Fall 2023 B561 Assignment 7

Muazzam Siddiqui, Radhika Agarwal, Pavan Kalyan Thota

Released: 1 December, 2023

Due: 8 December, 2023

Section 1 - Graded Questions:

1. 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.

   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 **(12.5 marks)**

2. Formulate the following query using the Venn diagram without counting condition. Create function or views to represent the sets used in the query.

   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. **(12.5 marks)**

3. Formulate the following query in object relational SQL queries. Use the relations, views, set operations and set predicates defined in assignment 5 with the same restrictions, i.e. You cannot use the Knows, companyLocation, and personSkill relations.

   Find the pid of each person who has the most job skills **(12.5 marks)**

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

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

   Write a trigger that maintains the number of tuples in this view. **(12.5 marks)**

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

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 $(R \bowtie S)$ **(12.5 marks)**

6. Suppose that we have an ordered file with r = 300,000 records stored on a disk with block size B = 4,096 bytes. The length of the record is 100 bytes.

a) Compute the number of block accesses required to search for a record. **(6.25 marks)**

b) Now suppose that the ordering key field of the file is V = 9 bytes long, a block pointer is P = 6 bytes long, and we have constructed a primary index for the file. Compute the number of block accesses required to search for a record using the primary index. **(6.25 marks)**

7. 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

a) R1(x); R2(y); R1(z); R2(x); R1(y) **(6.25 marks)**

b) R1(x); W2(y); R1(z); R3(z); W2(x); R1(y) **(6.25 marks)**

8. **Scenario:** A hospital manages patients, doctors, and nurses. Each patient has a unique patient ID, a name, and a date of birth. Each doctor has a unique doctor ID, a name, and a specialty. Each nurse has a unique nurse ID, a name, and a department.

Patients can be assigned to doctors for care. Each doctor can care for multiple patients. Nurses can be assigned to patients for care. Each nurse can care for multiple patients.

a) Create an Entity Relationship Diagram for the above scenario. **(6.25 marks)**

b) Convert the above entity relationship diagram to schema. **(6.25 marks)**

Section 2 - Practice Questions:

1. Write a PL/pgSQL function that takes an integer as an input and returns the sum of first n prime numbers.

2. Write a PL/pgSQL function that sorts a given array using selection sort

3. Formulate the following query using the Venn diagram without counting condition. Create function or views to represent the sets used in the query.

   Find the cname of each company who only employs persons who make less than 50000.

4. Formulate the following query in object relational SQL query. Use the relations, views, set operations and set predicates defined in assignment 6 with the same restrictions, i.e. You cannot use the Knows, companyLocation, and personSkill relations.

   Find, for each person, that person's pid and name along with the number of persons he or she manages.

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

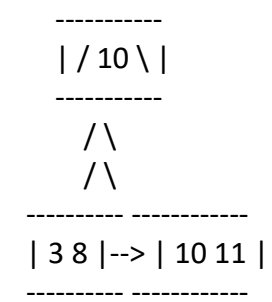
   Assuming that you use the sort-merge join algorithm to implement natural join operations, specify the time complexity to evaluate the relational algebra expression $(R \bowtie S) \bowtie T$. You can make the assumption that $B(R \bowtie S) \leq M^2$, where $B(R \bowtie S)$ is the number of blocks to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

6. Suppose that we have an ordered file with r = 300,000 records stored on a disk with block size B = 4,096 bytes. The length of the record is 100 bytes.

   a) Suppose we want to search for a record with a specific value for a secondary key—a nonordering key field of the file that is V = 9 bytes long. Compute the number of block accesses required to search for a record using the secondary key.

   b) Suppose that we construct a secondary index on that nonordering key field of the file. Compute the number of block accesses required to search for a record using the secondary index.

7. Consider the following B+-tree of order n=2 that indexes records, with keys 3, 8, 10, and 11



Show the contents of your B+-tree index after inserting records with keys 0, 7, 14, and 9 in that order.

8. Let x, y, and z be data objects. State if the following schedule is conflict-serializable or not conflict-serializable, and if this schedule is serializable, give a serial schedule with which that schedule is conflict-equivalent

R1(z); W2(x); R2(z); R2(y); W1(x); W3(z); W1(y); R3(x)