UNIVERSITY OF TORONTO

FACULTY OF APPLIED SCIENCE AND ENGINEERING

FINAL EXAMINATION, DECEMBER, 1999

Fourth Year -- Computer Engineering Program

CSC 470F - COMPUTER SYSTEM MODELLING AND ANALYSIS

Examiner: K. C. Sevcik

NAME _____ STUDENT #: _____

Open books and notes. Use of calculators is permitted.

Write all answers in the space provided on the examination paper.
Show your work where appropriate.
Attempt all parts of the test before spending too much time on any single part.

[NOTE: Parts III, V, and VI are worth more marks than the other four parts.]

PART	MARK
I.	/12
II.	/12
III.	/18
IV.	/11
V.	/22
VI.	/17
VII.	/8
Total	**/100**

PART I. [12 marks]

A. [4 marks]

1. In a few words, indicate the major advantage(s) of *Simulation* over *Queuing Analysis*.

2. In a few words, indicate the major advantage(s) of *Queuing Analysis* over *Simulation*.

B. [4 marks]

In planning a modelling study of a computer system, you need to choose between using a *CSim simulation* or a *Map analytic model*. For each of the situations mentioned below, circle either **CSim** or **Map**, to indicate which one is the better choice in the situation.

(i). The system is completely new (as opposed to being a modification of an existing system).

Csim **Map**

(ii). Lots of detailed measurement data can be obtained from an existing form of the system.

Csim **Map**

(iii). It is necessary to obtain an estimate of the full response time distribution (as opposed to just getting an estimate of the average).

Csim **Map**

(iv). It is better to get rough estimates quickly rather than taking longer to obtain more accurate answers.

Csim **Map**

C. [4 marks]

More generally, for each of the situations below, indicate whether *queueing analysis* or *simulation* would be the *better* modelling approach to take.

(i). Assess the efficiency of network protocol processing at a switching node in a wide-area network.

queueing analysis **simulation**

(ii). Assess the effectiveness of a particular cache replacement strategy for a 2nd level cache.

queueing analysis **simulation**

(iii). Estimate the execution rate of a complex of high-performance processors that do very large computational fluid dynamics calculations while interconnected by a special local area network.

queueing analysis **simulation**

(iv). Investigate alternative arrangements of workload components across a set of computer systems with different powers but the same underlying architecture.

queueing analysis **simulation**

PART II. [12 marks]

A. [4 marks]

In a queueing network with three centres and a single class of *terminal* type, the *service demands* (or *loadings*) are 5, 3 and 2, respectively, and the average think time is 10. Use *asymptotic bounds* to answer the following questions.

(i). What is the system *throughput* with one customer ($X(1)$)?

(ii). What is the highest achievable system throughput for many customers?

(iii). What is the *response time* with one customer ($R(1)$)?

(iv). Give a (reasonably tight) lower bound on $R(20)$, the response time with 20 customers.

B. [4 marks]

Consider a system with a single disk that supports both *Batch* jobs and *Interactive* users submitting transactions from workstations. In a one-hour observation interval (that is, 3600 seconds), the following measurements were taken:

Class	Parameter	Value
Interactive	Online Users	60
Interactive	Average think time	10 seconds
Batch	Throughput	180 jobs/hr
Interactive	Disk operations per transaction	5
Interactive	Disk operations per job	200
Both classes	Disk service time	20 msecs.
Total	Disk utilization	60%

(i). What is the disk utilization due to Batch jobs, $U_{B,disk}$?

(ii). What is the throughput of Interactive transactions, X_I?

(iii). What is the average response time for Interactive users, R_I?

(iv). Assuming that the disk is the bottleneck resource of the system, what is the greatest possible Batch throughput that can be achieved such that it remains feasible (with respect to asymptotic bounds) to keep average response time for Interactive users under 20 seconds?

C. [4 marks]
In a five centre single class *batch* type queueing network, *all five centers* (K=5) have service demand three seconds (D_k = 3, k=1,2,3,4,5).

(i). If the multiprogramming level is N=4, what is the average response time, $R(4)$?

(ii). If the multiprogramming level is N=4, what is the utilization of each center, $U_k(4)$?

