1. $if\ F = \emptyset\ then\ E_1\ else\ E_2$

$$\pi_A(E_1) - \pi_A((\pi_{(\ )}(F) \times E_1)) \cup \pi_B(\pi(F) \times E_2))$$

where $A = e1.*$ and $B = e2.*$.

2.

a. Query:

```
select c.cname,
        case when exists(
                select 1
                from worksFor w, hasManager m
                where w.cname = c.cname and w.salary > 55000
                        and m.eid = w.pid and m.mid in (
                                select k.pid2
                                from Knows k
                                where k.pid1 = w.pid
                        )
                ) then '1'
            else '0'
        end as t
from Company c;

Simplifying,

select c.cname, '1' as t
from Company c
where exists (
        select 1
        from worksFor w, hasManager m, Knows k
        where k.pid1 = w.pid and k.pid2 = m.mid
                and m.eid = w.pid  and w.cname = c.cname and w.salary >
55000
)
union
select c.cname, '0' as t
from Company c
where not exists (
        select 1
        from worksFor w, hasManager m, Knows k
        where k.pid1 = w.pid and k.pid2 = m.mid
                and m.eid = w.pid  and w.cname = c.cname and w.salary >
55000
);
```

Converting it to RA SQL:

```
select distinct q1.cname, 1 as t
from (
        select c.*
        from Company c, worksFor w, hasManager hm, Knows k
        where k.pid1 = w.pid and k.pid2 = hm.mid
                and hm.eid = w.pid  and w.cname = c.cname and w.salary >
55000
) q1
union
        select distinct q2.cname, 0 as t
        from (
                select c.*
                from Company c
                except
                select c.*
                from Company c, worksFor w, hasManager hm, Knows k
                where k.pid1 = w.pid and k.pid2 = hm.mid
                        and hm.eid = w.pid  and w.cname = c.cname and w.salary >
55000
) q2;
```

b. RA expression:

$\pi_{c.cname,t}(E_1) \cup \pi_{c.cname,t}(E_2)$

$where\ E_1 = \pi_{c.*}(\sigma_{k.pid1=w.pid \wedge k.pid2=hm.mid \wedge hm.eid=w.pid \wedge w.cname=c.cname \wedge w.salary>55000}(C \bowtie hM \bowtie K))$

$and\ E_2 = \pi_{c.*}(C - \sigma_{k.pid1=w.pid \wedge k.pid2=hm.mid \wedge hm.eid=w.pid \wedge w.cname=c.cname \wedge w.salary>55000}(C \bowtie hM \bowtie K))$

3.

a. exists (union)

Translating the **exists** in where clause:

```
select L1 (r1, …, rn)
from (R1 r1, … Rn rn), (S1 s1, …, S1 sm)
where C1 (r1, …, rn) and C2 (s1, …, sm, r1, …, rn)
union
select L1 (r1, …, rn)
from (R1 r1, … Rn rn), (S1 s1, …, S1 sm)
where C1 (r1, …, rn) and C3 (s1, …, sm, r1, …, rn)
```

RA expression:

$$\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn)\,\wedge\,C2(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right) \cup$$
$$\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn)\,\wedge\,C3(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right)$$

b. exists (intersect)

```
select L1(r1,...,rn)
from R1 r1, ..., Rn rn
where C1(r1,...,rn) and exists (select s1.*, s2.*, ..., sm.*
                                from S1 s1,..., S1 sm
                                where C2(s1,...,sm, r1,...,rn)
                                intersect
                                select s1.*, s2.*, ..., sm.*
                                from S1 s1,..., S1 sm
                                where C3(s1,...,sm, r1,...,rn));
```

Translating **exists** in where clause:

```
select distinct L1^q(r1,...,rn)
from (
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
        where C1(r1,...,rn) and C2(s1,...,sm, r1,...,rn)
        intersect
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
        where C1(r1,...,rn) and C3(s1,...,sm, r1,...,rn)
) q;
```

