University of California, Berkeley
College of Engineering
Computer Science Division – EECS

Fall 2018 Ion Stoica

First Midterm Exam *Solutions*

October 1, 2018
CS162 Operating Systems

Your Name:	
SID AND 162 Login (e.g s042):	
TA Name:	
Discussion Section Time:	

General Information:
This is a **closed book and one 2-sided handwritten note** examination. You have 80 minutes to answer as many questions as possible. The number in parentheses at the beginning of each question indicates the number of points for that question. You should read **all** of the questions before starting the exam, as some of the questions are substantially more time consuming.

Write all of your answers directly on this paper. ***Make your answers as concise as possible.*** If there is something in a question that you believe is open to interpretation, then please ask us about it!

Good Luck!!

QUESTION	POINTS ASSIGNED	POINTS OBTAINED
P1	**20**	
P2	**20**	
P3	**10**	
P4	**30**	
P5	**20**	
TOTAL	**100**	

NAME: _____

P1. (20 points total) True/False and Why? **CIRCLE YOUR ANSWER.**

For each question: 1 point for true/false correct, 1 point for explanation.
An explanation cannot exceed 2 sentences.

a) Ignoring context switching overhead, round robin scheduling always has better average response time than first-come-first-serve

TRUE FALSE

Why?

FALSE.
Consider a FIFO queue with processes in the increasing order of their burst times.

b) Context switches are always involuntary

TRUE FALSE

Why?

FALSE: A thread can call yield() to relinquish the CPU

c) Every thread has its own heap

TRUE FALSE

Why?

FALSE: Threads in the same process share the same heap.

d) Using the stream API, it is possible to read from a random location of a file

TRUE FALSE

Why?

TRUE: Yes, you can use seek() to read from anywhere in a file.

e) In Pintos, the default scheduler implements First-Come, First-Serve

TRUE FALSE

Why?

FALSE. The default is round robin

f) Assume a process opens a file, stores the file descriptor in a variable `fd`, and then calls `fork()`. If the parent closes `fd`, the child's `fd` wil be closed as well

TRUE FALSE

Why?

FALSE: Fork replicates the file descriptors, so the parent closing an fd won't affect the child.

g) Starvation implies a system is deadlocked

TRUE FALSE

Why?

FALSE: A process can make progress by starving another one.

h) Semaphores can be used to implement scheduling constraints

TRUE FALSE

Why?

TRUE: Use V() to signal, and P() to wait for signal from ready_list.

i) Implementing critical sections and mutual exclusion involves waiting

TRUE FALSE

Why?

TRUE: If one process is in the critical section, other processes which want to access the critical section must wait.

j) In Pintos, the ordering of members inside the thread struct does not matter

TRUE FALSE

Why?

FALSE: unsigned magic must be last to detect stack overflows

NAME: _____

P2. (20 points total) What The Fork.

Eleanor, a CS162 student, wants to create a multithreaded program that she can use to tell everyone how much she loves her favorite operating systems class. She wrote the program below, but she finds that different runs produce different outputs.

```
1. void* func1(void* args) {
2.     printf("is\n");
3.     return NULL;
4. }

5. void* func2(void* args) {
6.     printf("CS162\n");
7.     return NULL;
8. }

9. int main(void) {
10. pid_t pid;
11. pthread_t pthread;
12.  int *ret = (int*) malloc(sizeof(int));
13.  int status;
14.  pid = fork();
15.  if (!pid) {
16.      pthread_create(&pthread, NULL, func2, (void*) ret);
17.        pthread_join(pthread, NULL);
18.  } else {
19.      printf("the\n");
20.  }
21.  printf("best!\n");
22.  return 0;
23.}
```

a) (**10 points**) List all of the possible outputs that her program could display when run. Assume that calls to fork and pthread_create always succeed.

List all possible outputs, one in each column.

CS162 best! the best!	CS162 the best! best!	the best! CS162 best!	the CS162 best! best!		

b) **(10 Points)** Modify Eleanor's program such that the output will always be:

```
CS162
is
the
best!
```

Fill in the blanks to show your changes to the original program.

