Systems Breadth Exam: Spring 2007

Department of Computer Science, University of Kentucky

You must answer all the questions in the Core Material section. You must also answer all the questions in two of the other sections.

Answer every section on a different piece of paper (or pieces of paper). **Put your exam id on every piece of paper you submit.** Do **not** put your name on any piece of paper. Your exam id is

1 Core material

1. In a particular computer architecture for paging, 16-bit addresses are considered to have 8 bits for page number and 8 bits for offset.

   (a) Are the page-number bits the high-order or the low-order bits?

   (b) What is the reason that the page-number bits must be the ones you suggest?

2. What three things does the hardware (not the operating system) do when an interrupt occurs?

3. Write a program with two processes that use semaphores as their synchronization technique. This program should *always* deadlock when it is run. Other than deadlocking, the program does not have to have any obvious purpose. Use any programming-language notation you like (it can look like C, for instance) so long as it is clear. You may use a barrier if you like, but the deadlock should not be at the barrier.

4. I am building a special-purpose computer for a scientific application. I need to represent integers, but I guarantee they will always be in the range from 0 to 2,935,121.

   (a) At least how many bits long must my integer representation be? (You may not use a calculator for this question!)

   (b) How many bits would you recommend I use for my integers?

   (c) Explain your reason for the answer you gave to (b).

5. Consider this truth table:

A	B	C	output
0	0	0	1
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	0
1	0	1	0
1	1	0	1
1	1	1	1

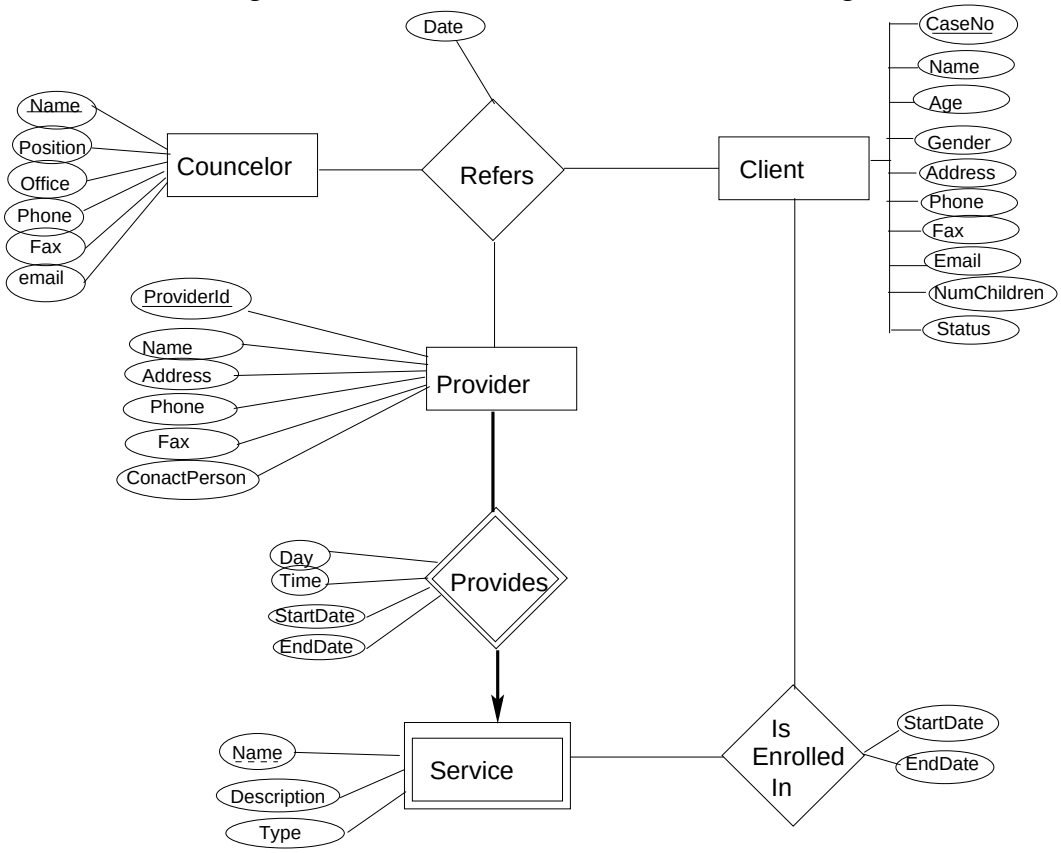
   (a) Write the function represented by this truth table as a Boolean combination of A, B, and C. Make your function as simple as you can.

   (b) Draw a circuit using only **AND**, **OR**, and **NOT** gates to implement this truth table. Try to minimize the number of gates you need.

2 Databases

1. Consider a database that keeps track of services provided by a municipal social-services agency. The agency is staffed with counselors, who meet with the clients (recipients of the services). The actual services are provided by a network of in-town providers. The database tracks the referrals that counselors issue for their clients to specific providers. Based on these referrals (or sometimes independently of the referrals), the clients enroll in the specific services.

   The E-R diagram of the database is described in this figure:



   (a) Describe the type of each relationship set in the diagram (one-to-one, many-to-one, total, . . .). In addition to constraints shown on the diagram, are there any *referential integrity* constraints that need to be added to the model? If so, show them on the diagram

and describe them in your answer.

(b) Translate the E-R diagram into a relational database. Indicate all primary keys and all foreign keys in your relations.