RA expression:

$$\pi_{L1(r1,\ldots,rn)}(\pi_{L1(r1,\ldots,rn)}(\sigma_{C1(r1,\ldots,rn)\,\wedge\,C2(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)) \cap (\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn)\,\wedge\,C3(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right))$$

c. exists (except)

```
select distinct L1(r1,...,rn)
from R1 r1, ..., Rn rn
where C1(r1,...,rn) and exists (select s1.*, s2.*, ..., sm.*
                                from S1 s1,..., S1 sm
                                where C2(s1,...,sm, r1,...,rn)
                                except
                                select s1.*, s2.*, ..., sm.*
```

```
                    from S1 s1,..., S1 sm
                    where C3(s1,...,sm, r1,...,rn));
```

Translating the exists in where clause:

```
select distinct L1^q(r1,...,rn)
from (
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
        where C1(r1,...,rn) and C2(s1,...,sm, r1,...,rn)
        except
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
        where C1(r1,...,rn) and C3(s1,...,sm, r1,...,rn)
) q;
```

RA expression:

$$\pi_{L1(r1,\dots,rn)}(\pi_{L1(r1,\dots,rn)}(\sigma_{C1(r1,\dots,rn)\wedge C2(s1,\dots,sm,r1,\dots,rn)}(R_1 \times \dots \times R_n \times S_1 \times \dots \times S_m)) - (\pi_{L1(r1,\dots,rn)}\left(\sigma_{C1(r1,\dots,rn)\wedge C3(s1,\dots,sm,r1,\dots,rn)}(R_1 \times \dots \times R_n \times S_1 \times \dots \times S_m)\right))$$

d. not exists (union)

```
select L1(r1,...,rn)
from R1 r1, ..., Rn rn
where C1(r1,...,rn) and not exists (select s1.*, s2.*, ..., sm.*
                    from S1 s1,..., S1 sm
                    where C2(s1,...,sm, r1,...,rn)
                    union
                    select s1.*, s2.*, ..., sm.*
                    from S1 s1,..., S1 sm
                    where C3(s1,...,sm, r1,...,rn))
```

Translating the **not exists** in the where clause:

```
select distinct L1^q(r1,...,rn)
from (
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn
        where C1(r1,...,rn)
        except
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
        where C1 (r1, …, rn) and C2 (s1, …, sm, r1, …, rn)
```

```
        union
        select L1 (r1, …, rn)
        from (R1 r1, … Rn rn), (S1 s1, …, S1 sm)
        where C1 (r1, …, rn) and C3 (s1, …, sm, r1, …, rn)
) q;
```

RA Expression:

$$\pi_{L1(r1,\ldots,rn)}(\pi_{L1(r1,\ldots,rn)}(\sigma_{C1(r1,\ldots,rn)}(R_1 \times \ldots \times R_n)) - (\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn) \wedge C2(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right) \cup \left(\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn) \wedge C3(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right)\right))$$

e. not exists (intersect)

```
select L1(r1,...,rn)
from R1 r1, ..., Rn rn
where C1(r1,...,rn) and not exists (select s1.*, s2.*, ..., sm.*
                                    from S1 s1,..., S1 sm
                                    where C2(s1,...,sm, r1,...,rn)
                                    intersect
                                    select s1.*, s2.*, ..., sm.*
                                    from S1 s1,..., S1 sm
                                    where C3(s1,...,sm, r1,...,rn))
```

Translating the **not exists** in the where clause:

```
select distinct L1^q2(r1,...,rn)
from (
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn
        where C1(r1,...,rn)
        except
        select distinct L1^q1(r1,...,rn)
        from (
                select L1(r1,...,rn)
                from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
                where C1(r1,...,rn) and C2(s1,...,sm, r1,...,rn)
                intersect
                select L1(r1,...,rn)
                from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
```

```
                where C1(r1,...,rn) and C3(s1,...,sm, r1,...,rn)
        ) q1
) q2
```

