Assignment Checkpoint 1

Y. Harsha Vardhan Reddy

February 6, 2025

1 Introduction

Reinforcement learning environments can be categorized as deterministic or stochastic. This report compares these two types in terms of states, actions, rewards, and objectives. Additionally, it includes visualizations and safety considerations to ensure optimal agent behavior.

2 Deterministic Environment

A deterministic environment is one where the outcome of each action is entirely predictable. The agent follows a predefined strategy where actions lead to the same result every time.

2.1 Environment Setup

- **State Space:** A 6x6 grid, represented as a matrix. Each state includes agent coordinates, pickup status, and drop-off location.
- **Action Space:** Six discrete actions—move up, down, left, right, pick up, and drop off.
- **Obstacles:** Static obstacles at specific coordinates: (1,1), (4,0), (2,5), and (4,4).
- **Pickup Locations:** Two randomly generated points ensuring they do not overlap with obstacles or the drop-off location.
- **Drop-off Location:** Fixed at (5,5).
- **Agent Start Position:** Fixed at (5,0).

2.2 Reward System

- Successful Delivery: +100
- First-time Pickup: +25

- Obstacle Collision/Out of Bound: -20
- Step Penalty: -1 (to encourage efficiency)

2.3 Step Function

- If a move leads to an obstacle or out of bounds, the agent reverts to the previous position and gets a penalty of -20.
- If the agent picks up an item, it gets +25 points, and the pickup location is removed from the list.
- If the agent reaches the drop-off point carrying an item, it gets +100 points.
- The episode terminates after 10 steps.

2.4 Visualization

Figure 1: Deterministic Environment Visualization

3 Stochastic Environment

A stochastic environment introduces randomness, meaning the same action can have different outcomes. The agent balances exploration and exploitation using an epsilon-greedy policy.

3.1 Environment Setup

- **State Space:** Defined as $6 \times 6 \times 4$, incorporating obstacles, robots, boxes, and drop-off locations.
- **Action Space:** Six discrete actions—move up, down, left, right, pick up, and drop off.
- **Multi-Agent Behavior:** Two robots operate independently.
- **Exploration vs. Exploitation:** 10% random actions, 90% best-known actions.

3.2 Reward System

- Successful Delivery: +100
- First-time Pickup: +25
- Obstacle Collision: -20
- Step Penalty: -1

3.3 Episode Termination

- If all robots successfully deliver their boxes.
- If the episode reaches 10 steps.

3.4 Visualization

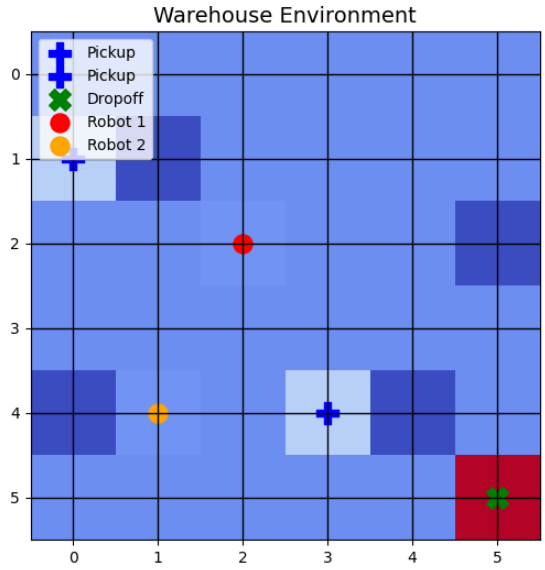

Figure 2: Stochastic Environment Visualization

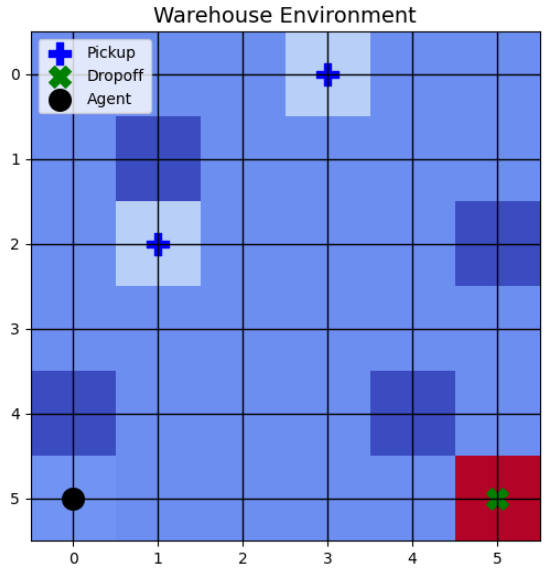

Figure 3: Stochastic Environment Visualization

4 Comparison: Deterministic vs. Stochastic

Feature	Deterministic	Stochastic
Action Outcome	Always predictable	Variable, probabilistic
Exploration	Not necessary	Required for learning
Multi-Agent	Single-agent	Multi-agent
Reward Optimization	Easier	Complex due to randomness

Table 1: Comparison of Deterministic and Stochastic Environments

5 Safety in AI

Ensuring safety in reinforcement learning environments is crucial:

- The agent cannot move into obstacles, preventing crashes and ensuring safety.
- Boundary checks prevent the agent from stepping out of the grid.
- Step penalties encourage efficient movement, reducing unnecessary actions.
- In the stochastic environment, robots operate independently while avoiding interference.
- The balance between exploration and exploitation ensures adaptability while preventing reckless movements.

6 Conclusion

This report compared deterministic and stochastic environments. The deterministic environment offers predictable decision-making, while the stochastic environment introduces uncertainty requiring adaptive learning. Ensuring safety remains a priority in both setups to prevent unintended failures.