National Exams December 2011
98-Comp-B3 – Data Base & File Systems

3 hours duration

Notes:

1. If doubt exists as to the interpretation of any question, the candidate is urged to submit with the answer paper a clear statement of any assumption made.
2. This is a Closed Book exam. A Casio or Sharp approved calculator is permitted.
3. Answer **five** questions as follow:
   a. **One** question from **questions 1 and 2** (only one question will be marked)
   b. **One** question from **questions 3 and 4** (only one question will be marked)
   c. **Three** questions from **questions 5, 6, 7, and 8** (only three questions will be marked)
4. All questions are of equal value. The marking scheme is as follows:

   Question 1: (a) 4 marks; (b) 4 marks; (c) 6 marks; (d) 6 marks

   Question 2: (a) [2 + 5 + 5 = 12 marks]; (b) [4 + 4 = 8 marks]

   Question 3: 20 marks

   Question 4: 20 marks

   Question 5: (a) 5 marks; (b) 7 marks; (c) 8 marks

   Question 6: (a) 3 marks; (b) 5 marks; (c) 7 marks; (d) 5 marks

   Question 7: (a) 7 marks; (b) 5 marks; (c) 4 marks; (d) 4 marks

   Question 8: (a) 6 marks; (b) 6 marks; (c) 8 marks
5. All answers should be clear, legible and brief.

Question 1

a. Explain the differences between RAID level 0 and level 1.

b. RAID levels 3 and 5 involve striping and storing parity information, but how is RAID level 5 differs from level 3?

c. Explain the rational for using fillfactors less than 1 in
   i. Sorted indices
   ii. B+ tree indices,
   iii. Hash indices.

d. Explain the difference between an equality search and a range search.

Question 2

a. Consider the B+ tree in Figure 1 below:
   i. Fill in the internal nodes without adding new keys.
   ii. Add key **bbb**. Show how the tree changes.
   iii. Delete the key **abc** from the result of (b). Show how the tree changes.

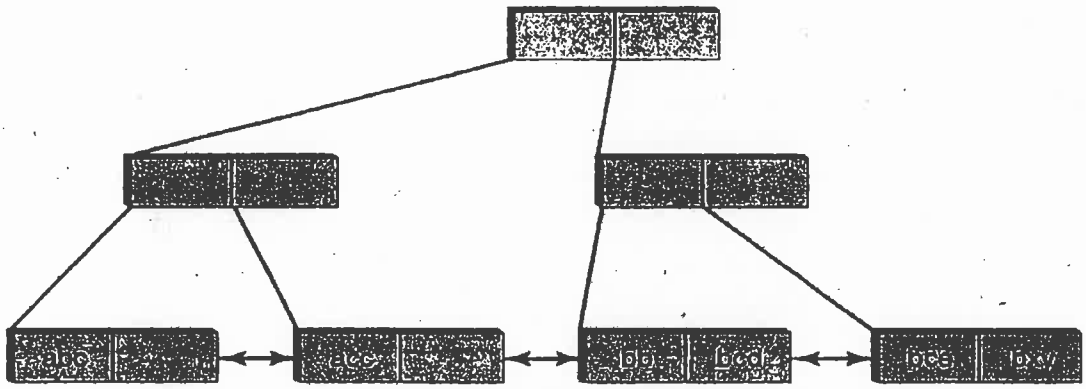

Figure 1

b. Consider the B+ tree in Figure 2. Suppose that it was obtained by inserting a key into a leaf node of some other tree, causing a node split. What were the original tree and the inserted key? Is the solution unique? Explain your answer.

