Assignment 5 RA Expressions

Sanket Pandilwar

We will use the following relation schemas and their respective abbreviations:

Student(sid, sname)	S, S_1, S_2 etc.
Book(bookno, title, price)	B, B_1, B_2, etc.
Buys(sid, bookno)	T, T_1, T_2 etc.
Cites(bookno, citedBookno)	C, C_1, C_2 etc.
Major(sid, major)	M, M_1, M_2 etc.

1. (a) Consult the lecture on set joins and semijoins. Using the techniques described in that lecture, develop a general RA expression for the all-but-two set semijoin.
Reference: Discussed this problem with one of the colleagues

Consider two relation A and B having common relationship with relation X. We want to find expression for "Find A that (X) 'all-but-two' B, where X might be a common relation between A and B relations."

General RA expression using these 3 relations is as follows:

Let E be the RA.

$$E = (B \times A) - (B \bowtie X)$$

Now to find all-but-two, we can find =2 and > 2 and difference will give us all-but-two tuples. Let E has a schema with two columns a and b. Where a references entry from relation A and b references entry from relation B. Now, All-but-two can be expressed in RA as:

$$\pi_{E_1.a}(E_1 \ltimes_{E_1.a=E_2.a \wedge E_1.b<>E_2.b} (E_2))$$
$$-\pi_{E_1.a}((E_1 \ltimes_{E_1.a=E_2.a \wedge E_1.b<>E_2.b} (E_2)) \ltimes_{E_1.a=E_3.a \wedge E_1.b<>E_3.b \wedge E_2.b<>E_3.b} (E_3))$$

(b) Apply this RA expression to the query "Find the bookno and title of each book that is bought by all but two students who major in CS."

By relating to previous problem, relation A is same as Book, X is same as buys and B is same as CS major students.

Applying above RA expression for this problem:

Let E be the RA.

$$E = \pi_{B.bookno,M.sid}((\pi_{M.sid}(\sigma_{M.major='CS'}(M)) \times B) - (\pi_{M.sid}(\sigma_{M.major='CS'}(M) \bowtie T)))$$

Now, All-but-two for this problem can be expressed in RA as:

$$\pi_{E_1.bookno}(\pi_{E_1.bookno}(E_1 \bowtie_{E_1.bookno=E_2.bookno \wedge E_1.sid<>E_2.sid} (E_2))$$
$$-\pi_{E_1.bookno}((E_1 \bowtie_{E_1.bookno=E_2.bookno \wedge E_1.sid<>E_2.sid} (E_2))$$
$$\bowtie_{E_1.bookno=E_3.bookno \wedge E_1.sid<>E_3.sid \wedge E_2.sid<>E_3.sid} (E_3)))$$

(c) Formulate the RA expression obtained in Problem 1b in SQL with relational operators. (So no SQL set predicates are allowed in your solution.)

```
with book_sid_relation as
(select b.bookno, s.sid from
(select sid from major where major='CS') s cross join book b
except
select b.bookno, s.sid from
(select sid from major where major='CS') s natural join buys b)

select b1.bookno from book_sid_relation b1
join book_sid_relation b2 on b1.bookno = b2.bookno
and b1.sid <> b2.sid
except
select b1.bookno from book_sid_relation b1
join book_sid_relation b2 on b1.bookno = b2.bookno and
b1.sid <> b2.sid join book_sid_relation b3 on b1.bookno=b3.bookno
and b1.sid<>b3.sid and b2.sid<>b3.sid
```

2. Consider two RA expressions E1 and E2 over the same schema. Furthermore, consider an RA expression F with a schema that is not necessarily the same as that of E1 and E2.

   (a)
   i. Write an RA expression, in function of E1, E2, and F, that expresses this if-then-else statement.

$$\pi_B(E_1 \times F) \cup (E_2 - \pi_B(E_2 \times F))$$

   ii. Then express this RA expression in SQL with RA operators. In particular, you can not use SQL set predicates in your solution.

```
select e1.* from E1 e1 cross join F
union
(select e2.* from E2 e2
except
select e2.* from E2 e2 cross join F)
```

   (b) Let A(x) be a unary relation that can store a set of integers A. Using the insights you gained from Problem 2a, solve the following problems:

i. Write an RA expression that expresses the above boolean SQL query.

$$\pi_B((B : True) \times F) \cup ((B : False) - \pi_B((B : False) \times F))$$

ii. Write a SQL query with relational operators, thus without set predicates, that expresses the above boolean SQL query.

```
select t.B from (select true as B) t cross join F
union
(select false as B
except
select q.B from (select false as B) q cross join F)
```

3. Let f : A -> B be a function from a set A to a set B and let g : B -> C be a function from a set B to a set C. The composition of the functions f and g, denoted g of f, is a function from A to C such that for x E A, g of f(x) is defined as the value g(f(x)). Represent f in a binary relation f with schema (A,B) and represent g in a binary relation g with schema (B,C).

