CMSC 508 - Spring 2018
Homework 1

1. For these instances of two relations in a banking database, indicate the following:

acctNo	type	balance
12345	savings	12000
23456	checking	1000
34567	savings	25

The relation Accounts

firstName	lastName	idNo	account
Robbie	Banks	901-222	12345
Lena	Hand	805-333	12345
Lena	Hand	805-333	23456

The relation Customers

a. **The attributes of each relation.**
   For Accounts: acctNo, type, balance.
   For Customers: firstName, lastName, idNo, account.

b. **The tuples of each relation.**
   For Accounts:
   (12345,saving, 1200)
   (23456,checking, 1000)
   (34567,saving, 25)
   For Customers:
   (Robbie, Banks, 901-222,12345)
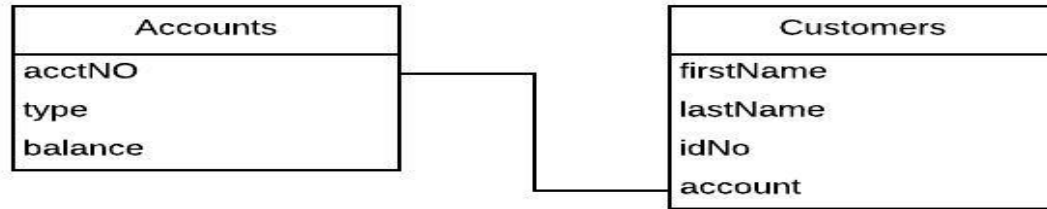   (Lena, Hand, 805-333,12345)
   (Lena, Hand, 805-333,23456)

c. **The relation schema for each relation.**
   Accounts: (acctNo, type, balance)
   Customers :(firstName, lastName, idNo, account).

d. **The database schema.**

e. A suitable domain for each attribute.

For Accounts:

acctNo: Number.

type: String (checking or saving).

balance: Number.

For customer:
First name: String.
Last name: String.
Id no: Number.
Account: Number.

2. **Please answer <u>in your own words</u>.**

   a. **Explain the concept of physical data independence, and its importance in database systems.**

   Data independence is the ability of to make changes to data characteristics without have to make changes to the programs that access the data. It is important because it saves time and potential errors caused by reducing modifications to data access software.

   b. **What is a transaction? Define ACID and provide a brief description of each.**

   A transaction is a logical unit of database processing that includes one or more access operations. A transaction may have several low level tasks or a very small unit of any program. There is a set of properties that guarantee that database transactions are processed reliably. These properties are called ACID properties.

   <u>Atomicity</u>: in atomicity database modifications follow either all or nothing rule. Atomicity property a transaction must be treated as an atomic unit. Either all of its operations are executed or none.

Consistency: in consistency the database remains in a consistent state before the start of the transaction and after the transaction is over whether successful or not. After the transaction is finished, its database must remain in a consistent state.

Durability: Durability refers to the guarantee that once the user has been notified of success, the transaction will persist, and not be undone. In any case all updates made on the database will persist even if the system fails and restarts.

Isolation: Isolation refers to the requirement that other operations cannot access or see the data in an intermediate state during a transaction. This constraint is required to maintain the performance as well as the consistency between transactions in a database.

c.   What is integrity? Describe two ways to enforce integrity in a database.

Integrity is the accuracy and consistency of data. Impose integrity in a database:

 Entity Integrity: No single component of a primary key value = null.

 Referential Integrity: If an attribute is a primary key in one relation and in a second relation then the values of the attribute in the second relation must be in the first relation or equal to null.

d.   Differentiate between an entity and an entity set.

Entity has real existence for example tuple. While an entity set is a set of entities that share the same entity type.

e.   Differentiate between a relation and a relationship. Can a relation represent a relationship?

Relations are the approach things are connected and Relationships are the actual connections. A relation is between two entities and the relationship how they are connected. A relation with can represent a relationship by a model. The relational model shows the relationships between entities.

f.   What is the difference between a super key, a candidate key, and a primary key?

Primary Key The attribute or combination of attributes that uniquely identifies a row or record in a relation. Primary key represent a candidate key that is chosen by the database designer as the principal means of identifying tuples within a relation.

Candidate Key A relation can have only one primary key. It may contain many fields or combination of fields that can be used as primary key. One field or combination of fields is used as primary key. The fields or combination of fields that are not used as primary key are called candidate key .

