CS 4226 Sample Answers of Written Assignment

Problem 1

You flip a biased coin many times. Each time, you get a head with probability p and a tail with probability $1-p$ independently. We define a random variable X to be the number of trials before you get the first head.

a) What is the distribution of X? Write down the probability distribution function $F(x) = P\{X \le x\}$?
b) What is the memoryless property? Explain it in words.
c) Show that the random variable X satisfies the memoryless property. State the memoryless property of X in a mathematical statement first, and then prove it.

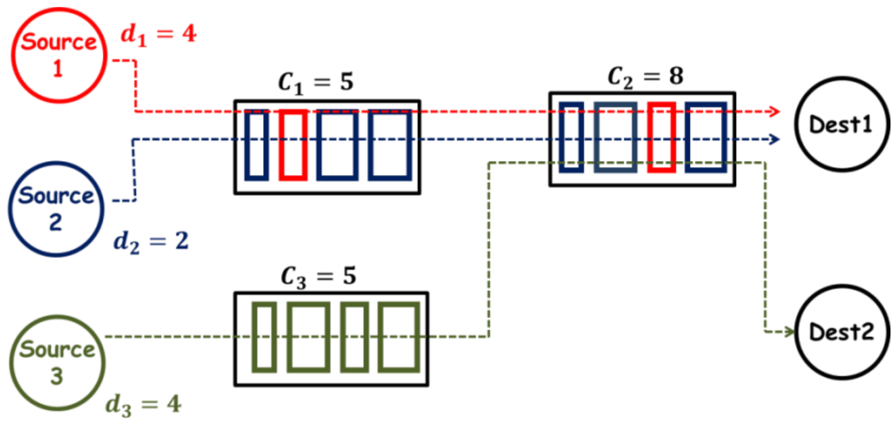
Sample answers:

a) X follows a geometric distribution. We first derive the probability that you need exactly i trials to get the first head, i.e., the first $i-1$ outcomes are all tails and the ith outcome is a head:
$$P\{X = i\} = (1-p)^{i-1}p$$
Then the distribution function can be calculated as:
$$P\{X \le x\} = \sum_{i=1}^{x} P\{X = i\} = \sum_{i=1}^{x} (1-p)^{i-1}p = p\sum_{i=1}^{x}(1-p)^{i-1} = p\frac{1-(1-p)^x}{1-(1-p)}$$
$$= 1-(1-p)^x$$
b) Memoryless property is that the future only depends on the history via the current state, but does not depend on the past.

c) We can show that $P\{X > m+n | X > m\} = P\{X > n\}$. Since $P\{X \le x\} = 1-(1-p)^x$, we have $P\{X > x\} = 1 - P\{X \le x\} = (1-p)^x$.

$$P\{X > m+n | X > m\} = \frac{P\{X > m+n, X > m\}}{P\{X > m\}} = \frac{P\{X > m+n\}}{P\{X > m\}} = \frac{(1-p)^{m+n}}{(1-p)^m} = (1-p)^n$$
$$= P\{X > n\}$$
The above derivation is very similar to the proof of memoryless property for the exponential random variables.

Problem 2



Consider the above topology with three flows and three links. The three flows have demands $d_1 = 4$, $d_2 = 2$ and $d_3 = 4$. The three links have capacities $C_1 = 5$, $C_2 = 8$ and $C_3 = 5$. Please identify the max-min fair allocation for the three flows and their bottleneck link(s).

a) Calculate the weighted max-min fair allocation $\boldsymbol{x} = (x_1, x_2, x_3)$ to the three flows when the weights satisfy $\phi_1 : \phi_2 : \phi_3 = 2:1:1$.
b) Calculate the weighted max-min fair allocation $\boldsymbol{x} = (x_1, x_2, x_3)$ to the three flows when the weights satisfy $\phi_1 : \phi_2 : \phi_3 = 1:2:2$.

