Reinforcement Learning Assignment3 Report

Ananya Hazarika, Kuhu Gupta, Nandika Soni, Meheeka Pandey- Team 1

4 MDP formulation

a. States and State Space

$$S = (Q_1, Q_2, G, T)$$

where:

- $Q_1, Q_2 \in \{0, 1, \ldots, 20\}$
- $G \in \{0, 1\}$
- $T \in \{0, 1, \ldots, 9\}$

b. Action and Action Space

$$A = \{0, 1\}$$

where:

- 0: Maintain current green light
- 1: Switch green light

c. Reward

$$R = -(Q_1 + Q_2)$$

d. State Transition Equations

1. **Arrivals**:

$$Q_i \to Q_i + 1 \quad \text{with probability } p_i$$

where:

- $p_1 = 0.245$, $p_2 = 0.35$
- $Q_i \leq 1810$

2. **Departures**:

$$Q_i \to Q_i - 1 \quad \text{with probability:}$$
$$\begin{cases} p_l + \frac{(p_h - p_l)}{\tau} \times T, & T < \tau \\ p_h, & T \geq \tau \end{cases}$$

where:

- $p_l = 0.2$, $p_h = 0.9$, $\tau = 10$

3. **Timer Update**:

$$T = \begin{cases} 0, & \text{if action } = 1 \\ \min(T + 1, 9), & \text{if action } = 0 \end{cases}$$

