# Introduction to Database Systems
Homework 2

Omar Shahbaz Khan and Eleni Tzirita Zacharatou

Fall 2023

You should submit the homework as a quiz on LearnIT. In this homework, you will practice ER diagram interpretation, ER diagram design, and SQL queries. Recommended readings are Chapters 3.1-3.3, 6.3-6.4, and 7.1-7.4 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!
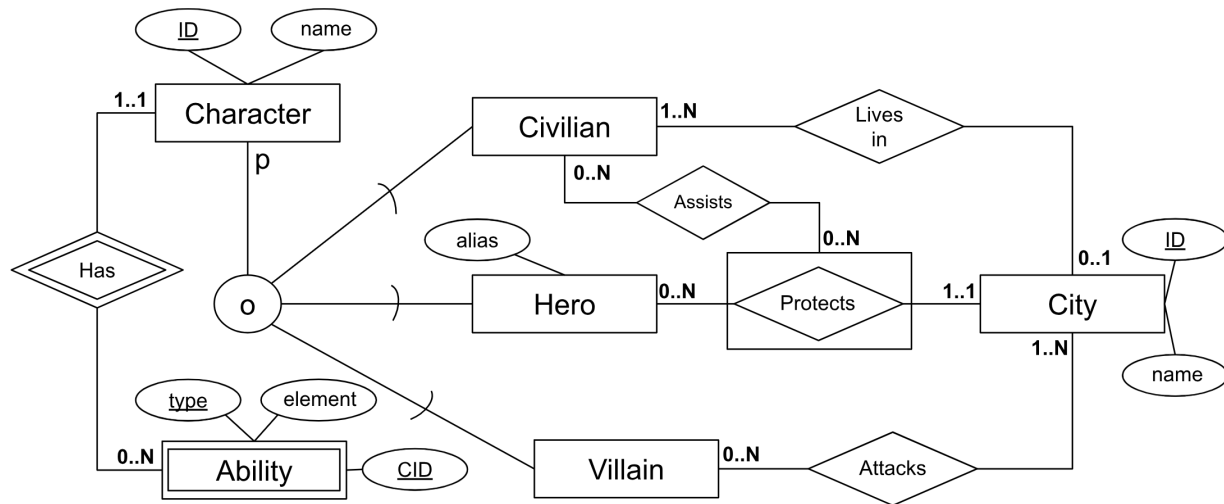
1) In this part, consider ONLY the ER diagram above. Choose the correct statements.

(a) Villains have an alias

(b) A civilian cannot have abilities

(c) A hero protects only one city

(d) A city is protected by one hero

(e) A hero can have multiple abilities with the same element

(f) A city has at least one civilian

(g) Civilians have to live in a city

(h) A civilian may assist in protecting the city

(i) Villains can attack multiple cities

(j) A character can have multiple abilities of the same type

(k) A hero can be a villain

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 strictly follow the requirements definition. For example, ID columns should only be used where specifically mentioned. Where ID columns are mentioned, 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 (thankfully!) NULL.

2. People are further divided 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! For enemies, the reason is registered.

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. 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 two, and each person may participate in multiple linkings with different people.

5. Each participation of a person in a linking is monitored by one member. The monitoring member never changes.

6. The WASP party has a list of roles, each with a unique ID and a unique title, as well as a monthly salary. Members take turns filling the roles: each member may fill multiple roles at any given time, and multiple members may fill the same role. However, members may be appointed to each role only once. The start and end dates of each appointment to a role are always known in advance and recorded.

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 also has an ID that serves as a primary key. For each such party, at any point in time (represented

by start and end dates), exactly one WASP member monitors the developments in the external party, and logs their observations in a text field called 'log'.

8. The WASP party keeps a list of allies, each ally is given a unique alias and a level of trust from 1 to 5. Additionally, how an ally may benefit the WASP party is disclosed in a field called value. An ally can provide multiple benefits. The ally is always affiliated with one or more members of the WASP party. An ally can also support other political parties, which typically impacts their trust levels.

9. The WASP party has sponsors. Each sponsor has an ID, name, address, and industry attribute. 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 can receive multiple grants from multiple sponsors. However, each grant is for a single person and each sponsor can give multiple sponsorships to each member, but at most one per day.

10. Each grant may be 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.

11. All parties and all enemies are two categories of opponents; opponents have an ID. Some WASP members may be assigned to oppose opponents. This opposition appointment has a start date and an end date (the latter may be unknown in the case of open-ended assignments).

---

2) Draw an ER diagram reflecting the requirements above, using the ER notation of the PDBM textbook, with amendments made in the lecture. If you draw the diagram by hand, please ensure that it is sufficiently readable.

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-DB.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)
```

- `continents` has **7** entries in the database.

- `countries` has **239** entries in the database.

- `countries_continents` has **243** entries in the database.

- `cities` has **4079** entries in the database.

- `empires` has **13** entries in the database.

- `countries_languages` has **986** entries in the database.

Most attributes are self-explanatory. The 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. Finally, upload a script with your queries. Queries should still adhere to the detailed guidelines given in Homework 1.

(a) The empire 'Danish Empire' consists of 3 countries. How many countries does the empire 'Great Britain' consist of?

(b) There are 4 countries that are present on more than one continent. How many of these countries are partially in Asia?

(c) In the countries of North America that have more than 10 million inhabitants, there are a total of 130,731,300 people who speak Spanish, according to the statistics in the database. What is the corresponding number for Europe?

(d) According to the database, two languages are spoken in all countries of 'Benelux'. How many languages are spoken in all countries of 'Great Britain'?

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

4 Passing grade

The maximum grade is 60 points. You get 20 points for the ER Diagram Interpretation, 20 points for the ER Diagram Design and Implementation, 4 points for each correct numerical answer for the SQL queries, and 4 points for your SQL query script. To pass the homework, you need to have *at least 10/60* points. However, the more you do, the more feedback you get and the more you learn.