Optimizing Strategies Against Uncertainty: A Comparative Study Of Advanced Learning Algorithms In Stochastic Environments With Visibility Constraints

MICHIGAN STATE UNIVERSITY

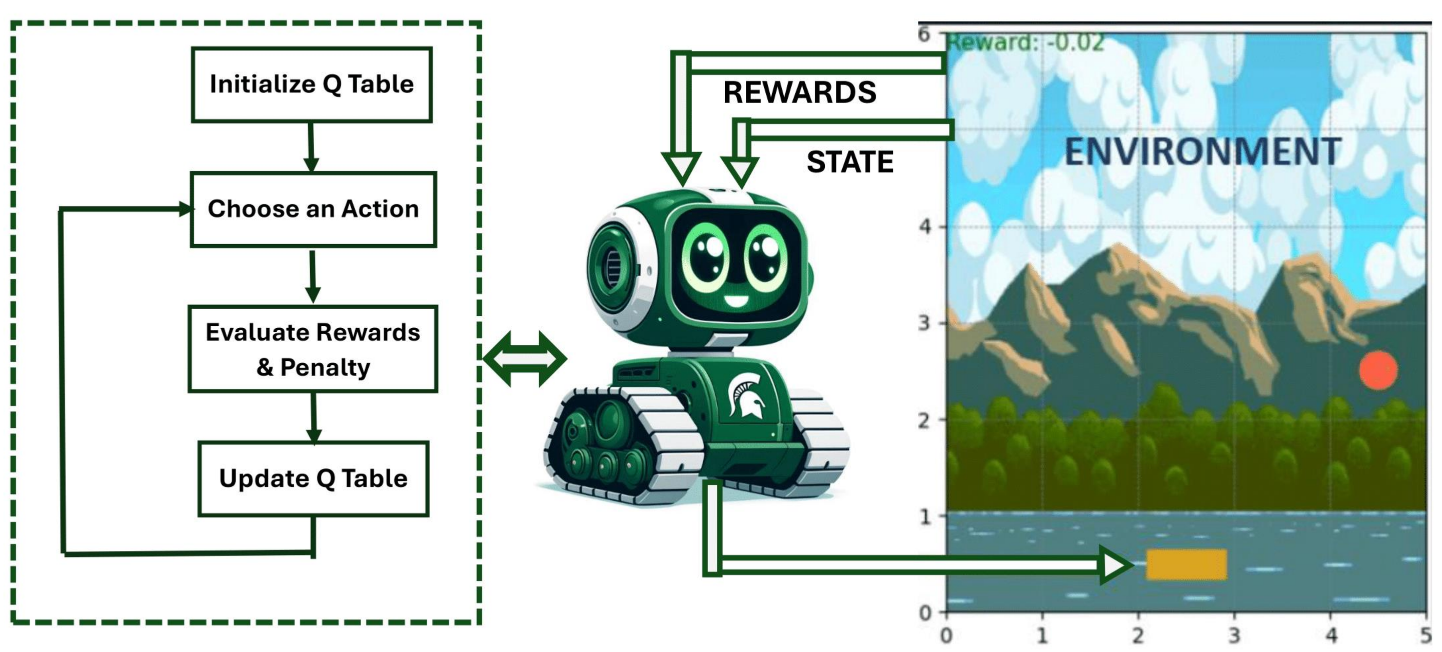
Author: MEGHA VISWANATH under the guidance of DR. SHAUNAK D BOPARDIKAR

