

COMP-421 Database Systems, Winter 2020

Written Assignment 1: Data Modelling / Relational Algebra

Due Date Jan 30, 11:59pm

This is an individual assignment. You need to solve these questions on your own. Use the discussion forum on mycourses for assignment 1 if you have any questions. TA office hours will be setup on the days before the due date.

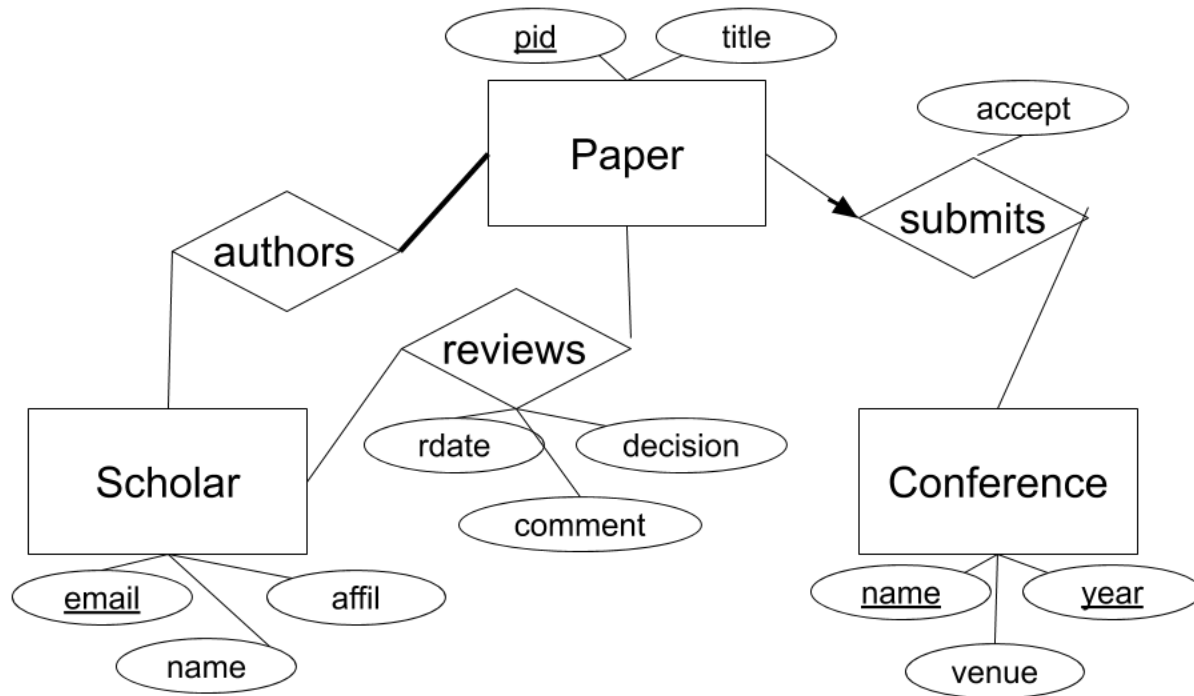
Ex. 1 — ER Modelling (45 Points)

A company is planning to use a database to manage its projects. The database should contain the project members, budget, schedule, and evaluations.

- Employees of the company are identified by their eid. Additionally, they have a name, an address, and a salary.
 - Each project has a unique project name. Additionally, it has an objective, a budget, and a max. enrollment that indicates the max number of employees can be involved in the project (e.g., 5).
 - Each project must have one project leader, but an employee can lead more than one projects.
 - Except for the leader, each project should have at least one member. An employee can only be the member of a single project group. In addition, an employee might not belong to any project group. The time since when an employee is enrolled to a project is stored.
 - A project might have some associated documents.
 - Each document has a title.
 - Documents have one or more versions. We need to store the version numbers and the date when it was created.
 - Each version of a document is always written by one or more employees (authors of a document can change between versions of the same document). Each employee might write more than an one document.
 - Each project must have a schedule, according to which a project is divided into a couple of steps. Each step has a begin time, an end time, a budget which is part of the total budget of the project, and a milestone objective.
 - At the end of each step of a project, the project leader will write a report to summarize the work in the step. The report should be stored.
 - At the end of each step of a project, the project must be evaluated according to a pre-defined schema. The schema describes the aspects on which the work in the step should be evaluated, and the points required to be considered for each aspect at the time of evaluation (points are plain text). There can be one or more aspects to evaluate a particular step in the project.
 - Evaluations are given by senior employees. The time since when an employee becomes senior is recorded.
 - According to the evaluation schema, an evaluation evaluates an aspect and gives a grade between 0 and 5, and some text. An evaluation is always written by only one senior employee. An aspect can have more than one evaluations from different senior employees. A senior employee can give several evaluations for different aspect of a project or different projects.
1. (30 Points) Design the E/R Model for the following specification. Indicate any constraints that you might not express in the E/R model. For each entity, decide on the primary keys. If you think there is no appropriate primary key, you can introduce an artificial key attribute.
 2. (15 Points) Transform your diagram into the relational model (in the form $R(\underline{a}, b, c)$. Indicate primary keys by underlining them. Indicate foreign keys by letting them point to the relation to which they relate, eg. $Q(\underline{d}, e, a)$ (a ref R) . DO NOT write SQL statements. Indicate if there are other constraints depicted in the ER model that you cannot (yet) describe in the relational model.

