Code:————————

Breadth Exam in Systems

September 12, 2005

The exam covers **Core area** and five **Optional Areas**. You must answer all questions in the core area and all questions in two of the optional areas you have chosen. This is a closed book/notes exam. Calculators and other electronic devices such as PDAs and cell phones are not allowed. If you have them, turn them off and put them away. You have **Three hours** to complete the exam. Leave at least $\frac{1}{2}$ inch margin on all four sides. Write on only one side of each sheet of paper. Start answering questions in each area on a separate sheet of paper. Write your code on each page at the top right corner. Use a black-ink pen or pencil only.

1 Core

1. Suppose there are N worker processes (0. . .N-1), each with a critical section where work is done. There is also a special process which executes the following code:

```
int p    = -1;    // GLOBAL - shared by all, initialized once
int lock = 0;     // GLOBAL - shared by all, initialized once

while (1) {
  while (lock == 1) { }   // empty loop body
  p = (p+1) % N;
}
```

   Each worker process has its own process ID mypid, and executes the following code:

```
while (1) {
  while (p != mypid) {}   // empty loop body
  lock = 1;
  // CRITICAL SECTION WHERE THE PROCESS DOES WORK
  lock = 0;
}
```

   Answer YES/NO to the following questions and provide a brief justification for your answer. Note that all N+1 processes are executing concurrently.

   (a) Does the solution guarantee mutual exclusion?

   (b) Does the solution guarantee that work will be done?

   (c) Does the solution guarantee fairness?

2. Consider a system with 64 MB of physical memory, 32-bit physical addresses, 32-bit virtual addresses, and 4 KB physical page frames.

   (a) Using a single-level paging scheme, what is the maximum number of page table entries for this system?

(b) Using a two-level paging scheme with a 1024-entry outer-page table (level-1), what would be the page offset of the page of the page table (level-2) accessed for the virtual address 00110000000100001110001101011011 ?

(c) Suppose a TLB is used with the two-level paging scheme described in part b, and the TLB has a 90% hit rate. If the TLB access time is 10 ns and memory access time is 100 ns, what is the effective memory access time of the system?

2 Databases

Problem 1

The Athens-2004 database stores information about different Olympic events and their results. The database schema is:

- Sport(SportId INT, Name Char(30), Type INT, Gender CHAR(3), FirstIntroduced INT);
- Competition(Sport INT, Code, CName Char(80), Gender CHAR(1)
- Results(Sport INT, Comp INT, Place INT, Team Char(20));

Notes. Competition.Sport is a foreign key on Sport. Results.Sport and Results.Comp form a foreign key on Competition. Sport.Type has values 1,2,3 with the following meaning:

Sport.Type value	meaning	Example sports
1	Individual competitions only	wrestling, boxing
2	Individual and pairs competitions only	tennis, table tennis
3	Individual and team competitions only	athletics, swimming, gymnastics
4	Team competitions only	soccer, basketball, volleyball

Sport.Gender can take values "W", "M" and "W/M" with the meaning that the given sport is a women-only, men-only, or a sport with competitions for both women and men. Competition.Gender can take values "W" and "M" — individual competitions are either men-only or women-only. Competition.CName is the name of an individual competition: e.g. "400 meters running" for athletics or "Men's team" for basketball.

1. Reverse engineering problem: draw the E-R diagram of the Athens-2004 database. Please include all entity sets and all relationship sets. Show primary keys/identifying attributes of each entity set on the diagram. Indicate the type of each relationship (many-to-many, one-to-many, one-to-one) – both on the diagram, and separately, in writing.

2. Write an SQL statement for each of the following queries.

   (a) Find all teams with at least one gold medal. Return the name of the team only. Each team name must appear no more than one in the answer.

   (b) For each team, find the number of its gold medals in team competitions. Output team name and the number of medals.

   (c) Find the team with the highest total number of medals.

3. Explain in English what the following SQL query will return.

```
(SELECT DISTINCT R.Team
   FROM Results)
EXCEPT
  (SELECT DISTINCT R.Team
   FROM Results R
   WHERE Place =1 AND
      Team NOT IN (SELECT R1.Team
                   FROM Results R1, Competitions C
                   WHERE R1.Sport = C. Sport AND R1.Comp = C.Code
                         AND C.Gender = "M" AND Place = 1
                   )
```

Problem 2

Consider the following transaction schedule.