INTRODUCTION

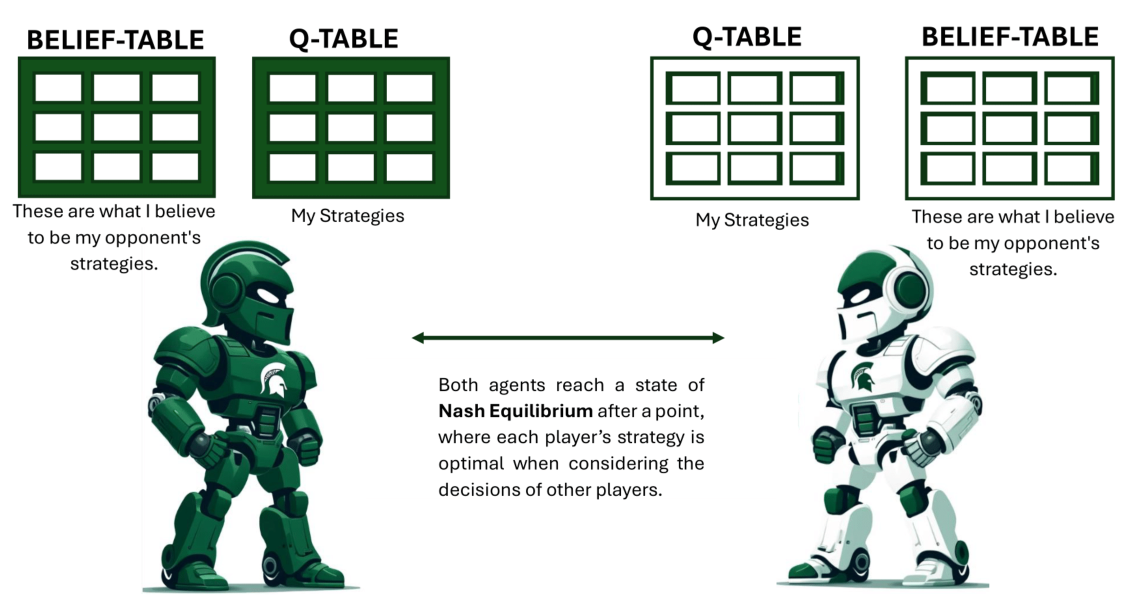
In exploring artificial intelligence, the game "Catch" offers a crucial testbed for refining learning algorithms, where an AI agent must catch falling objects with unpredictable paths to earn points. This study introduces additional complexity with visibility constraints, requiring strategic decisions based on limited views of object types. We evaluate Q-Learning and Deep Q-Networks (DQN) to manage this randomness and maximize scores autonomously. By incorporating a dual-agent scenario, we further test the scalability and adaptability of these models in dynamic, uncertain environments.

