



Optimising Reward Shaping for Sparse-Reward Environments: A Comparative Study of Potential-Based Functions in DQN and Double DQN on Grid-World Navigation

Niveditha S Nair

B.Tech Student, Chinmaya Vishwavidyapeeth, Deemed-to-be University, Ernakulam, Kerala, India

Pradeeba V

Chinmaya Vishwavidyapeeth, Deemed-to-be University, Ernakulam, Kerala, India

Niby Babu

Chinmaya Vishwavidyapeeth, Deemed-to-be University, Ernakulam, Kerala, India

Abstract:

Training a deep reinforcement learning agent on a sparse-reward task is a frustrating experience. The agent wanders around collecting zeros until it accidentally finds the goal, and by the time you have enough successful episodes to form a useful gradient, you have already burned through most of your training budget. This paper investigates how much that situation improves when you add a mathematically structured hint to the reward signal – specifically, a potential-based shaping function that gives the agent directional feedback without changing what the optimal policy actually is. We built a 15×15 grid navigation task in Gymnasium and tested three shaping designs: one based on Manhattan distance to the goal (MDWS), one that adds a time-pressure term to that (TPDS), and one that also factors in how close the agent is to obstacles (OPS). We ran each of these on both DQN and Double DQN – 40 training runs in total, five random seeds per configuration. The clearest result is that TPDS combined with DDQN converges in 210.20 ± 20.67 episodes on average, which is 52.93% faster than training DQN with no shaping at all (446.60 ± 2.06 episodes). Every comparison is statistically significant after Bonferroni correction. What surprised us was that despite these large differences in convergence speed, the final MSE loss and collision-avoidance behaviour were nearly identical across all eight configurations. Shaping made training faster; it did not make the learned policy better. That distinction matters when you are deciding whether to use shaping in a deployment pipeline.

Keyword:

Sparse Reward, Reward Shaping, Potential-Based Shaping, Deep Q-Network, Double DQN.