

**CPSC 304 Midterm 2**  
**Nov 10, 2004**  
**Total: 17 points**

**Question 1 (9 points)**

The donations relation was defined based on the following SQL statement:

```
CREATE TABLE donations
    (recipientName CHAR(20) NOT NULL,
    donorOrganization CHAR(20) NOT NULL,
    organizationType CHAR(20),
    amount REAL,
    PRIMARY KEY (recipientName, donorOrganization)
)
```

For each of the following relational calculus queries, determine if there exists an equivalent relational algebra statement. If your answer is positive, give such a statement; otherwise, just state that no such statement exists. (You may get part marks if you write down in English the correct meaning of each query.)

- a)  $\{ \langle N \rangle \mid \exists x, y (\langle N, x, \text{tobacco}, z \rangle \in \text{donations} \text{ and } \langle N, y, \text{tobacco}, r \rangle \in \text{donations} \text{ and } x \neq y) \}$

*(2 points) Recipients who received donations from at least 2 distinct tobacco organizations.*

$\pi_{A.recipientName} (\sigma_{A.recipientName = B.recipientName \wedge A.organizationType = tobacco \wedge B.organizationType = tobacco \wedge A.donorOrganization \neq B.donorOrganization} [ \rho A (\text{donations}) \times \rho B (\text{donations}) ])$

- b)  $\{ \langle N \rangle \mid \forall x (\langle r, x, \text{tobacco}, s \rangle \in \text{donations} \Rightarrow \langle N, x, \text{tobacco}, t \rangle \in \text{donations}) \}$

*(3 points) Recipients who received donations from every tobacco donor organization.*

$\pi_{recipientName, donorOrganization}(\text{donations}) / \pi_{donorOrganization} (\sigma_{organizationType = tobacco} (\text{donations}))$

- c)  $\{ \langle N \rangle \mid \exists x (\langle N, r, \text{tobacco}, x \rangle \in \text{donations} \text{ and } \forall y (\langle s, t, \text{tobacco}, y \rangle \in \text{donations} \Rightarrow x \geq y) \}$

*(4 points) Short version: Recipients who received the highest donation amount from a tobacco company. [note; this English version isn't entirely precise, because it could be interpreted "for each company", which isn't what we mean. See the Relational Algebra for an exact comparison]*

$\pi_{recipientName} (\sigma_{organizationType = tobacco} (\text{donations})) - \pi_{A.recipientName} (\sigma_{A.organizationType = tobacco \text{ and } B.organizationType = tobacco \text{ and } A.amount < B.amount} [ \rho A (\text{donations}) \times \rho B (\text{donations}) ] )$

- d)  
(Removed from the exam.)

**Question 2 (8 points)**

Based on the donations relation defined above, determine whether each of the following four pairs of SQL statements is equivalent. If the pair is equivalent, just say yes and no explanation is needed. If you do not think the pair is equivalent, construct an instance of the donations relation to illustrate the difference between the pair of statements.

- a)  
select distinct recipientName from donations A  
where not exists  
(select B.donorOrganization from donations B  
where recipientName = "Campbell"  
and A.donorOrganization  $\neq$  B.donorOrganization)

vs

select distinct recipientName from donations A  
where not exists  
( (select donorOrganization from donations  
where recipientName = "Campbell" )  
except  
(select donorOrganization from donations B  
where B.recipientName = A.recipientName))

(3 points) No. Consider the instance:

<i>recipientName</i>	<i>donorOrganization</i>	<i>...</i>
<i>John</i>	<i>A</i>	
<i>John</i>	<i>B</i>	
<i>Campbell</i>	<i>A</i>	
<i>Campbell</i>	<i>B</i>	

*Top query returns empty, while the bottom query returns John and Campbell.*

- b)

(select distinct recipientName from donations  
where amount  $\geq$  500)  
union  
(select distinct recipientName from donations  
where amount < 500)

vs

select distinct recipientName from donations

*No. Consider the instance:*

<i>recipientName</i>	<i>Amount</i>	<i>...</i>
<i>John</i>	<i>100</i>	
<i>Campbell</i>	<i>2000</i>	
<i>Mary</i>	<i>null</i>	

*Top query returns John and Campbell, whereas bottom query returns all 3 names.*

c)

select distinct recipientName from donations A, donations B  
where A.amount  $\geq$  1000  
and A.recipientName = B.recipientName  
and A.donorOrganization  $\neq$  B.donorOrganization

vs

select distinct recipientName from donations  
where amount  $\geq$  1000  
group by recipientName  
having count(donorOrganization)  $\geq$  2

*(2 points) No. Consider the instance:*

<i>recipientName</i>	<i>donorOrganization</i>	<i>Amount</i>	<i>...</i>
<i>Campell</i>	<i>CUPE</i>	<i>100</i>	
<i>Campbell</i>	<i>Canucks</i>	<i>2000</i>	

*Top query returns Campbell, whereas bottom query returns nothing.*

d)

```
select distinct recipientName from donations
where recipientName not in
  (select recipientName from donations
   where organizationType = tobacco)
```

vs

```
select distinct recipientName from donations A
where exists
  (select * from donations B
   where B.recipientName = A.recipientName
   and organizationType  $\neq$  tobacco)
```

*(2 points) No. Top query finds recipients who did not receive donations from any tobacco company, whereas the bottom query finds recipients who received at least one donation from a non-tobacco company.*

<i>recipientName</i>	<i>OrganizationType</i>	<i>...</i>
<i>John</i>	<i>tobacco</i>	
<i>John</i>	<i>entertainment</i>	
<i>Mary</i>	<i>sports</i>	

*Top query returns Mary, while bottom query returns John and Mary.*

**--- The End ---**