Student ID _____ Student Name _____ Total Mark _____

CMPUT 291 B2 Tamer Ozsu
File and Data Management
Final Exam
18 April 2000

1. Exam duration: 120 minutes.
2. Closed book, and no collaboration. No questions during the exam; if you are unsure, write down your assumptions.
3. All questions need to be answered.
4. Marked exam will be available on 25th of April between 1:00 and 4:00 PM. Closing date for appeals is 25th April 2000 at 4:00PM.
5. **Answer all questions on this sheet.**

Question 1 (20 pts) Answer the following questions with a few sentences (no longer than the allocated space).

(a) What are the three steps of the database design (data modeling) process? Define each by ***one*** sentence.

(b) What types of participation constraints can you have in an E-R model? Define each by ***one*** sentence.

(c) Given relation R(A, B, C, D, E) where (A,B) is the key, and the functional dependencies (A, B) → (C, D, E) and B → D, is R in Third Normal Form (3NF)? Justify your answer with ***one*** sentence.

Student ID _____ Student Name _____

(d) What is the main difference between relational calculus and relational algebra? Which one do DBMSs use ***internally***?

(e) What is the difference between logical data independence and physical data independence?

(f) What is the formula that determines the cost of reading from disk? Define the components of the formula.

(g) What is referential integrity? How do you represent it in relational model?

(h) What is the main difference of B-trees and B^+-trees? What effect does this difference have on the index search performance?

Student ID _____ Student Name _____

(i) What are the ways in which you could separate and identify the ***variable length*** fields of a record?

(j) Define internal fragmentation and external fragmentation.

Question 2 (15 pts; each 3 points; NO PARTIAL MARKS) You are designing a database for Edmonton SPCA (Society for the Prevention of Cruelty to Animals). The result is the following set of relations where the type of each relation's attribute is given following the attribute (e.g., ID: integer):

Animals(ID: integer, Name: string, Owner: integer, DateAdmitted: date, Type:string)

Adopter(SIN: integer, Name: string, Address: string, OtherAnimals: integer)

Adoption(AnimalID, SIN, AdoptDate: date, chipNo: integer)

where

(a) Animals stores information about the animals currently at the SPCA. Each is given an ID, and their names together with with the SIN of their previous owners (attribute Owner), and their date of admission is recorded. Type refers to the type of animal (dog, cat, etc)

(b) Adopter is the relation that holds information about animal adopters. The attributes are self-descriptive, except OtherAnimals which records the number of other animals that the adopter currently has at home.

(c) AnimalID in Adoption refers to the ID of Animals and has the same type. Similarly, SIN in Adoption refers to the SIN of Adopter and has the same type. Attribute chipNo stores the number on the microchip that is implanted on the animal for tracking.

Formulate the following queries in SQL:

Student ID _____ Student Name _____

(a) Retrieve the total number of dogs that were brought to the SPCA on 18 April 2000.

(b) List the name of the adopter who has adopted every type of animal.

(c) For each animal type, list the animal type and total number of adoptions on 14 June 1999.

(d) List the types of animals who have not had any adoptions.

(e) For each adopter who has made at least two adoptions, list their names and addresses.

Question 3 (9 pts; 3 pts each) Write the following queries in relational algebra.

(a) Find the chip number of the dog whose name is "Shina".

(b) Find the name of adopter who has adopted a dog Named 'Shina' on 14 May 1996.

(c) Find the names of cats who were brought to the SPCA on 18 April 2000.

Question 4 (10 pts) Give the SQL DDL statements to declare the four relations of Question 2. In addition to the data type and key specifications given in that question, the following have to be specified:

(a) All the referential integrity constraints have to be specified.

(b) No one can adopt more than three animals.

(c) If someone has left an animal at the SPCA before, he/she cannot adopt an animal after that date.

Student ID _____ Student Name _____

Student ID _____ Student Name _____

Question 5 (16 pts) Consider the following partial ER model for a publishing company. Answer the following questions about it, showing all the changes on the model itself.

(a) Books consist of many chapters, each of which can appear in only one book.

(b) Authors write chapters (many of them) and each chapter can have many authors.

(c) Authors live in houses; since they are generally poor, a number of them can share a house.

(d) Houses can be either condos or detached houses. For condos, we keep track of the condo fee and for houses we store the remaining mortgage.