PART III. [18 marks]

A. [6 marks]
In developing a simulation program for a particular situation, describe briefly how you would make the following choices:

(i). *time-driven vs. event-driven*

(ii). *deterministic vs. stochastic*

(iii). *static vs. dynamic*

B. [4 marks]

(i). Is the issue of proper warm-up in a simulation more significant in an open system under moderate load or in a closed system with moderate load? Why? (Assume that "moderate load" means that the bottleneck center is 50% to 80% utilized.)

(ii). Is the issue of properly dealing with end effects more significant in simulating a system with $first-come-first-served$ scheduling or in a system with $shortest-job-first$ scheduling? Why?

C. [8 marks]

Consider a stochastic event-driven simulation program to model message delivery times in a local area network. One hundred independent trials were run with different seeds for the pseudo-random number generator(s) and each trial calculated the average response to of 1000 message delivery times. Let t_i represent the average time observed in the *ith* independent trial, and let m and s^2 be the sample mean and sample variance respectively of $\{ t_1, t_2, \ldots, t_{100} \}$. Assume that the simulation results of the 100 trials are as follows:

$$m = \frac{1}{100} \sum_{i=1}^{100} t_i = 8 \text{ msecs.} \qquad s^2 = \frac{1}{99} \sum_{i=1}^{100} (t_i - m)^2 = 64$$

(i). Based on these experiments, what is **80% confidence interval** on the true average delivery time in the system?

Reminder: $Prob[m - \frac{s}{\sqrt{n}} T_\alpha \leq \mu \leq m + \frac{s}{\sqrt{n}} T_\alpha] = 1 - \alpha$

T_α	1.00	1.28	1.65	1.96	2.58
$1 - \alpha$	.68	.80	.90	.95	.99

(ii). If the experiment were repeated, but this time with each of the 100 independent trials based on observing 4000 message delivery times instead of 1000, how would you expect the width of the **80% confidence interval** to be different (if at all)?

(iii). Estimate the approximate probability that the *true average delivery time* is actually greater than 10 msecs.

(iv). Including the original 100 trails, approximately how many trials in total would you expect to have to run to obtain an answer that is correct within plus or minus one second with **95% confidence**?

PART IV. [11 marks]

A. [6 marks]

In solution algorithms for queueing network models that are based on *mean value analysis*, residence times at service centers are calculated from the equation:

$$R_k(n) = V_k \; S_k \left[1 + A_k(n)\right]$$

where $A_k(n)$ is the expected queue length found at center k by an arriving customer when the network population is n.

(i). In what way does this equation rely on the *memoryless property* of the service time distribution.

(ii). For each of *Open* and *Closed* queueing network models, indicate the correct relationship between the average queue lengths *seen at arrival instants*, $A_k(n)$, and the overall average queue length, $Q_k(n)$, by circling *one entry* in *each row*.

Open	$A_k(n) < Q_k(n)$	$A_k(n) = Q_k(n)$	$A_k(n) > Q_k(n)$
Closed	$A_k(n) < Q_k(n)$	$A_k(n) = Q_k(n)$	$A_k(n) > Q_k(n)$

(iii). In *exact mean value analysis*, what value is used for $A_k(n)$?

(iv). In *approximate mean value analysis* using proportional estimation, what estimate is used for $A_k(n)$?

(v). In *linearizer*, what estimate is used for $A_k(n)$?

B. [5 marks]

Below are listed five solution techniques for queueing network models. Here are some relative *advantages* and relative *disadvantages* that can be used to characterize the solution techniques relative to one another:

Relative Advantages:

A1- Provides an exact solution to the model.

A2- Can be extended conveniently to account for some characteristics of non-product form models.

A3- Lowest computational cost among the techniques.

A4- Applies equally well to product form and non-product form models.

Relative Disadvantages:

D1- Gives only an estimate of or bound on the true solution of the model.

D2- Susceptible to numerical problems of overflow and underflow.

D3- Highest computational cost among the techniques.

D4- Relies on several homogeneity assumptions for validity.

In the table below, indicate *one or two advantages* and *one or two disadvantages* that characterize the corresponding solution technique by entering their numbers in the appropriate boxes.

