

Introduction to Database Systems

Homework 2

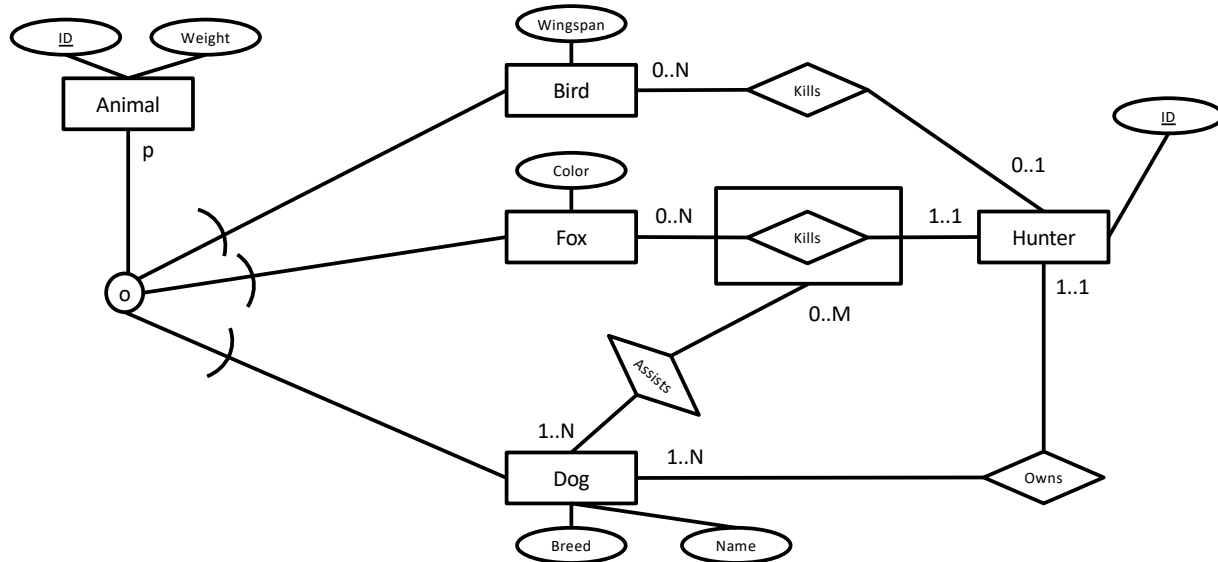
Björn Thór Jónsson

Fall 2021

This homework is to be submitted using a quiz on LearnIT. It exercises ER-diagram interpretation, ER-diagram design, and SQL queries. Recommended readings are Chapters 3 and 7 from the PDBM book.

1 ER Diagram Interpretation

Note: This exercise is intentionally designed to be confusing, to prepare you for the real world where you may encounter counter-intuitive or outright nonsensical database designs!



1) In this part, consider ONLY the ER diagram above. Choose the correct statements.

- (a) All foxes have a weight.
- (b) A hunter can kill a fox without the assistance of a dog.
- (c) A dog can kill a bird.
- (d) Hunters are animals.
- (e) All animals are dead.
- (f) A hunter can kill a fox with the assistance of at least one dog.
- (g) All foxes are dead.
- (h) A bird can be killed by many hunters.
- (i) All birds are dead.
- (j) Given a hunter it is possible to find the total weight of the animals he/she has killed.
- (k) A bird can be a dog.

2 ER-Diagram Design and Implementation

A political party, the Wily And Sinister Party or WASP, has found its support waning and decided to create a database to support its operations. The WASP have already hired consultants to interview potential users, coming up with the following requirements for the database, but they need you to design their database by creating an ER-diagram.

The following requirements were determined through interviews with WASP. Your database design should follow the requirements definition strictly. For example, ID columns should only be used where specifically mentioned. Where ID columns are mentioned, however, you can assume that they are a key.

