Assigment5

Pauravi Wagh

3/1/2020

1. a)

All but two case signifies that some condition is satisfied by all the tuples belonging to a particular set, except 2. For example, we need to find courses such that each course is taken by all but 2 students. So, here we will calculate the students who have **not** taken that particular course and from that ,we will find out how many courses have not been taken by atleast 2 students and how many courses have not been taken by at least 3 students. After subtracting the at least 2 cases from at least 3 cases, we will get courses taken by all but 2 students.
So, for generalizing, let us say we have a relation E that contains tuples that are obtained by subtracting a condition from all possible cases. So E will contain the excess part of the difference, from which we can calculate all but two.

So basically, we are checking atleast 2 and atleast 3 on the right ear of the venn diagram. Which means that we take at least 2 and at least 3 from the NOT ALL section of the venn diagram. We have a relation E from which we wish to extract tuples that satisfy the all but two case.

So, E is a set of values from the right ear of the Venn diagram. So, if A and B are two relations then basically E is $A \ominus B$

We get from relation E, atleast 2 values of S for a desired value of R
$\pi_{e_1.p_1 \ldots e_1.p_n}(E_1 \bowtie_{e_1.p_1 \theta e_2.p_1 \wedge \ldots \wedge e_1.p_n \theta e_2.p_n} (E_2))$

Now we get the relation from E, atleast 3 values of S for a desired value of R

$\pi_{e_1.p_1 \ldots e_1.p_n}(E_1 \bowtie_{e_1.p_1 \theta e_2.p_1 \wedge \ldots \wedge e_1.p_n \theta e_2.p_n} (E_2) \bowtie_{e_1.p_1 \theta e_3.p_1 \wedge e_2.p_1 \theta e_3.p_1 \wedge \ldots . \wedge e_2.p_n \theta e_3.p_n \wedge e_1.p_n \theta e_3.p_n} (E_3))$

Now, we will subtract the at least 3 case from atleast 2 case, so that we obtain all but 2 values.

$\pi_{e_1.p_1 \ldots e_1.p_n}(E_1 \bowtie_{e_1.p_1 \theta e_2.p_1 \wedge \ldots \wedge e_1.p_n \theta e_2.p_n} (E_2))$
$-\pi_{e_1.p_1 \ldots e_1.p_n}(E_1 \bowtie_{e_1.p_1 \theta e_2.p_1 \wedge \ldots \wedge e_1.p_n \theta e_2.p_n} (E_2) \bowtie_{e_1.p_1 \theta e_3.p_1 \wedge e_2.p_1 \theta e_3.p_1 \wedge \ldots . \wedge e_2.p_n \theta e_3.p_n \wedge e_1.p_n \theta e_3.p_n} (E_3))$

1. b)

```
with CSStudents as
    (select distinct m.sid
    from major m join major m1 on(m.sid = m1.sid and m1.major = 'CS')),
booksNotBoughtByCSStudents as
    (select b.bookno, s.sid
    from book b, CSStudents s
    except
    select t.bookno, t.sid
    from buys t),

atleast2 as
    (select distinct b1.bookno
    from booksNotBoughtByCSStudents b1 join booksNotBoughtByCSStudents b2
                                       on (b1.bookno = b2.bookno and b1.sid !=b2.sid)),
atleast3 as
```

```
(select distinct k.bookno
from
(select b1.bookno, b1.sid as s1, b2.sid as s2
from booksNotBoughtByCSStudents b1 join booksNotBoughtByCSStudents b2
                                    on (b1.sid!=b2.sid and b1.bookno = b2.bookno)) k
                                   join booksNotBoughtByCSStudents b3
                                    on (k.s1 !=b3.sid and k.s2 != b3.sid and b3.bookno = k.bookno))

select b.bookno, b.title
from book b natural join(select b.bookno
                            from atleast2 b
                            except
                            select b.bookno
                            from atleast3 b)p;
```

1. c)

Consider a relation CS for CS students

$\pi_{m_1.sid}(M_1 \bowtie_{m_1.sid=m_2.sid \wedge m_1.major='CS'} (M_2))$

Consider a relation K which obtains the student ids and booknos that they did not buy

$\pi_{bookno,sid}(\pi_{bookno}(B) \times \pi_{sid}(CS) - T)$

Consider a relation G which obtains atleast 2 students who did not buy the book.

$\pi_{bookno}(K_1 \bowtie_{k_1.sid<>k_2.sid \wedge k_1.bookno=k_2.bookno} (K_2))$

Consider a relation H which obtains atleast 3 students who did not buy the book.

$\pi_{k_1.bookno}(K_1 \bowtie_{k1.sid<>k_2.sid \wedge k_1.bookno=k_2.bookno} (K_2) \bowtie_{k_2.bookno=k_3.bookno \wedge k_1.sid<>k_3.sid \wedge k_2.sid<>k_3.sid} (K_3)$

Now to obtain books bought by all but 2 CS students we subtract books not bought by atleast 3 students from books not bought by atleast 2 students, we will intergrate.