Solution Technique	Advantages	Disadvantages
Asymptotic Bounds		
Convolution		
Exact Mean Value Analysis		
Approximate Mean Value Analysis		
Global Balance		

PART V. [22 marks]

A. [12 marks]

The figures below show the *probability density function*, $f(x)$, and the *cumulative distribution function*, $F(x)$, from which a sequence of pseudo-random samples are to be drawn. In the three parts of this question, you are to write a fragment of pseudo-code (about three to five lines of pseudo-code suffice in each case) that generates pseudo-random variates with the specified distribution based on the general method indicated, using a source of Uniform (0, 1) pseudo-random numbers. Your pseudo-code should leave the variable R_i holding the value of the pseudo-random number. In your pseudo-code, let $U_i^{(j)}$ represent the jth sample from the Uniform (0, 1) source used to generate R_i.

Probability Density Function:

$$f(x) = \begin{cases} 1/4 & 1 \le x \le 4 \\ 1/8 & 4 \le x \le 6 \\ 0 & \textit{otherwise} \end{cases}$$

Cumulative Distribution Function:

$$F(x) = \begin{cases} \dfrac{x-1}{4} & 1 \le x \le 4 \\ \dfrac{x+2}{8} & 4 \le x \le 6 \end{cases}$$

(i). Use the *method of decomposition* to produce pseudo-random variates:

(ii). Use the *method of rejection* to produce pseudo-random variates:

[Information repeated from previous page:]
Cumulative Distribution Function:

$$F(x) = \begin{cases} \frac{x-1}{4} & 1 \le x \le 4 \\ \frac{x+2}{8} & 4 \le x \le 6 \end{cases}$$

(iii). Use the *method of inversion* to produce pseudo-random variates:
[Note that the method of inversion should use only one U(0,1) sample to generate each variate.]

B. [10 marks]
In the table below, each row corresponds to one approach to generating U(0,1) pseudo-random numbers, and each column corresponds to one desirable property of such generators. In each of the 25 boxes of the table, circle "yes" if the corresponding method when used properly *has* the corresponding desirable property, and circle "no" otherwise.

Method	Tables	Devices	Mid-Square	Congru-ential	Shift-based
Efficient	yes no	yes no	yes no	yes no	yes no
Reproducible	yes no	yes no	yes no	yes no	yes no
Dense	yes no	yes no	yes no	yes no	yes no
Uniform	yes no	yes no	yes no	yes no	yes no
Independent	yes no	yes no	yes no	yes no	yes no

PART VI. [17 marks]

A. [2 marks]

Examine the *standard* MVA residence time equation and the *modified* one below:

$$R_{c,k}(\vec{N}) = D_{c,k}[1 + \sum_{j=1}^{C} Q_{j,k}(\vec{N}-1_c)] \qquad \text{(standard)}$$

$$R_{c,k}(\vec{N}) = V_{c,k}[S_{c,k} + \sum_{j=1}^{C} S_{j,k}\ Q_{j,k}(\vec{N}-1_c)] \qquad \text{(modified)}$$

Describe briefly the conditions under which the using the *modified* equation will lead to more accurate answers than using the *standard* one.

B. [6 marks]

Consider a queueing network model with a single batch type class with up to 3 customers, in which the service demands at the two centres are 3 and 2, respectively. The first table below shows the computation of the *normalization constants*, $G(n,k)$, n=1,2,3, k=1,2, for the system according to the *Convolution* algorithm. Using the values of the normalization constants, fill in entries in the second table below showing the value of the four performance measures for numbers of customers from 1 to 3.
(You may express you answers as fractions.)

n		D_1=3	D_2=2	D_1=3	D_2=2
		1	1	1	1
1	0	3	5	8	7
2	0	9	19	43	33
3	0	27	65	---	---

n	$X(n)$	$R(n)$	$Q_1(n)$	$Q_2(n)$
1				
2				
3				

C. [9 marks]

The figure below shows a two class queueing network with just one customer in each class. The interactive customer, I, alternately thinks for average time Z_I, and then visits the cpu for exactly one service time of average length $S_{I,cpu}$. Each batch customer, B, makes on average $V_{B,cpu}$ visits to the cpu and $V_{B,disk}$ visits to the disk before completing and being replaced by the next batch customer.