METHODS

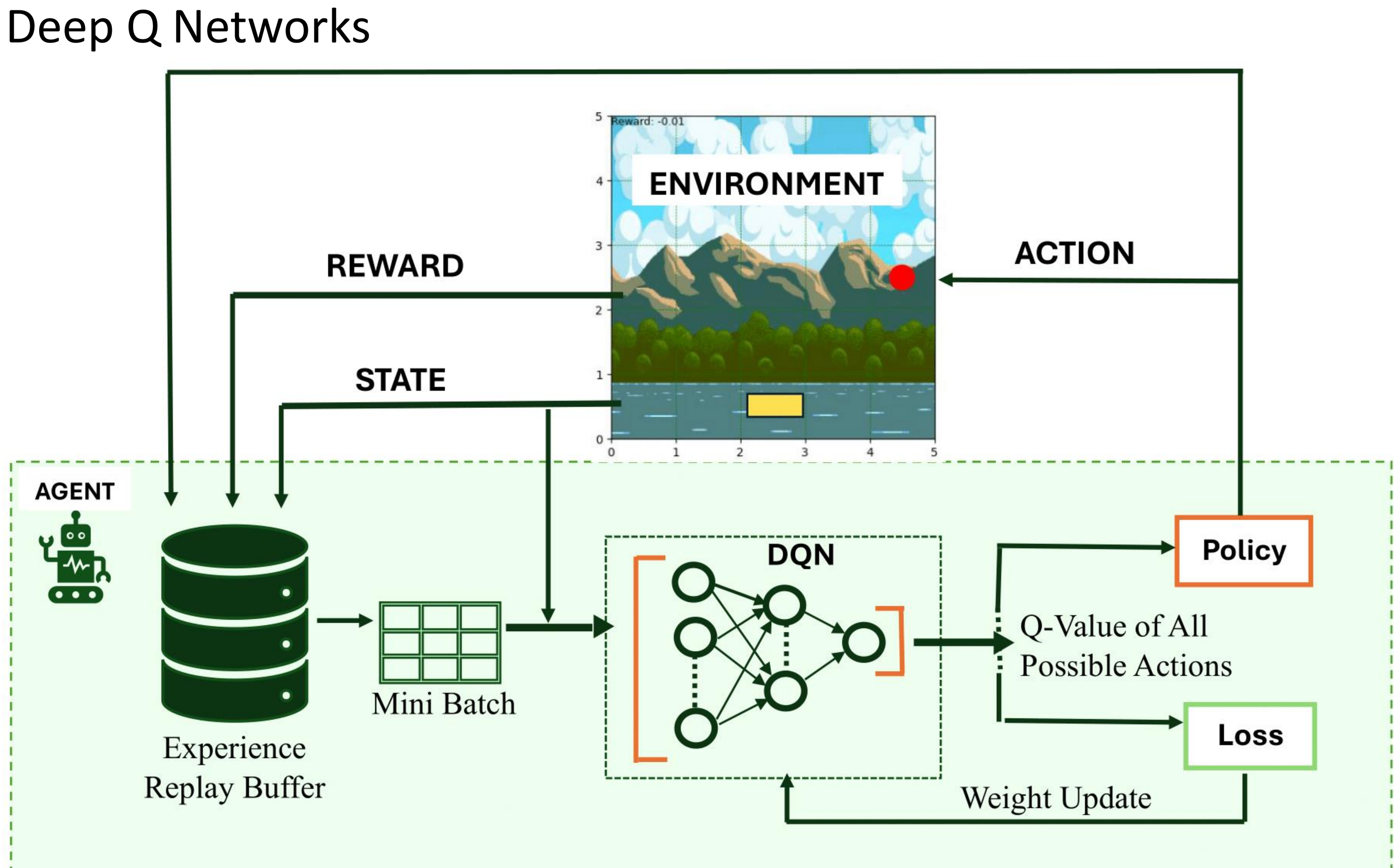
1. Q-Learning Algorithm



2. Fictitious Play for two-player zero sum stochastic game

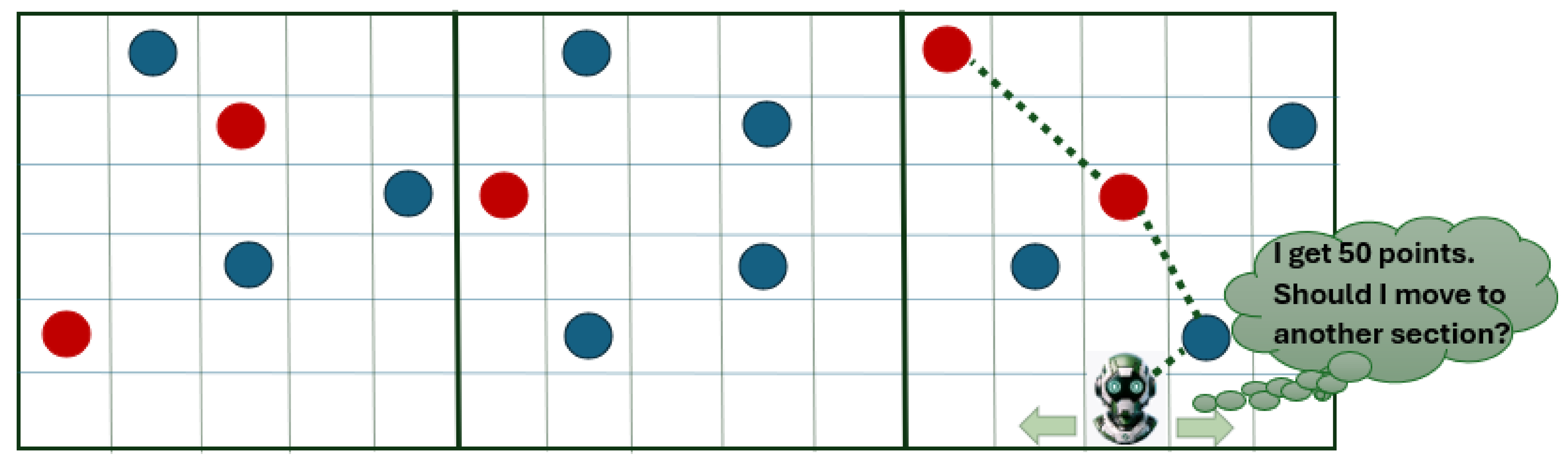


3. Deep Q Networks



In DQN, the employment of a neural network to approximate the