RA expression:

$$\pi_{L1(r1,\ldots,rn)}\Big(\Big(\pi_{L1(r1,\ldots,rn)}\Big(\sigma_{C1(r1,\ldots,rn)}(R_1 \times \ldots \times R_n)\Big)\Big) - (\pi_{L1(r1,\ldots,rn)}\Big(\pi_{L1(r1,\ldots,rn)}\Big(\sigma_{C1(r1,\ldots,rn) \wedge C2(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\Big)\Big) \cap \pi_{L1(r1,\ldots,rn)}\Big(\sigma_{C1(r1,\ldots,rn) \wedge C3(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\Big)\Big)\Big)$$

f. not exists (except)

```
select L1(r1,...,rn)
from R1 r1, ..., Rn rn
where C1(r1,...,rn) and not exists (select s1.*, s2.*, ..., sm.*
                        from S1 s1,..., S1 sm
                        where C2(s1,...,sm, r1,...,rn)
                        except
                        select s1.*, s2.*, ..., sm.*
                        from S1 s1,..., S1 sm
                        where C3(s1,...,sm, r1,...,rn))
```

Translating the **not exists** in the where clause:

```
select distinct L1^q2(r1,...,rn)
from (
        select L1(r1,...,rn)
        from R1 r1, ..., Rn rn
        where C1(r1,...,rn)
        except
        select distinct L1^q1(r1,...,rn)
        from (
                select L1(r1,...,rn)
                from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
                where C1(r1,...,rn) and C2(s1,...,sm, r1,...,rn)
                except
                select L1(r1,...,rn)
                from R1 r1, ..., Rn rn, S1 s1,..., S1 sm
                where C1(r1,...,rn) and C3(s1,...,sm, r1,...,rn)
        ) q1
) q2
```

RA Expression:

$$\pi_{L1(r1,\ldots,rn)}(\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn)}(R_1 \times \ldots \times R_n)\right) - \pi_{L1(r1,\ldots,rn)}(\pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn) \wedge C2(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right) - \pi_{L1(r1,\ldots,rn)}\left(\sigma_{C1(r1,\ldots,rn) \wedge C3(s1,\ldots,sm,r1,\ldots,rn)}(R_1 \times \ldots \times R_n \times S_1 \times \ldots \times S_m)\right)))$$

4. To prove,

   $\pi_{a,d}(R \bowtie_{c=d} S) = \pi_{a,d}(\pi_{a,c}(R) \bowtie_{c=d} \pi_d(S))$

   Given, R (a, b, c) and S (d, e)

   Consider three attributes (a, b, c) that belong to R and (d) that belong to S and c = d.

$$= \pi_{a,d}(R \bowtie_{c=d} S)$$
$$= \{a, d | \exists a\ \exists b\ \exists c\left(R(a,b,c)\right) \wedge\ \exists d(S(d,e)) \neq \emptyset\}$$
$$= \{a, d | \exists a\ \exists c\ (R(a,b,c)\ \wedge \exists d\ (S(d,e)) \wedge c = d)\}$$
$$= \{a, d | \{a, c | R(a,b,c)\} \wedge \{d | S(d,e)\} \wedge c = d\}$$
$$=\ \pi_{a,d}\left(\pi_{a,c}(R) \bowtie_{c=d} \pi_d(S)\right)$$

   Attribute, c exists in R and d in S, such that the join condition is satisfied.

5.

   b. The RA SQL query can be expressed in standard notation as below:

$$\pi_{c.cname,c.headquarter}(C \bowtie_{w.cname=c.cname} (\sigma_{w.salary>40000}(W)) \bowtie_{w.pid=p.pid} (\sigma_{p.city='Cupertino'}(P)))$$

   c. (c1) Pushing projections over Joins,

$$\pi_{c.cname,c.headquarter}(C \bowtie_{w.cname=c.cname} (\pi_{w.pid,w.cname}\left(\sigma_{w.salary>40000}(W)\right)) \bowtie_{w.pid=p.pid} (\pi_{p.pid}(\sigma_{p.city='Cupertino'}(P)))\ )$$

