Homework #13

Name _____ Sec ____

Questions:	Answers:

1. Use the relations given below to evaluate the relational expressions or explain why the expression is invalid.

a) $\sigma_{A<4}Q \times U$
b) $\pi_A\sigma_{B<A}Q$
c) $\pi_{BE}(R \times \pi_{CE}S)$
d) $Q \cup \pi_B R$
e) $\pi_{CDE}(\sigma_{B<2}R \times \rho_{C\leftarrow E}U) - S$

Q:

A	B
5	1
6	1
4	2
3	4

R:

B	C
1	4
2	4
2	5
3	6
3	9

S:

C	D	E
4	1	1
4	2	1
3	3	2
2	4	2

U:

C	D
1	2
2	4

2. Use the relations given above in problem 1 to evaluate the relational expressions or explain why the expression is invalid.

a) $\pi_E S \mid\times\mid U$

b) $Q \mid\times\mid R$

c) $\sigma_{B<2}Q \mid\times\mid R \mid\times\mid \rho_{C\leftarrow B}U$

3. Using the database instance below and for each query to the right, give: (i) a relational algebra expression; (ii) an SQL query*; and (iii) a Datalog query and rule**.

*Execute SQL queries against the SnoopyDatabase, and hand in screenshots of your queries and the results of running your queries. (For details about how to do this, see **http://students.cs.byu.edu/~cs236ta/sharedLib/homework/SQLite3Essentials.html**.)

**Enter rules and execute Datalog queries against the SnoopyDatabase and hand in screenshots of your queries and the results of running your queries. (Use the Datalog interpreter at: http://students.cs.byu.edu/~cs236ta/sharedLib/homework/DatalogInterpreter.html)

SNAP

StudentID	Name	Address	Phone
12345	C. Brown	12 Apple St.	555-1234
67890	L. Van Pelt	34 Pear Ave.	555-5678
22222	P. Patty	56 Grape Blvd.	555-9999
33333	Snoopy	12 Apple St.	555-1234

CR

Course	Room
CS101	Turing Aud.
EE200	25 Ohm Hall
PH100	Newton Lab.

CDH

Course	Day	Hour
CS101	M	9AM
CS101	W	9AM
CS101	F	9AM
EE200	Tu	10AM

a) List the names of students whose phone number is 555-1234.

Name
C. Brown
Snoopy

i)

b) Find the names and corresponding course numbers of all students who have a class in the Turing Aud.

Name	Course
C. Brown	CS101
L. Van Pelt	CS101
Snoopy	CS101

i)

c) Find the name and phone number of students taking any of the immediate prerequisites of CS120.

Name	Phone
C. Brown	555-1234
L. Van Pelt	555-5678
Snoopy	555-1234

i)

EE200	**W**	**1PM**
EE200	**Th**	**10AM**
PH100	**Tu**	**11AM**

CSG

Course	StudentID	Grade
CS101	**12345**	**A**
CS101	**67890**	**B**
EE200	**12345**	**C**
EE200	**22222**	**B+**
EE200	**33333**	**B**
CS101	**33333**	**A-**
PH100	**67890**	**C+**

CP

Course	Prerequisite
CS101	**CS100**
EE200	**EE005**
EE200	**CS100**
CS120	**CS101**
CS121	**CS120**
CS205	**CS101**
CS206	**CS121**
CS206	**CS205**

4, Let R and S be relations in a relational database. Prove that the join operator is commutative (i.e. Prove: R |× | S = S |× | R.)

Do your proof by showing, step by step, how to derive the right-hand side from the left-hand side. Use the formal definitions for relational operators and laws for predicate calculus. Justify each step.

5. Let R and S be relations in a relational database. Prove: If all attributes mentioned in condition C are in the schema of R, then $\sigma_C(R \times | S) = \sigma_C R \times | S$.

Example using the class database: $\sigma_{Name = \text{'Snoopy'}}(SNAP \times | CSG) = \sigma_{Name = \text{'Snoopy'}} SNAP \times | CSG$.

Do your proof by showing, step by step, how to derive the right-hand side from the left-hand side. Use the formal definitions for relational operators and laws for predicate calculus. Justify each step.

6. Use the laws of relational algebra proved in Problems 4 & 5 to optimize the following expressions. As part of the optimization, when the expression has multiple joins, order the joins left-to-right so that the expression will be optimal for the particular instance of the database in Problem #3. a) $\sigma_{\text{Prerequisite='CS100'}}\sigma_{\text{Grade='A'}}$ (CP |× | CSG) b) $\sigma_{\text{Name='C. Brown'}}\sigma_{\text{Room = 'Turing Aud.'}}\sigma_{\text{Day='F'}}$ $\sigma_{\text{Hour='9 AM'}}$ (SNAP |× | CSG |× | CR |× | CDH)	
7. Assuming a simple double-loop implementation of the join operator and a simple scan implementation for the selection operator, give the number of comparisons required to execute the following expressions for the database instance in Problem 3. a) the unoptimized expression in problem 6a b) your optimized expression for problem 6a c) the unoptimized expression in problem 6b d) your optimized expression for problem 6b Assume that the comparison operator executes in one microsecond. Give the time difference between the estimated execution time of e) the expressions in (a) and (b);	

f) the expressions in (c) and (d). As an interesting thought question, suppose the database instance were populated instead actual current data for BYU. With 30,000 students for the SNAP relation and whatever the numbers are for the other relations, what would the time differences be? (Think about this question, but you do not need to give an answer.)	