1. (15 points.) Use Sharpe to define a Markov model and obtain the following performance measures of an M/M/3/10 system:

(a) average customer turned-away probability
(b) average population (including both being served and waiting)
(c) average throughput
(d) average response time per customer (for those not being rejected)
(e) cumulative customers served over the time interval [0, 7200] seconds.

To get the response time per customer, you can apply Little’s law. Repeat the same procedure for the following λ/µ ratios: 1, 2.5, 5, 7.5 and 10, with λ fixed at 5 customers per second. Organize the outputs obtained from Sharpe in a table format provided below. Submit sharpe code and output.

<table>
<thead>
<tr>
<th>λ/µ</th>
<th>turned-away prob.</th>
<th>population</th>
<th>throughput</th>
<th>response time</th>
<th>customers served over [0,7200]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7.5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. (15 points.) Consider a system with 3 CPUs and 2 memory modules that requires at least 1 CPU and 1 memory module to be functioning for the system to be functioning. Use SHARPE to calculate the system reliability \( R(t) \) at \( t=12000 \) hours. Assume that the MTTFs of a CPU and a memory module are 7200 and 14400 hours, respectively. The MTTRs of a CPU and a memory module are 240 and 120 hours, respectively. Show the SHARPE program listing and output. Consider the following three cases separately. For each case, you need to compute \( R(t) \) at \( t=12000 \) hours. Submit sharpe code and output. (Hint: for each case, define a Markov model that follows the definition of \( R(t) \).)

(a) Each component has an independent repair facility.
(b) Each subsystem (cpu or memory) has an independent repair facility that can repair failed components within the subsystem one at a time.
(c) The whole system shares a repair facility which repairs failed components one at a time. However, it always repairs failed memory modules first. If no memory module fails, it repairs failed CPU modules.

3. (15 points.) Suppose that a network switch center has \( n = 3 \) slots to accommodate incoming high and low priority clients, with arrival rates of \( \lambda_h \) and \( \lambda_l \) and departure rates of \( \mu_h \) and \( \mu_l \), respectively. A high priority client must always occupy one full slot. A low priority client, on the other hand, can lower its quality of service (QoS) by occupying only one half of a slot, if necessary. When a low priority client occupies a full slot, we call it a low priority, high QoS client; when it occupies one half of a slot, we call it a low priority, low QoS client. The system adopts the following admission control policy:

- If there is at least one full slot available, an incoming client always occupies an empty full slot, regardless of its priority class.
- If no slot is available, an incoming high priority client can lower the QoS of two low priority, high QoS clients, if they are available, after which the high priority client occupies one full slot and the two low priority, high QoS clients both become low priority, low QoS clients, with each occupying one half slot.
- If no slot is available, an incoming low priority client can lower the QoS of one low priority, high QoS client, if it is available, after which each occupies one half slot.
• A client will turn away if none of the above cases is applicable.
• A low priority, low QoS client occupying one half slot can immediately become a low priority, high QoS client occupying one full slot upon a client’s departure.

Answer the following questions:

(a) Draw a Markov state transition diagram for modeling the above admission control policy. Use the representation \((a, b, c)\) where \(a\) stands for the number of low-priority, low QoS clients, \(b\) stands for the number of low-priority, high QoS clients, and \(c\) stands for the number of high-priority clients. Organize the Markov model so that when a high priority client arrives, the transition goes right; and when a low priority client arrives, the transition goes down. Label the transition rate of each transition clearly.

(b) Assume \(\lambda_h = \mu_h = 5\) and \(\lambda_l = \mu_l = 15\). Write Sharpe code and assign rewards to states of the Markov model to compute the average number of high-priority clients in the system. Submit Sharpe code and output.

(c) Continued from (b), write Sharpe code to compute the throughput of low-priority, low QoS clients. Submit Sharpe code and output.

4. (15 points.) Consider a client-server system with a fixed number \(m\) of client workstations that are connected by an Ethernet network to a database server. The server consists of a single disk and a single CPU. This leads to a closed QNM shown in the bottom figure. The Ethernet network is being modeled as a load-dependent server with service rate defined as follows:

\[
\mu_{\text{net}}(1) = \frac{1}{N_p L_p B} + S \times C(1)
\]

and

\[
\mu_{\text{net}}(k) = \frac{1}{N_p L_p B} + S \times C(k+1)
\]

for \(k > 1\).

where \(C(k) = \frac{1-A(k)}{A(k)}\) is the average number of collisions per request and \(A(k) = (1-\frac{1}{k})^{k-1}\) is the probability of a successful transmission and \(k\) is the number of workstations that desire the use of the network. Note that you should use the above two equations to separately define \(\mu_{\text{net}}(1)\) and \(\mu_{\text{net}}(k)\) for \(k > 1\). In your sharpe code, you can use the \texttt{bind} construct to define \(\mu_{\text{net}}(1)\) directly and use the \texttt{func} construct to define the function \(\mu_{\text{net}}(k)\) for \(k > 1\).

Assume that \(N_p = 7\) (number of packets per request); \(B=10\text{Mb/sec}\) (bandwidth of the Ethernet); \(S = 51.2 \times 10^{-6}\) second (slot duration); \(L_p = 1518\) bits (average packet length); \(\mu_{\text{client}} = 0.1/\text{sec}\); \(\mu_{\text{CPU}} = 16.7/\text{sec}\); and \(\mu_{\text{disk}} = 18.5/\text{sec}\). Write a sharpe program to calculate the system throughput of the server subsystem and the population at the server subsystem for \(m = 15\).