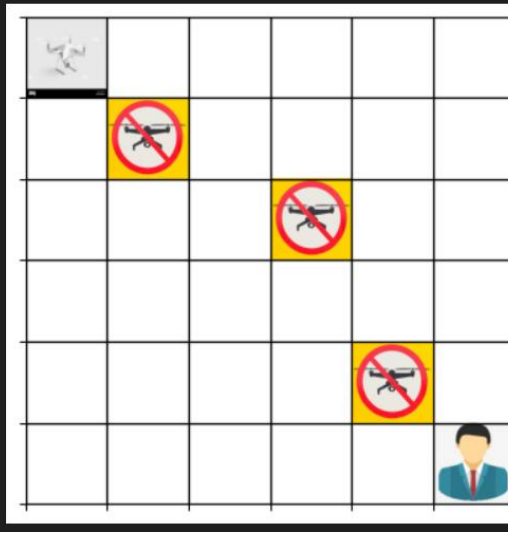
Action: DOWN, State: (0, 1, False), Reward: -20, Done: False



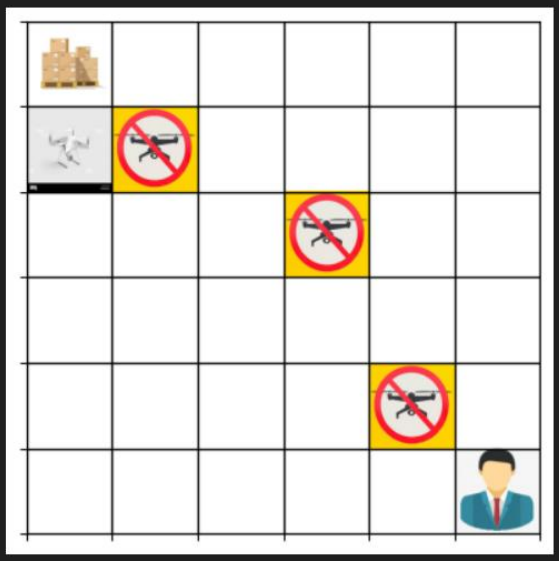
Action: UP, State: (0, 1, False), Reward: -1, Done: False



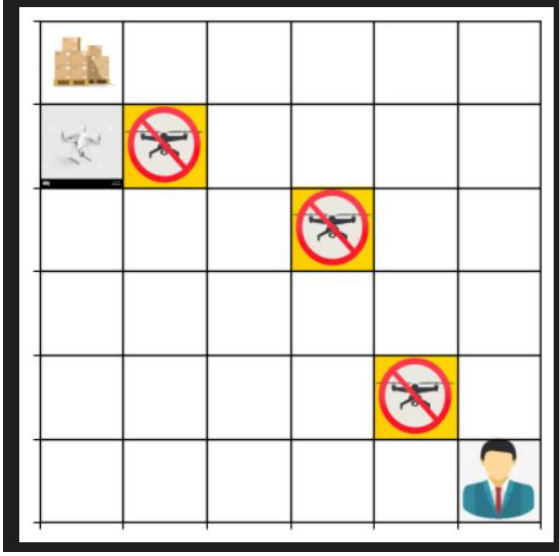
Action: LEFT, State: (0, 0, False), Reward: -1, Done: False



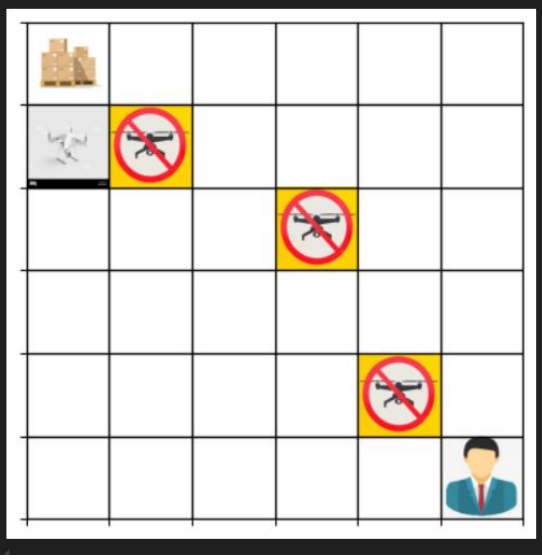
Action: DOWN, State: (1, 0, False), Reward: -1, Done: False



