CSC 246: Concepts and Facilities of Operating Systems for Computer Scientists (Section 001)

Midterm Exam – Spring 2023

Name Qizheng Jin Unity ID qjin3

Please read the following instructions carefully:

- This exam is closed book and closed notes.
- Your work must be individual. Cheating will be punished according to university academic integrity code.

About this exam:

- The total point value of the exam is 100.
- Problem 1 is a list of multiple-choice questions.
- Problem 2 is a list of short answer questions.
- Problems 3-5 are problems require longer answers.

Problem	Possible Points
1. Multiple Choice	12
2. Short Answer	20
3. Processes	20
4. Scheduling	20
5. Threads and Synchronization	28
Total	100

1. Multiple Choice Questions (12 pts, 2 pts each)

Mark the **best** answer for each of the following questions.

B a) A microkernel is a kernel _____.

A. containing many components that are optimized to reduce resident memory size
B. that is compiled to produce the smallest size possible when loading the operating system into main memory
C. that is stripped of all nonessential components
D. has very limited or no structure at all

B b) Which of the following would you expect to be stored in a Process Control Block for some process, P?

A. All the executable code for P
B. A copy of all the CPU Registers for P
C. A copy of the stack for P
D. A copy of the text section for P

D c) _____ saves the state of the currently running process and restores the state of the next process to run.

A. The CPU scheduler
B. Swapping
C. A trap instruction
D. A context switch

D d) Which is NOT shared by threads in a process?

A. Program Counter
B. Heap
C. Opened files
D. Code

A e) ____ is the number of processes that are completed per time unit.

A. CPU utilization
B. Response time
C. Throughput
D. Turnaround time

A f) A multilevel feedback queue scheduler generally assigns a long quantum to _____

A. high priority processes.
B. low priority processes.
C. new processes.
D. old processes.

2. Short Answer (20 pts, 5 pts each)

Provide brief answers to the following questions.

a) For inter-process communication, explain why shared memory is typically faster than message passing during the communication.

It can directly gain the information from the memory, but the communication needs to convey the information through a pipe between child and parent process, which takes more time.

b) Processes (or threads) can be in one of three states: Running, Ready, or Waiting. For each of the following examples, write down which state the process (or thread) is in:

1) Spin-waiting for a variable to become zero.

Running

2) Having just completed an I/O, waiting to get scheduled again on the CPU.

Ready

3) Blocking on a condition variable waiting for some other thread to signal it.

Waiting

4) Scanning through the buffer cache within the kernel in response to read().

Running

c) For Round Robin scheduling, it's important to choose a proper quantum length. Explain (1) what is the issue if we make the quantum length too long, and (2) what is the issue if we make the quantum length too short?

If it is too long, ~~there will~~ it will make no difference with other kind of scheduling because the Ready Queue won't work effectively and processes can't be arranged fair enough.

If it is too short, the processes will change very quickly, and it takes more resources to switch between the processes, which is a kind of waste.

d) Consider the following flawed solution to the critical section problem.

```
int turn = 0;

ThreadA() {
  for ( int i = 0; i < 999;
        i++ ) {
    while ( turn == 1 ) {
    }
    critical Section;
    turn = 1;
    remainder Section;
  }
}
```

```
ThreadB() {
  for ( int i = 0; i < 1000;
        i++ ) {
    while ( turn == 0 ) {
    }
    critical Section;
    turn = 0;
    remainder Section;
  }
}
```

There are three requirements that a correct solution to the critical section problem must satisfy. Which of these isn't satisfied by this code? Briefly explain.

After some loops, it will cause deadlock. When turn=1, it will stuck in the while loop in Thread A, because it is

It should give out the signal that a thread has finished its critical section, like Peterson algorithm.

3. Processes (20 pts)

Considering the following C code. Pretend this is an implementation of a UNIX shell. Like any shell, the job of this program is to repeatedly read commands from the user and execute them. Each user command gives the name of an application. The shell is supposed to run the application, wait for it to finish and then read the next command from the user.

```
while (true) {
      get_command(command, parameters);
      exec(command, parameters); //this is the exec() system call
}
```

a) Using the above code, what problem will happen after a user issues a command? Explain. (5 pts)

If the command causes a dead loop, it will stuck in the exec( ) and can't come out, and the shell won't stop it automatically.

b) Using a few lines of code change/addition (hint: using fork()), you can fix the problem. Show how this can be fixed (10 pts)

```
while true {        pid_t pid;
    get_command (command, parameters);
    fork( );
    exec (command, parameters);
    if ( pid!=0) { printf " Can't execute the command ");
                   exit (0);
}.                }.
```

If parent process can't get signal from child process, then terminate the shell

c) A UNIX shell also supports background execution of a command. That is, once the application is started, the shell is ready to receive the next command, without waiting for the finish of the previous command. Modify your code for problem b) to support background execution. (5 pts)

```
while true {        pid_t pid
    get_command (command, parameters);
    fork( );
    exec (command, parameters);
    while (exec( )!=0)) get_command (command, parameters);.
    if (pid!=0) { printf " Can't execute the command ");
                  exit (0);
                }.
}.
```

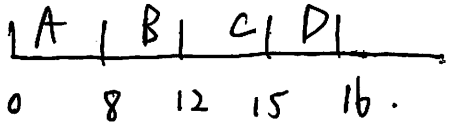
4. Scheduling (20 pts)

The table below shows four processes that are to be scheduled on a single CPU. Each process has an arrival time before which it cannot run, a CPU burst time - the number of time units it needs to run, and a priority number (assume smaller number represents higher priority). For SJF (shortest job first) scheduling, assume that the scheduler knows the length of each CPU burst. Break any ties by choosing the process that appears earlier in the table. For each scheduling, draw the Gantt chart and calculate the average waiting time of the four processes.