   (a) Write an RA expression that computes the function g of f. I.e., your expression should compute the binary relation {(x, g of f(x)) | x E A}.

   $$\pi_{f.a,g.c}(f \bowtie_{f.b=g.b} g)$$

   (b) Let y be a value in C. Write an RA expression that computes the set x E A | g of f(x) = y. I.e., these are the values in A that are mapped by the function g of f to the value y.

   $$\pi_{f.a}(f \bowtie \pi_{g.b}(\sigma_{g.c=y}(g)))$$

4. Let f : A -> B be a function from a set A to a set B. We say that f is a one-to-one function if for each pair x1 and x2 of different values in A (i.e., x1 != x2) it is the case that f(x1) != f(x2). Represent f by a relation f with schema (A,B).
   Write an RA expression that returns the value "true" if f (as stored in f) is a one-one-one function, and returns the value "false" otherwise.

   Let Q be the RA.

   $$Q = \pi_{F_1.*}(\sigma_{F_1.a \neq F_2.a \wedge F_1.b=F_2.b}(F_1 \times F_2))$$

   Then the expression is

   $$\pi_E((E : False) \times Q) \cup ((E : True) - \pi_E((E : True) \times Q))$$

5. Let f : A -> B be a function from a set A to a set B. We say that f is an onto function if for each value y in B, there exists a value x in A such that f(x) = y. Represent f by a relation f with schema (A,B).
   Write an RA expression that returns the value "true" if f (as stored in f) is an onto function, and returns the value "false" otherwise.

   Let Q be the RA.

   $$Q = \pi_{F.b}(\sigma_{F.a=\Phi}(F)) - \pi_{F.b}(\sigma_{F.a\neq\Phi}(F))$$

   Then the expression is

   $$\pi_E((E : False) \times F) \cup ((E : True) - \pi_E((E : True) \times F))$$

6. A graph G is a structure (V, E) where V is a set of vertices and wherein E is a set of edges between these vertices. Thus E subset of V x V . A path in G is a sequence of vertices (v0, v1, . . . , vn) such that for each i belongs to [0, n-1], (vi, vi+1) belongs to E. We call n the length of this path. Represent E by a binary relation E(source,target). A pair (s, t) is in E if s and t are vertices in V and (s, t) is an edge in E. Think of s as the source of this edge and t as the target of this edge. We say that two vertices v and w in V are connected in G by a path of length n if there exists a path (v0, v1, . . . , vn) such that v = v0 and w = vn.
   Write an RA expression that returns the set of pairs (v, w) that are connected by a path of length at most n. (You may assume that n >= 1.)

   Assumptions and Approach:

   (a) Let E be the relation which contains source and target vertices of length 1.
   (b) To find vertices of length 2 we can join E relation twice on condition that "first's target vertex would be second's source vertex and first's source vertex should not be same as second's target vertex" since that would mean it came back to same vertex. which is not acceptable path for us.
   (c) To find vertices with path length n, we can join E relation n times on above condition recursively.
   (d) So if we want vertices with path of length at most n, we can use union to add all the set of vertices with path length one with path length two and so on...till n path length

   Let C be the generalized condition which is used between two relations E1 and E2 such that:

   $$C = \{E_1.target = E_2.source \wedge E_1.source \neq E_2.target\}$$

$$\pi_{E_1.source,E_1.target}(E_1)$$
$$\cup$$
$$\pi_{E_1.source,E_2.target}(E_1 \bowtie_C E_2)$$
$$\cup$$
$$\pi_{E_1.source,E_3.target}((E_1 \bowtie_C E_2) \bowtie_C E_3)$$
$$\cup$$
.
.
.
$$\cup$$
$$\pi_{E_1.source,E_n.target}(...((E_1 \bowtie_C E_2) \bowtie_C E_3).... \bowtie_C E_n)$$

7. Find the sid and name of each student who majors in CS and who bought a book that cost more than 10. (Assignment 2, Problem 1.)

   Consider the following expressions:

$$\begin{aligned} E_1 &= \pi_{sid}(T \ltimes \pi_{bookno}(\sigma_{price>10}(B))) \\ E_2 &= \pi_{sid}(\sigma_{major='CS'}(M) \ltimes E_1) \end{aligned}$$

   Then the expression is

$$\pi_{sid,sname}(S \ltimes E_2).$$

```
with e1 as
(select sid from buys natural join (select bookno from book where pri
e2 as
(select m.sid from major m join e1 on (m.sid = e1.sid and m.major='CS
select distinct sid,sname
from student natural join e2;
```

8. Find the bookno, title, and price of each book that cites at least two books that cost less than 60. (Assignment 2, Problem 3.)

