CSI2132 Assignment 3 2006

Total of 45 marks

Instructions
This is an individual assignment. Submit via BlackBoard Learn. Completion is optional, but it is (of course) very important to go through the solutions in order to practice for the final examination.

A. Relational Algebra

Consider the following extended schema about Kids (Campers) that register for Camps in Adventures and Science in Engineering at uOttawa.

$Camper(\underline{CName}$: `string`, Age : `string`, $Email$: `string`, $tshirt$: `string`, Fee : `real`$)$
$Camp(\underline{CampID}$: `int`, $CampTitle$: `string`, $EmpID$: `int`, $StartDate$: `date`, $Year$: `date`$)$
$Signup(\underline{CName}$: `string`, $\underline{CampID}$: `string`$)$
$Mentor(\underline{EmpID}$: `int`, $Name$: `string`, $EmploymentDate$: `date`, $Salary$: `currency`$)$

1. Provide the relational algebra statement to find the names, fees and emails of all Campers who attended both the GirlScience and the GirlTech camps in 2016 (3 marks).

2. Provide the relational algebra statement to find the names of the mentors who ran the MakerTech Camps for Campers that where older than 10 years old. Display their names together with the start date and year of the Camps they attended (3 marks).

3. Provide the relational algebra statement to find the names, t-shirt sizes and emails of the 7 year old Campers who never attended a Camp that was led by a Mentor named Sandy who earns a salary of $500.00 (4 marks).

B. Normal Forms

Consider a Book relation denoted by R with attributes $ABCD$. A sample instance of R is shown below:

BookTitle (A)	Distributor (B)	Topic (C)	Price (D)
Java Solutions	Prentice Hall	Programming	119.00
Topology	Independent	Math	120.00
Database Design	Addison W	Databases	115.00
C++	Wiley	Programming	119.00
C++	Addison W	Programming	129.00

1. This relation may have some data anomalies. Explain what the three anomalies are and give an example of each, using the Book relation (2*3 = 6 marks).

2. Suppose that the following sets of functional dependencies hold over R. For each one of the following sets of functional dependencies, determine the highest normal form and motivate your answer (2*3= 6 marks).

   (a) C $\longrightarrow$ D, C $\longrightarrow$ A, B $\longrightarrow$ C

   (b) A $\longrightarrow$ B, BC $\longrightarrow$ D, A $\longrightarrow$ C

   (c) AB $\longrightarrow$ C, AB $\longrightarrow$ D, C $\longrightarrow$ A, D $\longrightarrow$ B

C. Physical database design

Reconsider the following relational schema about Kids (Campers) that register for Camps in Adventures in Science and Engineering.

$Camper(\underline{CName}$: `string`, Age : `string`, $Email$: `string`, $tshirt$: `string`, Fee : `real`$)$
$Camp(\underline{CampID}$: `int`, $CampTitle$: `string`, $EmpID$: `int`, $StartDate$: `date`, $Year$: `date`$)$
$Signup(\underline{CName}$: `string`, $\underline{CampID}$: `string`$)$
$Mentor(\underline{EmpID}$: `int`, $Name$: `string`, $EmploymentDate$: `date`, $Salary$: `currency`$)$

Suppose that the Camper table is organized as a heap file, and that it contains the records of a total of 20,000 kids (i.e. current and past Campers). A disk block has the capacity to store 1,000 records and the buffer pool contains 10 slots. On average, a Camper registers for two Camps in a year and attends the Camps for four years in a row.

1. Explain how you would use any *one* of the two different heap file implementations, as discussed in class, in order to organize the pages of the Camper table on disk (5 marks).

2. Assume that you wish to execute a query that displays all the personal information about the Campers in your database (i.e. `SELECT * FROM Camper`). Explain the exact process that is followed i) to locate the data on disk, ii) to transfer the data into the buffer and iii) to deal with potential buffer sizing issue (8 marks).

D. Storage and Indexing

Consider the following table that contains the information about the prices (in $) and ratings of products sold in a convenience store.

ProductID	Type	Price	Rating
120	Gum	1	5
121	Camera	50	2
122	Candy	8	4
123	7Up 24 Box	18	5
124	Kitkat	2	3
125	Coke 24 Box	18	5
126	Nachos	6	4
127	Hat	73	5
128	Jacket	94	5
129	Boots	99	4
130	Backpack	27	3
131	Camera	52	5
132	Walking Stick	58	2
133	Aero	5	4
134	Cheese	14	2
135	Carpet	89	4

Assume that the DBMS constructs a B+ tree with an order d of 2 on the Price attribute. You decide to store the data entries using Alternative 3.

1. Show the final tree after you have inserted **all** the data in the order they are shown in the table. That is, do not use the bulk loading algorithm (5 marks).

2. Suppose that the products with the following ProductIDs are removed from the tree, in this order: 131, 128, 127 and 129. Show the resultant tree after you have deleted these four entries (5 marks).