> **Superkey** is a combination of attributes that can be uniquely used to identify a database record. A table might have many superkeys. Candidate keys are a special subset of superkeys that do not have any extraneous information in them.

3. Draw an entity-relationship diagram for each of the following data sets. Remember to identify keys and to label relationships with their functionality. If you have to make any assumptions in order to decide what to draw, then state those assumptions. Think carefully about what is an entity, what is a relationship and what is an attribute. Sometimes things sound like one but are better represented in another way.

   a. Doug's Dog Walking Service has grown from a part-time venture for earning beer money to a big business. (It seems like everyone in Richmond has a dog these days.) Doug needs a database to keep track of his clients and their beloved pet(s). Data is to be stored about clients, dogs, and walking history (track record, if you will). This data includes the animals name, date of birth, breed of dog, owner, and the dog's exercise history. Data to be stored about owners includes name, address, and phone number. Data stored about the exercise history is: which dog went for a walk, date walked, length of the walk, and the charge to the owner. Doug is expanding his business and has hired some employees as dog walkers, so he'd like you to include some employee information, too. Data stored about each employee is their name, address, and phone number, along with which dogs they exercised, and when they did so.
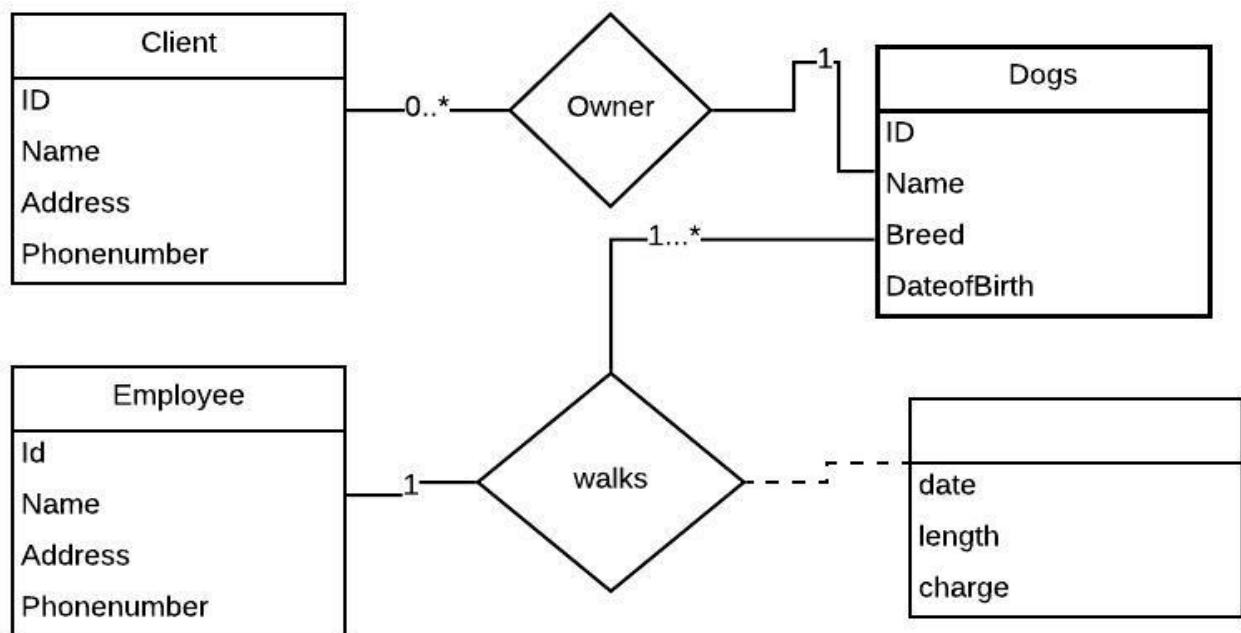
b. Ultimate Hair salon has offered you free life-time haircuts if you design its database schema. Given the rising cost of personal care and your impeccable sense of style, you agree. Here is the information that you gathered: Clients are identified by their unique customer loyalty card number, and we also store their names and age. Stylists are identified by their employee ID number, and we also store their names and specialty. Each patient has one preferred stylist, and we want to know how long the client has been with her preferred stylist. Each stylist has at least one client.
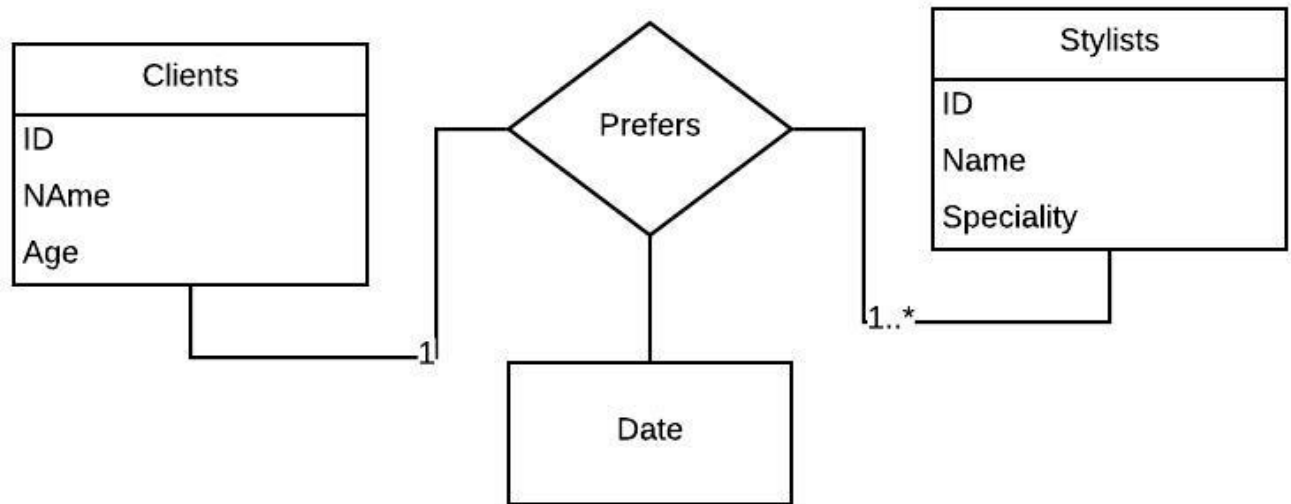
4. For the following E/R diagram, provide the relational database schema that you would create for the database.
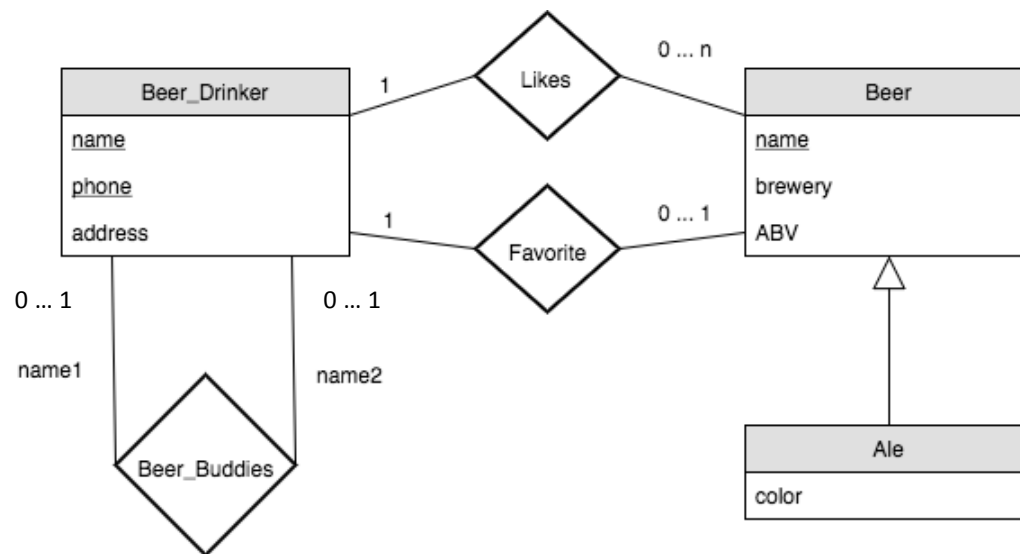
Beer_Drinker (ID , name , address , beerID)
Beer_Bodies (Beer_Drinker_1_ID, Beer_Drinker_2_ID)
Beer (ID , brewery, ABV, AleID)
Ale (ID , color)
Like (Beer_Drinker_ID, BeerID)