

COMS4111 Fall 2015

Homework 1

Due date: by **Monday October 5 at 5pm**. Your submission should be a single PDF. It is also ok to have some parts of this homework handwritten (e.g., drawings of the ER diagrams, relational algebra expressions), then take **clean** pictures of these pages and incorporate them in your final single PDF document.

Problem 1 (2.6 pts)

You are about to design a database for a car insurance company to keep information about customers, cars, insurance policies and payments. Here is what you have been told about this application.

Customers of this company are persons that are identified by a customer id (cid) and have a name and address. Cars are identified by their VIN number and they have attributes such as license plate number, make, model and year. Each car has associated with it zero or any number of recorded accidents that include information about the place and date of accident. A car is owned by one person. A person may own zero or any number of cars.

Each insurance policy covers one or more cars owned by the same person, and has a series of premium payments associated with it. Each premium payment is for a particular period of time (e.g., a month, a quarter, etc.), it has an associated dollar amount, due date, and the date when the payment was actually received.

Draw an ER diagram to capture the situation described above. Make sure you capture as many integrity constraints as possible and explicitly mention those that cannot be captured by the ER diagram (if any). Make sure that your E/R diagrams are clear and unambiguous (e.g., your bold lines look unmistakably bold, your arrowheads look unmistakably like arrowheads, etc.). You will not receive the benefit of the doubt in the presence of ambiguity.

Problem 2 (2.6 pts)

Map your ER diagram of Problem 1 to a relational database by defining the schema of the database. You may use either the actual SQL schema definition syntax or, if you prefer, you may use the compact form of schema definition (e.g., see slides #29, #31, #35) that does not include the data types of attributes. If you do use the actual SQL syntax just use INT or STRING for all attributes (you will not be graded on the choice of data types).

Problem 3 (1 pt., 0.25 pts each question)

Consider the two union compatible tables T1 and T2 shown below.

Table T1			Table T2		
A	Q	R	A	B	C
10	a	5	10	b	6
15	b	8	25	c	3
25	a	6	10	a	5

Show the results of the following operations:

- a) $\Pi_{A,R}(\sigma_{Q=a} T1)$
- b) $T2 - (T1 - (T1 \cap T2))$
- c) $\Pi_{Q,B} (T1 \bowtie_{T1.A = T2.A \text{ AND } T1.R < T2.C} T2)$
- d) $T1 \bowtie T2$

Problem 4 (3.8 pts, 0.95 pts each question)

Consider the following relational schema (primary keys are underlined):

- *Person* (ssn, pn, age, city); a person's name and age, and the city he/she lives in
- *Company* (cn, city); a company's name and the city it is located in
- *WorksFor* (ssn, cn, comp); the company a person works for and his/her compensation from this particular company (a person may work for several companies)

Express the following queries in relational algebra. You may use *P*, *C*, and *W* as shorthand for *Person*, *Company*, and *WorksFor*, respectively.

- a) Find the names of people who do not live in NYC and work in both SF and LA.
- b) Some companies are located in unloved locations. Find the names of companies that do not employ anyone living in the same city the company is located in.
- c) Some companies don't hire young people. Find the names of each company where all of its employees are older than 30 years old.
- d) Some companies are really big ... Find the names of the companies that employ someone from every single city recorded in the Person table.