Your name:

Database Design: Practice Midterm 1
CMSC424

- Numbers in square brackets indicate points (total 50 points = 12.5% of the course grade).
- **Show your reasoning. Write partial solutions.** You will get a fair amount of the credit if I think that you know the concepts.
- Unless otherwise specified, a *Yes/No* answer without accompanying explanation will not get any points.
- Do not write anything other than your name above the line at top. Try to keep your answers within the allotted space.

1. **[10 pts]** Circle true or false: **1 pt** for correct answer. **0.5 pts** will be deducted for incorrect.

(i) SQL EXCEPT operation eliminates duplicates.	TRUE FALSE
(ii) FALSE AND UNKNOWN = FALSE.	TRUE FALSE
(iii) Dynamic SQL typically requires use of a language-specific preprocessor to compile/prepare SQL statements, whereas Embedded SQL doesn't.	TRUE FALSE
(iv) Using Prepared Statements leaves the system vulnerable to SQL Injection attacks.	TRUE FALSE
(v) In an E/R model, attributes can also be attached to relationships.	TRUE FALSE
(vi) Under the "multi-set/bag" semantics, {a, b} × {c, c} = {(a, b), (c, c), (a, c), (b, c)}	TRUE FALSE
(vii) UNKNOWN OR TRUE = UNKNOWN.	TRUE FALSE
(viii) On the instructor table: instructor(ID, name, dept_name, salary), ID (which is the primary key) can take the value of null if the ID of the instructor is unknown or undisclosed.	TRUE FALSE
(ix) Relational algebra is a non-procedural language.	TRUE FALSE
(x) The keyword 'where' in SQL maps to the σ operation in Relational Algebra.	TRUE FALSE

2. **[2 pts]** What is a good primary key for the following relation, that represents information about US presidents ? Explain your reasoning.

   `President(name, start_date, end_date, vice_president, preceded_by, succeeded_by)`

Your name:

3. **[3 pts]** Consider two create table statements: *(a) create table R (a integer primary key);*
(b) create table S (b integer primary key, c integer references R(a) on update set to null); .
What will happen when a tuple in R is updated or deleted?

4. **[3 pts]** Consider a view V defined on relation $R(a, b)$ where b is a numeric attribute, as follows:
create view V **as select** a, **sum**(b) **as** sumb **from** R **group by** a;
Discuss whether we should allow inserts, deletes, and updates into the view, and how to propagate them to underlying relation (the answer may be different for different operations).

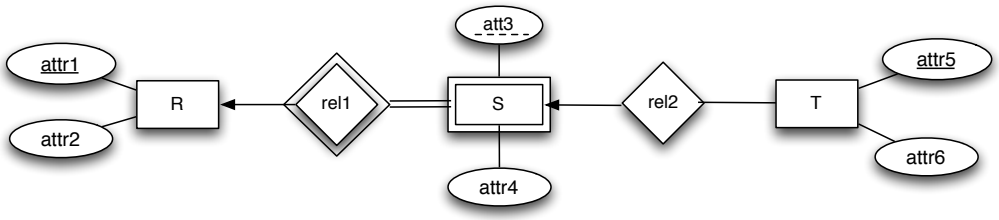
5. **[3 pts]** Briefly explain the notion of "triggers" in SQL, including an example of where it might be used.

Your name:

6. **[3 pts]** Given a table called T with three integer attributes called a, b, and c, write a SQL query that returns the sum of all the values of c for every unique combination of a and b.

7. **[3 pts]** What is "specialization" in E/R modeling? Illustrate through an example.

8. **[3 pts]** Which set of relations do we end up with when converting the converting the following E/R diagram to a relational schema?



Your name:

9. **[4 pts]** On the relation shown below, compute the answers to the two queries listed.

A	B	C
α	a	1
β	b	2
α	c	1
α	a	3
γ	c	3
γ	a	2
β	b	2
α	c	2
β	c	2

(i) select A, max(C) from R where C != null group by A;

(ii) select A from R r1 where exists
(select * from R r2 where r2.A = r1.A and r2.C < r1.C);

10. **[4 pts]** Given the relations, $R(a, b), S(b, c)$, where *a, b, c* are integer attributes, describe in words what are the results of following relational algebra expressions.

(a) $\sigma_{R.b \neq S.b}(R \times S) - (R \times S)$

(b) $R - \pi_{R.a,R.b}(\sigma_{R.b=S.b}(R \times S))$

Your name:

11. **[4 pts]** Give an example of a relation and a SQL query that uses a self-join over that relation that returns exactly 3 rows.

12. **[4 pts]** The following two queries are not equivalent (they don't always produce identical results) because of NULLs. Identify and explain the problem. Schemas are: $R(a,b,d), S(c,d)$. Assume a is the primary key for R.

Query I	**Query II**
select a **from** R **where** R.b = (**select count**(S.c) **from** S **where** R.d = S.d)	**select** a **from** R, S **where** R.d = S.d **group by** R.a **having** R.b = **count**(S.c);

Your name:

13. **[4 pts]** Suppose we have three relations $r(A, B)$, $s(B, C)$, and $t(B, D)$, with all attributes declared as not null. Consider the expressions

    - r natural left outer join (s natural left outer join t), and
    - (r natural left outer join s) natural left outer join t

    (a) Give instances of relations r, s and t such that in the result of the second expression, attribute C has a null value but attribute D has a non-null value.

    (b) Is the above pattern, with C null and D not null, possible in the result of the first expression? Explain why or why not.