Ex. 2 — Testing your E/R decoding skills. (16 points)

Consider the following E/R diagram used by a publications management system used for academic conferences (16 points)



For each of the below statement, indicate whether you agree or disagree and briefly state your reasons (one liners are sufficient). You will lose half of the points if you do not provide reasons for your choices.

- (A) Model does not help track co-authors of a given paper.
- (B) Not all papers are required to be submitted to a conference.
- (C) Model does not allow us to compute the number of papers a scholar has published in a specific year.(accepted to a conference).
- (D) We can have scholars who have not authored any papers.
- (E) Model by itself does not prevent a scholar from being a reviewer of their own paper.
- (F) Multiple reviews of the same paper by the same scholar can be tracked separately as long as they are on different dates.
- (G) Model allows for reviews of papers that are not submitted to any conference.
- (H) Same paper can be submitted to two different conferences as long as the year of those conferences are different.

Ex. 3 — Relational Algebra (39 Points)

Look at the following relational schema of a banking database (with some example tuples for each relation)

Customer	(<u>custid</u> , name, city, streetaddr, province) <i>12131, 'Joe Smith', 'Montreal', '412 Rue Minne', 'PQ'</i>
Account	(<u>acctid</u> , custid, atype, startdate, balance, branchid), (custid ref Customer, branchid ref Branch) <i>515512, 12131, 'checking', '2014-12-20', 1432.56, 31</i>
Branch	(<u>branchid</u> , mgrid, city, streetaddr, province), (mgrid ref Employees.empid) <i>31, 902, 'Ottawa', '450 Le Passfail', 'ON'</i>
Employees	(empid, name, branchid, salary, city, streetaddr, province), (branchid ref Branch) <i>902, 'Megan Crowley', 31, 6456.22, 'Ottawa', '31 Walkaway Apt 302', 'ON'</i>
Transactions	(<u>tid</u> , acctid, transtype, transdate, transamount, branchid), (acctid ref Account, branchid ref Branch) <i>612312334, 12131, 'deposit', '2016-01-24', 200, 48</i>

Some useful information.

- A customer can have multiple accounts, an account however has only one customer.
- Transactions can be conducted at any branch, irrespective of where the account is located.
- An employee can be living in a city, but working in a branch in a different city.
- Cities can have same name across different provinces (but not within the same province), and should be considered different. For example Delta, ON and Delta, BC are two different cities. Keep this in mind while answering questions that involve cities to be compared.
- The combination (name, street_addr, city, province) can be considered unique for an individual (helps identify employees who are also a customer, etc.). Therefore a person cannot have two different names/addresses.

Express the following queries in relational algebra using the notations and operators taught in class:

1. (3 Pts) List the city and street address of all the bank branches in the province of 'PQ'.
2. (3 Pts) Give the customer ids of all customers who has a 'savings' account with amount less than 500.
3. (4 Pts) Find the account ids of all the 'savings' accounts owned by the customer 'Madea Fakename'.
4. (5 Pts) Find the customer ids of all customers who had opened both a 'checking' and 'savings' account this year with the bank.
5. (5 pts) Find the account ids of all accounts that had only a 'deposit' transaction and no 'withdraw' transaction on '2016-12-22'.
6. (6 Pts) Find the name of the branch managers and the complete address of the branch where they work if they have a checking account in the branch where they work.
7. (6 Pts) Find the transaction ids for all the 'withdraw' transactions performed on any accounts held by 'Mone L Aunderer' which was carried out on a branch which is not the account's branch.
8. (7 Pts) List all the 'savings' account ids and the name and customer ids of their customers which had only one transaction this year.