# Example Examination

Allocated Time: 100 minutes
Maximum Points: 250

STUDENT NAME: _____

**General Instructions:**

- This test is a *closed book* exam (besides one cheat sheet).
- Write your answers in the space provided, use the back sides of sheets, if necessary.
- If a question seems vague, make reasonable assumptions, and answer the question under those assumptions. *Make sure to show your work and to state any assumptions you make!*
- The points for each question are as indicated and roughly indicate the number of minutes you should need to answer the question.
- Good Luck!

| PROBLEMS: | MAXIMUM SCORE | YOUR SCORE |
|---|---|---|
| problem 1: | 60 | |
| problem 2: | 30 | |
| problem 3: | 70 | |
| problem 4: | 30 | |
| problem 5: | 60 | |
| total: | 250 | |

**Problem 1: SQL Queries.** [60pts]

Consider the following database:

Suppliers(*sid* : *integer*, sname: string, address: string);
Parts(*pid* : *integer*, pname: string, color: string);
Catalog(*sid* : *integer*, *pid* : *integer*, cost: real);

Give an SQL query expression for each of the queries below. Whenever possible, use one expression per query without views and temporary variables.

(1) Find the snames of suppliers who supply every part.
(2) Find the pnames of parts supplied by Widget-Guys Suppliers and by no one else.
(3) Find the sid of suppliers who charge more for some part than the average cost of that one part (averaged out over all suppliers who supply that part).
(4) Find the sid of suppliers who supply only red parts.

**Problem 2: Constraints in SQL.** [30pts]

Consider the following relational schema and briefly answer the questions below:

Emp(*eid* : *integer*, ename: string, age: integer, salary: real);
Works(*eid* : *integer*, *did* : *integer*, pct-time: integer);
Dept(*did* : *integer*, budget; real, managerid: integer);

1. Define a table constraint on Dept that will ensure that all managers have age $> 30$.

2. Define an assertion on Dept that ensures that all managers have (age $> 30$).

3. Compare the assertion in #2 with the table constraint in #1 above, and explain which one is better and why?

**Problem 3: ER modeling.** [70pts]

Given below information about a university environment.

3.1. Draw an ER diagram for each of the statements below that most appropriately takes into account *the* assertion given in that one statement. Use as many different types of ER constructs, such as, inheritance, aggregation, keys, 1-1 versus 1-n relationships, relationships with different arity, etc., whenever possible, represent the semantics of the problem. Indicate if a given assertion cannot be captured by ER constructs, and/or if additional constraints or functional dependencies are necessary to capture them. [45pts]

- A professor always has a unique name and an address. A professor may teach none, one or more courses in a given term. All courses are taught by only one professor.

- A student can be both a university employee and also a student. A student has a name and a gpa, and a university employee has a name and a salary.

- A student is either a graduate student or an undergraduate student, but not both. An undergraduate student has an MQP and an IQP project, whereas a graduate student has a major research area.

- Only graduate students can be teaching assistants, whereas both graduate and undergraduate students can be graders.

- All teaching assistants must be graduate students.

3.2. Now combine the above smaller ER diagrams from 3.1 to form one complete ER design that most appropriately takes into account *all* assertions. [25pts].

**Problem 4: Mapping from ER models to SQL.** [30pts]

Give the SQL table definitions for problem 3.1.c., explaining the use of keys, foreign keys, etc.. Justify your particular choice for doing this mapping (versus alternate options).

**Problem 5: About functional dependencies and normalization. (60 pts)**

Answer each question "yes" or "no", plus a brief explanation justifying each answer.

1. The set of FDs { AB $->$ C and A $->$ B } is functionally equivalent to the set { A $->$ C and A $->$ B }.

2. Given two subsets of attributes A and B of the relational scheme R. If A is a superkey of R, then the functional dependency A $->$ B always holds for R.

3. Given the set of FDs F = { A $->$ BC and B $->$ D }. Then the set { A $->$ C and A $->$ BD } is a minimal cover of F.

4. Given the set of FDs F = { A $->$ BC and B $->$ A }. Then the set { A $->$ C, A $->$ B, B $->$ A } is a minimal cover of F.

5. Given the set of FDs F = { A $->$ BCD and B $->$ CA }. Then the attribute C in the dependency A $->$ BCD is extraneous.

6. What are the main reasons for transforming a relational design into 3NF format (as opposed to no normalization at all)?

7. A 3NF design is better than a BCNF design, because a BCNF design does not preserve dependencies.

8. During normalization, it is more important to preserve dependencies than to achieve a lossless join decomposition.