2. Consider the following relational database describing a very large fictional castle:

Floors(Name CHAR(10) , Purpose CHAR(40))
Rooms(Name CHAR(20), Floor CHAR(10), Internal INT)
Connections(Room1 CHAR(20),Room2 CHAR(20))
Furnishings(ID INT, Name CHAR(80), Type CHAR(20), Room CHAR(20))
Inhabitant(Name CHAR(40), Room CHAR(20))

The castle consists of various floors, each of which has a stated purpose (such as "Guest Lodgings" or "Secret hideout"). Each floor contains a number of rooms. Rooms may be connected to each other, even if they are on different floors. The fact that room A is connected to Room B does not imply that Room B is connected to Room A. (Fictional castles follow fictional rules!) An *internal room*, marked by Rooms.Internal = 1, is a room without windows; an *external room*, marked by Rooms.Internal = 0, is a room with windows giving it an outside view. The database also keeps track of furnishings (such as pieces of art and furniture) and their locations in the castle.

The foreign keys in the databases are straightforward. For example, Rooms.Floor is a foreign key referencing Floors.

Write the following queries in SQL.

(a) Find all pairs of connected rooms located on different floors. Report the names of the rooms and the names of their respective floors.

(b) Find external rooms that are reachable from the room called "Old Duke's study" via exactly two intermediate rooms. (You have to go from Old Duke's study to room 1, then to room 2, and then to the room that you are reporting). Return the names of these two-step reachable rooms and their floors.

(c) The master of ceremonies of a dinner party wants to ensure there are enough chairs in the "Olde Grande Dining Hall". He needs

to find all the chairs in the rooms adjacent to this dining hall. Because the chairs need to be returned back to their original locations after the dinner, "adjacent" means "connected in both directions". Write a query that returns the list of chairs in rooms adjacent to this dining room; for each, report the name of the chair and its location.

(d) The custodian of the castle needs to clean all the rugs on the 49th floor of the castle. (Its floor name is "49th".) For each internal room on this floor that contains a rug (which is a type of a furnishing), find an external room (a room with a window from which this rug could be hung) that is separated by at most one room (room-with-rug $\rightarrow$ room1 $\rightarrow$ room-with-window) from the room with the rug. However, the path to the room with the window must not go through any inhabited room; it is impolite to carry a dirty rug to or through someone's bedroom. Report the room with the rug along with the external room where that rug can be hung from the window.

(e) For each floor, find the room that is connected to the largest number of other rooms *on the same floor*. Report floor names and room names.

(f) Find the floor with the largest number of inhabitants (that is, people assigned to rooms on this floor). Report the floor name and the number of inhabitants.

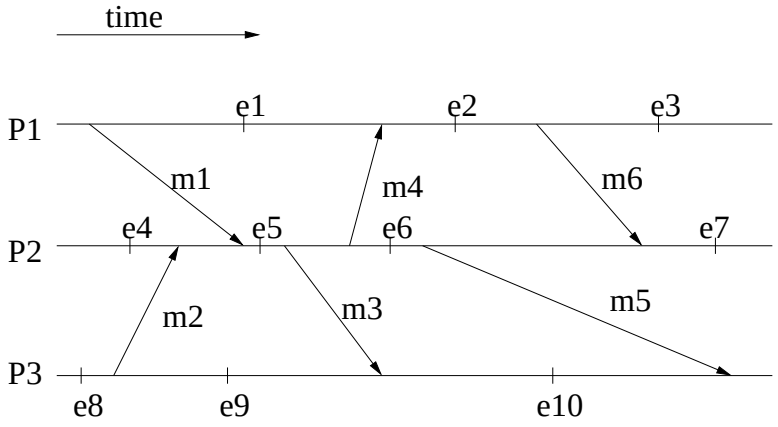
3 Distributed Operating Systems

1. **Lamport Clocks**

   Three processes P1, P2 and P3 communicate with each other as shown by messages m1, m2, m3, m4, m5 and m6 in the figure below. Additional internal events (e1, e2 and e3 for P1; e4, e5, e6 and e7 for P2; e8, e9 and e10 for P3) happen at these processes.

   (a) Trace an implementation of Lamport's logical clocks for these processes. Be specific about the time of each event, including the time at which each message is sent and received.

   (b) Do Lamport's clocks capture causality? Justify your answer using the example.



2. **Global State**

   (a) State the Chandy-Lamport Global State Recording Algorithm. What good is the global state recorded by this algorithm when the distributed computation might not have passed through this recorded global state?

   (b) Redesign the Chandy-Lamport algorithm for a distributed system where communication channels are non-FIFO. To simplify the design, assume that every process has a physical clock and physical clocks at all processes are perfectly synchronized.

4 Networks

1. **TCP**

   (a) Consider a gigabit ethernet network with exactly one client machine and one server machine (no other machines). Draw a graph that shows (approximately) what the congestion window size of the TCP connection is over time (the x axis is time, the y axis is the congestion window size) for the following workloads:

      i. The client downloads a huge file from the server.

      ii. The client downloads multiple tiny web pages (2KB), opening a new TCP connection for each web page.

   Annotate your graphs to explain what is going on.

   (b) If TCP loses an acknowledgment, is a retransmission guaranteed to occur? Explain.

2. **Routing protocols**

   This problem deals with routing protocols, in particular with the *metrics* associated with links in the network graph for use in selecting the "best" path.

   (a) What metric is used by the RIP protocol?

   (b) Why does OSPF allow multiple metrics to be associated with a link?

   (c) Explain some of the difficulties associated with metrics that attempt to characterize the current load on a link in order to achieve load balancing. (Instantaneous queue length is an example of such a metric, but it is not the only one.) Give at least one example of a specific problem that can arise when using such metrics.