

#####

Assignment # 6

OZKAN EMRE OZDEMIR – 2/28/2016

#####

1. $\{a, b, c\}$ is a relation that contains the tuples a, b, and c. In the following cases the tuples have arity of

1. Calculate the following:

a. $(\{1, 2, 3\} \cup \{5, 7, 11\}) \cap \{2, 4, 6, 8, 10\} = \{1, 2, 3, 5, 7, 11\} \cap \{2, 4, 6, 8, 10\} = \{2\}$

b. $(\{1, 2, 3\} \cap \{2, 4, 6, 8, 10\}) \cup (\{5, 7, 11\} \cap \{2, 4, 6, 8, 10\}) = \{2\}$

#####

2. Use formal notation to write an algebraic example of the following SQL:

a. SELECT Column1, Column3 FROM MyTable WHERE Column2 = Column3

$\pi_{\text{Column1, Column3}}(\sigma_{\text{Column2} = \text{Column3}}(\text{MyTable}))$

b. Reverse the order of projection and selection in your algebraic formulation. What happened?

$\sigma_{\text{Column2} = \text{Column3}}(\pi_{\text{Column1, Column3}}(\text{MyTable}))$

When we reverse the order of projection and selection, we project only C1 and C3.

#####

3. $\pi_{c1, c2}(\sigma_{\varphi1}(\sigma_{\varphi2}(\pi_{c1, c2, c3, c5}(R))))$

Where

- $\varphi1: C1 = C5;$
- $\varphi2: C5 = \text{"Test"};$
- $R: \text{MyTable};$

a. Write a SQL statement that declares the intent of the algebraic notation

SELECT C1, C2 FROM (SELECT C1, C2, C3, C5 FROM R WHERE C5 = 'Test') WHERE C1 = C5

b. Simplify the algebraic statement. Simplification means minimize the number of parentheses and terms.

$\pi_{c1, c2}(\sigma_{\varphi^3}(R))$ where $\varphi^3: C1 = 'Test'$;

#####

4. SELECT * FROM T1 JOIN T2 ON T1.C1 = T2.C1

a. Write out an equivalent in relational algebra using the join operator

$T1 \bowtie_{\varphi} T2$ where $\varphi : T1.C1 = T2.C1$

b. Write out an equivalent in relational algebra without using the join operator

$\sigma_{\varphi}(T1 \times T2)$ where $\varphi : T1.C1 = T2.C1$

#####

5. $\pi_{S.C1, R.C2}(\sigma_{\varphi^1}(R) \bowtie_{\varphi^2} S)$

where

• $\varphi^1 = (R.C2 = 'A')$

• $\varphi^2 = (R.C1 = S.C2)$

Write out equivalent SQL and test this SQL using relations R and S that you create for this example. The relations R and S in RelationalAlgebraAndSQL.pdf and RelationalAlgebraAndSQL.sql don't quite work because their column types do not match for this assignment.

SELECT S.C1, R.C2 FROM R JOIN S ON R.C1 = S.C2 WHERE R.C2 = 'A'

#####

6. Download, install, and setup the Hadoop VM according to the segment titled "Setup Virtual Machine". Take a screen shot of your virtual machine analogous to the screenshot in the slide titled "Setup VM (07)" except that your name should be typed into the last line of the console.

My current wifi connection is not strong enough to download the 2.4 GB file right now. I will complete it and send the screenshot via email tonight.

#####