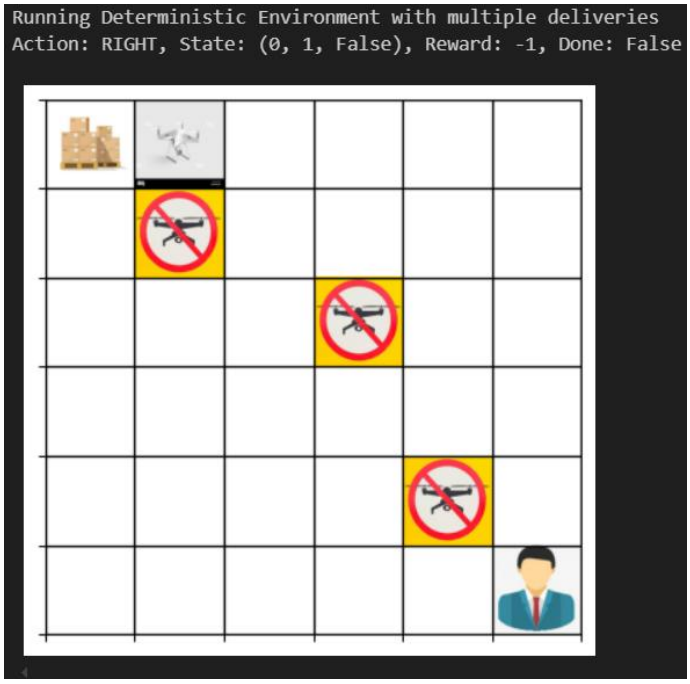
Action: DROPOFF, State: (1, 0, False), Reward: -1, Done: False

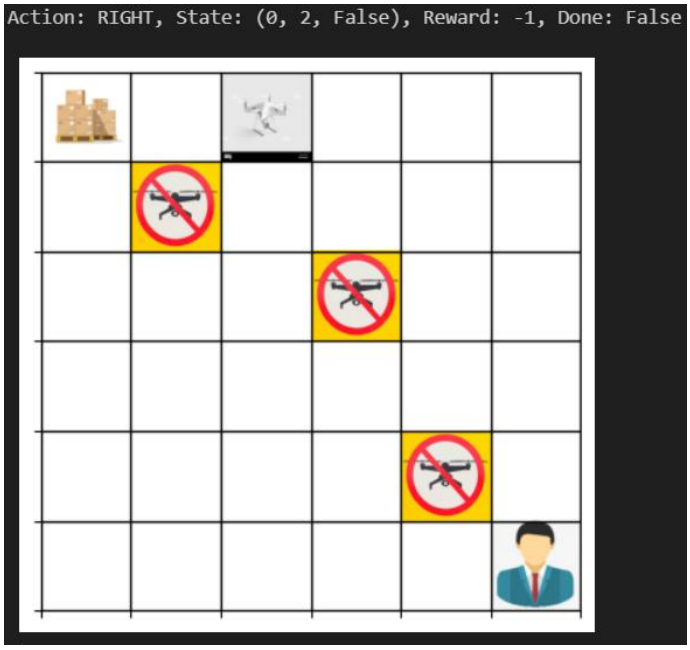


Action: RIGHT, State: (1, 0, False), Reward: -20, Done: False

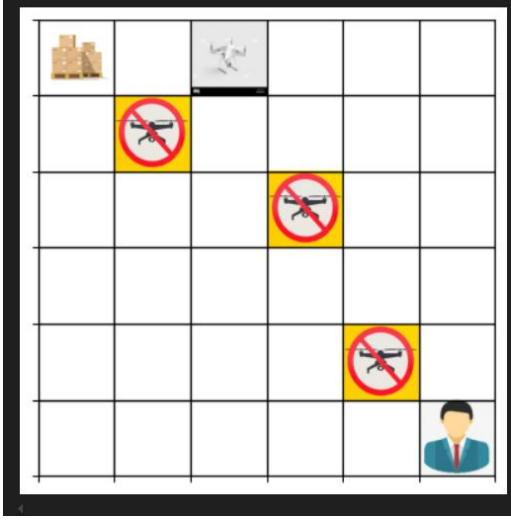


Multiple deliveries visualization logic:

