1. (16%) Consider two-process solutions to the critical-section problem. The processes are numbered P0 and P1. For convenience, when presenting Pi, we use Pj to denote the other process; that is, j = 1 - i. The processes share two variables:
   var flag: array[0..1] of boolean;
   turn: 0..1;
Initially, flag[0]=flag[1]=false, and the value of turn is immaterial (but is either 0 or 1). The structure of Pi is shown for each of the following 4 proposed solutions. You are to check correctness of each solution. (不倒豮)

Solution 1:
   repeat
       flag[ i ] := true;
       turn := j;
       while (flag[ j ] and turn = j) do no-op;
       ....critical section......
       flag[ i ] := false;
   until false;
(a) Is mutual exclusion preserved? (Just answer yes or no.)
(b) Is the progress requirement satisfied? (Just answer yes or no.)

Solution 2:
   repeat
       flag[ i ] := true;
       turn := i;
       while (flag[ j ] and turn = i) do no-op;
       ....critical section......
       flag[ i ] := false;
   until false;
(c) Is mutual exclusion preserved? (Just answer yes or no.)
(d) Is the progress requirement satisfied? (Just answer yes or no.)
Solution 3:

\[
\text{repeat} \\
\text{flag[ i ] := true; } \\
\text{turn := j; } \\
\text{while (flag[ i ] and turn = j) do no-op; } \\
\text{......critical section......} \\
\text{flag[ i ] := false; } \\
\text{until false; }
\]

(e) Is mutual exclusion preserved? (Just answer yes or no.)
(f) Is the progress requirement satisfied? (Just answer yes or no.)

Solution 4:

\[
\text{repeat} \\
\text{flag[ j ] := true; } \\
\text{turn := i; } \\
\text{while (flag[ j ] and turn = i) do no-op; } \\
\text{......critical section......} \\
\text{flag[ j ] := false; } \\
\text{until false; }
\]

(g) Is mutual exclusion preserved? (Just answer yes or no.)
(h) Is the progress requirement satisfied? (Just answer yes or no.)

2. (4%) A semaphore S is an integer variable that, apart from initialization, is accessed only through two standard atomic operations: wait and signal. The classical definitions of the two operations are:

\[
\begin{align*}
\text{wait}(S): & \quad \text{while } S \leq 0 \text{ do no-op;} \\
& \quad S := S - 1; \\
\text{signal}(S): & \quad S := S + 1;
\end{align*}
\]

What are the actions in \text{wait}(S) operation that must be executed without interruption?
3. (5%) In a uniprocessor environment, semaphores can be implemented by interrupt enabling/disabling. Suppose the following is the implementation of signal(S) operation for semaphore S:

   <disable interrupt>;
   S := S + 1;
   <enable interrupt>;

   What is the corresponding implementation of wait(S) operation for semaphore S?

4. (5%) Processes that want to communicate must have a way to refer to each other. They can use either direct communication or indirect communication. List five differences between these two different approaches.

5. (3%) What is the priority inversion problem? This problem can be solved by the priority inheritance protocol. Explain how this protocol works.

6. (6%) In the original UNIX operating system designs, system calls are purposely made un-preemptible. Why? To support real-time process scheduling, however, long-duration system calls should be made preemptible. Describe an approach that can safely and effectively achieve this goal.

7. (6%) Draw a diagram to explain the inverted page table design. Compare it with the normal page table design and list its several advantages and disadvantages.

8. (5%) They are two places that a swap space can reside. Swap space can be carved out of the normal file system, or it can be in a separate disk partition. List the advantages and disadvantages of each approach.
9. (2%, 1%, 2%, 2%) (a) What is the disadvantage if without applying Amdahl's law? (b) What is the meaning of CPI = 1.5? (c) Can the CPI be smaller than 1? Why? (d) Do we need to know the ISA while design a good compiler? Why?

10. (5%) Please draw the formats of five MIPS addressing modes.

11. (4%) Compare the number of gate delays for the critical paths of two 16-bit adders, one using ripple carry and the other using two-level carry lookahead.

12. (3%) Prove that the Booth’s algorithm work for multiplication of two’s complement signed integers.

13. (6%) According to the following figure, what are the values of a(1), a(2), …, d(7) in the table?

<table>
<thead>
<tr>
<th>Instruction</th>
<th>RegDst</th>
<th>ALUSrc</th>
<th>Mem to Reg</th>
<th>Reg Write</th>
<th>Mem Read</th>
<th>Mem Write</th>
<th>Branch</th>
<th>ALUOp1</th>
<th>ALUOp0</th>
</tr>
</thead>
<tbody>
<tr>
<td>R-format</td>
<td>a(1)</td>
<td>a(2)</td>
<td>a(3)</td>
<td>a(4)</td>
<td>a(5)</td>
<td>a(6)</td>
<td>a(7)</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>lw</td>
<td>b(1)</td>
<td>b(2)</td>
<td>b(3)</td>
<td>b(4)</td>
<td>b(5)</td>
<td>b(6)</td>
<td>b(7)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>sw</td>
<td>c(1)</td>
<td>c(2)</td>
<td>c(3)</td>
<td>c(4)</td>
<td>c(5)</td>
<td>c(6)</td>
<td>c(7)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>beq</td>
<td>d(1)</td>
<td>d(2)</td>
<td>d(3)</td>
<td>d(4)</td>
<td>d(5)</td>
<td>d(6)</td>
<td>d(7)</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
14. Why it may cause the stale data problem during DMA I/O data transfer? How to overcome it? Give three different approaches to resolve this problem and explain your reasons clearly. (6%)

15. What is called split transaction protocol used in I/O data bus design? (3%)

16. There are three ways to schedule the branch delay slot in order to reduce or eliminate the control hazard. Give simple examples to explain their principle clearly and briefly. (6%)

17. It is well known that multi-level cache design is one of the most important ways to upgrade the CPU performance. Suppose that we have a processor with a base CPI of 1.0, assuming all references hit in the primary cache, and a clock rate of 1000 MHz. Assume a main memory access time of 200 ns, including all miss handling. Suppose the miss rate per instruction at the primary cache is 5%. How much faster will the machine be if we add a secondary cache that has a 20 ns access time for either a hit or a miss and is large enough to reduce the miss rate to main memory to 2%? What are the global miss rate as well as local miss rate for this two level cache machine? (10%)