UNIVERSITY OF EDINBURGH

FACULTY OF SCIENCE AND ENGINEERING

COMPUTER SCIENCE FOUR

MODELLING AND SIMULATION

Wednesday 22 April 1998

16:00 to 17:30

Examiners:

G. Birtwistle (External)

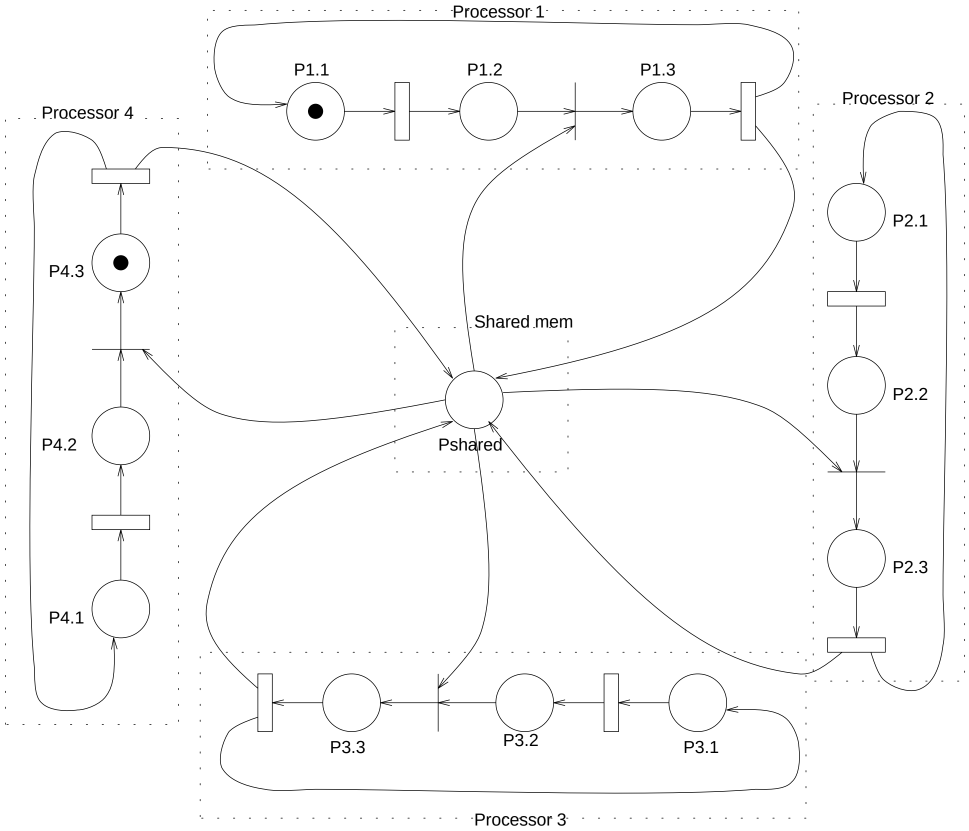
G. Brebner (Chairman)

INSTRUCTIONS TO CANDIDATES

Answer any TWO questions.

All questions carry equal weight.

1. a) The Generalised Stochastic Petri-net (GSPN) below models accesses to shared memory in a four processor system.



i) Draw the reachability graph for the Petri-net, given the initial marking shown.

ii) How would you go about solving the Markov process, and what could this solution tell you?

iii) The initial marking is changed to: one token in each of the places

$\{$P1.1, P2.1, P3.1, P4.1$\}$

and one token in place Pshared. How would the GSPN above arbitrate between concurrent accesses to shared memory, and how could a priority scheme be introduced into the model? [*12 marks*]

b) Why use simulation when Markov analysis gives exact results? [*3 marks*]

c) Sketch the code for a discrete event simulation model of the GSPN in (a) with the marking of part (a)iii), using either DEMOS or simjava. You should make clear how mutual exclusion is represented in the simulation model, and how you would obtain results. [*10 marks*]

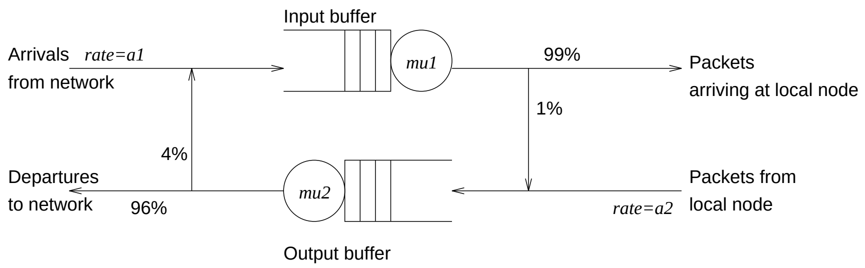
2. You have just started working for a company which makes network routing chips, and have been asked to design part of the next generation system (codename WEBROUTE 2). Your manager is concerned that too much silicon area in the current system (WEBROUTE 1) is taken up with packet buffers, which she says are "way too big" and "hardly ever used". However, the engineer who designed the first system says he made them that way "just in case" lots of packets arrived at the same time, and thinks you might have to increase buffer size in the next system as it will have to cope with heavier traffic.