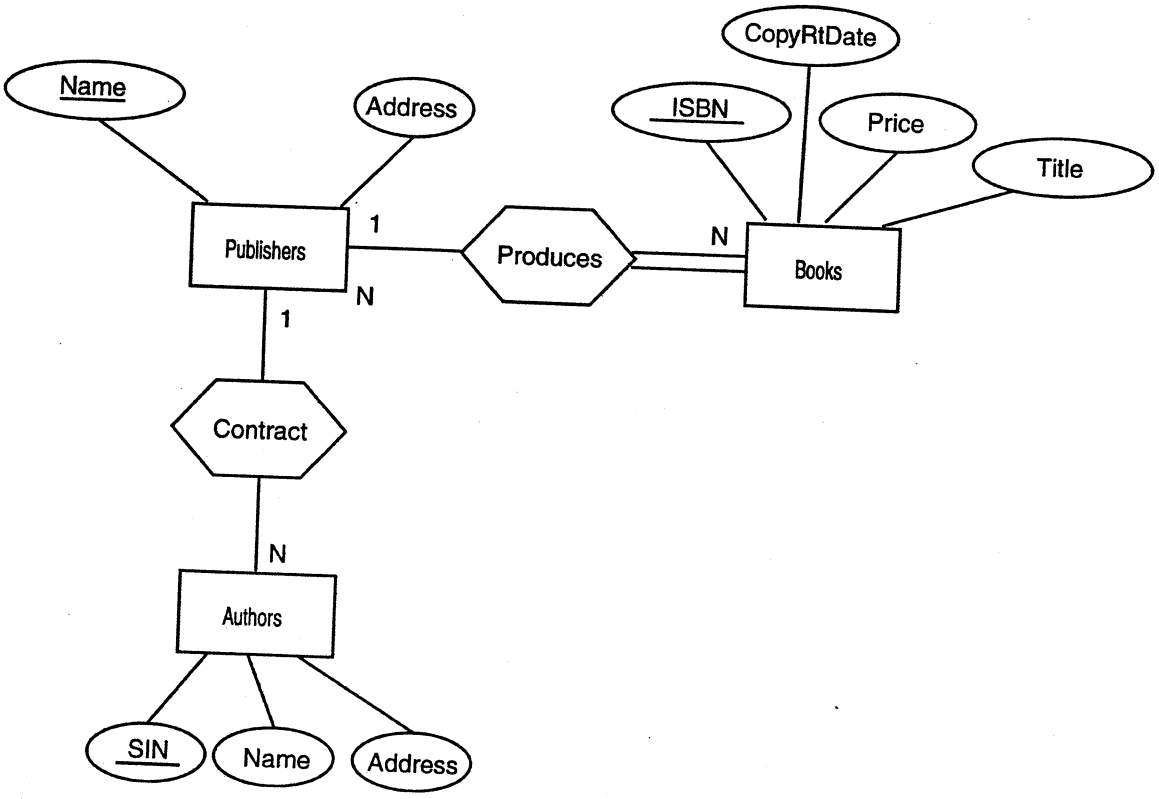

Figure 1. Figure for Question 5

Student ID _____ Student Name _____

Question 6 (30 pts). Consider the B tree index shown in Figure 2. All promotions are from the left. For each case below, you need to draw a separate tree.

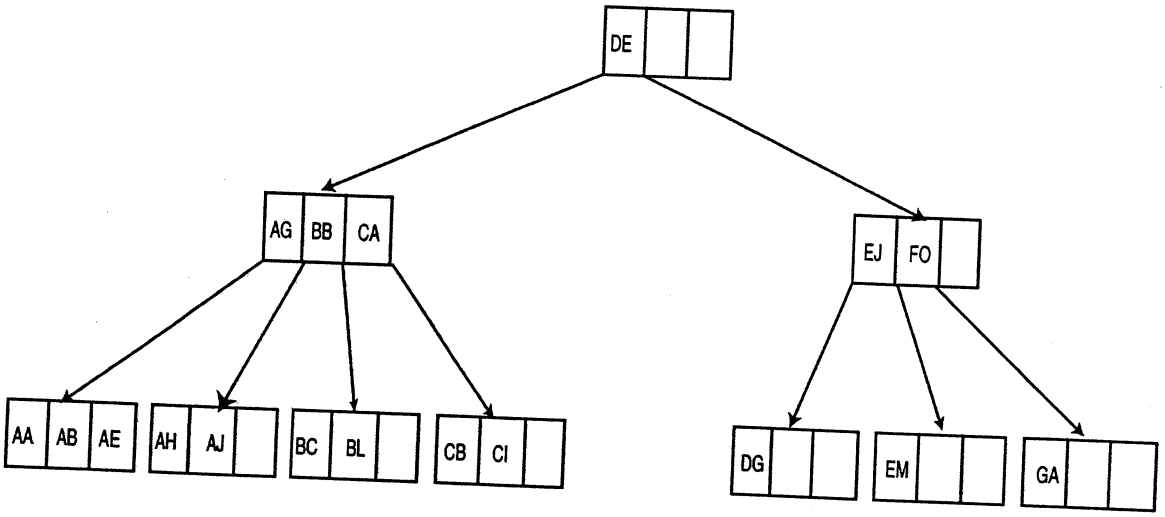

Figure 2. Figure for Question 6

(a) Show the tree that would result from inserting a data entry with key AI into this tree.

(b) Show the B tree that would result from inserting a data entry with key AC into the *original* tree.

(c) Show the B tree that would result from deleting the data entry with key AH from the *original* tree, assuming that the left sibling is checked for possible re-distribution; concatenation, if needed, can be done with either left or right sibling.

(d) Show the B tree that would result from starting with the *original* tree, inserting the data entry DA and then deleting the data entry with key DG, assuming that the left sibling is checked for possible re-distribution; concatenation, if needed, can be done with either left or right sibling.

(e) Show the B tree that would result from successively deleting the entries with keys CB, CI, DG, GA, and FO from the original tree, assuming that the left sibling is checked for possible re-distribution; concatenation, if needed, can be done with either left or right sibling.

Student ID _____ Student Name _____