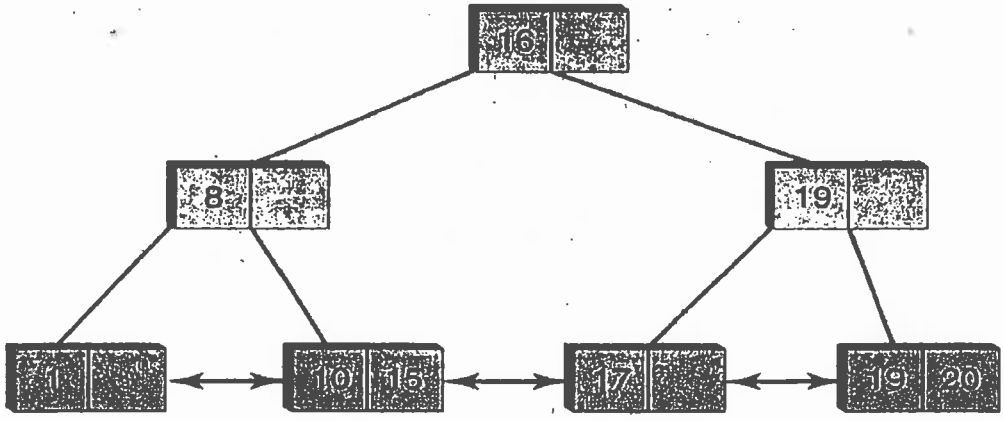

Figure 2

Question 3

Reservation system – A trip reservation consists of a sequence of flight reservations, where each flight reservation refers to a specific flight. Sometimes another flight is substituted for a booked flight because of equipment problems, weather delays, or customer preference. The passenger may reserve a seat for each flight. A trip reservation is made on some date; the passenger must purchase a ticket within a certain number of days or the reservation becomes void. The airlines use record locators to find a particular trip reservation quickly and unambiguously. A trip is reserved by an agent, who either works for an airline or a travel agency. The frequent flyer account may be noted for a passenger. The owner of the frequent flyer account must be the passenger. Multiple payments may be made for a trip, such as two credit-card charges. Payment may also be made by cash or check.

Design a conceptual schema for the trip reservation system using an ER diagram. Be sure to indicate all key and cardinality constraints and any assumptions you make. Identify any constraints you are unable to capture in the ER diagram and briefly explain why you could not express them.

Question 4

A real estate firm keeps track of the houses for sale and customers looking to buy houses. A house for sale can be *listed* with this firm or with a different one. Being, "listed" with a firm means that the house owner has a contract with an agent who works for that firm. Each house on the market has price, address, owner, and a list of features, such as the number of bedrooms, bathrooms, and type of heating, appliances, size of garage, and the like. This list can be different for different houses, and some features can be present in some houses but missing in others. Likewise, each customer has preferences that are expressed in the same terms (the number of bedrooms, bathrooms, etc.). Apart from these preferences, customers specify the price range of houses they are interested in. Perform conceptual design for this enterprise.

Draw an ER diagram for the real agent database. Be sure to indicate the various attributes of each entity and relationship set; also specify the key and participation constraints for each relationship set.

Question 5

Consider the relations below:

Passengers (*Name, Address, Age*)
Reservations (*Name, FlightNum, Seat*)
Flights (*FlightNum, DepartCity, DestinationCity, DepartureTime, ArrivalTime, MinutesLate*)

Express each of the following queries in SQL:

a. Get the names of passengers who had a reservation on a flight that was more than 30 minutes late.
b. Get the names of passengers who had reservations on **all** flights that were more than 60 minutes late.
c. Get the names of pairs of passengers who are the same age.

Question 6

Consider the following relations containing airline flight information:

Flights(*flno:* **integer**, *from:* **string**, *to:* **string**, *distance:* **integer**, *departs:* **time**, *arrive:* **time**)
Aircraft(*aid:* **integer**, *aname:* **string**, *cruisingrange:* **integer**)
Certified(*eid:* **integer**, *aid:* **integer**)
Employees(*eid:* **integer**, *ename:* **string**, *salary:* **integer**)

The key attributes are underlined. The Flights relation lists flight number, origin, destination, distance, and departure and arrival times. The Aircraft relation shows aircraft id, name and cruising range. The Certified relation gives employee id, and aircraft id. The Employees relation describes pilots (and other kinds of employees); every pilot is certified for some aircraft (otherwise he or she would not qualify as a pilot). Note that only pilots are certified to fly.

Now, write the following queries in **relational algebra**:

a. Find the *eids* of pilots certified for some Bombardier aircraft.
b. Find the names of pilots certified for some Airbus aircraft.
c. Find the *aids* of all aircraft that can be used on non-stop flights from Ottawa to Boston.
d. Identify the flights that can be piloted by every pilot whose salary is more than $160,000.00 Canadian.

Question 7

Consider a relation schema with attributes ABCGWXYZ and the set of dependencies:

$$\mathbf{F} = \{XZ \rightarrow ZYB,\ YA \rightarrow CG,\ C \rightarrow W,\ B \rightarrow G,\ XZ \rightarrow G\}.$$

Solve the following problems using the appropriate algorithms.

a. What is the minimal cover?
b. Is the dependency XZA → YB implied by **F** and why?
c. Is the decomposition into XZYAB and YABCGW lossless and why?
d. Is the decomposition in (c) above dependency preserving and why?

Question 8

a. Given below is a schedule that is produced by a non-strict two-phase locking concurrency control:

$\mathbf{w_1(x)\ r_2(x)\ w_2(x)\ commit_2\ commit_1}$

(i) Is the schedule serializable and why?
(ii) Is the schedule in commit order and why?

b. Give an example of a schedule at the READ COMMITTED isolation level in which a lost update occurs. Explain your example.

c. What happens to the schedule below at a REPEATABLE READ isolation level? Explain your answer.

$\mathbf{r_1(x)\ r_1(y)\ w_1(x)\ r_2(y)\ r_2(x)\ w_1(y)\ commit_2\ commit_1}$