Below is the state transition diagram for the five possible states of the queueing network. Each legal transition between states is indicated by an arrow with a number in brackets next to it for identification purposes.

(i). Assume that the scheduling discipline at the cpu is **first-come-first-served**, and answer the six questions in the table.

Which two arcs have transition rate $1 / Z_I$?		
Which two arcs have transition rate $1 / S_{B,disk}$?		
Which two arcs have transition rate $1 / S_{I,cpu}$?		
Which two arcs have transition rate $p / S_{B,cpu}$?		
Which two arcs have transition rate $(1-p) / S_{B,cpu}$?		
Which arc has transition rate 0 ?		

(ii). Now assume that the scheduling discipline is **(preemptive) priority, with the class I customer having higher priority**, *and answer the following questions.*

Which state disappears (i.e., has probability zero of occurrence)?		
Which two arcs have different transition rates than under fcfs?		
What are the new rates associated with these two arcs?		

PART VII. [8 marks]

In the last lab assignment in the course, you built both an analytical model (in Map) and a simulation model (in Csim) of a symmetric multiprocessor system. Two of the parameters of the system were: $\bar{P}$, the average duration of a pause, and B, the number of cycles to send a message on the bus.

A. [2 marks]
Explain briefly how the parameter $\bar{P}$ was used:

(i). in your CSim simulation program:

(ii). in your Map model:

B. [2 marks]
Explain briefly how the parameter B was used:
(Note: If it is used more than once, indicate all uses.)

(i). in your CSim simulation program:

(ii). in your Map model:

C. [2 marks]
In the lab, you used both the CSim program and the Map model to predict the execution rates in various cases. In a specific case, call the rate predicted by the Map model X_{map}, and the rate predicted by the Csim model X_{csim}. Over all the cases you investigated in the lab, pick the most appropriate single word below to complete the statement:

X_{map} is **[BLANK]** greater than X_{csim}

CIRCLE THE MOST APPROPRIATE WORD TO REPLACE "BLANK" ABOVE:

always **usually** **sometimes** **seldom** **never**

D. [2 marks]
Based on your knowledge of the assumptions made in each of the two modelling approaches, explain briefly why the relationship identified in part C holds.

frame invis ht 1.2 wid 3.5 left solid bot solid label left "$f(x)$" left .2 label bot "x" ticks left out at .125 "1/8", .25 "1/4" ticks bot out from 0 to 7 by 1 coord x 0,7 y 0,.3 line solid from 1,0 to 1,.25 line solid from 1,.25 to 4,.25 line solid from 4,.25 to 4,.125 line solid from 4,.125 to 6,.125 line solid from 6,.125 to 6,0 frame invis ht 1.2 wid 3.5 left solid bot solid label left "$F(x)$" left .2 label bot "x" ticks left out at .25 "1/4", .5 "1/2", .75 "3/4", 1.0 "1" ticks bot out from 0 to 7 by 1 coord x 0,7 y 0,1.1 line dotted from 0,1.0 to 6,1.0 line dotted from 0,.75 to 4,.75 line solid from 1,0 to 4,.75 line solid from 4,.75 to 6,1 line solid from 6,1 to 7,1

frame invis ht 1.2 wid 3.5 left solid bot solid label left "$f(x)$" left .2 label bot "x" ticks left out at .125 "1/8", .25 "1/4" ticks bot out from 0 to 7 by 1 coord x 0,7 y 0,.3 line solid from 1,0 to 1,.25 line solid from 1,.25 to 4,.25 line solid from 4,.25 to 4,.125 line solid from 4,.125 to 6,.125 line solid from 6,.125 to 6,0 frame invis ht 1.2 wid 3.5 left solid bot solid label left "$F(x)$" left .2 label bot "x" ticks left out at .25 "1/4", .5 "1/2", .75 "3/4", 1.0 "1" ticks bot out from 0 to 7 by 1 coord x 0,7 y 0,1.1 line dotted from 0,1.0 to 6,1.0 line dotted from 0,.75 to 4,.75 line solid from 1,0 to 4,.75 line solid from 4,.75 to 6,1 line solid from 6,1 to 7,1