1 TripleQLearning Policy Results

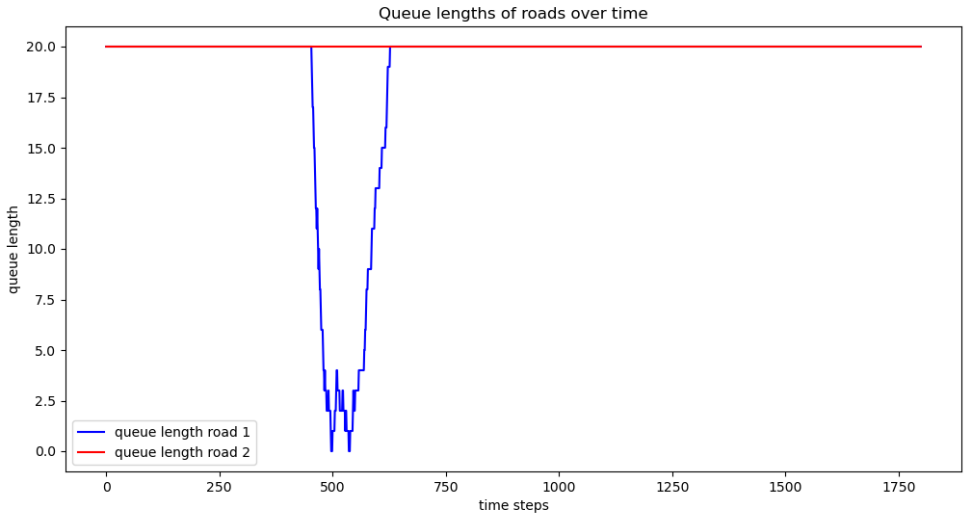

Figure 1: Queue lengths of roads over time using Triple Q-Learning

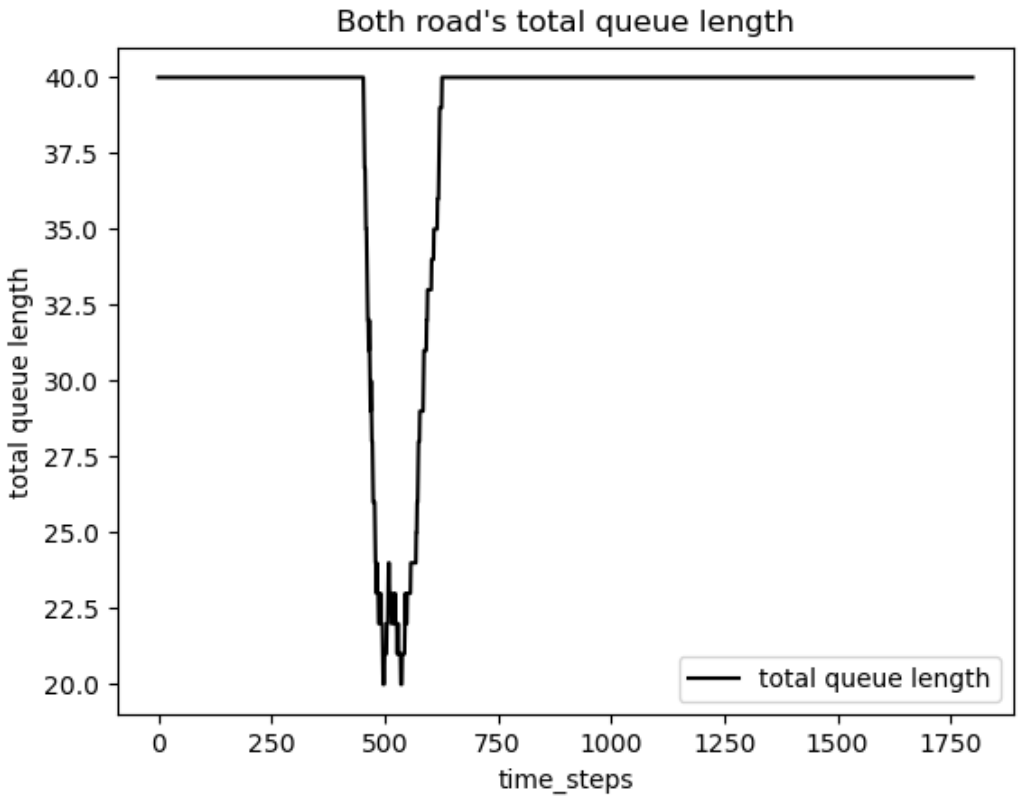

Figure 2: Combined queue length of both roads over time using Triple Q-Learning

2 Modified SARSA Policy Results

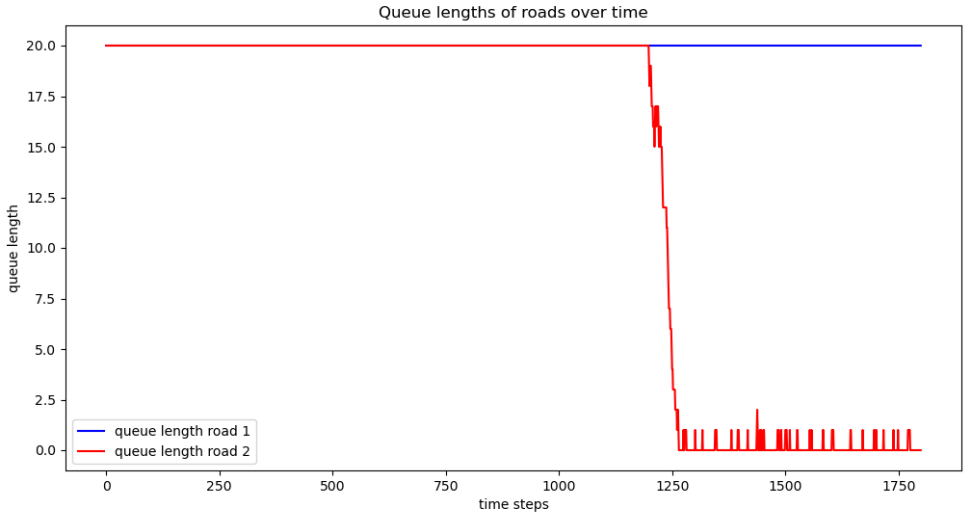

Figure 3: Queue lengths of roads over time using Modified SARSA

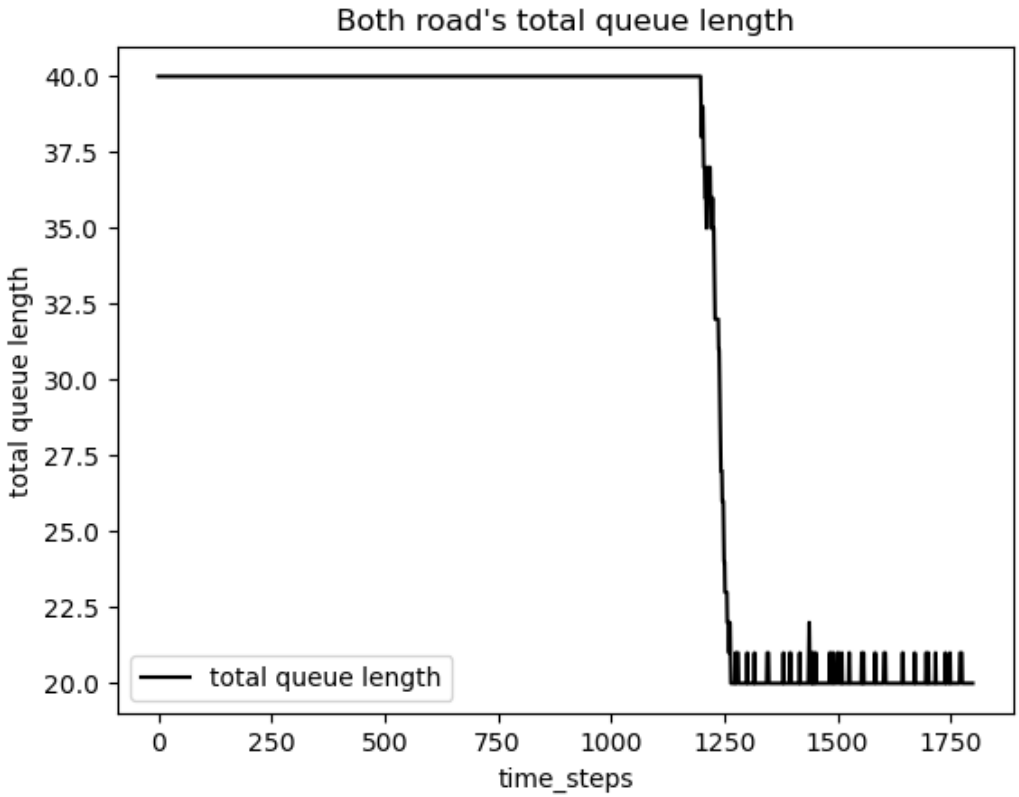

Figure 4: Combined queue length of both roads over time using Modified SARSA