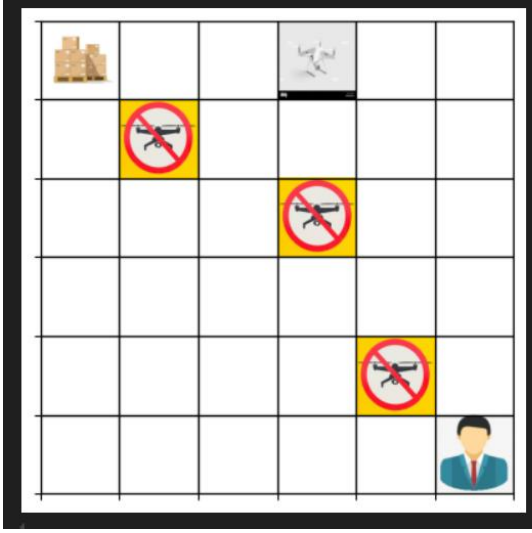




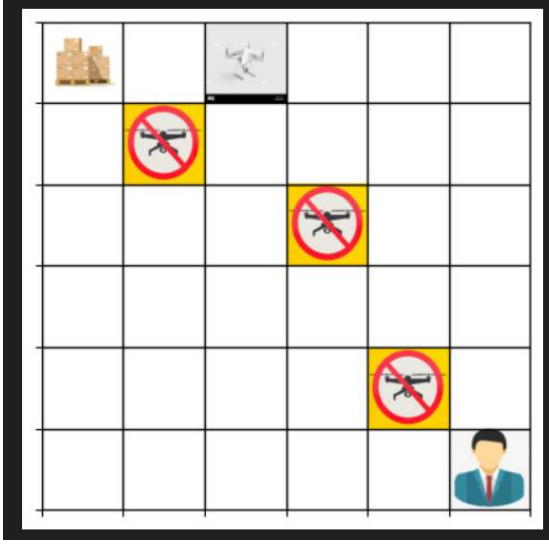
Action: PICKUP, State: (0, 2, False), Reward: -1, Done: False



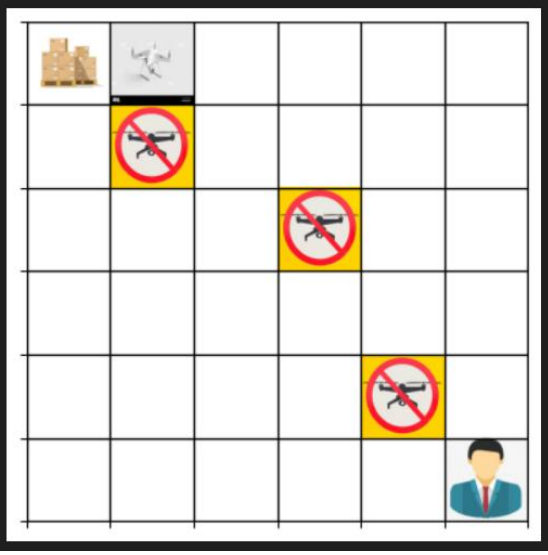
Action: RIGHT, State: (0, 3, False), Reward: -1, Done: False



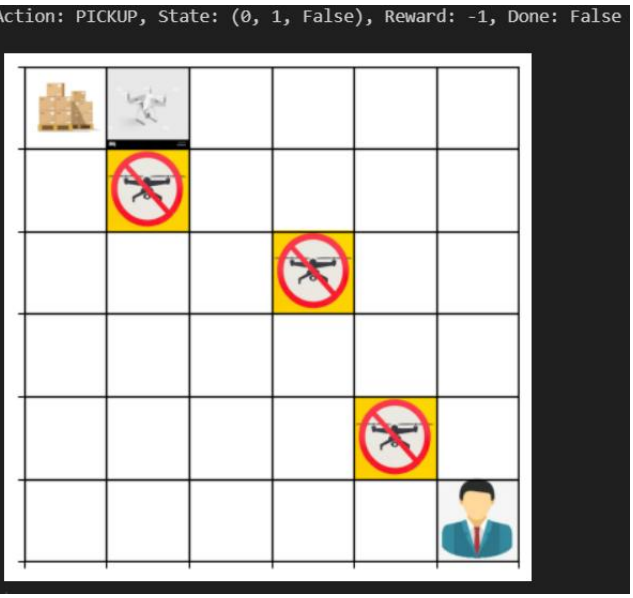
Action: LEFT, State: (0, 2, False), Reward: -1, Done: False



