Homework-1

Q-Learning Formula

The Q-Learning formula for updating the Q-values for each state-action pair is given by:

$$Q(s,a) = Q(s,a) + \alpha \times \Big[R(s,a) + \gamma \times \max_{a'} Q(s',a') - Q(s,a)\Big]$$

where:

- s is the current state,
- a is the action taken,
- $R(s,a)$ is the reward received after taking action a in state s,
- γ is the discount factor,
- α is the learning rate,
- $\max_{a'} Q(s',a')$ is the highest Q-value for the next state s'.

Initial Q-Table

The initial Q-table for states $S1$, $S2$, $S3$, and actions $A1$, $A2$:

Action	$A1$	$A2$
$S1$	0.5	-0.2
$S2$	-0.1	0.4
$S3$	0.2	0.0

Given Trajectories

Update the Q-table using the following trajectories:

- Trajectory 1: $(S1, A1, +1)$, $(S2, A2, +2)$, $(S3, A1, +3)$
- Trajectory 2: $(S1, A2, +0)$, $(S2, A1, -1)$, $(S3, A2, +4)$
- Trajectory 3: $(S2, A2, +2)$, $(S3, A1, +0)$, $(S1, A1, +1)$
- Trajectory 4: $(S3, A2, -2)$, $(S2, A1, +3)$, $(S1, A2, +0)$

Parameters:

- Learning rate α: 0.1
- Discount factor γ: 0.9

Bonus Question: Introducing the "Jetpack" Cell

In the enhanced task, we introduce a new type of cell called the "Jetpack" cell. This cell provides the agent with an ability to navigate through hazard blocks without receiving the usual negative rewards. This simulates the agent acquiring a jetpack that allows it to fly over hazards safely.

Code Modifications for "Jetpack" Cell

To incorporate the "Jetpack" cell into your reinforcement learning environment, you will need to make several modifications:

1. **State Representation Update:** Include information about whether the agent currently has the "Jetpack" ability as part of the state representation.
2. **Reward Function Modification:** Update the reward function to avoid applying the negative reward for moving through hazards if the "Jetpack" ability is active.
3. **Transition Dynamics Update:** Integrate logic for the acquisition and use of the "Jetpack" ability within the environment's dynamics.
4. **Learning Algorithm Adjustment:** Ensure the Q-learning algorithm properly accounts for the new state space and reward dynamics introduced by the "Jetpack" cell.

Modify the provided Q-learning code to accommodate the introduction of the "Jetpack" cell. Update the state representation, reward function, and implement logic for handling the "Jetpack" ability. Test your modified code to assess how the "Jetpack" cell influences the agent's learning and decision-making in environments with hazard blocks.

Note: Participants who successfully solve the bonus question will possibly be rewarded with a small prize! Enjoy the challenge and sweeten your success!