McMaster University
CS 3DB3, Fall 2017
Assignment 1
Due: October 2, 2017 at 10:00am

September 11, 2017

Relational Model (15 marks)

1. (3 marks) State the two properties of data independence provided in a DBMS. Why are they important? Given an example situation of when data independence is useful.

2. (12 marks) Consider the relations in a banking database given by the schemas below.

   - BRANCH (branchName, address, city, sales, manager)
   - CUSTOMER (cID, firstName, lastName, address, city, birthDate, phoneNum, totalAssets)
   - LOAN(loanNum, branchName, amount, duration, interest)
   - BORROWER(cID, loanNum)
   - ACCOUNT(acctNum, branch, balance, type)
   - DEPOSITER(cID, acctNum)

   BRANCH provides information on the branch name, it's street address, city, total sales thus far for the year, and the branch manager. CUSTOMER records customer information such as a customer ID (cID), their name, street address, city, birth date, telephone number, and total assets recorded. LOAN records information about loans, specifically, the loan number, the branch that opened the loan, the amount, duration period, and yearly interest rate. BORROWER associates a customer to a loan. ACCOUNT provides details about bank accounts such as the account number, the home branch, the current balance, and the type of account. Finally, DEPOSITOR associates a customer to an account.

   Given these schemas, answer the following questions:

   a) Identify a primary key for each relation. For each key, briefly state the assumptions or conditions under which each key would be valid.

   b) Given your choice of primary keys in (a), define four referential integrity constraints. State the appropriate primary to foreign key references.

E-R Modeling (35 marks)

3. You will prepare an E-R diagram describing the schema for the Hamilton Street Railway (HSR) transit system. *Hamilton Street Railway (HSR)* is a public transport agency providing bus service in the Hamilton area. The head of HSR has hired you to design a database that tracks the activity and operations of its bus fleet so they can analyze the utilization of buses and its resources in serving customers.

HSR would like to record information about all persons associated with the transit system. A person may be a passenger, a bus driver, or a maintenance personnel. A person has a unique identifier, first name, last name, gender, date of birth, street address, city address, province, and occupation. A person has at least one contact phone number, and if there is more than one (e.g., home, work, cell number), they are all recorded. A passenger has a type (Child, Senior, or Student), and a separate relation stores the fares for each type of passenger. A bus driver has a salary (in CAD), his/her years of service, and any driving infractions incurred. For driving infractions, the database should record the date of occurrence, type of infraction (e.g., collision, speeding), demerit points incurred, and the financial penalty in CAD dollars. A driver may have zero to many infractions during the course of their employment (but at most one infraction per day), and an infraction is associated to exactly one driver. For maintenance personnel, the database records the employee's area specialization (e.g., engine, tire, tune-up, etc), their level (technician, senior technician, Supervisor, Manager), years of service, and salary (in CAD). A maintenance personnel can should work on (fix) at least one bus, and a bus can be fixed by many maintenance personnel. The date the fix was applied should be recorded.

HSR would like you to record information about buses, routes and stops. A bus consists of a unique identifier, years in operation, the number of seats, manufacturer, advertising revenue, and the fuel type. A bus can be driven by any driver. A route is served by many buses, and a bus serves exactly one route. A route and a stop each consist of a unique identifier and a name. There are two relationships between routes and stops: (1) a route contains many stops and a stop may belong to many routes; and (2) a schedule relationship that records the arrival time at a stop for each bus on a route. The arrival times record bus arrivals from 6am to 11pm each day between May 1 and May 7, 2017.

Sites are commonly visited locations such as a local supermarket, library, stadium, and parks. For each site, the database records its name, address, phone number, capacity, and category. HSR would also like to associate bus stops to sites so passengers have an easy way to arrive to sites. There is at least one stop associated to a site, and a stop can serve multiple sites. Some (not all) sites can host events such as concerts and sporting meets. For each event hosted by a site, the database records the name of the event, the event date and time, and the expected number of participants. For this initial design, we will assume that each event is held exactly once at a site. Finally, there is a relationship that associates routes to sites. Each site is served by at least one route, and a route can serve zero or more sites.

   a) `ER.pdf`: Draw the E-R diagram capturing the described requirements. You may use any drawing tool of your choice, but please ensure your E-R diagram is clearly readable, and the notation you use is clear and consistent (i.e., notation from the lecture slides or textbook). Handwritten models will not be accepted.

   b) `ERDesc.txt`: Give a brief (one sentence) description of each of your entities and relationships, and any constraints that exist. For example, X is a weak entity with attributes

2

(a, b, c), and has a many-one relationship with Y.

c) `HSR.ddl`: Provide the corresponding DB2 'CREATE TABLE' statements describing the relational schema. Please include all your statements in an executable script (`HSR.ddl`) that can be run on the DB2 command line, in a single command. Ensure that your script runs on the CAS DB2 servers. Scripts that do not execute on these servers will not be marked.

Grading

This assignment is worth 12% towards your final grade.

Submission

All files are to be submitted using the Avenue system. Please ensure you submit all files with the correct names, as described below:

1. For Q1-Q2: Please include all your answers in a text file `relation.txt`.

2. For Q3: Submit files `ER.pdf, ERDesc.txt, HSR.ddl`.