Sample answers:

a) $\boldsymbol{x} = \left(\frac{10}{3}, \frac{5}{3}, 3\right)$. C_1 is the bottleneck for flow 1 and 2, C_2 is the bottleneck for flow 3.
b) $\boldsymbol{x} = (2,2,4)$. C_2 is the bottleneck for flow 1 and 3.

Problem 3

Under an M/M/1 system that uses First-In-First-Out (FIFO) scheduling policy. Suppose the arrival rate is λ and the service rate is μ.

a) If a packet arrives and finds that the queue is empty, but the server is busy, what is this packet's mean sojourn time?
b) If a packet arrives and finds that there are n other packets in the queue, what is this packet's mean sojourn time?
c) Can you derive the mean sojourn time for a typical packet?

Sample answers:

a) Since the system uses FIFO, this packet's sojourn time equals the remaining service time of the packet in the server, plus this packet's own service time. Because under M/M/1, the service time are exponentially distributed with mean $1/\mu$ and because exponential random variable has the memoryless property, the remaining service time of the packet in the server will have the same exponential distribution. Therefore, this packet's mean sojourn time equals $\frac{1}{u}+\frac{1}{\mu}=\frac{2}{\mu}$.

b) Similarly, this packet's sojourn time equals $(1+n+1)\frac{1}{\mu}=\frac{n+2}{\mu}$.

c) By part a) and b), we know that given that a packet sees l packets in the system (server + queue) upon arrival, its own mean sojourn time is $\mathrm{E}[W|L=l]=\frac{l+1}{\mu}$. To calculate the mean sojourn of the packets, we can "divide and conquer" by conditioning on how many packets exist in the system upon a packet's arrival.

$$\mathrm{E}[W]=\sum_{l=0}^{+\infty}\mathrm{E}[W|L=l]\,\mathrm{P}\{L=l\}$$

For M/M/1, we know that $\mathrm{P}\{L=l\}=\rho^l(1-\rho)$, therefore, we have

$$\mathrm{E}[W]=\sum_{l=0}^{+\infty}\mathrm{E}[W|L=l]\,\mathrm{P}\{L=l\}=\sum_{l=0}^{+\infty}\frac{l+1}{\mu}\rho^l(1-\rho)=\frac{1-\rho}{\mu}\sum_{l=0}^{+\infty}(l+1)\,\rho^l$$

$$=\frac{1-\rho}{\mu}\left[\sum_{l=0}^{+\infty}l\rho^l+\sum_{l=0}^{+\infty}\rho^l\right]=\frac{1-\rho}{\mu}\left[\sum_{l=0}^{+\infty}l\rho^l+\frac{1}{1-\rho}\right]$$

As we derived in the tutorial that

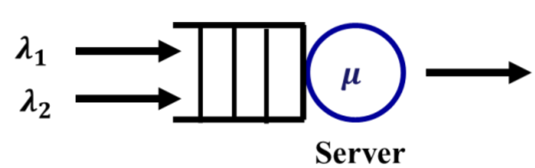
$$\sum_{l=0}^{+\infty}l\rho^l=\frac{\rho}{(1-\rho)^2}$$

After substituting the above into $\mathrm{E}[W]$, we obtain

$$\mathrm{E}[W]=\frac{1-\rho}{\mu}\left[\frac{\rho}{(1-\rho)^2}+\frac{1}{1-\rho}\right]=\frac{1}{\mu}\left[\frac{\rho}{1-\rho}+1\right]=\frac{1}{\mu}\left[\frac{1}{1-\rho}\right]=\frac{1}{\mu-\lambda}$$

So the above gives the same mean sojourn time result for M/M/1.