   (c2) Using semi-joins rule,

$$\pi_{c.cname,c.headquarter}(C \bowtie_{w.cname=c.cname} (\pi_{w.pid,w.cname}\left(\sigma_{w.salary>40000}(W)\right)) \ltimes \sigma_{p.city='Cupertino'}(P)\ )$$

   Therefore, the optimized RA expression is:

$$\pi_{c.cname,c.headquarter}(C \bowtie_{w.cname=c.cname} (\pi_{w.pid,w.cname}\left(\sigma_{w.salary>40000}(W)\right)) \ltimes \sigma_{p.city='Cupertino'}(P))$$

6.

b. The RA SQL query can be expressed in standard notation as below:

$$\pi_{c.cname,c.headquarter}(\pi_{c.cname,c.headquarter}(C \bowtie_{c.cname=w.cname} W) - (\pi_{c.cname,c.headquarter}\left(\pi_{c.cname,c.headquarter,w.pid}\left(\sigma_{w.salary\geq 50000}(C \bowtie_{c.cname=w.cname} W)\right) - \pi_{c.cname,c.headquarter,w.pid}(\sigma_{ps.skill="Programming"}\left(C \bowtie W \bowtie_{w.pid=ps.pid} pS\right)\right)) \cup (\pi_{c.cname,c.headquarter,w.pid}(\sigma_{w.salary\geq 50000}(C \bowtie_{c.cname=w.cname} W)) - \pi_{c.cname,c.headquarter,w.pid}(\sigma_{ps.skill="Networks"}(C \bowtie W \bowtie_{w.pid=ps.pid} pS)))))$$

c. (c1) Pushing selection over joins,

$$\pi_{c.cname,c.headquarter}(\pi_{c.cname,c.headquarter}(C \bowtie_{c.cname=w.cname} W) - (\pi_{c.cname,c.headquarter}\left(\pi_{c.cname,c.headquarter,w.pid}\left(\sigma_{w.salary\geq 50000}(C \bowtie_{c.cname=w.cname} W)\right) - \pi_{c.cname,c.headquarter,w.pid}\left(C \bowtie W \bowtie_{w.pid=ps.pid} (\sigma_{ps.skill="Programming"}(pS)\right)\right)) \cup (\pi_{c.cname,c.headquarter,w.pid}(C \bowtie_{c.cname=w.cname} (\sigma_{w.salary\geq 50000}(W))) - \pi_{c.cname,c.headquarter,w.pid}(C \bowtie W \bowtie_{w.pid=ps.pid} (\sigma_{ps.skill="Networks"}(pS))))))$$

(c2) Converting natural joins to Semi-joins,

$$\pi_{c.cname,c.headquarter}(\pi_{c.cname,c.headquarter}(C \ltimes W) - (\pi_{c.cname,c.headquarter}\left(\pi_{c.cname,c.headquarter,w.pid}\left(\sigma_{w.salary\geq 50000}(C \ltimes W)\right)\right) - \pi_{c.cname,c.headquarter,w.pid}\left(C \ltimes W \ltimes (\sigma_{ps.skill="Programming"}(pS)\right)\right)) \cup (\pi_{c.cname,c.headquarter,w.pid}(C \ltimes (\sigma_{w.salary\geq 50000}(W))) - \pi_{c.cname,c.headquarter,w.pid}(C \ltimes W \ltimes (\sigma_{ps.skill="Networks"}(pS)))))$$

7.

b. The RA SQL query can be expressed in standard notation as below:

$$\pi_{p.pid,p.city}(\pi_{p.*}(P) - \pi_{p.*}(\sigma_{w.cname='Amazon'}(\pi_{p.*}(P) \bowtie_{p.pid=k.pid1} K \bowtie_{k.pid2=p1.pid} P_1 \bowtie_{p1.pid=w.pid} W))$$

c. (c1) Eliminating attributes,

$$\pi_{p.pid,p.city}(\pi_{p.pid,p.city}(P) - \pi_{p.pid,p.city}(\sigma_{w.cname='Amazon'}(\pi_{p.pid,p,city}(P) \bowtie_{p.pid=k.pid1} K \bowtie_{k.pid2=p1.pid} P_1 \bowtie_{p1.pid=w.pid} W))$$