1. For people, the database should keep track of their ID, name, address, phone number, date of birth (DOB), and date of death (DOD). The default value of DOD is NULL!
2. People are further divided up into members of WASP and enemies; each person is a member, an enemy, or possibly both. For members, the starting date of membership is registered. Note that WASP members cannot leave the party, even in death!
3. For each WASP member, a list of their assets is maintained, which could be used to achieve the WASP agenda. Each asset of a person is identified with the name of the asset, but text fields must also be maintained to a) describe the asset in more detail and b) describe how the asset could potentially be used.
4. For each WASP enemy, some WASP members may be assigned as opponents. The opponent appointment has a start date and an end date (the latter may be unknown, in case of open-ended assignments).
5. Furthermore, events and relationships between various people (e.g. accidents, business partnership, marriage, mortal enemies, ...) are registered in a fairly flexible way. The WASP party uses the general term 'linkings'. Each linking is an entity that is assigned an ID, name, type and a more detailed description. Multiple people may participate in each linking, but the minimum number of people per linking is one, and each person may participate in multiple linkings with different people.
6. The WASP party has a list of roles, each with a unique ID and a unique title. Members take turns filling the roles: each member may fill multiple roles at any given time and multiple members may fulfill the same role. Members may, however, only be appointed to the same role once. The start and end dates of each appointment to a role are always known in advance and recorded, as well as the monthly salary for the role.
7. The WASP party keeps track of other political parties, both domestically and abroad. For each party, the country and name are a unique identifier, but the party is given an ID as well that serves as a primary key. For each such party, at each time (represented

by start and end dates), exactly one WASP member monitors the developments in the external party.

8. The WASP maintains a list of sponsors. For each sponsor, an ID is registered, as well as their name, their address and their industry.
9. Each sponsor may give grants to a number of WASP members. For each grant, the date the grant is awarded is registered, as well as the amount and a text field called 'payback' that describes what the sponsor expects in return. Each sponsor may give multiple grants to multiple members, and each member may receive multiple grants from multiple sponsors. Each grant, however, is to a single person and each sponsor can give multiple sponsorships to each member, but at most one per day.
10. Each grant is reviewed by one member of WASP. The date of the review is decided when the grant is registered, typically one year in the future. At review time, the grant is assigned a numerical grade from 1 to 10, depending on how well the WASP member executed the payback.

2) Draw an ER diagram reflecting the requirements above, using the ER-notation of the PDBM textbook, with amendments made in lectures. The diagram may be hand-drawn, if it is drawn clearly enough.

3 SQL

In this homework you will work with a fictional database of countries, cities and languages. To start working with the database, run the commands in `HW2.sql` found in LearnIT using the PostgreSQL DBMS on your laptop. It is recommended to use `psql` for this purpose.

The database contains a variety of information on countries in the following schema:

```
continents(Continent)
countries(Code, Name, Region, ..., Population, ...)
countries_continents(CountryCode, Continent, Percentage)

cities(ID, Name, CountryCode, District, Population)
empires(CountryCode, Empire)
countries_languages(CountryCode, Language, IsOfficial, Percentage)
```

Most attributes are self-explanatory. Primary and foreign keys are correctly defined, but you must study the DDL commands to understand the details of these. Some additional notes are in order:

- Some countries are present on more than one continent, and therefore have two entries in `countries_continents`; the `Percentage` attribute refers to the percentage of the population that lives on that continent.
- The table `empires` lists the constituent countries of some (fictional) empires. Countries that are not present in this table are not considered part of any empire.
- The `Percentage` data for languages in `countries_languages` also refers to the percentage of the population that speaks the language. The data is not complete, as the sum of percentages for countries is not 100.0 in all cases; this may be due to rounding errors or due to missing data.
- The data has various other errors, partly by design and partly because it is based on a publicly available dataset that has some errors in it.
- In `cities`, the `District` attribute refers to the region of the country where the city is located (rather than a district of the city).

3) Answer each of the following questions using a single SQL query on the homework database. Enter the numerical answer of each query into the LearnIT quiz. Queries should still adhere to the detailed guidelines given in Homework 1.

- (a) The empire 'Great Britain' consists of 4 countries. How many countries does the empire 'Danish Empire' consist of?
- (b) There are 4 countries that are present on more than one continent. How many of these countries are partially in Europe?
- (c) In the countries of North America that have more than 1 million inhabitants, there are a total of 164,688,674 people that speak Spanish, according to the statistics in the database. What is the corresponding number for South America?
- (d) According to the database, one language is spoken in all the countries of the 'Danish Empire'. How many languages are spoken in all the countries of 'Benelux'?

Note: This is a *division* query; points will only be awarded if division is attempted.