$\pi_{bookno,title}(B \bowtie (G - H))$

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

Answer:
$\pi_{E_1}(E_1 \times F) \cup \pi_{E_2}(E_2 - \pi_{E_2}(E_2 \times F))$

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

Answer:

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

2. b. i.
Answer:
$\pi_{AIsNotEmpty}((AIsNotEmpty : true) \times \pi_{()}(A)) \cup ((AIsNotEmpty : false) - \pi_{AIsNotEmpty}((AIsNotEmpty : false) \times \pi_{()}(A)))$

2. b. ii.
Answer:

```
select t.A_isNotEmpty
from (select true as A_isNotEmpty) t
cross join (select distinct row() from  A)k

union

(select t.A_isNotEmpty
from (select false as A_isNotEmpty) t
except
select t.A_isNotEmpty
from (select false as A_isNotEmpty) t
  cross join (select distinct row() from A) k)
```

3.a
Answer:
The RA for the binary relation
$\{(x, g \circ f(x)) | x \in A\}$ is

$\pi_{A,C}(F \bowtie G)$

3.b
Answer:
The RA expression for
$\{X \in A | g \circ f(x) = y\}$ is

$\pi_A(F \bowtie_{f.b=g.b \wedge g.c=y} (G))$

4.
Answer

Consider a Relation F as

$F_1 \bowtie_{f_1.a \neq f_2.a \wedge f_1.b=f_2.b} (F_2)$

The RA for the one-to-one function is

$\pi_B((B : false) \times F) \cup ((B : true) - \pi_B((B : true) \times F))$

5.
Answer:
Let the sub-expression G be
$\pi_{f.b}(F) - \pi_{f_1.b}(F_1 \bowtie_{f_1.a=f_2.a} (F_2))$

The RA expression is

$\pi_B((B : false) \times G) \cup ((B : true) - \pi_B((B : true) \times G))$

6.
Answer:

We have a relation E(source, target), which has connected edges in it.
We have a path P(v,w), where v and w are two vertices connected by a path of length between 1. . . N.
We realize that path of length 2 can be obtained by joining E with E as follows:

$\pi_{e_1.source, e_2.target}(E_1 \bowtie_{e_1.target=e_2.source} E_2)$

The above relation gives us a path of length 2. Let us consider this as L_2. Let's say we collect all the paths in the above defined relation P. So P now has all the paths upto length 1 and 2. We can achieve this by taking a union of E with L_2. $L_2 \cup E$

Then to calulate path of length 3, we just need to combine E with L_2 as follows:

$\pi_{e.source, l_2.target}(E \bowtie_{e.target=l_2.source} L_2)$

We say this to be L_3.We can continue this for path of length 4 where,

$\pi_{e.source, l_3.target}(E \bowtie_{e.target=l_3.source} L_3)$.

If we carefully observe the pattern here, then the path of length n can be obtained by simple joining paths of length n-1 with E. To generalize,

$\pi_{e.source, l_{n-1}.target}(E \bowtie_{e.target=l_{n-1}.source} L_{n-1})$

So, to find sets of vertices with path lenghts of atmost one, we can just union all the previously obtained paths.
so,

$P \cup \pi_{e.source, l_{n-1}.target}(E \bowtie_{e.target=l_{n-1}.source} L_{n-1})$

The above expression should give us paths of length at most n,

where P = $E \cup L_1 \cup L_2 \cup .... \cup L_{n-1}$

7.Find the sid and name of each student who majors in CS and who bought a book that cost more than $10.

Answer:
$\pi_{sid,sname}(S \bowtie (\pi_{sid}(\sigma_{major='CS'}(M)) \bowtie \pi_{sid}(T \bowtie_{t.bookno=b.bookno \wedge b.price>10} (B))))$

8.Find the bookno, title, and price of each book that cites at least two books that cost less than $60.

Answer:

Consider R which selects the books having price less than 60 as
$\pi_{bookno}(\sigma_{price<60}(B))$

Consider E which selects only those books whose citedbooknos form a part of the R relation as
$\pi_{c.bookno,c.citedbookno}(C \bowtie_{c.citedbookno=r.bookno} R)$

$\pi_{bookno,title,price}(B \bowtie \pi_{e_1.bookno}(E_1 \bowtie_{E1.bookno=E2.bookno \wedge E1.citedbookno \neq E2.citedbookno} E_2))$

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

Answer: Consider relation R which selects sids of students who have math as their major
$\pi_{sid}(\sigma_{major='Math'}(M))$

Consider relation E which selects books bought by Math Majors as
$\pi_{bookno}(T \bowtie R)$

Now all the books apart from the books selected in the E clause will be selected
$\pi_{bookno,title,price}(B \bowtie \pi_{bookno}(\pi_{bookno}(B) - E))$

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.

Answer:

Consider a relation R that connects book prices with books bought by the student

$\pi_{t.sid,t.bookno,b.price}(T \bowtie_{t.bookno=b.bookno} \pi_{bookno,price}(B))$

Consider a relation G that selects the not most expensive books for each student

$\pi_{r_1.sid,r_1.bookno}(R_1 \bowtie_{r1.sid=r2.sid \wedge r1.price<r2.price} (R_2))$

Now we find the student's sid and bookno of the most expensive book bought by the student
Consider it to be K

$\pi_{sid,bookno}(T - G)$

Consider a relation L which gives us the book title, price and sid of the student who bought his most expensive book.

