

A5

Problem 1

The original query:

```
SELECT w.wid, w.wname
FROM Westerosi w, Predecessor p
Where w.wid = p.succid
AND w.wid NOT IN(
    SELECT w.wid
    FROM Westerosi w, Predecessor p, Ofhouse oh1, OfHouse oh2
    WHERE w.wid = p.succid AND oh1.wid = p.succid AND oh2.wid = p.predid
    AND oh1.wages <= oh2.wages);
```

We can convert NOT IN to an except:

```
SELECT w.wid, w.wname
FROM Westerosi w, Predecessor p
WHERE w.wid = p.succid
EXCEPT
SELECT w.wid, w.wname
FROM Westerosi w, Predecessor p, Ofhouse oh1, OfHouse oh2
WHERE w.wid = p.succid AND oh1.wid = p.succid AND oh2.wid = p.predid
    AND oh1.wages <= oh2.wages;
```

Which yields the following relational algebra expression

$$\pi_{wid, wname}(\sigma_{wid=succid}(W \times P)) - \pi_{wid, wname}(\sigma_{wid=succid \wedge oh1.wid=p.succid \wedge oh2.wid=p.predid \wedge oh1.wages \leq oh2.wages}(W \times P \times Oh1 \times Oh2))$$

I will optimize the query on the left of the "-" separately from the right for ease, and then put them back together at the end.

Left:

$$\begin{aligned} & \pi_{wid, wname}(\sigma_{wid=succid}(W \times P)) \\ &= \pi_{wid, wname}(W \bowtie_{wid=succid} P) \\ &= \pi_{wid, wname}(\pi_{wid, wname}(W) \bowtie_{wid=succid} \pi_{succid}(P)) \\ &= \pi_{wid, wname}(W \bowtie \pi_{(succid:wid)}(P)) \end{aligned}$$

Right:

$$\begin{aligned} & \pi_{wid, wname}(\sigma_{wid=succid \wedge oh1.wid=p.succid \wedge oh2.wid=p.predid \wedge oh1.wages \leq oh2.wages}(W \times P \times Oh1 \times Oh2)) \\ &= \pi_{w.wid, w.wname}(\sigma_{w.wid=p.succid}(\sigma_{oh1.wid=p.succid}(\sigma_{oh2.wid=p.predid}(\sigma_{oh1.wages \leq oh2.wages}(W \times P \times Oh1 \times Oh2))))) \\ &= \pi_{w.wid, w.wname}(W \bowtie_{w.wid=p.succid} (P \bowtie_{oh1.wid=p.succid \wedge oh2.wid=p.predid} (Oh1 \bowtie_{oh1.wages \leq oh2.wages} Oh2))) \\ &= \pi_{w.wid, w.wname}(W \bowtie \pi_{p.succid:wid}(P \bowtie_{oh1.wid=p.succid \wedge oh2.wid=p.predid} (Oh1 \bowtie_{oh1.wages \leq oh2.wages} Oh2))) \\ &= \pi_{w.wid, w.wname}(W \bowtie \pi_{p.succid:wid}(P \bowtie \pi_{oh1.wid:succid, oh2.wid:predid}(Oh1 \bowtie_{oh1.wages \leq oh2.wages} Oh2))) \end{aligned}$$

So, we have the following optimized RA expression:

$$= \pi_{wid, wname}(W \bowtie \pi_{(succid:wid)}(P)) - \pi_{w.wid, w.wname}(W \bowtie \pi_{(p.succid:wid)}(P \bowtie \pi_{oh1.wid:succid, oh2.wid:predid}(Oh1 \bowtie_{oh1.wages \leq oh2.wages} Oh2)))$$

Resulting in the optimized query:

```
SELECT w.wid, w.wname
FROM Westerosi w NATURAL JOIN (SELECT p.succid as wid FROM Predecessor p) as s
EXCEPT
SELECT w.wid, w.wname
FROM Westerosi w NATURAL JOIN
  (SELECT p.succid as wid
   FROM Predecessor p NATURAL JOIN
     (SELECT oh2.wid as predid, oh1.wid as succid
      FROM OfHouse oh1 JOIN OfHouse oh2 ON (oh1.wages <= oh2.wages)
     ) as q
  ) as s;
```

Problem 2

```
SELECT H.HNAME, H.KINGDOM
FROM HOUSE H
WHERE H.HNAME in
    (SELECT OH.HNAME
     FROM OFHOUSE OH
     WHERE OH.WAGES < 60000
          AND OH.WID = SOME
            (SELECT WS.WID
             FROM WESTEROSISKILL WS
             WHERE WS.SKILL = 'Archery'));
```

is equivalent to:

```
SELECT DISTINCT H.HNAME, H.KINGDOM
FROM HOUSE H, OFHOUSE OH, WESTEROSISKILL WS
WHERE H.HNAME = OH.HNAME
    AND OH.WAGES < 60000
    AND OH.WID = WS.WID
    AND WS.SKILL = 'Archery';
```