- You must use `pthread_create` at least once.
- You may not call `printf` or the helper functions (`func1`/`func2`) directly.
- Write at most one statement per line. You may not need all lines

```
1. void* func1(void* args) {
2.    printf("is\n");
3.    return NULL;
4. }
5. void* func2(void* args) {
6.    printf("CS162\n");
7.    return NULL;
8. }
9. int main(void) {
10.   pid_t pid;
11.   pthread_t pthread;
12.   int *ret = (int*) malloc(sizeof(int));
13.   int status;

      _____
14.   pid = fork();
15.   if (!pid) {
            _____
16.         pthread_create(&pthread, NULL, func2, (void*) ret);
17          pthread_join(pthread, NULL);
            pthread_create(&pthread, NULL, func1, (void*) ret);
            pthread_join(pthread, NULL);
            exit(0);                    // return 0 also works
            _____
18.   } else {

            wait(&status);
            _____
19.         printf("the\n");

            _____
20.   }

      _____
21.   printf("best!\n");
22.   return 0;
23.}
```

NAME: _____

P3. (10 points total) Networking

The following program implements a simple network server that accepts client connections and writes back the string "Hello!", and then closes the connection. Assume no failures occur in any function calls and that reads and writes to sockets will always result in the entire message being received or sent.

```
#define BUFFER_SIZE 4096

int main() {
    struct sockaddr address;
    int addrlen = sizeof(address);
    /* socket initialization code omitted */
    int serverfd = socket(PF_INET, SOCK_STREAM, 0);
    bind(serverfd, &address, sizeof(address));
    listen(serverfd, 5);
    char buffer[BUFFER_SIZE];
    char *string = "Hello!";
    while (1) {
        int newsockfd = accept(serverfd, &address, &addrlen);
        write(newsockfd, string, strlen(string));
        close(newsockfd);
    }
    close(serverfd);
    return 0;
}
```

a) **(5 points)** On receiving the message, the server writes it to an existing log file in the local directory named "log.txt". Assume that messages are no larger than 4096 bytes. Fill in the code below to accomplish this task. Please write at most one statement per line.

```
char buffer[BUFFER_SIZE];
char *string = "Hello!";
int fd = open("log.txt", O_WRONLY | O_APPEND);
while (1) {
    int newsockfd = accept(serverfd, &address, &addrlen);

      size_t bytes_read = read(newsockfd, buffer, BUFFER_SIZE);

      write(fd, buffer, bytes_read);

    write(newsockfd, string, strlen(string));
    close(newsockfd);
}
close(fd);
```

b) **(5 points)** Instead of writing to the log, we want to print messages to the console. Add only a single statement to your solution in part a) in order to print client messages to standard output instead of to the log. You may assume that part a) was implemented correctly, and you may not modify any of part a).

```
while (1) {
      int newsockfd = accept(serverfd, &address, &addrlen);

      dup2(STDOUT_FILENO, fd); OR
      dup2(1, fd); OR
      dup2(fileno(stdout), fd)

      /* solution to part a */
      write(newsockfd, string, strlen(string));
      close(newsockfd);
}
```

A solution with `fd = 1` was awarded partial credit; this does not close `fd`, thus it results in a resource leak.

NAME: _____

P4. (30 points total) CPU Scheduling.

Consider the following single-threaded processes, and their arrival times, CPU bursts and their priorities (a process with a higher priority number has priority over a process with lower priority number):

Process	Burst Time	Arrival Time	Priority
A	3	1	1
B	3	2	3
C	4	2	2
D	1	4	4

You will consider three scheduling algorithms: (1) round robin (RR), (2) shortest remaining time first (SRTF), and (3) priority scheduling. The following apply:

- All schedulers are preemptive.
- The time quanta of the RR scheduler is 1.
- If a process arrives at time x, it can be scheduled immediately.
- If two processes arrive at the same time, they are inserted in the ready queue in the lexicographical order. For example, if B and C arrive at the same time, B is inserted first, and C second in the ready queue.
- In the case of the RR scheduler, a new arriving process is inserted at the *end* of the ready queue. When the RR quantum expires, the currently running thread is added at the end of the ready list before any newly arriving threads.
- The RR scheduler ignores the process priorities.
- The SRTF scheduler uses priorities to break ties, i.e., if two processes have the same remaining time we schedule the one with the highest priority.

NAME: _____

a) **(15 points)** Given the above information, fill in the following table. You may not need all the spaces.

Time	Round Robin	SRTF	Priority
1	A	A	A
2	A	A	B
3	B	A	B
4	C	D	D
5	A	B	B
6	B	B	C
7	D	B	C
8	C	C	C
9	B	C	C
10	C	C	A
11	C	C	A
12			
13			
Average completion time	(5+8+10+4) / 4 = 6.75	(3+6+10+1) / 4 = 5	(11+4+8+1) / 4 = 6

Duplicate of table for your convenience:

Process	Burst Time	Arrival Time	Priority
A	3	1	1
B	3	2	3
C	4	2	2
D	1	4	4

b) **(15 points)** Additionally, now assume there are two locks L1 and L2, respectively. Furthermore assume the following:

- Process A acquires L1 and L2 in its first unit of time. It releases L1 in its last unit of time and releases L2 in its second unit of time
- Process B acquires L1 in its first unit of time and releases L1 in its last unit of time
- Process D acquires L2 in its first unit of time and releases L2 in its last unit of time
- Processes will sleep if acquiring lock fails. (Still scheduled, does not consume a burst)
- Priority donation is implemented.