(c2) Pushing selections over joins,

$$\pi_{p.pid,p,city}(\pi_{p.pid,p,city}(P) - \pi_{p.pid,p,city}(\pi_{p.pid,p,city}(P) \bowtie_{p.pid=k.pid1} K \bowtie_{k.pid2=p1.pid} P_1 \bowtie_{p1.pid=w.pid} \sigma_{w.cname='Amazon'}(W))$$

8.

b. The RA SQL query can be expressed in standard notation as below:

$$\pi_{c.cname}\Big(\big(\big(\pi_{w.cname}(W \bowtie_{w.cname=cl.cname} cL)$$
$$- \pi_{w.cname}(\sigma_{ps.skill='Programming'}(C \bowtie_{c.cname=w.cname} W \bowtie_{w.pid=ps.pid} pS))\big)$$
$$\cap \big(\pi_{w.cname}(W \bowtie_{w.cname=cl.cname} cL)\big)$$
$$- \pi_{w.cname}(\sigma_{ps.skill='AI'}(C \bowtie_{c.cname=w.cname} W \bowtie_{w.pid=ps.pid} pS))\big)$$
$$- \pi_{w.cname}(\sigma_{cl.city='Sunnyvale'}(W \bowtie_{w.cname=cl.cname} cL))\Big)$$

c. (c1) Pushing selections over joins

$$\pi_{c.cname}\Big(\big(\big(\pi_{w.cname}(W \bowtie_{w.cname=cl.cname} cL)$$
$$- \pi_{w.cname}(C \bowtie_{c.cname=w.cname} W \bowtie_{w.pid=ps.pid} (\sigma_{ps.skill='Programming'}(pS)))\big)$$
$$\cap \big(\pi_{w.cname}(W \bowtie_{w.cname=cl.cname} cL)\big)$$
$$- \pi_{w.cname}(\sigma_{ps.skill='AI'}(C \bowtie_{c.cname=w.cname} W \bowtie_{w.pid=ps.pid} pS))\big)$$
$$- \pi_{w.cname}\big(W \bowtie_{w.cname=cl.cname} (\sigma_{cl.city='Sunnyvale'}(cL))\big)\Big)$$

(c2) Using semi-joins

$$\pi_{c.cname}\Big(\big(\big(\pi_{w.cname}(W \ltimes cL) - \pi_{w.cname}(C \ltimes W \ltimes (\sigma_{ps.skill='Programming'}(pS)))\big)$$
$$\cap \big(\pi_{w.cname}(W \ltimes cL)\big) - \pi_{w.cname}(\sigma_{ps.skill='AI'}(C \ltimes W \ltimes pS))\big) - \pi_{w.cname}\big(W \ltimes (\sigma_{cl.city='Sunnyvale'}(cL))\big)\Big)$$

9.

b. The RA SQL query can be expressed in standard notation as below:

$$\pi_{ps.pid}(\sigma_{hm.eid \neq hm1.eid \wedge ps.skill='AI'}\big(pS \bowtie_{ps.pid=hm.mid} hM \bowtie_{hm.mid=hm1.mid} hM_1\big))$$

c. (c1) Cascading selections

$$\pi_{ps.pid}(\sigma_{hm.eid \neq hm1.eid}\big(\sigma_{ps.skill='AI'}(pS) \bowtie_{ps.pid=hm.mid} hM \bowtie_{hm.mid=hm1.mid} hM_1\big))$$

(c2) Attribute elimination,

$$\pi_{ps.pid}(\sigma_{hm.eid \neq hm1.eid}\big(\pi_{ps.pid}(\sigma_{ps.skill='AI'}(pS)) \bowtie_{ps.pid=hm.mid} hM \bowtie_{hm.mid=hm1.mid} hM_1\big))$$