Problem 4



Two packet flows, one TCP with arrival rate $\lambda_1 = 200$ per second and one UDP with arrival rate $\lambda_2 = 500$ per second, go through a common link, which can be modelled as a single-server queueing system with a queue of infinity length. For the TPC packet flow, half of the packets have a length of 50 bytes and another half of the packets have a length of 100 bytes. All UDP packets have the same length of 30 bytes. The link has a processing rate of $5 * 10^4$ bytes per second.

a) What is the utilization of the link?
b) If the average number of packets in the system is 1000, what is the average queueing delay for these packets?

Sample answers:

a) We know that utilization ρ satisfies $\rho = \frac{\lambda}{\mu} = \lambda \mathrm{E}[S]$. Notice that λ and μ are the arrival and service rates, however, the inter-arrival times and the service times are not necessarily exponentially distributed. The total arrival rate to the system is $\lambda = \lambda_1 + \lambda_2 = 700$ packets per second. The remaining task is to calculate the mean service time $\mathrm{E}[S]$ for all the packets, regardless TCP or UDP packets. When we don't know how to calculate it directly, we think about using the law of total probability as follows. Given a packet with length l bytes, it takes $\frac{l}{5*10^4}$ second for the server to complete it, i.e., the service time for the packet. For all the packets, a ratio or percentage of $\frac{\lambda_1}{\lambda_1+\lambda_2}$ of them are TCP packets and the remaining are UDP packets. Therefore, we can calculate the expected service time by

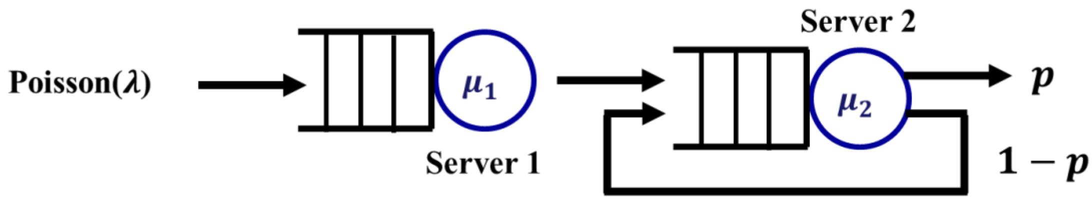
$$\begin{aligned}
\mathrm{E}[S] &= P\{long\ TCP\} * S_{long\ TCP} + P\{short\ TCP\} * S_{shor\ TCP} + P\{UDP\} * S_{UDP} \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{1}{2} * S_{long\ TCP} + \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{1}{2} * S_{shor\ TCP} + \frac{\lambda_2}{\lambda_1+\lambda_2} * S_{UDP} \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{1}{2} * \frac{100}{5*10^4} + \frac{\lambda_1}{\lambda_1+\lambda_2}\frac{1}{2} * \frac{50}{5*10^4} + \frac{\lambda_2}{\lambda_1+\lambda_2} * \frac{30}{5*10^4} \\
&= \frac{\lambda_1}{\lambda_1+\lambda_2} * \frac{75}{5*10^4} + \frac{\lambda_2}{\lambda_1+\lambda_2} * \frac{30}{5*10^4} = \frac{200*75+500*30}{\lambda*5*10^4}.
\end{aligned}$$

$$\Rightarrow \rho = \frac{\lambda}{\mu} = \lambda \mathrm{E}[S] = \frac{200*75+500*30}{5*10^4} = \frac{3}{5} = 60\%$$

b) We are given that $\mathrm{E}[L] = 1000$ packets and we want to know $\mathrm{E}[D]$. By using the Little's law, we know that $\mathrm{E}[W] = \frac{E[L]}{\lambda}$. Since $W = D + S$, we know that

$$\mathrm{E}[D] = \mathrm{E}[W] - \mathrm{E}[S] = \frac{\mathrm{E}[L]}{\lambda} - \frac{1}{\mu} = \frac{\mathrm{E}[L]}{\lambda} - \frac{\rho}{\lambda} = \frac{1}{\lambda}(\mathrm{E}[L] - \rho) = \frac{1}{700}(1000 - 0.6)$$

Problem 5



Consider the tandem queue structure where the arrival is Poisson with rate λ and the service times are exponential with rate μ_1 and μ_2. After getting served by Server 2, each job has a probability of p to leave the system and a probability of $1-p$ to join the queue of Server 2 again. Suppose that the system has run a long time and is stable.

a) When you take a random observation of the system, what is the probability that both queues of the system are empty?

b) When you take a random observation of the system, what is the probability that both queues have the same number of jobs?