Process	Arrive	Burst	Priority
A	0	8	5
B	3	4	3
C	5	3	6
D	9	1	1

3<5.
3<6.

a) First come first served scheduling: (5 pts)

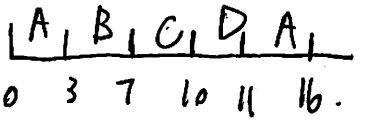


Turnaround: (8+9+10+7)/4 = 8.5

Waiting: (0+5+7+6)/4 = 4.5

b) Shortest job first (SJF) scheduling: (5 pts)

5

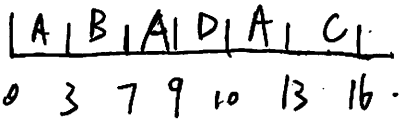


Turnaround: (16+4+5+2)/4

Waiting: (8+0+2+1)/4 = 11/4

c) Preemptive priority scheduling: (5 pts)

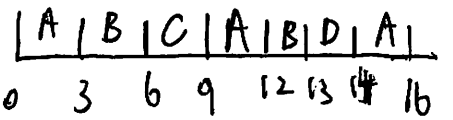
5-2=3.



Turnaround: (13+4+11+1)/4

Waiting: (5+0+8+0)/4 = 13/4

d) Round-robin scheduling (quantum = 3): (5 pts)



Turnaround: (16+10+4+5)/4

Waiting (8+6+1+4)/4 = 19/4

RQ: ABCABD.A

5. Threads and Synchronization (28 pts)

a) Consider the following three threads. Pretend that the individual assignment statement will each execute atomically.

```
1  threadP(){
2    a = 1;
3    b = 2;
4  }
5
6  threadQ(){
7    b = 1;
8  }
9
10 threadT(){
11   a = 2;
12 }
```

1) List all the possible interleavings of the execution of statements in these three threads. For example, "a = 1; b = 2; b = 1; a = 2" is one possible execution order. (6 pts)

① a=1 b=2 b=1 a=2. ④ a=1 b=2 a=2 b=1

② b=1 a=1 b=2 a=2 ⑤ b=1 a=2 a=1 b=2.

③ a=2 a=1 b=2 b=1 ⑥ a=2 b=1 a=1 b=2

2) If all interleavings are as likely to occur, what is the probability to have a = 1 and b = 1 after all threads complete execution? (2 pts)

① X ② X ③ ✓ ④ X ⑤ X ⑥ X ∴ $\frac{1}{6}$

3) Revise the original code using semaphores to guarantee after all threads complete execution, it always has a = 1 and b = 1. Show clearly how many semaphores are used, what are their initial values, and where you insert the acquire () and release () operations (using line numbers). (5 pts)

T → P → Q

sem TDone = 0 sem PDone = 0.

~~Thread X {~~

```
thread P() {
  acquire(TDone);
  a=1;
  b=2;
  release(PDone);
}.

thread Q() {
  acquire(TDone);
  acquire(PDone);
  b=1;
}

thread T() {
  a=2
  release(TDone);
}
```

b) Consider two threads A and B that perform two operations each. Let the operations of thread A be A1 and A2; let the operations of thread B be B1 and B2. We require that threads A and B each perform their first operation before either can proceed to the second operation. That is, we require that A1 be run before B2 and B1 before A2. Consider the following solutions based on semaphores for this problem (the code run by threads A and B is shown in two columns next to each other). For each solution, explain whether the solution is correct or not. If it is incorrect, you must also point out why the solution is incorrect. (15 pts, 5pts each)

$A_1 \longrightarrow B_2$. 安. $B_1 \Rightarrow A_2$ 安.

1)
```
sem A1Done = 0;
sem B1Done = 0;

ThreadA(){                          ThreadB(){
A1;                                 B1;
acquire( B1Done );                  acquire( A1Done );
release( A1Done );                  release( B1Done );
A2;                                 B2;
}                                   }
```

若 A_1, B_1, B_1 A_1 B_2 A_2

Incorrect. In Thread B, it will release B_1 after get signal from A1Done, but it will only release (A1Done) after it get signal from B1Done, which causes a dead lock.

2)
```
sem A1Done = 0;
sem B1Done = 0;

ThreadA(){                          ThreadB(){
A1;                                 B1;
acquire( B1Done );                  release( B1Done );
release( A1Done );                  acquire( A1Done );
A2;                                 B2;
}                                   }
```

A_1 B_1 A_2/B_2 B_2/A_2.

correct

3)
```
sem A1Done = 0;
sem B1Done = 0;

ThreadA(){                          ThreadB(){
A1;                                 B1;
release( A1Done );                  release( B1Done );
acquire( B1Done );                  acquire( A1Done );
A2;                                 B2;
}                                   }
```

A_1 B_1 B_2/A_2.

correct

This is one scratch page.