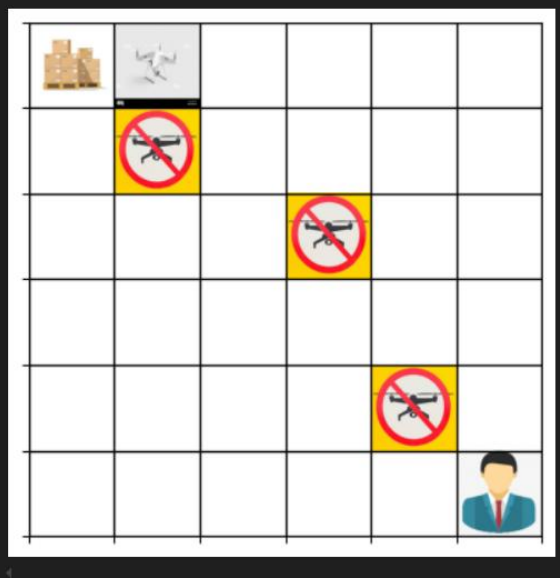
Action: LEFT, State: (0, 1, False), Reward: -1, Done: False

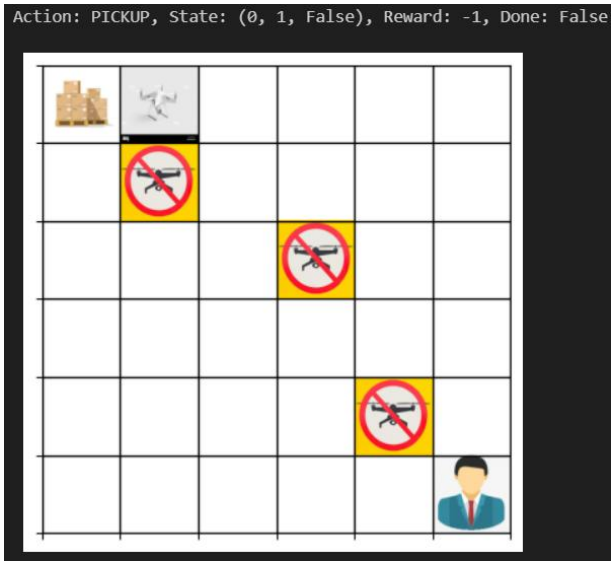


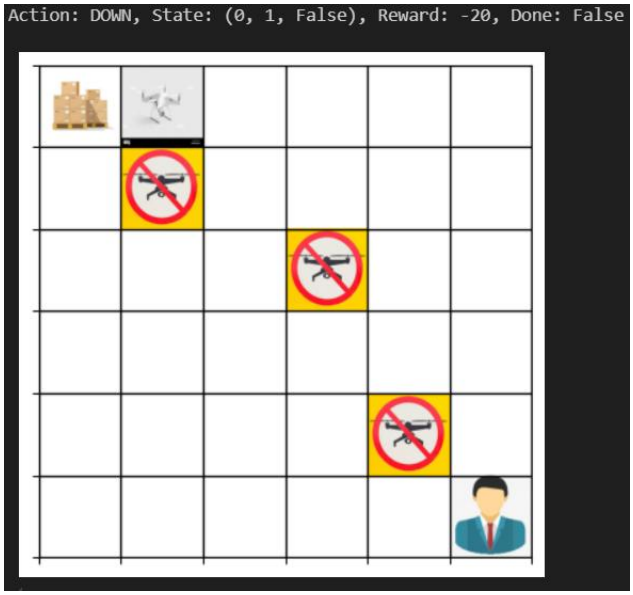
Action: PICKUP, State: (0, 1, False), Reward: -1, Done: False



Action: DOWN, State: (0, 1, False), Reward: -20, Done: False







3. How did you define the stochastic environment?

- The stochastic environment was defined by introducing random deviations in the drone's movement. This was implemented in the _applying_stochastic_conditions method which affects movement with a 10% probability and also makes sure that new position remains within grid boundaries. When the drone attempts to move (UP, DOWN, LEFT, RIGHT), there is a 10% chance that instead of moving in the intended direction, it will move in a random adjacent direction.
  The possible deviations are:
  - (0, 1) → Move right
  - (1, 0) → Move down
  - (0, -1) → Move left
  - (-1, 0) → Move up
- The drone does not always move exactly where it intends and if the drone wants to move RIGHT then there is a 10% chance that it may instead move UP, DOWN, RIGHT or LEFT**.** This increases uncertainty which forces the agent to adapt and plan paths more cautiously.

4. What is the difference between the deterministic and stochastic environments?

1. Deterministic Environment:

- The drone moves exactly in the direction specified by the chosen action (UP, DOWN, LEFT, RIGHT).
- The environment is fully predictable, if the drone chooses to move right, it will always move right.
- The agent can plan optimal paths without worrying about unexpected deviations.
- No randomness is introduced in drone's movement.