$\pi_{k.sid,b.title,b.price}(B \bowtie_{b.bookno=k.bookno} (K))$

Now we can extract student sid, name, along with book title and price of the most expensive book.

$\pi_{s.sid,s.sname,l.title,l.price}(S \bowtie_{s.sid=l.sid} (L))$

11. Find the booknos and titles of books with the next to highest price.

Answer:
Consider relation R which selects all the books that are not the highest priced

$\pi_{B1.*}(B_1 \bowtie_{b_1.price<b_2.price} (B_2))$

Now we subtract from R all the books that are not highly priced in R.

$\pi_{bookno,title,price}(R) - \pi_{r_1.bookno,r_1.title,r_1.price}(R_1 \bowtie_{r_1.price<r_2.price} (R_2))$

12. Find the bookno, title, and price of each book that cites a book which is not among the most expensive books.

Answer:

Consider a relation R which selects the most expensive book.
$\pi_{bookno}(B - \pi_{B_1.*}(B_1 \bowtie_{b_1.price<b_2.price} (B_2)))$

Now Consider a relation E which contains all the books that do not cite the most expensive book.

$\pi_{bookno}(\pi_{bookno,citedbookno}(C) - \pi_{c_1.bookno,c_1.citedbookno}(C_1 \bowtie_{c_1.citedbookno=r.bookno} (R)))$

Now we use this relation to find the booknos, title and proce of books that do not cite a book which is amongst the most expensive ones.

$\pi_{bookno,title,price}(B \bowtie E)$

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.

Answer:

Consider a relation R which selects students having a single major

$\pi_{sid}(M) - \pi_{m_1.sid}(M_1 \bowtie_{m_1.sid=m_2.sid \wedge m_1.major \neq m_2.major} (M_2))$

Consider a relation E that selects books less than 40

$\pi_{bookno}(\sigma_{price<40}(B))$

Consider a relation G that selects students who have not bought books less than 40.

$\pi_{sid}(S) - \pi_{sid}(T \bowtie E)$

Now we select the snames and sids of students who have a single major and have not bought a book that costs less than 40.

$\pi_{sid,sname}(S \bowtie R \bowtie G)$

14. Find the bookno and title of each book that is bought by all students who major in both CS and in Math.

Answer:

Consider a relation R that is a set of CS and Math majors.
$\pi_{sid}(M_1 \bowtie_{m1.sid=m2.sid \wedge m1.major='CS' \wedge m2.major='Math'} M_2)$

Consider a relation E where we find out the books not bought by all the students who major in Cs and Math.

$\pi_{bookno}((R \times \pi_{bookno}(B)) - T)$

Now we subtract from the relation books to eliminate books that have not been bought by all cs and math students.

$\pi_{bookno,title}(B \bowtie (\pi_{bookno}(B) - E))$

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.

Consider relation F which collects all the students who bought books greater than or equal to 70

$\pi_{t.sid}(T \bowtie_{b.bookno=t.bookno \wedge b.price>=70} (B))$

Consider a relation E which selects all the students except those who did not buy books less than 30

$\pi_{sid}(S - \pi_{t.sid}(T - \pi_{t_1.sid}(T_1 \bowtie_{t_1.bookno=b.bookno \wedge b.price<30} (B))))$

We select the student satisfying the condition as

$\pi_{s.sid,s.sname}(S \bowtie (\pi_{e_1.*}(E_1 \times F) \cup \ (E_2 - \pi_{e_2.*}(E_2 \times F))))$

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.

Answer:

Consider a relation R that contains students having a common major.

$\pi_{m_1.sid_1,m_2.sid_2}(M_1 \bowtie_{m_1.sid \neq m_2.sid \wedge m_1.major=m_2.major} (M_2))$

Consider a relation K that connects students of common major in every possible way.

$\pi_{r_1.sid_1,r_2.sid_2}(R_1 \times R_2)$

Consider a relation E1 that first joins K relation to buys relation on sid1. Thus, we have data where we know sid1 has bought the book, but we are not sure about sid2. So, we find a similar data for sid2 in relation E2, where we know what books have been bought by sid2 and we do not know anything about sid1. So, in order to eliminate the students who bought common books, we subtract the two sets from each other separately. Let E1 be $\pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_1=t.sid} (T))$ and E2 be $\pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_2=t.sid} (T))$

So E1-E2(let us call it H1) is,
$\pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_1=t.sid} (T)) - \pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_2=t.sid} (T))$

And E2-E1(let us call it H2) is,
$\pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_2=t.sid} (T)) - \pi_{k.sid_1,k.sid_2,t.bookno}(K \bowtie_{k.sid_1=t.sid} (T))$

Now, H1 and H2 contains the data of students who did not buy common books. We now need to extract from them the ones with common major.

$\pi_{r.sid_1,r.sid_2}(R \bowtie \pi_{sid_1,sid_2}(H_1 \cup H_2))$