Which has the following RA form:

$$\begin{aligned} & \pi_{h.name, h.kingdom}(\sigma_{h.hname=oh.hname \wedge oh.wages < 60000 \wedge oh.wid=ws.wid \wedge ws.skill='Archery'}(H \times OH \times WS)) \\ &= \pi_{h.name, h.kingdom}(\sigma_{h.hname=oh.hname}(\sigma_{oh.wages < 60000}(\sigma_{oh.wid=ws.wid}(\sigma_{oh.wid=ws.wid}(\sigma_{ws.skill='Archery'}(H \times OH \times WS)))))) \\ &= \pi_{h.hname, h.kingdom}(H \bowtie \pi_{oh.hname}(\sigma_{oh.wages < 60000}(OH \bowtie \pi_{ws.wid}(\sigma_{ws.skill='Archery'}(WS)))) \end{aligned}$$

Resulting in the optimized query:

```
SELECT DISTINCT H.HNAME, H.KINGDOM
FROM HOUSE H NATURAL JOIN (
  SELECT OH.HNAME
  FROM OFHOUSE OH NATURAL JOIN (
    SELECT WS.WID
    FROM WESTEROSISKILL WS
    WHERE WS.SKILL='Archery'
  ) AS B
  WHERE OH.WAGES < 60000
) AS A;
```

Problem 3

the original query:

```
SELECT Distinct w.wid
FROM Westerosi w
WHERE w.wloaction = 'Winterfell' AND EXISTS(
    SELECT 1
    FROM Ofhouse h, WesterosiSkill w1
    WHERE w.wid = h.wid
        AND w.wid = w1.wid
        AND h.wages = 50000
        AND NOT w1.skill= 'Swordsmanship'
);
```

Can be translated to ra sql:

```
SELECT Distinct w.wid
FROM Westerosi w, Ofhouse h, WesterosiSkill ws
WHERE w.wlocation = 'Winterfell'
    AND w.wid = h.wid
    AND w.wid = w1.wid
    AND h.wages = 50000
    AND NOT w1.skill= 'Swordsmanship';
```

And is equivalent to the following RA expression

$$\begin{aligned} & \pi_{w.wid}(\sigma_{w.wlocation='Winterfell' \wedge w.wid=h.wid \wedge w.wid=ws.wid \wedge h.wages=50000 \wedge ws.skill \neq 'Swordsmanship'}(W \times H \times WS)) \\ &= \pi_{w.wid}(\sigma_{w.wlocation='Winterfell'}(\sigma_{w.wid=h.wid}(\sigma_{w.wid=ws.wid}(\sigma_{h.wages=50000}(\sigma_{ws.skill \neq 'Swordsmanship'}(W \times H \times WS)))))) \\ &= \pi_{w.wid}(\sigma_{w.wlocation='Winterfell'}(W \bowtie_{h.wid=ws.wid} \pi_{h.wid}(\pi_{h.wid}(\sigma_{h.wages=50000}(H) \bowtie_{h.wid=ws.wid} \pi_{ws.wid}(\sigma_{ws.skill \neq 'Swordsmanship'}(WS)))))) \\ &= \pi_{w.wid}(\sigma_{w.wlocation='Winterfell'}(W \bowtie \pi_{h.wid}(\pi_{h.wid}(\sigma_{h.wages=50000}(H) \bowtie \pi_{ws.wid}(\sigma_{ws.skill \neq 'Swordsmanship'}(WS)))))) \\ &= \pi_{w.wid}(\sigma_{w.wlocation='Winterfell'}(W \bowtie \pi_{h.wid}(\pi_{h.wid}(\sigma_{h.wages=50000}(H) \bowtie \pi_{ws.wid}(\sigma_{ws.skill \neq 'Swordsmanship'}(WS)))))) \end{aligned}$$

Which yields the following optimized RA SQL query

```
SELECT DISTINCT w.wid
FROM Westerosi w NATURAL JOIN(
  SELECT h.wid
  FROM OfHouse h NATURAL JOIN (
    SELECT ws.wid
    FROM WesterosiSkill ws
    WHERE ws.skill <> 'Swordsmanship'
  ) as skill
  WHERE h.wages = 50000
) as wages
WHERE w.wlocation = 'Winterfell';
```

