

CMSC424
Spring 2008

Database design
Midterm 1
Closed Book

Pop

Name _____

Weight: 12.5% of final grade.

Total points possible: 110 (10 bonus points hidden throughout the exam)

Budget your time carefully - you only get 1 hr 15 min for this exam.

Unless otherwise requested, please explain the answers to the questions.

Note: for convenience E-R diagram symbols are listed on the last page

Honor Pledge

The University of Maryland Code of Academic Integrity requests that you write by hand and sign the following statement pledging your commitment to academic integrity. Please do so in the blank space below the text of the honor pledge.

I pledge on my honor that I have not given or received any unauthorized assistance on this examination.

Signature _____

A. Miscellaneous questions [20 pts]

- [5 pts] In the schema customer(SSN, name, birth date, address, income), the following represent possible keys:
 - SSN
 - birth date, address, name
 - name, address
 - address, income

For each one indicate for each one if it is one or more of the following: **candidate key**, a **super key**. Which you would choose as a **primary key**. Briefly explain why.

- [5 pts] What is a foreign key?
- [5 pts] What is a weak entity set?
- [5 pts] In the E-R model we talk about entities, entity sets, relationships, and attributes, while in the relational model we refer to tables, tuples, rows, and columns. Indicate the correspondence between these terms.

B. Modeling [30 pts]

- [20 pts] Construct an E-R diagram for a car insurance company whose customers own one or more cars each. Each car has associated with it zero or more recorded accidents. For each customer we also want to model the neighborhood the customer lives in (affects insurance rates). The following information needs to be retrievable from this database: make and year of the car, customers' address and age, date and place of accidents, and insurance rate associated with a given neighborhood.

Please add any additional attributes that might be necessary in this diagram.

I expect you to clearly indicate primary keys, participatory constraints, and the cardinality of the relationships in the diagram. Clearly describe the assumptions you made in constructing this diagram.

- [10 pts] Convert the E-R diagram to a relational schema. Please underline the primary keys, and explain the correspondence between the tables you created and the objects in the E-R diagram.

C. Relational algebra and SQL [50 pts]

- [5 pts] Construct the Cartesian product of the following three tables:

<u>A1</u>	<u>A2</u>	<u>B1</u>	<u>B2</u>	<u>B3</u>	<u>B4</u>	<u>C1</u>	<u>C2</u>	<u>C3</u>
1	red	red	17	1972	VW	Bob	NYC	5000
2	blue	blue	8	1968	Honda	Elaine	DC	18000

2. [5 pts] Construct the natural join, left outer join and right outer join tables for the following tables:

<u>name</u>	<u>street</u>	<u>city</u>	<u>name</u>	<u>branch name</u>	<u>salary</u>
Coyote	Toon	Hollywood	Coyote	Mesa	1500
Rabbit	Tunnel	Carrotville	Rabbit	Mesa	1300
Smith	Revolver	Seattle	Gates	Redmond	1500
Williams	Seaview	Seattle	Brin	Mountainview	3000

3. [15 pts] Consider the following relational schema with the keys underlined:

country(name, code, capital, province)

city(name, country, province, population)

borders(country1, country2, length)

encompasses(country, continent, percentage)

- a) [5 pts] Does this schema allow you to represent the different spellings of a country's capital (e.g. the capital of China can be Beijing or Peking)? Why?
- b) [10 pts] Write the following queries in the relational algebra:
- find the names of all countries in Europe that do not border Switzerland.
- find the names of all cities in the world with population higher than every city in Peru.
4. [15 pts] In the relational schema from point 3, write the following queries in SQL:
- a) [5 pts] Find all cities located in the same province as the capital city in France
- b) [5 pts] Find all countries not completely contained in Europe that border a country in Europe.
- c) [5 pts] Increase the population of all cities in France by 10%.
5. [20 pts] Using the relational schema from point 3, answer the following questions:
- a) [5 pts] Write an SQL query that returns the average length of borders shared by Venezuela with its neighbors.
- b) [10 pts] Rewrite the following query so it doesn't use sub-queries:
select distinct co.name
from country co
where exists (select ci.name
from city ci
where ci.country = co.name and ci.population > 1000000)
- c) [5 pts] Is the following query correct? If yes, explain what it does. If no, explain what you think it is intended to do and how you would fix it. What assumption does this query make about the "borders" table?

```
select co.name, sum(co2.population)
from country co, country co2, borders bo
where co.code = bo.country1
and co2.code = bo.country2
```