Q-value function obviates the need for the exhaustive state-action pair tables that were utilized in the previous two models based on Q-Learning.

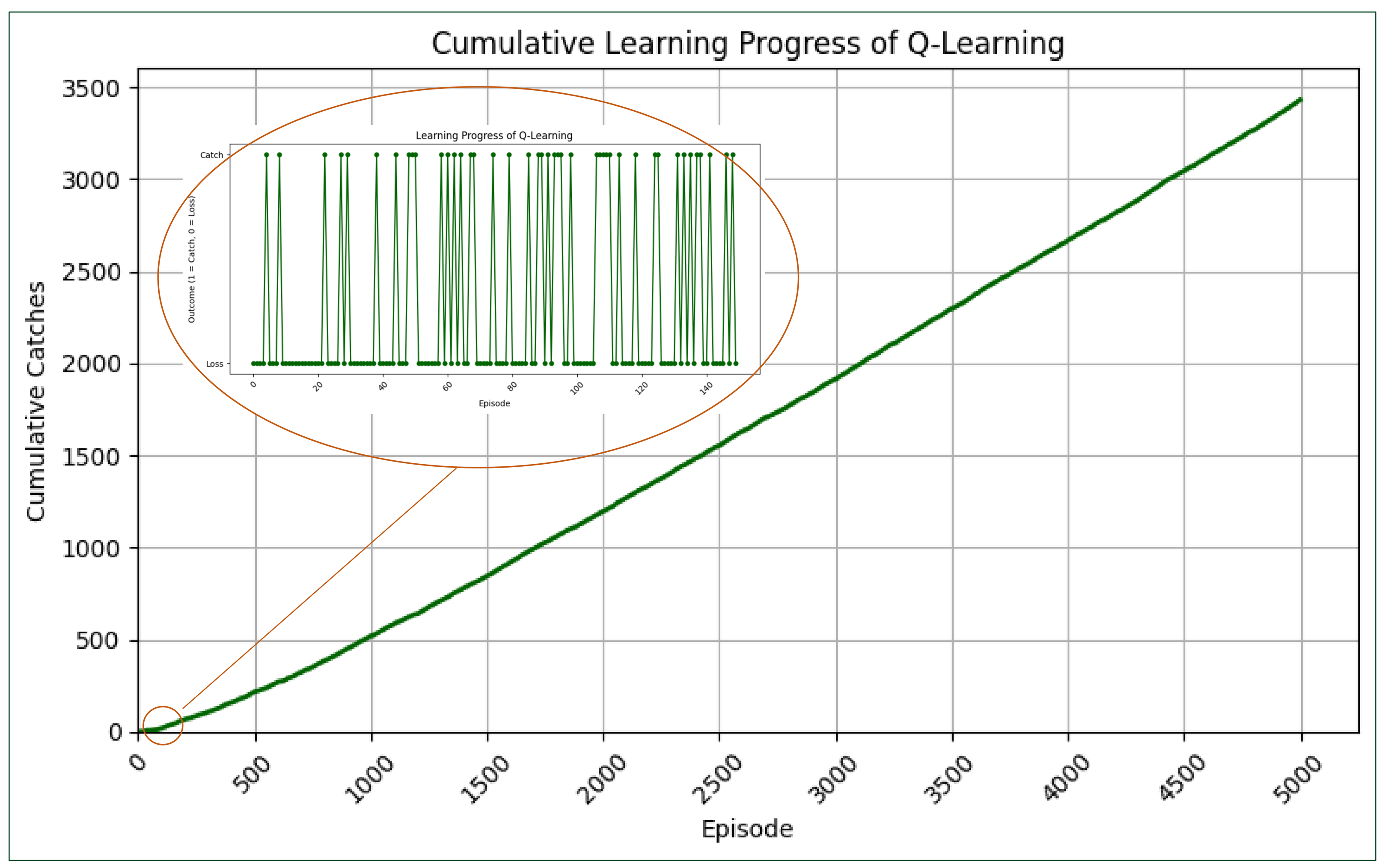
4. Evaluating Neural Network Strategies in Partially Observable Game Fields