T1	T2	T3	T4
R(A)			
	R(B)		
R(C)			
		R(B)	
			W(B)
	W(C)		
		R(A)	
W(A)			
		R(C)	
	W(B)		
			W(A)

1. Identify all conflicts in this schedule, specify their type.
2. Is this schedule serializable? Explain your answer.

Problem 3

Consider the following transaction scheduling protocol:

> **XLock-only.** Transactions must acquire exclusive (X) locks on objects they want to write to. Each object can have at most one exclusive lock associated with it. Transactions do not need locks to read objects, but they cannot read an object if another transaction has an X lock on it. All locks are released when transaction is committed or aborted.

Will transaction schedules formed according to the XLock-Only protocol be serializable? Explain your answer (if the answer is ”yes”, give a proof, if the answer is no, give a counterexample).

3 Distributed/Advanced Operating Systems

1. For each of the following statements, state whether it is true of false. **Justify your answer.**

   (a) Lampot's mutual exclusion algorithm will work correctly even if local clock time is used for timstamping messages.

   (b) Lamport's time is useful for tracking causal dependency among events of a distributed computation

   (c) Chandy-Lamport's global snapshot collection algorithm will work correctly even if channels are not FIFO.

   (d) A program that works correctly under a Distributed Shared Memory System that implements Processor Consistency model will work correctly under a Distributed Shared Memory System that implements Pipelined Random-Access Memory Consistency model.

   (e) Existence of a cycle in the Resource Allocation Graph is necessary and sufficient for the existence of deadlock in the system.

2. (a) Describe two applications of Lamport's time in detail. Give one use of vector time.

   (b) Assume that the clocks of all machines in a distributed system are perfectly synchronized. Present an efficient coordinated checkpointing algorithm for distributed computations in such a system.

   (c) Describe one advantage and one disadvantage of (i) Sender-initiated load sharing algorithms, and (ii) Receiver-initiated load sharing algorithms. Discuss in detail under which situations each of these types of algorithms will be efficient.

4 Networks

Problem 0. This problem deals with *subnetting* (or, more generally, classless addresses) in IP networks

1. Explain the purpose of subnetting. (**Note:** do not say what subnetting *does*—see the next part—this part of the problem is about *why* subnetting is used.)
2. Explain how it works, and what protocol mechanisms are used to implement it.
3. What property of the assignment of addresses to networks is required in order for subnetting to be effective?

Problem 1. This problem deals with flow control and congestion control.

1. Explain the difference between flow control and congestion control.
2. Say what protocol(s), if any, are responsible for these functions in today's Internet, and describe, in as much detail as possible, how each of these functions is implemented.

Problem 2. Consider a simple link-layer sliding-window protocol with a 3-bit sequence number space. The receive window size is 1. (**Note:** show your work.)

1. What is the maximum number of frames the sender can have outstanding at any one time?

2. Suppose this protocol is in use over an 80 Mbps channel (that's 8×10^7 bits/sec) with a 25msec propagation delay in each direction, and that every frame is 1000 bytes. What is the absolute throughput (in bits/sec) if no packets are lost in transit?
3. For the same channel, what is the absolute throughput if one frame in every 10,000 is lost?
4. For these channel conditions, what *one* change could you make to the protocol to drastically improve throughput?

5 Compilers

1. (a) Describe the problem of register allocation and its importance in code generation and optimization.

   (b) In the short program below, there are six temporaries `a, b, c, d, e` and `g`.

```
a = 1;
b = 13;
c = 10 + a;
d = a + c;
e = c + d;
g = b + 7;
c = g + e;
g = c + e;
b = c + 5;
return b + g;
```

   i. What is the fewest number of registers required for this program so there is no need for spilling? Justify your answer by using for example the interference graph and its coloring.

   ii. Show the program after register allocation.

2. Define the concept of unambiguous grammars. Provide an ambiguous Context Free Grammar for the language of arithmetical expressions and show two different parse trees for `a + b + c`.

6 Programming Languages

1. Show how to use call-by-name to code a single procedure that can be used to (1) add elements 1 . . . 10 of an array, (2) add elements 3 . . . 12 of an array, (3) add 1^2, 2^2, . . ., 10^2, all by just changing the actual parameters. (This trick is called *Jensen's Device*.)

2. Algol-like languages usually use a *central stack* at runtime; the stack holds chunks of data known as *frames*. 1) What is stored in a stack frame? Give a complete list. (2) If stack frames were stored in well-known places in memory instead of a stack, what ability would the programming language lose? (3) Can CLU iterators be implemented using a single stack? Explain why or why not. (4) Can Ada tasks be implemented using a single stack? Explain why or why not.