NATIONAL EXAMINATIONS May 2015

98-COMP A-5 OPERATING SYSTEMS

3 Hours Duration

NOTES:

1. If doubts exist as to the interpretation of any question, the candidate is urged to submit with the answer paper a clear statement of any assumption made.

2. Provide justifications for your answers. Show all your work.

3. CLOSED BOOK. Candidates may use one of the two pocket calculators, the Casio approved model or Sharp approved model. No other aids.

4. The candidate has to answer **any five questions** (each question has multiple parts).

5. Total Marks = 100.

6. This exam has got 6 pages (including this page).

1 [20 marks].

Consider a **preemptive** short term scheduling strategy in which the priority of a process may change dynamically with time. (Larger priority numbers imply higher priority). At any point in time the highest priority process is run on the system. Ties are broken in favour of the process that entered the ready to run queue first. If a running process is preempted then its time of entry into the ready to run queue is the time at which the preemption was made.

The priority of all processes is set to **P** when they enter the ready to run queue (upon arrival or after being preempted). When a process is waiting in the ready to run queue its priority changes at a rate **a**. That is,

> Priority of a process in the ready to run queue = $P + at^2$
> (where t is the time elapsed (in seconds) after the process entered the ready to run queue).

When a process is selected to run on the CPU its priority is set to **Q**. As it starts running on the CPU its priority changes at a rate **b**. That is,

> Priority of the process running on the CPU = $Q + bt'$
> (where t' is the time elapsed (in seconds) after the process started running on the CPU).

The parameters P, Q, a, and b can be set to give many different scheduling policies. Once chosen the values of these parameters become fixed and cannot change.

(i) Determine P, Q, a, and b that will produce the First Come First Served policy.

(ii) Determine P, Q, a, and b that will produce the Last Come First Served policy.
Under this policy whenever a process arrives on the system it preemptively captures the CPU. That is, if the CPU is free the process is allocated the CPU; if the CPU is busy then the executing process is preempted and the CPU is allocated to the process that just entered the ready to run queue. Whenever a process completes its CPU burst and the CPU is free, the process (in the ready to run queue) that was preempted most recently is allocated the CPU.

2. [20 marks]

(a) Discuss whether increasing the system memory can improve the CPU utilization in the following scenarios. If your answer is NO for any one of these scenarios describe what change needs to be made for the improvement of CPU utilization.

(i) when the CPU utilization is 12% and the paging disk utilization is 5%

(ii) when the CPU utilization is 10% and the paging disk utilization is 94%
(iii) when both the CPU and paging disk utilizations are 60%

(b) Consider a multiprogrammed system that uses multiple partitions (of variable size) for memory management. A linked list of holes called the free list is maintained by the operating system to keep track of the available memory in the system. At a given point in time the free list consists of holes with sizes:

100K, 42K, 205K, 180K, 70K, 91K, 125K, and 150K

The free list is also ordered in the sequence given above: the first hole in the list is of size 100K words which is followed by a hole of size 40K words and so on. Jobs with different memory requirements arrive on the system in the following order:

	Arrival Time	Memory Requirement
Job 1	t1	120K
Job 2	t2	104K
Job 3	t3	204K
Job 4	t4	88K

[Given t1 < t2 < t3 < t4]

Explain how memory allocation would be performed in the given situation for (i) the best fit and (ii) the first fit policy. [For each policy determine which hole is allocated to each job after it arrives on the system].

3. [20 marks]

(a) Discuss how the three requirements associated with the solution to the critical section problem are satisfied when a *monitor* is used for protection of shared data.

(b) Briefly describe the "Readers-Writers" problem that is well known in the context of concurrent processes. Provide the algorithm for a single monitor-based solution to the "Readers-Writers" problem.

(c) Distinguish between deadlock and starvation in the context of shared data being accessed by concurrent processes. Can starvation occur if the entry to the critical section for the shared data is guarded by a semaphore?

4. [20 marks]

(a) Using examples discuss the four necessary conditions for the occurrence of a deadlock.

(b) Consider a system comprising multiple resources. Multiple processes run concurrently on the system and can access these resources. A resource can be used by only one process at a time.

