Final Exam Practice Problems
Spring 2022

This set of practice problems is released as a makeup assignment. Problems carry different mark.

Consider the following database schema for practice problems 1, 2, 9 & 10.

- `Person (`<u>`pid`</u>`, pname, city, birthYear)`
- `Knows (`<u>`pid1, pid2`</u>`)`
- `Company (`<u>`cname, city`</u>`)`
- `Works (`<u>`pid`</u>`, cname, salary)`
- `JobSkill (`<u>`skill`</u>`)`
- `PersonSkill(`<u>`pid, skill`</u>`)`

Problem 1.   Using the SQL aggregate function(s), find the pid and pname of each person who lives in 'Chicago' and who knows at least one person who has at least 3 job skills. **(5 marks)**

Problem 2.   Use the method of Venn diagrams to formulate the following query in pure SQL. You need to use views and parameterized views to specify the relevant sets. **(10 marks)**

    a. Find the pid and pname of each person who knows all the persons who (a) work at Apple, (b) make at most 60000, and (c) are born before 2000.
    b. Find the pairs (p1, p2) of different person pids such that the person with pid p1 and the person with pid p2 knows the same number of persons.

Problem 3.   Let A(x) be the relation schema for a set of positive integers. (The domain of x is INTEGER.) Write a SQL statement that produces a table which, for each $x \in A$, lists the tuple $(x, x^{1/3}, x^y, 10^x, x!, \log_2 x)$. **(5 marks)**

Problem 4.   Consider two relations R(A, B) and S(B, C), two constant a and c, and a view with the following definition: **(5 marks)**

```
SELECT r.A, s.C
FROM R r, S s
WHERE r.A != a AND r.B = s.B AND s.C != c
```

Suppose we want to maintain this view as a materialized view under the insertion of tuples in R and/or S. Define triggers and their associated trigger functions on the relations

R and S that implement this materialized view. Make sure that your trigger functions act in an incremental fashion and that no duplicates appear in the materialized view.

Problem 5. Consider the relations R(A, B), S(B,C), and T(C, D). Assume that R, S, and T are stored in B(R), B(S), and B(T) blocks, respectively. Furthermore, assume that you have a buffer of (approximate) size M **(15 marks)**

a. Assuming that you use the block nested-loop join algorithm to implement natural join operations, specify the time complexity to evaluate the relational algebra expression $(R \bowtie S) \bowtie T$. You can make the assumption that $B(R \bowtie S) \leq M^2$, where $B(R \bowtie S)$ is the number of blocks to store $(R \bowtie S)$

b. Assuming that you use the sort-merge join algorithm to implement natural join operations, specify the time complexity to evaluate the relational algebra expression $(R \bowtie S) \bowtie T$. You can make the assumption that $B(R \bowtie S) \leq M^2$, where $B(R \bowtie S)$ is the number of blocks to store $(R \bowtie S)$

Problem 6. Write a PL/pgSQL function that takes an integer as an input and returns true if the integer is a prime number, false otherwise **(5 marks)**

Problem 7. Write a PL/pgSQL function that sorts a given array using selection sort **(5 marks)**

Problem 8. Consider a relation PC(P, C) which indicates that person P is a parent of person C. Furthermore, assume that there are two unary relations Male(P) and Female(P) that specify the gender of a person P. **(15 marks)**

a. Write a program that defines the predicate Ancestor(x, y) which specifies that x is an ancestor of y.

b. Write a program that defines the predicate Ancestor_Male_Female(x, y, z) which specifies that x is an ancestor of a male descendant y, and y is an ancestor of a female descendant z

Problem 9. Specify an Entity-Relationship Diagram that models the database schema provided with this assignment. **(5 marks)**

Problem 10. Specify the node and relationship types of a Property Graph for this database schema. In addition, specify the properties, if any, associated with each such type **(5 marks)**

Problem 11. Using the Property Graph model in the previous problem, formulate the following queries in the Cypher query language **(5 marks)**

a. Find the types of the relationships associated with Person nodes.

b. Find each person (node) whose name is 'John' and has a salary that is at least 50000.

Problem 12.   Consider the following B+-tree of order n=2 that indexes records, with keys 3, 8, 10, and 11 **(5 marks)**

```
 -----------
| / 10 \ |
 -----------
    /\
    /\
--------- ------------
| 3 8 |--> | 10 11 |
--------- ------------
```
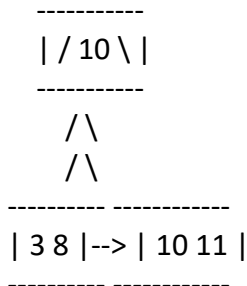
Show the contents of your B+-tree index after inserting records with keys 0, 7, 14, and 9 in that order.

Problem 13.   Let x, y, and z be data objects. State which of the following schedules are conflict-serializable or not conflict-serializable, and for each schedule that is serializable, give a serial schedule with which that schedule is conflict-equivalent **(15 marks)**
   a. R1(x); R2(y); R1(z); R2(x); R1(y)
   b. R1(x); W2(y); R1(z); R3(z); W2(x); R1(y)
   c. R1(z); W2(x); R2(z); R2(y); W1(x); W3(z); W1(y); R3(x)