   You have decided to do some performance analysis of the network buffers to come up with numbers to guide your design.

   a) Packets arrive at an average of 250/second at the network input buffer and each packet takes an average of 0.002 seconds processing time before it is sent on. The buffer size at present is 100 packets. Packets arriving when the buffer is full are lost. Assuming inter-packet delays and service times approximate to an exponential distribution (*you may wish to refer to the equations in table 1 overleaf*):

      i) How would you represent this queue using Kendall's notation?
      ii) What is the traffic intensity for this queue?
      iii) Sketch the underlying Markov chain.
      iv) How many packets would be lost per second if the input buffer size were reduced to 2? *[9 marks]*

   b) The following queueing network models both input and output buffers. 1 percent of incoming packets are badly addressed and are sent via the output buffer back to the network. 4 percent of the output traffic is destined for the local node and goes straight back into the input queue, the rest is sent out onto the network.



      i) What are the *traffic equations* for this network?
      ii) Use the traffic equations to calculate the traffic intensity at both service centres in terms of μ_1, μ_2 and the arrival rates. *[7 marks]*

(QUESTION 2 CONTINUED ON NEXT PAGE)

c) The engineer is expressing skepticism that these figures are reliable in the face of bulk and train arrivals, and mutters darkly about mutual exclusion being an important feature which is not incorporated into your queueing models.

i) What are bulk and train arrivals, and is he right to be skeptical?

ii) Show in detail how you would build a model of a queueing system which addresses each of his concerns.

iii) How would you validate your model?

[*9 marks*]

Performance Measures	$M/M/1$	$M/M/c$	$M/M/1/K$
Traffic Intensity ρ	$\frac{\lambda}{\mu}$	$\frac{\lambda}{c \times \mu}$	$\frac{\lambda}{\mu}$
Utilisation U (per server)	ρ	ρ	$\rho(1 - \frac{(1-\rho)\rho^K}{1-\rho^{K+1}})$
Prob. system is idle $\boldsymbol{\pi_0}$	$1-\rho$	$\left(1 + \frac{(c\rho)^c}{c!(1-\rho)} + \sum_{n=1}^{c-1} \frac{(c\rho)^n}{n!}\right)^{-1}$	$\frac{1-\rho}{1-\rho^{K+1}}$
Prob. buffer non-empty B	ρ^2	$\frac{(c\rho)^c}{c!(1-\rho)}\boldsymbol{\pi_0}$	
Mean no. in system N	$\frac{\rho}{1-\rho}$	$c\rho + \rho B/(1-\rho)$	$\frac{\rho}{1-\rho} - \frac{(K+1)\rho^{K+1}}{1-\rho^{K+1}}$
Mean no. in buffer N_b	$\frac{\rho^2}{1-\rho}$	$\rho B/(1-\rho)$	$\frac{\rho}{1-\rho} - \rho\frac{1+K\rho^K}{1-\rho^{K+1}}$
Mean response time R	$\frac{1}{\mu(1-\rho)}$	$\frac{1}{\mu}\left(1 + \frac{B}{c(1-\rho)}\right)$	$\frac{N}{\lambda(1 - \frac{(1-\rho)\rho^K}{1-\rho^{K+1}})}$

Table 1: Queueing theory performance measures

The meaning of B in the last three rows is as defined in the fourth row.

3. a) Why simulate software systems when you could just instrument and measure their performance directly? [*3 marks*]

b) The following is part of a simjava simulation model of a client-server system:-

```
class client extends sim_entity {
  void body() {
    while (true) {
      sim_hold(100.0);
      sim_schedule(server_id,0.0,REQ);
      sim_wait_for(sim_type_p(ACK));
    }
  }
}
class server extends sim_entity {
  void body() {
    int client_id;
    while (true) {
      sim_get_next(ev);
      client_id = ev.scheduled_by();
      sim_hold(10.0);
      sim_schedule(client_id,0.0,ACK);
    }
  }
}
```

The main program creates 3 clients and 1 server.

i) Sketch a sample tracefile and timing diagram for the system. [*7 marks*]

ii) Give code fragments to show how you would extract the wait time of each client and the utilisation of the server. [*4 marks*]

iii) If the deterministic delays above were replaced by random delays, how and why would you reduce the variance of the output results? [*4 marks*]

iv) Each client's hold time is a negative exponential distribution with mean 50.0, and the server's hold time is a normal distribution with mean 10.0 and variance 0.1. Sketch the probability density functions for both distributions, and a sample timing diagram for the first 200.0 time units of the simulation run. [*7 marks*]