Once a resource is allocated it needs to be released by the process before it can be used by another process.

Each process is associated with a unique process identification number (PID). If process A has requested for a resource that is currently held by process B then process A is made to wait for process B to release the resource if and only if the PID of process A is higher than that of process B. Otherwise, process A is terminated and restarted after some time.

Can a deadlock occur on the system? Provide clear justifications for your answer.

(c) Discuss the difference between a safe, unsafe and a deadlock state.

5 [20 marks].

(a) What is real time system? Discuss why preemptive scheduling is an important in the context of any real time system.

(b) Consider a moving head hard disk which consists of a single platter (surface) with 120 tracks on it. The tracks are numbered 0 to 119. The disk has just completed a request at track 50 and is currently serving a request at track 54. The queue of pending requests in FIFO order is:

39, 68, 95.

(i) What is the total head movement (in number of tracks) needed to satisfy all these requests when the Shortest Seek Time First (SSTF) algorithm is used for disk scheduling?

(ii) What is the total head movement (in number of tracks) needed to satisfy all these requests when the LOOK algorithm is used for disk scheduling?

(iii) In what order should the requests be served to minimize the total head movement?

[Assume that no further requests arrive on the system during the service of the above requests.]

(c) Briefly explain the difference between load time and execution time address binding in the context of memory management.

6 [20 marks]

(a) Consider the following arrivals on a system. Each process has a single CPU burst and does not perform any I/O.

Process	Arrival Time (seconds)	Execution Time(seconds)
Proc1	0	16
Proc2	2	8
Proc3	4	18
Proc4	6	2

Find the mean process turnaround time when a variant of the Round Robin strategy is used for CPU scheduling. In this strategy the time slice allocated to a process is based on its execution time. For processes with execution times that are below 4 seconds a time slice of 2 seconds is used. For processes with execution times equal to or greater than 4 seconds but below 12 seconds a time slice of 3 seconds is used. For all other processes a time slice of 4 seconds is used.

(b) Consider the following page reference string (reference string) on a demand paged virtual memory system:

311,312,313,314,312,311,315,316,312,312,311,313,317,316,313,312,312,311,313,316

(i) What is the minimum number of page faults for this reference string that can occur on the system?

(ii) Determine the number page faults for this reference string that can occur on the system when 3 frames are allocated to the program and the Least Recently Used (LRU) page replacement strategy is sued.

7 [20 marks].

(a) Different methods exist for storing information on the disk. Consider a file currently consisting of 125 blocks. Assume that the directory is available in main memory.

For each of the following cases (A-C) compute the minimum number of disk operations that are required when <u>contiguous allocation</u> is used. Assume that there is no room for the file to grow in the beginning but there is room to grow in the end.

(A) The contents of block 100 are copied to block 2.
(B) The contents of block 110 are exchanged with the contents of block 101.
(C) A new block is inserted after block 90. The contents of this new block are the same as that of block 89. [Note that the file will now contain 126 blocks]

Consider each case (A-C) separately. Note that each disk operation corresponds to the reading of a block from the disk <u>or</u> the writing of a block to the disk. While computing the number of disk operations, ignore the disk operations that may be required for the location and maintenance of free space. Since the directory is in main memory any operation on the directory is not counted as a disk operation.

ASSUME: The length of the file is known to the system.

(b) With the help of examples discuss how fragmentation can occur on a disk. Describe a technique for controlling disk fragmentation.

(c) Briefly discuss how multiple disks can be used for improving system reliability.

(d) Discuss briefly the goals of protection for the file system used in a multi-user computer system.

Approximate Marking Scheme

1 20 marks

2(a) 7 marks
2(b) 13 marks

3(a) 5 marks
3(b) 9 marks
3(c) 6 marks

4(a) 8 marks
4(b) 7 marks
4(c) 5 marks

5(a) 5 marks
5(b) 10 marks
5(c) 5 marks

6(a) 8 marks
6(b) 12 marks

7(a) 8 marks
7(b) 4 marks
7(c) 4 marks
7(d) 4 marks