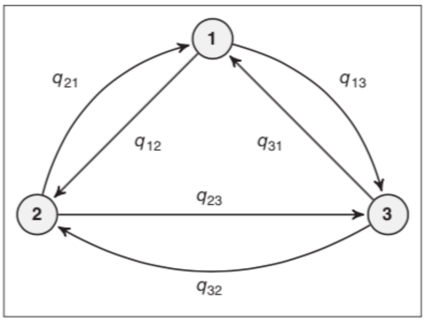
SI252 Reinforcement Learning 2025/03/13

Homework 2

Professor: Ziyu Shao *Due: 2025/03/23 11:59pm*

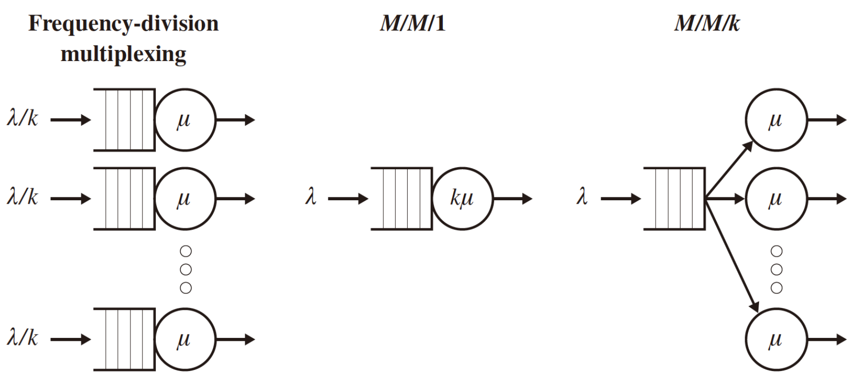
1. Six problems from Harvard Textbook BH (the first edition): Chapter 11, Problems 2, 4, 6, 7, 8, 14

2. Given a three-state CTMC with transition rates between states shown in the following diagram:



   Find holding time for each state, transition probability matrix of embedded chain, and the generator matrix.

3. **Three Server Organizations**: in the following figure, we show three data center systems with the same arriving rate λ and the same total service rate $k\mu$: FDM, M/M/1 and M/M/k. Please discuss the pros and cons of each system. Regarding the performance of average delay, which system is the best? Show your analysis (CTMC & Queueing Theory) and simulation results.



4. Consider a continuous-time Markov chain $\{X_t, t \geq 0\}$ with a finite discrete state space $\mathcal{X}$ and a transition rate matrix $Q = (q_{i,j}, i, j \in \mathcal{X})$. Let

$$p_{i,j}(t) = P(X_{t+s} = j | X_s = i) = P(X_t = j | X_0 = i).$$

Show the following results:

(a) $s, t \geq 0$

$$p_{i,j}(t+s) = \sum_{k \in \mathcal{X}} p_{i,k}(t) p_{k,j}(s).$$

(b) Kolmogorov's Backward Equation:

$$p'_{i,j}(t) = \sum_{k \neq i} q_{i,k} p_{k,j}(t) - v_i p_{i,j}(t).$$

where $v_i = -q_{i,i}$.

(c) Kolmogorov's Forward Equation:

$$p'_{i,j}(t) = \sum_{k \neq j} q_{k,j} p_{i,k}(t) - v_j p_{i,j}(t).$$

where $v_j = -q_{j,j}$.

5. Consider a continuous-time Markov chain $\{X_t, 0 \leq t \leq T\}$ with a finite discrete state space $\mathcal{X}$. It's transition rate matrix is time-varying and is denoted by $Q_t = \{Q_t(x, y), x, y \in \mathcal{X}\}$. The state distribution in time $t \in [0, T]$ is denoted by $\pi_t = \{\pi_t(x), x \in \mathcal{X}\}$. Now we consider the reverse time process of continuous-time Markov chain $\{X_t, 0 \leq t \leq T\}$, and denote it as $\{\overline{X}_t, 0 \leq t \leq T\}$. Then we have $\overline{X}_t = X_{T-t}$. Show the following results:

(a) The reverse time process $\{\overline{X}_t, 0 \leq t \leq T\}$ is a continuous-time Markov chain.

(b) Denote the time-varying transition rate matrix of such reverse time process as $R_t = \{R_t(x, y), x, y \in \mathcal{X}\}$, then

$$R_t(x, y) = \frac{\pi_t(y)}{\pi_t(x)} Q_t(y, x)$$