CS2102 Tutorial 2: RA & SQL 1 Wk 4, Sem 2, 2021/22

Questions 1a, 2, and 3 will be discussed.

1. This question considers a binary relational algebra operator called the **division operator** denoted by /.

   Consider two relations R and S where the set of attributes in the schema of R and S are $\{A_1, \cdots, A_m, B_1, \cdots, B_n\}$ and $\{B_1, \cdots, B_n\}$, respectively, where $m \geq 1$ and $n \geq 1$. That is, the set of attributes in S is a proper subset of the set of attributes in R. Assume that the attributes that are in R but not in S are ordered as $(A_1, \cdots, A_m)$ in the schema of R. Let L denote the list of attributes in the schema of R. The division of R by S (denoted by R/S) computes the largest set of tuples $Q \subseteq \pi_{A_1,\cdots,A_m}(R)$ such that for every tuple $(a_1, \cdots, a_m) \in Q$,

   $$\pi_L(\{(a_1, \cdots, a_m)\} \times S) \subseteq R$$

   Q is also referred to as the **quotient** of R/S, and its schema is $(A_1, \cdots, A_m)$.

   The following example illustrates R/S and T/U.

R

A	B
a	1
a	2
b	1
c	1
c	2
c	3
d	2
d	3

S

B
1
2

R / S

A
a
c

T

A	B	D	C
a	100	x	2
a	100	x	4
a	200	y	2
a	500	z	4
b	100	x	1
b	500	z	1
c	100	x	7
e	500	z	0

U

C	A
1	b
2	a
4	a

T / U

B	D
100	x

   (a) Consider again Question 5 in Tutorial 1 to find the restaurants that sell all the pizzas that Maggie likes and don't sell any pizza that Ralph likes. Write a relational algebra expression for this query that uses the division and natural join operators.

   (b) Given relations R(A,B) and S(B), write a relational algebra expression to compute the division of R by S using only the basic relational operators (i.e., σ, π, ρ, $\times$, $\cap$, $\cup$, $-$).

2. Consider a relational database for a company that consists of the following two tables.

   - Offices (office_id, building, level, room_number, area)
   - Employees (emp_id, name, office_id, manager_id)

   The database satisfies the following constraints:

   - `Offices` stores information about the office rooms in the company. Each room has a unique identifier `office_id` (which is the primary key of `Offices`) and information on its building name, floor level, room number and floor area.
   - {`building`, `level`, `room_number`} is a candidate key of `Offices`.
   - `Employees` stores information about the employees in the company.
   - `emp_id` is the primary key of `Employees`.
   - The name of each employee must be a non-null value.
   - Each employee must be assigned to exactly one office identified by `office_id`.
   - Each employee may be managed by at most one manager.
   - If an employee is managed by someone, the `emp_id` of his/her manager is recorded in `manager_id`.
   - A record in `Offices` cannot be removed if there's some employee assigned to that office.
   - A manager in `Employees` cannot be removed if there's some other employee managed by that manager.
   - Any modification to `office_id` in `Offices` is propagated to other database records.
   - Any modification to `emp_id` in `Employees` is propagated to other database records.

   Write SQL statements to create the database schema with appropriate attribute domains and constraints.

3. Consider the following relational database for an online book seller:

   - Books (isbn, title, authors, year, edition, publisher, number_pages, price)
   - Customers (cust_id, name, email)
   - Carts (cust_id, isbn)
   - Orders (order_id, order_date, cust_id)
   - Order_items (order_id, isbn)

   The database satisfies the following constraints:

   - `Books` record information about the books available for sale in an online shop. Each book has a unique identifier `isbn` and information on its title, authors, publisher, publication year, edition, number of pages, and selling price. The title and authors must have non-null values. The value of edition must be non-null with one of the following values: paperback, hardcover, or ebook. The selling price must have a positive value. If the number of pages is known, it must be a positive value.
   - `Customers` store information about the shop's customers. Each customer has a unique identifier `cust_id`, a name and an email address. The name must have a non-null value.
   - `Carts` store information about the books in customers' shopping carts. Each shopping cart record indicates a book that a customer is interested to order.
   - When a customer checks out with a non-empty shopping cart, a new record for the customer's order is recorded in the `Orders` table which has the following information: a unique identifier `order_id`, the date of the order, and the customer identifier. In addition, each book in the customer's shopping cart is added to the `Order_items` table and the customer's shopping cart is emptied.

   (a) Write SQL statements to create the database schema with appropriate constraints. Assume that attributes `isbn`, `title`, `authors`, `edition`, `publisher`, `name`, `email` have text domain; attributes `year`, `number_pages`, `cust_id`, and `order_id` have integer domain; attribute `price` has numeric domain; and attribute `order_date` has date domain.

   (b) Suppose that the schema of `Orders` is changed with `order_date` being replaced by `order_timestamp`, where each customer has at most one order at any timestamp. How would this change affect your answer for part (a)?

   (c) For each of the following additional constraints on the database, state whether the constraint can be expressed using the SQL constructs that you have learned. If it is possible, add this constraint to your database schema in part (a).

      1. If a book is a hardcover edition, its selling price must be at least 30.
      2. If a book has both hardcover and paperback editions (for the same book title and authors), the selling price for the hardcover edition must be higher than the selling price for the paperback edition.
      3. If the number of pages in a book is more than 1000, the edition of the book must be a ebook or its price must be at least 100.

4. All the books published by 'Acme' from 2010 onwards have only ebook edition.

(d) Consider the following additional constraints on the database:

1. If a customer is deleted from `Customers`, remove all the customer's records from `Carts` and `Orders`.
2. If a book is deleted from `Books`, remove all the records from `Carts` that reference this book, and for each of the records in `Order_items` that reference this book, change its `isbn` value to the default value of '000'.
3. If an order is deleted from `Orders`, remove all the records from `Order_items` that reference this order.
4. Any modification of cust_id in `Customers` is propagated to other records in the database.
5. Any modification of isbn in `Books` is propagated to other records in the database.
6. Any modification of order_id in `Orders` is propagated to other records in the database.

Add these additional constraints to your answer for part (a).