University of Ottawa
CSI 2132 and 2532 – Final Examination
Professor(s): Herna L. Viktor and Iluju Kiringa

17 April 2014
09:h30-12h30
Duration: 3 hrs

Closed book; no aid allowed, except one double-sided letter-size "cheat sheet". Answer all questions in **ink**. **Good luck! / Bonne Chance!**

Family name: _____

First name: _____

Student number: _____

There are 7 questions and a total of 100 points.

This exam must contain 12 pages,
including this cover page.

1 – EER Diagram	/ 15
2 – The Relational Model and SQL	/ 15
3 – Relational Algebra and Relational Calculus	/ 10
4 – Normal Forms	/ 15
5 – Physical database design	/ 15
6 – Storage and Indexing: B+ trees	/ 15
7 – Storage and Indexing: Extendible Hashing	/ 15
Total	/ 100

1 EER Diagram — 15 points

Consider the following description of the GoGoGreen electric scooter rental company that is based in Ontario. The headquaters is in Ottawa and GoGoGreen currently has ten branches throughout the province.

- Each branch has an inventory of electric scooters, all of which are recorded in a catalog.
- For each scooter, we record the unique registration number, the make (e.g. Suzuki), the model (e.g. Dolce Vita), the year of manufacturing (e.g. 2015), the rental price per hour (e.g. $12.00 per hour), the main color (e.g. red) and the battery duration (e.g. 100 hours).
- A branch may have many scooters of the same make and model, but the registrations number is unique throughout the company.
- Before renting a scooter from the company, a customer must first register as a member of a local branch. The data held on a member include her membership number, name, address, phone number and the date she first registered at the branch. We also record credit card details.
- Once registered, a member is free to rent a scooter for a duration from (minimum) 1 hour to (maximum) 48 hours.
- The data held on each scooter rental is the rental contract number, the date, the membership number, the registration number of the scooter, the daily rental price, as well as the time rented out and returned. A status attribute is used to indicate whether a scooter is available or not.
- For each branch we record the branch number, name, address and phone number.
- Each branch has a manager, who is responsible for the day to day running of the branch. A branch also has at least two additional staff members, who are full-time employees. For each staff member, we record her name, address and social insurance number.

Draw an EER diagram for the GoGoGreen database, using the description given above. Show **all your assumptions** clearly.

Draw your EER diagram here.

... Cont'd

2 The Relational Model and SQL — 15 points

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

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

Answer the following questions.

Part A — 2 points Explain what a secondary key is and identify one in the above database schema.

Part B — 6 points Give the SQL statement to find the names and employment dates of the Mentors who earn less than any Mentor named Randy. Display this information together with the titles of the Camps these unfortunate Mentors ran (were responsible for), as well as the years they were Mentors.

... Cont'd

Part C — 4 points Give the SQL statement to display the names, fees paid and ages of the youngest Camper(s) for all of those Camps that had an enrollment that was higher than 20. That is, these popular Camps were attended by at least 21 Campers.

Part D — 3 points Campers have to be at least 6 years old to enroll in a Camp. Provide the SQL code fragment to enforce this constraint.

... Cont'd

3 Relational Algebra and Relational Calculus — 10 points

Reconsider the following extended relational schema about Kids (Campers) that register for Camps in Adventures in Science in 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`, $\underline{Name}$: `string`$)$
$Mentor(\underline{EmpID}$: `int`, $Name$: `string`, $EmploymentDate$: `date`, $Salary$: `currency`$)$

Part A — 7 points Provide the relational algebra statement to find the names, tshirt sizes and emails of the 10 year old Campers who never attended a Camp that was led by a Mentor named Sandy who earns a salary of $500.00.

Part B — 1 point State whether the following statement is **true** or **false**. Relational algebra is a declarative language, since we specify exactly what steps need ot be taken when a query is executed.

Part C — 2 points Explain the difference between tuple relational calculus and domain relational calculus.

4 Normal Forms — 15 points

... Cont'd

5 Physical database design — 15 points

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

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

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.

Part A — 5 points Explain how you would use *one* of the two different heap file implementations, as discussed in class, in order to organize the pages of the Camper table on disk.

... Cont'd

Part B — 10 points 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.

... Cont'd

6 Storage and Indexing: B+ trees — 15 points

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.

Part A — 10 points Show the final tree after you have inserted **all** the data using the bulk loading algorithm.

... Cont'd

Part B — 5 points 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.

... Cont'd

7 Storage and Indexing: Extendible Hashing — 15 points

... End of Final Examination