CS330 Prelim.   **NAME:** _____
March 11, 2003
Start Time: 7:30pm
End Time: 9:30pm   **Cornell NETID:** _____

**I herewith state that I understand and will adhere to the Cornell Code of Academic
Integrity.**

Signature

**Maximum number of points possible: 100. This exam counts for 20% of your overall grade.
Questions vary in difficulty. Do not get stuck on one question. Good luck!**

PART A. Very Short Answers. (4 points each, 16 points total)

Answer each question with yes/no or a number.

Assume we have two relations R(A,B) and S(B,C). All three attributes (A, B, and C) are integer
attributes. Assume that Relation R contains the following tuples: (1,2), (2,3), and (3,4). Assume
that Relation S contains the following tuples (2,2), (2,3) and (5,1).

   A.1) Is A a key for R. Answer (yes/no):

   A.2) Is (B, C) a superkey for S. Answer (yes/no):

   A.3) Does every relation have at least one key? Answer (yes/no):

   A.4) What is the maximum number of superkeys that a relation with five attributes can have?
   Answer (a number):

Part B: Short Anwers (5 points each, 40 points total)

B.1) Explain <u>in at most three sentences</u> what a foreign key is.

B.2) Explain the two main concepts behind RAID <u>in one sentence each</u>, and describe for each concept <u>in one sentence</u> why it is useful.

B.3) Why do we normalize relations? Give three reasons and explain **one** reason <u>in at most two sentences</u>.

Part B: Short Anwers -- continued. (5 points each, 40 points total)

B.4) Assume we have a relation R(A,B,C), and we would like to create a B+-tree index with a composite search key on attributes A and B. Does it make a difference whether we create the index on (A,B) or on (B,A) (Yes/No answer)? If so, why, if not, why not --- explain in at most two sentences.

B.5) Consider two relations R(A,B,C) and S(B,C), i.e., R has three attributes A, B, and C, and S has two attributes B and C. Express the relational division R/S using relational operators without the division operator.

B.6) What is the difference between a clustered and an unclustered index, and why is this difference important?

Part B: Short Anwers -- continued. (5 points each, 40 points total)

B.7) Explain what an index-only plan is, and give an example of an index-only plan. Your example should include a relation and a query.

B.8) Assume that you have four students, Joe, Art, Bob, and Cal. Assume that Joe has created the SAILORS relation. The following series of SQL commands is executed:

Joe: GRANT SELECT ON Sailors TO Art WITH GRANT OPTION
Art: GRANT SELECT ON Sailors TO Bob WITH GRANT OPTION
Bob: GRANT SELECT ON Sailors TO Art WITH GRANT OPTION
Joe: GRANT SELECT ON Sailors TO Cal WITH GRANT OPTION
Cal: GRANT SELECT ON Sailors TO Bob WITH GRANT OPTION
Joe: REVOKE SELECT ON Sailors FROM Art CASCADE

Draw the resulting authorization graph.

PART C: Queries. (6 points each, 24 points total). Do not get stuck on this question.

Consider the following schema:

Suppliers(<u>sid: integer</u>, sname: string, address: string)
Parts(<u>pid: integer</u>, pname: string, color: string)
Catalog(<u>sid: integer, pid: integer</u>, price: real)

The Suppliers relation describes suppliers of parts. The Parts relation contains information about each part. The Catalog relation lists the prices in dollars charged for parts by suppliers. (The keys are underlined: sid is a key for Suppliers, (sid,pid) is a key for Catalog, and pid is a key for Parts.)

Write the following query in **relational algebra** (6 points):

    C.1) Find the names of suppliers who supply exactly two parts.

Write the following query **in SQL** (6 points):

    C.2) Find the names and catalogs of suppliers that supply two red parts and two blue parts.

PART C: Queries – continued. (6 points each, 24 points total).

Suppliers(<u>sid: integer</u>, sname: string, address: string)
Parts(<u>pid: integer</u>, pname: string, color: string)
Catalog(<u>sid: integer, pid: integer</u>, price: real)

Write the following queries **in SQL** (6 points each):

C.3) Find the pnames and pids of red parts that are supplied by the most suppliers.

C.4) Find the names and pids of suppliers that supply all parts that cost strictly less than 200 dollars.

PART D: ER Diagrams (10 points total)

Cornell has offered you free tuition for the spring semester on top of your usual consulting fee (of $2500 per day – your discount rate, since you are just about to finish college) if you design a new database schema for its University database. Looking forward to spending Christmas in Maui, you decide to accept the challenge. Talking with the people in the administration, you gather the following information:

- The database contains information about professors and courses.
- Professors have a SSN, a name, and a research specialty. Each professor can be uniquely identified by her SSN.
- Each course can be uniquely identified by a course number, and we also want to store the course's name.
- For each course offering, we need to record the professor who teaches the course.
- Each course is taught by exactly one professor.

D.1) Draw an ER diagram that captures the above information (5 points)

D.2) Write SQL statements to create the corresponding relations that capture all the constraints expressed in the ER-diagram. (5 points)

PART E. Normal Forms (10 points total)

E.1) Consider a relation R with five attributes ABCDE. Now assume that R is decomposed into two smaller relations ABC and CDE. Define S to be the relation (ABC NaturalJoin CDE).

Assume that the above decomposition is lossless join, but not dependency preserving. You do not know any additional information about the decomposition. Which of the following statements can you infer to be *always* true: (1) R = S, (2) R ⊆ S, (3) R ⊂ S, (4) R ⊇ S, (5) R ⊃ S, (6) none of the above. List all true statements if more than one statement can be inferred to be true. (4 points)

E.2) Consider a relation R with five attributes ABCDE. You are given the following functional dependencies:
A → B, BC → E, ED → A
Find all the keys for the relation. In what normal form is relation R? (6 points)

This page will be used for grading your prelim. Do not write anything on this page.

SECTION	QUESTION	SCORE	SECTION TOTAL
Part A Very Short Answers	**A.1** (Max: 4 points)		(Max: 16 points)
	A.2 (Max: 4 points)		
	A.3 (Max: 4 points)		
	A.4 (Max: 4 points)		
Part B Short Answers	**B.1** (Max: 5 points)		(Max: 40 points)
	B.2 (Max: 5 points)		
	B.3 (Max: 5 points)		
	B.4 (Max: 5 points)		
	B.5 (Max: 5 points)		
	B.6 (Max: 5 points)		
	B.7 (Max: 5 points)		
	B.8 (Max: 5 points)		
Part C Queries	**C.1** (Max: 6 points)		(Max: 24 points)
	C.2 (Max: 6 points)		
	C.3 (Max: 6 points)		
	C.4 (Max: 6 points)		
Part D ER	**D.1** (Max: 5 points)		(Max: 10 points)
	D.2 (Max: 5 points)		
Part E Normalization	**E.1** (Max: 4 points)		(Max: 10 points)
	E.2 (Max: 6 points)		
Total (Max: 100 points)			