2. Stochastic Environment:

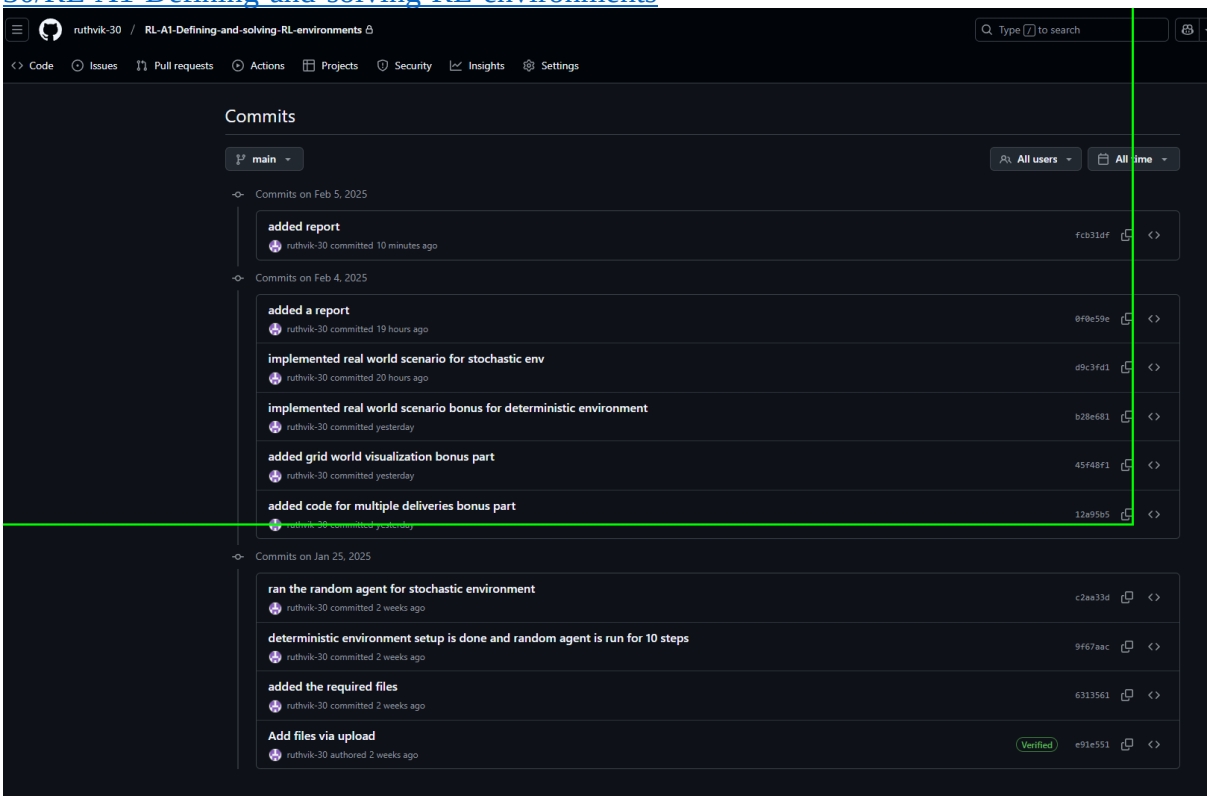
- The drone has a 10% chance of deviating from the chosen direction.
- Actions do not always result in the intended movement, if the drone chooses to move RIGHT, there is a 10% chance it may instead move UP, DOWN, RIGHT or LEFT**.**
- The agent must adapt to occasional movement errors which makes it difficult to plan paths optimally.
- The _applying_stochastic_conditions() function randomly modifies the movement when triggered.

5. Safety in AI:

To keep the AI environment safe, we make sure that the drone only takes actions that are allowed (up, down, right, left, pickup, dropoff). It can't fly into no-fly zones or move outside the grid so it always stays within the defined space. The drone also follows clear rules like only picking up packages at the warehouse and dropping them off at the correct location of the customer. Even in the stochastic environment where some randomness is added, we control it so the drone doesn't move unpredictably or break the game. These checks help the drone navigate safely while still making decisions on its own.

Bonus Tasks:

- **Git Expert:**
  Link to my private GitHub project and commit history: https://github.com/ruthvik-30/RL-A1-Defining-and-solving-RL-environments



References:

- Part I is based on my (Ruthvik Vasantha Kumar) Assignment 3 submission for CSE 574, Fall 2024.