   Consider the following expressions:

$$\begin{aligned} E_1 &= \pi_{bookno}(C \bowtie_{B.bookno=C.citesbookno \wedge B.price<60} (B)) \\ E_2 &= \pi_{bookno}(C \bowtie_{B.bookno=C.citesbookno \wedge B.price<60} (B)) \end{aligned}$$

   Then the expression is

$$\pi_{B.*}((B \bowtie_{E_1.bookno=B.bookno} E_1) \bowtie_{E_1.bookno=E_2.bookno \wedge E_1.citedbookno<>E_2.citedbookno} (E_2))$$

```
with e1 as
(select c.bookno, c.citedbookno,b.title,b.price
from cites c join book b
on (b.bookno = c.citedbookno and b.price<60))
select distinct c.bookno,c.title,c.price
from e1 a join e1 b
on a.bookno=b.bookno and a.citedbookno<>b.citedbookno
join book c on c.bookno=a.bookno;
```

9. Find the bookno, title, and price of each book that was not bought by any Math student. (Assignment 2, Problem 2.)

$$\pi_{B.*}(B) - \pi_{B.*}(B \bowtie_{B.bookno=T.bookno} (T \bowtie (\sigma_{M.major='Math'}(M)))))$$

```
select b.bookno,b.price,b.title from book b
except
select b.bookno,b.price,b.title
from book b natural join buys t natural join
(select m.sid from major m where m.major='Math') e1
```

10. Find the sid and name of each student along with the title and price of the most expensive book(s) bought by that student. (Assignment 2, Problem 4.)

    Consider the following expressions:

$$E_1 = \pi_{S.sid,S.sname,B.title,B.price}(S \bowtie T \bowtie B)$$

Then the expression is

$$\pi_{E_1.*}(E_1) - \pi_{E_1.*}(\sigma_{E1.sid=T.sid \wedge B.bookno=T.bookno \wedge !(E1.price>=B.price)}(E_1 \times T \times B))$$

```
with e1 as
(select distinct s.sid,s.sname,bo.title,bo.price
from student s natural join buys bu natural join book bo)

select distinct e.sid,e.sname,e.title,e.price from e1 e
except
select distinct e.sid,e.sname,e.title,e.price
from e1 e cross join buys bu1 cross join book bo1
where e.sid=bu1.sid and bo1.bookno=bu1.bookno
and not (e.price>=bo1.price);
```

11. Find the booknos and titles of books with the next to highest price. (Assignment 2, Problem 6.)

    Consider the following expressions:

$$E = \pi_{B.*}(B \bowtie_{B.price<B_1.price} (B_1))$$

Then the expression is

$$\pi_{E_1.bookno,E_1.title}(E_1) - \pi_{E_2.bookno,E_2.title}(E_2 \bowtie_{E_2.price<E_3.price} (E_3))$$

```
with E as
(select b.*
from book b join book b1 on b.price < b1.price)
select bookno, title from E
except
select e.bookno, e.title
from E e join E e1 on e.price < e1.price;
```

12. Find the bookno, title, and price of each book that cites a book which is not among the most expensive books. (Assignment 2, Problem 7.)

    Consider the following expressions:

    $$E = \pi_{B.bookno,B.title,B.price,C.citedbookno}(B \bowtie C)$$

    Then the expression is

    $$\pi_{E_1.bookno,E_1.title,E_1.price}(E_1 - (\pi_{E_1.bookno,E_1.title,E_1.price}(E_2 \bowtie B - ((E_3 \bowtie_{E_3.citedbookno=B_1.bookno} B_1) \bowtie_{B_1.price<B2_price} (B_2)))))$$

```
with E1 as
(select distinct b.bookno,b.title,b.price,c.citedbookno
from book b natural join cites c)
select q1.bookno,q1.title,q1.price
from (select distinct e1.bookno,e1.title,e1.price from E1 e1
except
select q.* from
(select distinct e2.bookno,e2.title,e2.price
from E1 e2 join book b on b.bookno=e2.citedbookno
except
select distinct e3.bookno,e3.title,e3.price from E1 e3 join book b on
join book b1 on b1.price > b.price)q)q1;
```

13. Find the sid and name of each student who has a single major and such that none of the book(s) bought by that student cost less than 40. (Assignment 2, Problem 8.)

