CMPT 354
Database Systems

Simon Fraser University
Summer 2016

Instructor: Oliver Schulte

Assignment 4: System Implementation, XML.

Instructions: Check the instructions in the syllabus. The university policy on academic dishonesty and plagiarism (cheating) will be taken very seriously in this course. *Everything submitted should be your own writing or coding.* You must not let other students copy your work. Discussions of the assignment is okay, for example to understand the concepts involved. If you work in a group, put down the name of all members of your group. On your assignment, put down your **name**, the number of the assignment and the number of the course. Spelling and grammar count.

For the due date please see our course management server https://courses.cs.sfu.ca .

Storage and Indexing. 40 points total.

Consider the following instance of the Students relation, sorted by *gpa*.

sid	name	age	gpa
53831	Madayan	11	1.8
53832	Guldu	12	2.0
53688	Smith	19	3.2
53666	Smith	18	3.4
53650	Jones	19	3.8

For the purposes of this question, assume that these tuples are stored in a sorted file in the order shown. Each page can store up to three data records. So the first three tuples are on page 1, the fourth is on page 2 etc.

I. (20) Explain what the data entries in each of the following indexes contain. If the order of entries is significant, say so and explain why. For definition of terms, refer to the text.
    1. An index on *gpa* using Alternative (1).
    2. A clustered index on *gpa* using Alternative (2).
    3. An unclustered index on *name* using Alternative (3).

II. (20) Consider a delete operation specified using an equality condition on a key. Assuming that no record qualifies, what is the cost for the three file organizations: heap file, sorted file, unclustered hash index? Present your analysis using the same parameters as in the lectures:

    B = number of data pages
    R = number of records per page
    D = time to read or write disk page

Query Evaluation. 28 Points total.

I. (10) Consider the following SQL query.

```
SELECT S.sname
FROM Sailors S, Reserves R, Boats B
WHERE S.sid = R.sid AND R.bid = B.bid AND B.colour = 'green'
INTERSECT
SELECT S.sname
FROM Sailors S, Reserves R, Boats B
WHERE S.sid = R.sid AND R.bid = B.bid AND B.colour = 'blue'
```

Write a query evaluation plan (relational algebra tree) for evaluating this query. (You do not have to annotate the nodes with access methods, just specify the relational operator for each node.)

II. (18) Consider the following schema with the Sailors relation:

Sailors(sid: integer, sname: string, rating: integer, age: real)

For each of the following indexes, list whether the index matches the given selection conditions. If there is a match, list the primary conjuncts.

(a) A hash index on the search key <Sailors.sid, Sailors.rating>

- $\sigma_{rating = 10 \text{ AND } sid=500}$ (Sailors)
- $\sigma_{rating > 10 \text{ AND } sid=500}$ (Sailors)

(b) A B+-tree on the search key <Sailors.sid, Sailors.rating>

- $\sigma_{sid< 500 \text{ AND } rating=10}$ (Sailors)
- $\sigma_{sid= 500 \text{ AND } rating > 10}$ (Sailors)
- $\sigma_{sid< 500}$ (Sailors)
- $\sigma_{rating > 10}$ (Sailors).

XPath Questions. 18 Points total.

Refer to: http://www.w3schools.com/xsl/xpath_examples.asp
For the below questions, use the XML document from the above page, which contains an XML structure for a bookstore. Please test your queries to ensure that your XPath queries work. The following website seems to evaluate Xpath expressions correctly http://www.online-toolz.com/tools/xpath-editor.php . Write Xpath expressions for the following queries.

(a) Select all books published after 2003.

(b) Display the language of the third book.

(c) Display only the book titles that are categorized for cooking.

Final Exam Question. 10 points total.

Design a question for the final exam.

The purpose is to start you thinking about the course material for the final exam. I will put the best question on the final. This will basically be graded on a pass/fail basis, with some higher points for special creativity, and lower points for lack of effort.

What to Submit

X-Path Question: Please submit a pdf showing a screenshot with each of your queries and the result of the query. Call this xpath.pdf .

All other questions: Write out the answers in a word processing program and submit a pdf with them. Call this main.pdf .

Please submit both pdfs on courses.cs.sfu.ca.

A note on the relational algebra tree (query evaluation question): We'd prefer you use a drawing program, but we'll accept hand-drawn trees as long as they are easy to read after scanning. Remember that by using a drawing program, you make it easy to modify your answer as you get better and better ideas.