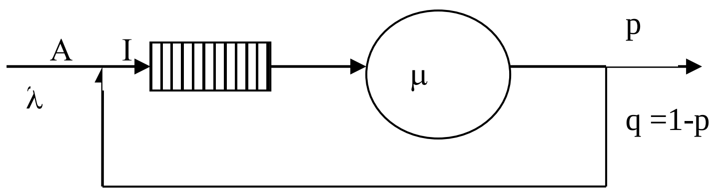
Dept. of ECE **Spring 2004**
Univ. of Conn. **KRP**

<u>Homework Set # 5 (Due March 17, 2004)</u>

1. Consider the two node tandem network considered in the class. Draw the state transition diagram and verify that the product-form solution is satisfied by the global balance equations.
2. Consider the M|M|1 queue with FCFS service discipline as shown in Fig. 1:



We want to show that the actual input process at I is not a Poisson Process.

a) Show that the inter-arrival time τ_I satisfies: $P(\tau_I > t) = exp(-\lambda t).\ P(Y>t)$, where Y is the time to the next feedback as measured from the time of the last input to the queue.

b) Show that the pdf of Y is given by:

$$f_Y(t) = \mu q \exp[-(\mu - \lambda)t]$$

c) Compute $P(Y>t)$ and conclude that the density of τ_I is the sum of two exponential densities (i.e., hyper-exponential density).

d) You can actually show that the departure process is Poisson with rate λ. Read the following paper: P.J. Burke, "Proof of a Conjecture on the Inter-arrival Time Distribution in an M|M|1 Queue with Feedback," *IEEE Trans. on Communications*, Vol. 24, pp. 175-178, 1976.

3. Bertsekas and Gallagher, Problem 3.59.
4. Suppose the capacity assignment problem is formulated as one of minimizing the response time as follows:

$$\min_{\mu_{ij}} R$$

$$\textit{suject to a budget constra}\text{int}: \sum_{ij} \mu_{ij} c_{ij} = C$$

a) Show that the optimal μ_{ij}^* are given by: $\mu_{ij}^* = \lambda_{ij} s + \left(\frac{C_e}{c_{ij}}\right) \frac{\sqrt{\lambda_{ij} c_{ij}}}{\sum_{mn} \sqrt{\lambda_{mn} c_{mn}}}$ where C_e is the excess dollars given by: $C_e = C - \sum_{ij} \lambda_{ij} c_{ij} s$. What is the optimum response time? What happens when c_{ij} = c for all links?

b) Suppose that each link is modeled as an infinite server node. Obtain an expression for the response time, and repeat the optimization process in part (a).

5. Show that the cardinality of the set of feasible states $\underline{n} = (n_1, n_2, ..., n_M)$ for a closed network with *M* nodes and population *N* is given by $\begin{pmatrix} N+M-1 \\ M-1 \end{pmatrix}$.
6. Balanced Throughput Bounds
   a. Derive an explicit, closed form expression for the throughput of a closed queuing network when the population in the network, *N*=1.
   b. Suppose that a single class closed queuing network consists of *M* single server nodes with balanced workload, i.e., $\frac{v_i s_i}{\mu_i} = constant$. Using the convolution algorithm, show that the throughput *X(N)* is given by $X(N) = \frac{NM}{(N+M-1)R_0}; R_0 = \sum_{i=1}^{M} \frac{v_i s_i}{\mu_i}$.
   c. Using the result in part (b), show that the throughput satisfies the bounds: $\frac{N}{R_0 + (N-1)\max_i \frac{v_i s_i}{\mu_i}} \leq X(N) = \frac{NM}{(N+M-1)R_0}$.

   (Hint: you may want to refer to the paper by Lazowska et al. in CACM, 1982 and/or [LAZ84]).
7. Show that the MVA equations with Schweitzer-Bard approximation imply that the throughput *X(N)* can be obtained by solving the scalar nonlinear equation:
$$f(x) = \sum_{i=1}^{M} \frac{v_i s_i x}{[\mu_i - \frac{N-1}{N} v_i s_i x]} - N = 0.$$
   a. Show that each term in the above equation is a convex function of both *x* and μ. (recall that a function *f(x)* is convex if $\frac{d^2 f}{dx^2} > 0$.)
   b. A convex function *f(x)* satisfies the so-called Jensen inequality: $f(\sum_{i=1}^{M} \alpha_i x_i) \leq \sum_{i=1}^{M} \alpha_i f(x_i); \alpha_i \geq 0; \sum_{i=1}^{M} \alpha_i = 1$. (Hint: use $\alpha_i = \frac{v_i s_i}{\mu_i R_0}$. Also, see the paper by Pattipati et al. in the July 1990 issue of *JACM*, pp. 643-673).
8. Consider a central server model (i.e., a CPU shared by several I/O devices) with two I/O channels and three CPUs. The average CPU time per program (i.e., $\frac{v_1 s_1}{\mu_1}$) is 500ms, and the average time per program on the I/O devices is 175 ms and 100ms, respectively. Compute the average system throughput for the degree of multi-programming *N* varying from 1 to 5. Now vary the number of CPUs from one to three and study the effect on the average system throughput.

9. An interactive system workload measurement showed that the average CPU time over all visits per terminal request was 4.6 seconds, and the average disk time per request was 4 seconds. Since the response time for the existing system was intolerably large, two alternative systems were proposed. Compared to the existing system (denoted by **ex**), one of the proposed systems (denoted by $\mathbf{S_1}$) had a CPU 0.9 times as fast and the disk was twice as fast. The alternative system (denoted by $\mathbf{S_2}$) had a CPU 1.5 times as fast as that in **ex** and a disk twice as fast. Decide whether a change from **ex** to $\mathbf{S_1}$ can be recommended (assuming first that $\mathbf{S_2}$ is intolerably expensive). Next, assuming that $\mathbf{S_2}$ is affordable, show how much reduction in response time is possible as a function of the number of terminals, i.e., *N*.
10. Flow control optimization
    a. For the sliding window flow control problem, show that as $\lambda \to \infty, \dfrac{X(N)}{R(N)}$ is maximized when *N*=*M-1*.
    b. Fopr the problem in (a), show that $when\ \lambda = \mu, \dfrac{X(N)}{R(N)}$ is maximized when *N*=*M*.