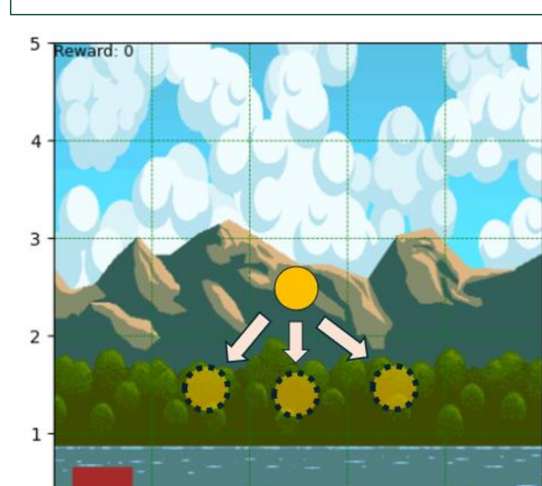


This model utilizes two neural networks in a restricted view game environment. The first network assesses potential points from visible objects, while the second decides whether to cross into another section based on the anticipated net gain or loss, enhancing the AI's strategic decision-making under uncertainty.

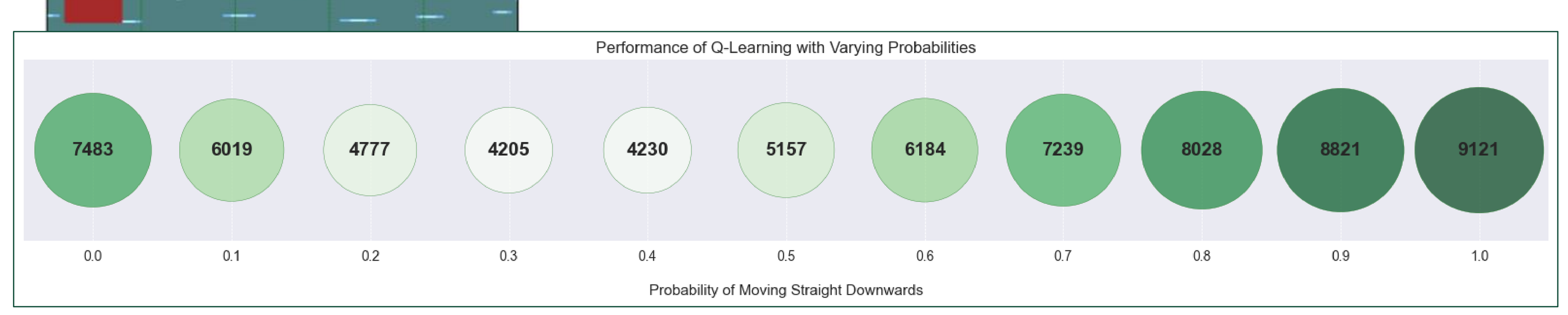
RESULTS

In a 5x6 game environment, our Q-Learning AI, trained over numerous episodes, progressed from initial frequent misses to markedly improved catch rates, adapting to the stochasticity inherent in 0.1 directional probabilities.

