RL Simulator Report

Algorithm: SARSA
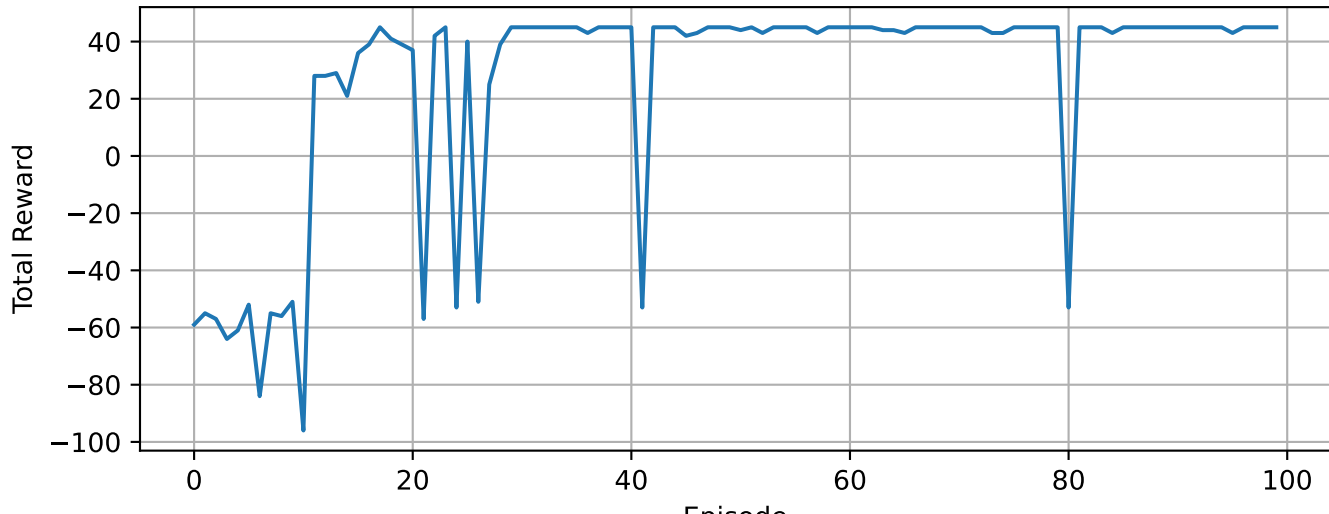
Episodes Trained: 100

⬚ A	→ B	↓ C	↑ D
↓ E	⬚ F	↓ G	⬚ H
→ I	→ J	↓ K	↓ L
⬚ M	→ N	→ O	⬚ P

Q-Value Heatmap (Best Actions)

State Visit Frequency
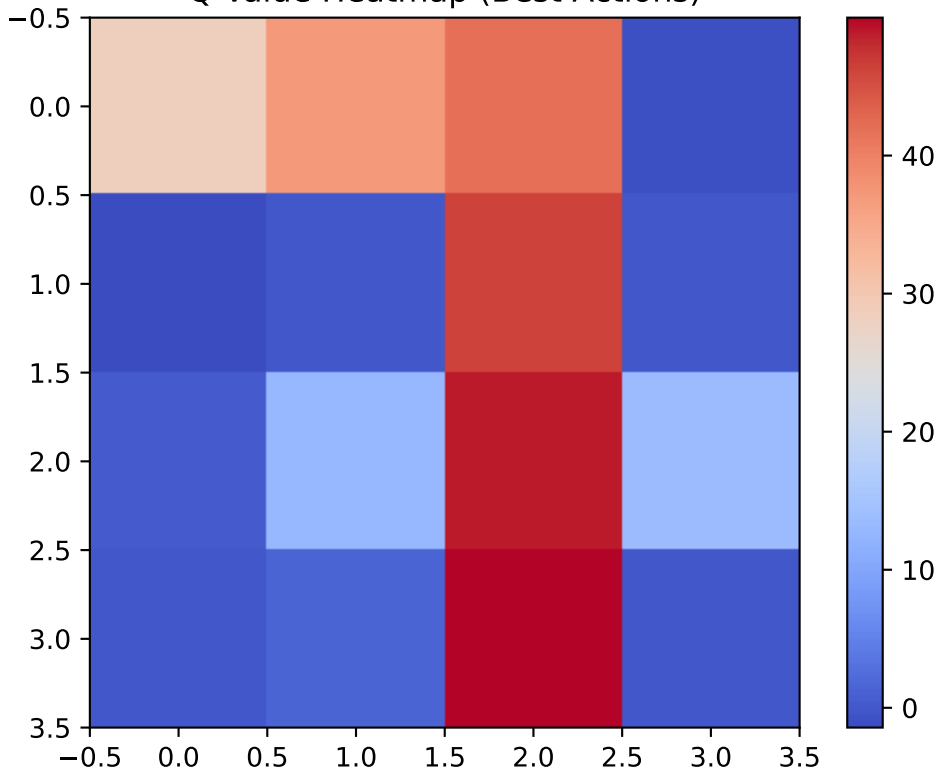
Reward vs Episode

Last Episode Path

A → B → C → G → K → O → P