Sample answers:

a) We fist calculate the effective arrival rates to the queues and the utilizations as follows.

$$\begin{cases} \lambda_1 = \lambda \\ \lambda_2 = \lambda_1 + \lambda_2 * (1-p) \end{cases} \Rightarrow \begin{cases} \lambda_1 = \lambda \\ \lambda_2 = \dfrac{\lambda}{p} \end{cases} \Rightarrow \begin{cases} \rho_1 = \dfrac{\lambda}{\mu_1} \\ \rho_2 = \dfrac{\lambda}{p\mu_2} \end{cases}$$

We know that the distribution of the number of jobs in the system follows

$$\mathrm{P}\{L_1 = i, L_2 = j\} = \rho_1^i(1-\rho_1)\rho_2^j(1-\rho_2).$$

When a queue is empty, the corresponding system either has 0 or 1 job. So the probability that both queue are empty equals

$$\mathrm{P}\{Q_1 = Q_2 = 0\} = \mathrm{P}\{0 \; job \; in \; both \; queues\} = \mathrm{P}\{0 \; or \; 1 \; job \; in \; both \; subsystems\}$$

$$\begin{aligned} &= \mathrm{P}\{L_1 = 0, L_2 = 0\} + \mathrm{P}\{L_1 = 0, L_2 = 1\} + \mathrm{P}\{L_1 = 1, L_2 = 0\} + \mathrm{P}\{L_1 = 1, L_2 = 1\} \\ &= (1-\rho_1)(1-\rho_2) + \rho_1(1-\rho_1)(1-\rho_2) + (1-\rho_1)\rho_2(1-\rho_2) \\ &+ \rho_1(1-\rho_1)\rho_2(1-\rho_2) = [1+\rho_1+\rho_2+\rho_1\rho_2](1-\rho_1)(1-\rho_2) \\ &= (1-\rho_1^2)(1-\rho_2^2) = \left[1-\left(\frac{\lambda}{\mu_1}\right)^2\right]\left[1-\left(\frac{\lambda}{p\mu_2}\right)^2\right] \end{aligned}$$

b) The probability that both queues have the same number of jobs equal the probability that they are both empty, plus that they have the same number of non-zero jobs.

$$\begin{aligned}
\mathrm{P}\{Q_1 = Q_2 > 0\} &= \sum_{q=1}^{+\infty} \mathrm{P}\{Q_1 = Q_2 = q\} = \sum_{q=1}^{+\infty} \mathrm{P}\{L_1 = L_2 = q+1\} \\
&= \sum_{q=1}^{+\infty} \rho_1^{q+1}(1-\rho_1)\rho_2^{q+1}(1-\rho_2) = (1-\rho_1)(1-\rho_2)\sum_{q=1}^{+\infty}(\rho_1\rho_2)^{q+1} \\
&= (1-\rho_1)(1-\rho_2)\frac{(\rho_1\rho_2)^2}{1-\rho_1\rho_2}
\end{aligned}$$

$$\begin{aligned}
\mathrm{P}\{Q_1 = Q_2\} &= \mathrm{P}\{Q_1 = Q_2 = 0\} + \mathrm{P}\{Q_1 = Q_2 > 0\} \\
&= (1-\rho_1^2)(1-\rho_2^2) + (1-\rho_1)(1-\rho_2)\frac{(\rho_1\rho_2)^2}{1-\rho_1\rho_2}
\end{aligned}$$