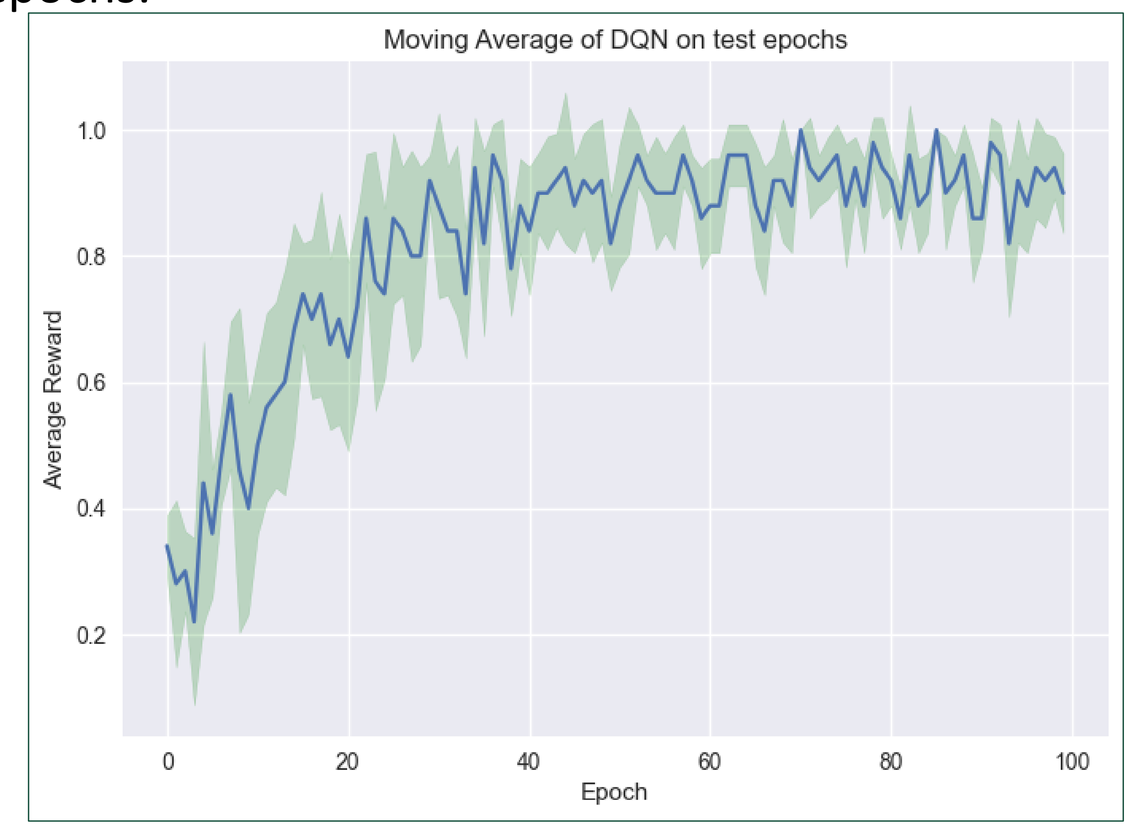




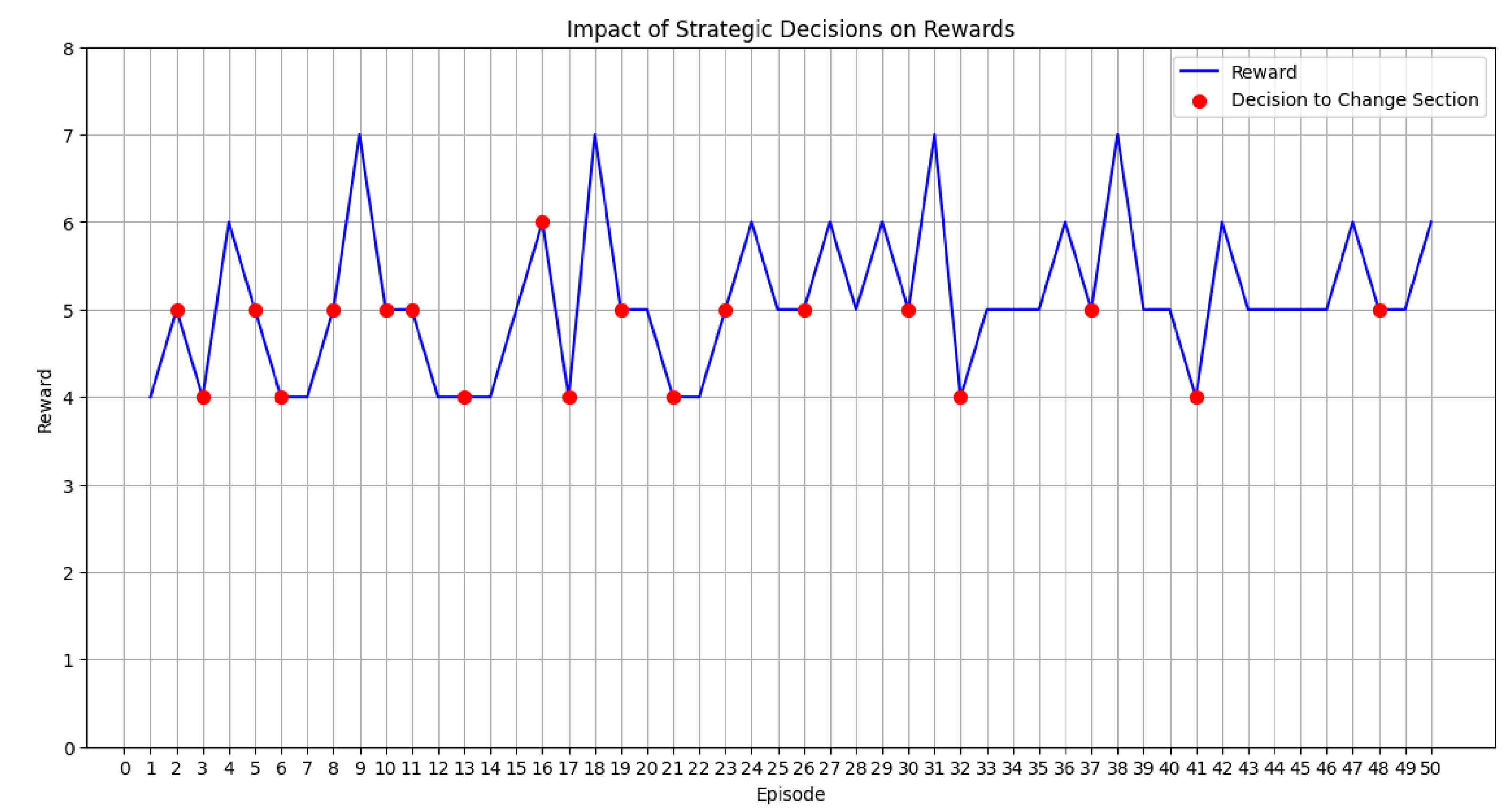
What if the probability of moving ball changes? The visualization below show that the agent performs well when objects fall mostly downward or purely sideways, while a mix of movements leads to more losses, revealing how the agent adapts to different patterns of unpredictability.



The DQN moving average plot below displays steady improvement and stabilization of average reward over time, reflecting the agent's learning consistency and gradual refinement of strategy throughout the test epochs.



The plot shows that the AI agent learns to minimize section changes, only doing so after rewards dip below 5 points, indicating increased strategic decision-making over time.



CONCLUSIONS

This study highlights that while Q-Learning and Fictitious Play excel in smaller environments, their use becomes limited in larger ones due to the exponential growth of the state space. Conversely, Deep Q-Networks (DQN) effectively address the challenge of an expanding state space through neural network function approximation and experience replay, offering a scalable solution for stochastic and uncertain environments. While DQN requires more computational power, leveraging High-Performance Computing Clusters (HPCC) can expedite its processes. To understand the convergence of strategy and Q-function beliefs to a stationary equilibrium in Fictitious Play the paper "Fictitious Play in Zero-Sum Stochastic Games" provides all details.

REFERENCES

1. M.O. Sayin, F. Parise, and A. Ozdaglar for their influential paper "Fictitious Play in Zero-Sum Stochastic Games."