Relational Algebra Practice- *Solutions*

November 23, 2015

Since we didn't get to cover relational algebra (RA) on a problem set, we're providing this notebook so you can get some practice in before the final exam. Solutions will be posted in a separate notebook- try doing these on your own first, then take a look at the solutions to check your understanding!

In particular, you should understand:

- How to go from SQL query $\rightarrow$ RA expression
- How to go from RA expression $\rightarrow$ SQL query
- How to optimize an RA expression by commuting operators

Note that some of the problems here will be slightly more involved than what would be on the exam!

Consider relations $R(A, B)$, $S(B, C)$, $T(C, D)$ and $U(D, E)$ for the below examples.

1 SQL $\rightarrow$ RA

Let's go through some examples where we'll translate SQL to Relational Algebra. For each of the below queries, translate them from SQL into RA.

1.1

```
SELECT DISTINCT *
FROM R
WHERE R.A = 2;
```

$\sigma_{A=2}$(R(A,B))

1.2

```
SELECT DISTINCT S.B
FROM S
WHERE S.C = 4;
```

$\Pi_B(\sigma_{C=4}(\text{S(B,C)}))$

1.3

```
SELECT DISTINCT R.A, S.C
FROM R, S
WHERE R.B = S.B;
```

$\Pi_{A,C}((\text{ R(A,B) }) \bowtie_B (\text{ S(B,C) }))$

1.4

```
SELECT DISTINCT R.A, T.D
FROM R, S, T
WHERE R.B = S.B AND S.C = T.C AND R.A = 2 AND S.B = 0;
```

$\Pi_{A,D}((\ (\ \sigma_{A=2}(\text{R(A,B)})\ )\ \bowtie_B\ (\ \sigma_{B=0}(\text{S(B,C)})\ )\ )\ \bowtie_C\ (\ \text{T(C,D)}\ ))$

1.5

```
SELECT DISTINCT R.A
FROM R
WHERE R.B = 0 OR R.B = 2;
```

$\Pi_A((\ \sigma_{B=0}(\text{R(A,B)})\ )\ \bigcup\ (\ \sigma_{B=2}(\text{R(A,B)})\ ))$

1.6

```
SELECT DISTINCT R.A
FROM R
WHERE R.B <> 2;
```

$\Pi_A((\ \text{R(A,B)}\ )\ \text{-}\ (\ \sigma_{B=2}(\text{R(A,B)})\ ))$

1.7

```
SELECT DISTINCT R.B, U.E
FROM R, S, T, U
WHERE R.B = S.B AND S.C = T.C AND T.D = U.D
  AND (S.C = 2 OR T.D = 4) AND U.D <> 2;
```

$\Pi_{B,E}((\ \text{R(A,B)}\ )\ \bowtie_B\ (\ (\ (\ (\ \sigma_{C=2}(\text{S(B,C)})\ )\ \bowtie_C\ (\ \text{T(C,D)}\ )\ )\ \bigcup\ (\ (\ \text{S(B,C)}\ )\ \bowtie_C\ (\ \sigma_{D=4}(\text{T(C,D)})\ )\ )\ )\ \bowtie_D\ (\ (\ \text{U(D,E)}\ )\ \text{-}\ (\ \sigma_{D=2}(\text{U(D,E)})\ )\ )\ ))$

2 RA $\rightarrow$ SQL

Now we'll go through some examples where we'll translate Relational Algebra to SQL

2.1

$\sigma_{B=0}(\Pi_B(S(B,C)))$

```
SELECT DISTINCT S.B
FROM S
WHERE S.B = 0;
```

2.2

$\Pi_{A,E}(\sigma_{A=2}(\sigma_{C=0}(R(A,B)\ \bowtie_B\ (S(B,C)\ \bowtie_C\ (T(C,D)\ \bowtie_D\ U(D,E))))))$

```
SELECT DISTINCT R.A, U.E
FROM R, S, T, U
WHERE R.B = S.B AND S.C = T.C AND T.D = U.D
  AND S.C = 0 and R.A = 2;
```

2.3

$\Pi_{A,C}(((\sigma_{B=0}(R(A,B))) \bowtie_B (S(B,C) \bowtie_C (\sigma_{C=0}(T(C,D))))$

```
SELECT DISTINCT R.A, T.C
FROM R, S, T
WHERE R.B = S.B AND S.C = T.C
  AND T.C = 0 AND R.B = 0;
```

2.4

$((\sigma_{A=2}(R(A,B))) \bigcup (\sigma_{A=4}(R(A,B)))) \bowtie_B ((\sigma_{C=2}(S(B,C))) - (\sigma_{B=1}(S(B,C))))$

```
SELECT DISTINCT R.A, R.B, S.C
FROM R, S
WHERE R.B = S.B AND (R.A = 2 OR R.A = 4) AND S.C = 2 AND S.B <> 1;
```

3 Optimization of RA Expressions

In this section, we'll optimize RA expressions, i.e. reduce the total IO cost of executing them.

3.1

$\Pi_D(T(C,D) \bowtie_D U(D,E))$

$(\Pi_D(\mathrm{T(C,D)})) \bowtie_D (\Pi_D(\mathrm{U(D,E)}))$

3.2

$\sigma_{A=2}(\Pi_{A,C}(R(A,B) \bowtie_B S(B,C))$

$\Pi_{A,C}((\sigma_{A=2}(\mathrm{R(A,B)})) \bowtie_B (\mathrm{S(B,C)}))$

3.3

$\sigma_{C=0}(\Pi_{A,C}(\sigma_{B=0}(((\mathrm{R(A,B)}) \bowtie_B (\mathrm{S(B,C)})) \bowtie_C (\mathrm{T(C,D)}))))$

$\Pi_{A,C}((\Pi_{A,C}(\sigma_{C=0}((\sigma_{B=0}(\mathrm{R(A,B)})) \bowtie_B (\sigma_{B=0}(\mathrm{S(B,C)}))))) \bowtie_C (\sigma_{C=0}(\mathrm{T(C,D)})))$

3.4

$\sigma_{C=0}(\Pi_C(\sigma_{D=2}(\sigma_{A=3}((\mathrm{R(A,B)}) \bowtie_B ((\mathrm{S(B,C)}) \bowtie_C (\mathrm{T(C,D)}))))))$

$(\Pi_C((\sigma_{A=3}(\mathrm{R(A,B)})) \bowtie_B (\sigma_{C=0}(\mathrm{S(B,C)})))) \bowtie_C (\Pi_C(\sigma_{C=0}(\sigma_{D=2}(\mathrm{T(C,D)})))))$