    Consider the following expressions:

    $$E = \pi_{S.sid,S.sname,M.major}(S \bowtie M)$$

    Then the expression is

    $$\pi_{sid,sname}(\pi_{E.sid,E.sname}(E) - \pi_{E_1.sid,E_1.sname}(E_1 \bowtie_{E_1.sid=M_1.sid \wedge E_1.major<>M_1.major} M_1) - \pi_{E_2.sid,E_2.sname}(E_2 \bowtie \pi_{T.sid}(\sigma_{B.price<40}(T \bowtie B))))$$

```
with E as
(select s.sid,s.sname,m.major from student s join major m
on s.sid=m.sid)
select e.sid,e.sname from E e
except
select e.sid,e.sname
from  E e join major m2 on e.sid=m2.sid and e.major <> m2.major
except
select e.sid,e.sname
from E e join buys bu on e.sid=bu.sid join book bo
on bu.bookno=bo.bookno and bo.price<40;
```

14. Find the bookno and title of each book that is bought by all students who major in both CS and in Math. (Assignment 2, Problem 9.)

    Consider the following expressions:

$$E = \pi_{M_1.sid}(\pi_{M_1.sid}(\sigma_{M_1.major='Math'}(M_1)) \bowtie \pi_{M_2.sid}(\sigma_{M_2.major='CS'}(M_2)))$$

    Then the expression is

$$\pi_{B.bookno,B.title}(B \bowtie (\pi_{T.bookno}(T) - (\pi_{B_1.bookno}((B_1 \times E) - (B_2 \bowtie T_1)))))$$

```
with E as
(select m1.sid from major m1 join major m2
on m1.major='Math' and m2.major='CS' and m1.sid=m2.sid)

select b1.bookno,b1.title from book b1 join
(select t.bookno from buys t
except
(select q1.bookno from
(select * from book b cross join E
except
select * from book b natural join buys)q1)
)q2
on b1.bookno=q2.bookno;
```

15. Find the sid and name of each student who, if he or she bought a book that cost at least than 70, also bought a book that cost less than 30. (Assignment 2, Problem 10.)

$$\pi_{sid,sname}(S) - \pi_{S_1.sid,S_1.sname}((T_1 \bowtie_{T_1.bookno=B_1.bookno \wedge B_1.price>70} B_1) \bowtie (S_1))$$

$$-\pi_{S_2.sid,S_2.sname}(((T_2 \bowtie S_2) \bowtie_{T_2.bookno=B_2.bookno \wedge B_2.price>=30} B_2) \bowtie_{T3.sid=S2.sid}$$

$$(T_3 \bowtie_{T_3.bookno=B_3.bookno \wedge B_3.price<=30} B_3))$$

```
select s.sid, s.sname
from student s
except
select s1.sid, s1.sname
 from buys t1 join book b1 on t1.bookno=b1.bookno
and b1.price > 70 join student s1 on t1.sid=s1.sid
union
select distinct s2.sid, s2.sname
from student s2 join buys t2 on t2.sid=s2.sid
join book b2 on t2.bookno=b2.bookno and b2.price >= 70
join buys t3 on t3.sid=s2.sid join book b3 on t3.bookno=b3.bookno and
```

16. Find each pair (s1, s2) where s1 and s2 are the sids of students who have a common major but who did not buy the same set of books. (Assignment 2, Problem 11.)

    Consider the following expressions:

$$
\begin{array}{rcl}
E & = & \pi_{m1.sid,m2.sid}(M_1 \bowtie_{M_1.major=M_2.major \wedge M_1.sid<>M_2.sid} (M_2)) \\
E_1 & = & \pi_{T_1.sid,T_1.bookno,S_1.sid}(T_1 \times S_1) \\
E_2 & = & \pi_{S_2.sid,T_2.bookno,T_2.sid}(S_2 \times T_2)
\end{array}
$$

    Then the expression is

$$\pi_{E.*}(E \cap (\pi_{E_1.sid1,E_1.sid2}(E_1 - E_2) \cup \pi_{E_2.sid1,E_2.sid2}(E_2 - E_1)))$$

```
with E as (select m1.sid,m2.sid
from major m1 join major m2
on m1.major = m2.major and m1.sid<>m2.sid),
E1 as
(select t1.sid as sid1, t1.bookno, s.sid as sid2
from buys t1 cross join Student s),
E2 as
(select s.sid as sid1, t2.bookno, t2.sid as sid2
from Student s cross join buys t2)

select * from E

intersect

(select distinct sid1, sid2
from
(select e1.* from E1 e1
except
select e2.* from E2 e2) q1

union
```

```
select distinct sid1, sid2
from
(select e2.* from E2 e2
except
select e1.* from E1 e1) q2
)
```