Given the above information, fill in the following table. You may not need all the spaces. If it helps, you may specify when a process is put to sleep/blocked

Note: For Round Robin, you may assume blocked threads are not kept on the ready queue (The second set of round robin answers makes this assumption)

Time	Round Robin	SRTF	Priority
1	A | A	A	A
2	A | A	A	B (blocked)
3	B | B (blocked)	A	A
4	C | C	D	D
5	A | A	B	A
6	B | D	B	B
7	D | C	B	B
8	C | B	C	B
9	B | C	C	C
10	C | B	C	C
11	B | C	C	C
12	C | B		C
13			
Average completion time	(5+11+10+4)/4 = 7.5 (5+11+10+3)/4 =7.25	(3+6+1+10) / 4 = 5	(5+7+11+1) / 4 = 6

NAME: _____

Depending on the interpretation of how blocked threads are handled, there are two possible solutions to round robin.

Each column is worth 5 points, 4 points were awarded for correctness of scheduling, 1 point was awarded for correct calculation of average completion time (based on your order)

P5. (20 points total) Concurrency.

Consider the following piece of code, which can be executed by one or more threads:

```
x = x + 1
y = x + y
```

Assumptions:

- Before any thread starts running, both x and y are initialized to 1. After that x and y are *only* updated by the threads executing the above code segment.
- A thread can be preempted at any point during its execution.
- Each of the two instructions is atomic.

a) **(5 points)** Assume two copies of the above piece of code run concurrently in two threads. Write down all possible values after both threads finish, one per column (number of outputs might be smaller than the number of columns).

x	3	3						
y	6	7						

b) **(5 points)** Same question, but now assume that three copies of the code segment in three threads.

x	4	4	4	4				
y	13	12	11	10				

c) **(5 points)** Assume *10* copies of the same program run concurrently. What is the *maximum* value of y?

111
More generally let n be the number of concurrent threads. Then, y = n^2 + n + 1

Maximum y occurs when all x = x + 1 executes first, and all y = x + y execute last:
x = x + 1 (x = 2)
...
x = x + 1 (x = n + 1)
y = x + y (y = (n + 1) + 1)
y = x + y (y = 2*(n+1) + 1)
...
y = x + y (y = n*(n+1) + 1 = n^2 + n + 1)

Grading:

- We subtracted only one point to people which were off by one iteration, either did the math for 9 or 11 threads. Thus, these solutions got 4 point.
- We gave 1 point for solutions that showed a few iterations, but their answer was far from the correct one.
- We didn't give any partial credit to answers that were way off, and, obviously for "no answer".

d) **(5 points)** Assume *10* copies of the same program run concurrently. What is the *minimum* value of y?

66

More generally let n be the number of concurrent threads. Then, y = n^2 + n + 1

Maximum y occurs when all x = x + 1 executes first, and all y = x + y execute last:
x = x + 1 (x = 2)
y = x + y (y = 3)
x = x + 1 (x = 3)
y = x + y (y = 6)
x = x + 1 (x = 4)
y = x + y (y = 10)
x = x + 1 (x = 5)
y = x + y (y = 15)
...
x = x + 1 (x = n + 1)
y = x + y (x = n + 1

y = 1 + 2 + 3 + 4 + 5 + .. + n+1 = (n+1)*(n+2)/2

Grading:

- We subtracted no points to people who wrote the series: 1 + 2 + 3 + … 11, but did the math wrong, i.e., wrote an answer other than 66. These solutions got full points.
- We subtracted only one point to people which were off by one iteration, either did the math for 9 or 11 threads. Thus, these solutions got 4 points.
- We gave 1 point for solutions that showed a few iterations, but their answer was far from the correct one.
- We didn't give any partial credit to answers that were way off, and, obviously for "no answer".

Note: very few students gave the answers 2 and 3. We assume this was because they thought this was a trick question, and gave the results after the first thread was executed. Obviously, this was not a trick question. Two more observations here. First, "2" is never possible because "y = x +y" is executed after at least one "x = x + 1" statement is executed. Third, even if you thought this was a trick question, why assume only one

thread was executed? Why not assume zero threads were executed, and then the minimum number was "1". For those reasons we gave no partial credit for these answers.

P6. (0 points) Potpurri

a) I am glad that there aren't anime references in this exam

TRUE TRUE

Why?

b) Name the one pop culture reference in this exam

The Good Place

c) Draw your TA