Problem 4

```
SELECT W.WID
FROM WESTEROSI W
WHERE NOT EXISTS (
    SELECT 1
    FROM HOUSEALLYREGION HA
    WHERE HA.REGION = 'IronIslands'
        AND HA.HNAME NOT IN (
            SELECT H.HNAME
            FROM OFHOUSE H
            WHERE H.WID = W.WID
                AND H.WID in (
                    SELECT WS.WID
                    FROM WESTEROSISKILL WS
                    WHERE WS.SKILL = 'Archery')));
```

find the wids of westerosi with archery as a skill and who belong to the only house who is equivalent to:

```
SELECT W.WID
FROM WESTEROSI W
WHERE NOT EXISTS (
    SELECT HA.HNAME
    FROM HOUSEALLYREGION HA
    WHERE HA.REGION = 'IronIslands'
    EXCEPT
    SELECT H.HNAME
    FROM OFHOUSE H, WESTEROSISKILL WS
    WHERE H.WID = W.WID AND H.WID = WS.WID AND WS.SKILL = 'Archery');
```

which is equivalent to

```

SELECT W.WID
FROM (
    SELECT W.WID
    FROM WESTEROSI W
    EXCEPT
    (SELECT HA.HNAME, W.WID
    FROM HOUSEALLYREGION HA, WESTEROSI W
    WHERE HA.REGION = 'IronIslands'
    EXCEPT
    SELECT H.HNAME, WS.WID
    FROM OFHOUSE H, WESTEROSISKILL WS
    WHERE H.WID = WS.WID AND WS.SKILL = 'Archery'
    ) AS W) as W;

```

Which is written in RA as:

$$\begin{aligned}
 & \pi_{wid}(\pi_{w.wid}(W) - \pi_{w.wid}(\pi_{ha.hname,w.wid}(\sigma_{ha.region='IronIslands'}(HA \times W)) - \pi_{oh.hname,oh.wid}(\sigma_{oh.wid=ws.wid \wedge ws.skill='Archery'}(OH \times WS)))) \\
 &= \pi_{wid}(\pi_{w.wid}(W) - \pi_{w.wid}(\pi_{ha.hname,w.wid}(W \bowtie \sigma_{ha.region='IronIslands'}(HA)) - \pi_{oh.hname,oh.wid}(OH \bowtie \pi_{ws.wid}(\sigma_{ws.skill='Archery'}(WS))))
 \end{aligned}$$

- reduced size of both cross products.

Which is equivalent to the following optimized ra sql:

```
WITH IRONISLANDS AS (SELECT HA.HNAME FROM HOUSEALLYREGION HA WHERE HA.REGION='IronIslands'),
    ARCHERS AS (SELECT WS.WID FROM WESTEROSISKILL WS WHERE WS.SKILL='Archery')
SELECT W.WID
FROM (
    SELECT W.WID
    FROM WESTEROSI W
    EXCEPT
    SELECT W.WID
    FROM (
        SELECT HA.HNAME,W.WID
        FROM WESTEROSI W NATURAL JOIN IRONISLANDS HA
        EXCEPT
        SELECT OH.HNAME, OH.WID
        FROM OFHOUSE OH NATURAL JOIN ARCHERS
    ) AS W
) AS W;
```

Problem 5

Find the wname and wlocation of each westerosi whose wages are strictly greater than 50000 and belongs to a house which has a kingdom in kings landing and has some skill.

```
SELECT DISTINCT W.WNAME, W.WLOCATION
FROM WESTEROSI W, OFHOUSE OH, HOUSEALLY HA, WESTEROSISKILL WS
WHERE W.WID=OH.WID AND WS.WID = W.WID AND OH.HNAME = HA.HNAME AND HA.REGION='KingsLanding';
```

Which can be written in relational algebra as

$$\begin{aligned} & \pi_{w.wname, w.wlocation}(\sigma_{w.wid=oh.wid \wedge ws.wid=w.wid \wedge oh.hname=ha.hname \wedge ha.region='KingsLanding'}(W \times OH \times HA \times WS)) \\ &= \pi_{w.wname, w.wlocation}(\sigma_{w.wid=oh.wid}(\sigma_{ws.wid=w.wid}(\sigma_{oh.hname=ha.hname}(\sigma_{ha.region='KingsLanding'}(W \times OH \times HA \times WS)))))) \\ &= \pi_{w.wname, w.wlocation}(W \bowtie \pi_{ws.wid}(WS \bowtie \pi_{oh.wid}(\sigma_{oh.wages > 50000}(OH \bowtie \pi_{ha.hname}(\sigma_{ha.region='KingsLanding'}(HA)))))) \end{aligned}$$

which can be converted into sql:

```
WITH KINGSLANDING AS (SELECT HA.NAME FROM HOUSEALLYREGION HA WHERE HA.REGION='KingsLanding')
SELECT DISTINCT W.WNAME, W.WLOCATION
FROM WESTEROSI W NATURAL JOIN (
    SELECT WS.WID
    FROM WESTEROSI WS NATURAL JOIN(
        SELECT OH.WID
        FROM OFHOUSE OH NATURAL JOIN KINGSLANDING
        WHERE OH.WAGES > 50000
    ) AS A
) AS B;
```