Practical 1

1) Investigating an e-commerce site

A Web-based brokerage company runs a site consisting of a web server, an application server and a database server. The database server has two processors and three disks. The application server converts business functions from the web server into transactions that are submitted to the database.
On average, each user visit to the web site generates 3 consecutive business functions and each business function generates 2.5 database transations. An average transaction generates 2 access to disk A, 3 accesses to disk B, 3 accesses to disk C and a total of 4 CPU bursts which are evenly distributed across the two processors.
Average service times per access at the three disks are 30, 40 and 25 milliseconds respectively, while average service times at the database processors are 50 milliseconds per burst. The average service demand of each business function is 120 milliseconds at the web server and 150 milliseconds at the application server. On average, a user thinks for 20 seconds between each business function.
For all parts you should state which operational laws you are using and explain your reasoning.

a.) *Which is the bottleneck device in the database?*

3 marks

b.) *What is the minimum average residence time per transaction, regardless of the number of users using the system?*

3 marks

c.) *What is the maximum possible utilisation of disk C and when would this be achieved?*

3 marks

d.) *How many users would be using the system if the average response time per business function is 8 seconds and the throughput at disk C is 0.002 (accesses/millisecond)?*

3 marks

e.) *What revised service time per access at disk A would you recommend to achieve a system response time per business function of 5 seconds when there are 100 users on the system? Which other devices, if any, would need to be speeded up to achieve this target?*

2) Loss probabilities within a communication network

In a communication network, if a packet arrives at a buffer which is already full (all places occupied) the packet must be discarded. This leads to a degradation of service or the need to retransmit, depending on the type of packet.
Construct a Markov process of a buffer within a communication network which has 5 places. Packets arrive at the buffer from the incoming link at a rate of 4 packets per second, packets depart, in first-in-first-out order, down the outgoing link. The departure rate down the out-going link is 5 packets per second. The *loss rate* is the rate, in steady state, at which packets are discarded because the buffer is full.

a.) *Draw the state transition diagram of the Markov process.*

2 marks

b.) *Construct the infinitesimal generator matrix corresponding to the Markov process.*

2 marks

c.) *Using* `xmaple` *or otherwise, solve the global balance equations for the process and give the steady state probability distribution.*

d.) *What is the loss rate? Explain how you arrived at your answer.*

3 marks

e.) *If the number of places in the buffer increased from five to six what is the effect on the loss rate, and what is the effect on the average residence time in the buffer? How does this compare with increasing the bandwidth of outgoing link to 8 packets per second?*

5 marks

14 marks total

3) Analysing an automated manufacturing system

In an automated widget factory, widgets are assembled from two parts, an A part and a B part. A parts are processed by machine 1 while B parts are processed by machine 2; machine 3 then assembles one A part and one B part to make one widget. A single robot transports parts from a conveyor belt to the appropriate machine; it is also responsible for moving completed A parts from machine 1 to machine 3, and completed B parts from machine 2 to machine 3. Machine 1 cannot start to load the next A part until the current one has been moved to machine 3; similarly for machine 2 and B parts. Machine 3 cannot accept parts of either kind until it has completed the assembly of the previous widget. There are always A and B parts available from the conveyor belt. If both machine 1 and machine 2 need to use the robot at the same time they are equally likely to acquire it. Loading parts from the conveyor belt, or transferring them to machine 3 takes 10 seconds on average. The mean duration of processing A parts at machine 1 is 125 seconds, while the mean duration of processing B parts at machine 2 is 200 seconds. Assembling a widget from A and B parts takes 100 seconds on average.

a.) *Draw a GSPN model to represent the system as described above, taking care to either give a meaningful label to each place and transition, or alternatively provide a separate table explaining the role of each place and transition. You should also include the firing rate or firing probability for each transition.*

4 marks

Note that there is more than one correct way of representing the system as a GSPN.

Machine 1 is getting old and rather temperamental: for approximately 1 part in 20 it jams during processing and needs to be repaired. On average, the repair time is 800 seconds. The processing of that part then continues. An SPNP model of the system, including the breakdown and repair of machine 1, can be found in the file `pract1-one1.c` which is available from the web page `http://www.inf.ed.ac.uk/teaching/courses/ms/`. To run this model you simply make a copy of it in your own file space and then type

```
make -f /opt/spnp-1/obj/Makerun SPN=pract-one1
```

The model will prompt you to set some of the parameters of the model. Once all the parameters have been assigned values, the model will be solved. Various files will be produced as discussed in lecture note 6. In particular you should be able to derive the information you need for parts b.) and c.) from `pract-one1.out`.

b.) *What is the utilisation of each of the Machines 1, 2 and 3?*

3 marks

c.) *What is the throughput in terms of widgets per second?*

2 marks

Since machine 1 soon blocks the whole system when it is out of action ways of avoiding this problem are being investigated. In particular, it has been decided to experiment with placing a buffer between machine 1 and machine 3. Now, a completed A part is transferred to the buffer, not to machine 3 directly. When machine 3 is ready to assemble another widget it loads an A part from the buffer, taking 2 seconds on average. The buffer only has limited capacity because space is limited within the workcell; initially the buffer capacity is set at 3.
An SPNP model representing the system with the additional buffer can be found in the file `pract-one2.c`.

d.) *What is the effect of introducing the buffer on the throughput of Machine 3? Does the size of the buffer matter?*

4 marks

e.) *What other steps could be taken to improve the throughput of the system? Which do you think are likely to be most effective?*

7 marks

20 marks total

This is the first of two practicals. The deadline for the complete practical is 16:00 on Thursday 6 November 2008. Submissions should handed in to the ITO. This is an individual practical: whilst discussion with your classmates is encouraged, the solutions that you submit must be your own work.

Kousha Etessami ⟨kousha@inf.ed.ac.uk⟩. October 10, 2008.