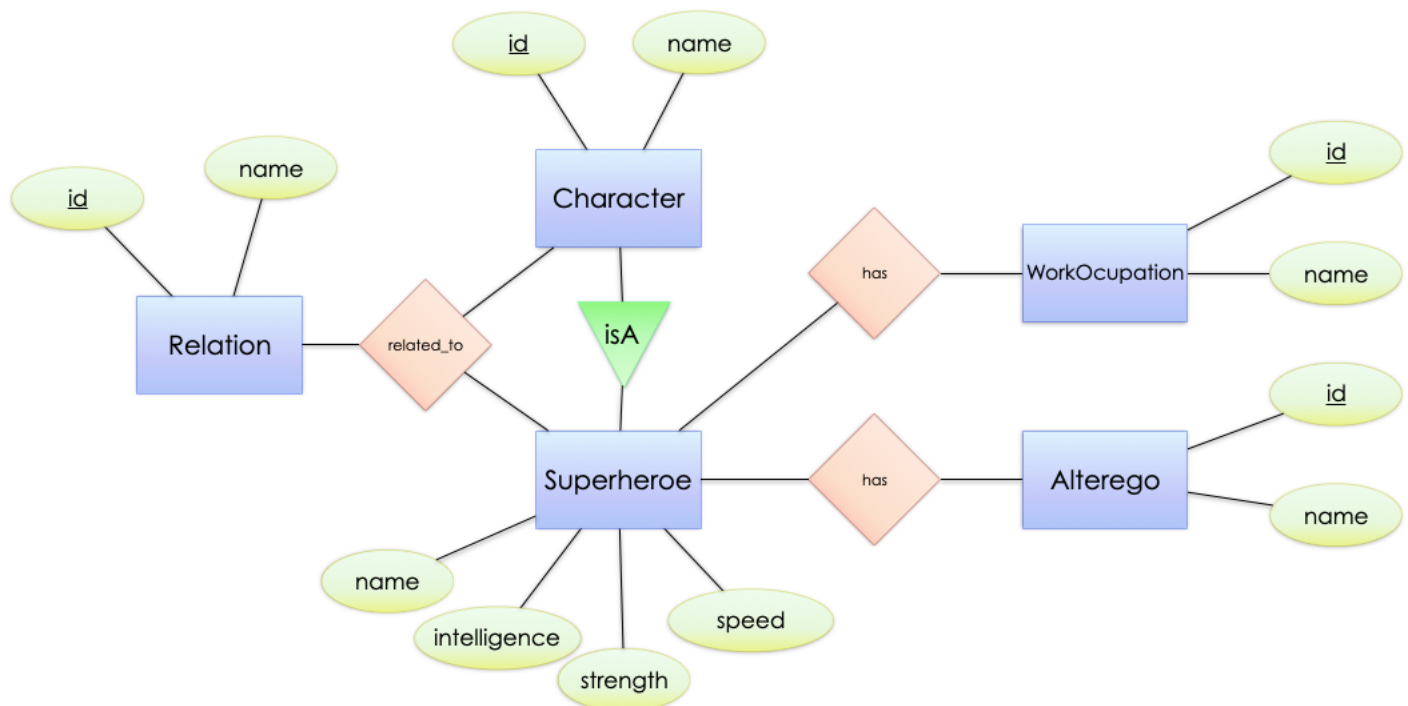


Laboratorio 5 - LDD y Acceso Programático

Profesores: Claudio Gutiérrez
Matías Toro
Auxiliares: Stefan Malewski
Cristian Salazar
José Vera
Florencia Yáñez

Considere el siguiente modelo entidad relacion que corresponde a una base de datos de super-heroes:



En este laboratorio usted deberá crear las tablas necesarias correspondientes al modelo entidad relación y luego importar los datos a partir de un archivo csv.

La creación de tablas e importación deberá ser hecha en el servidor del curso en el esquema **superheroes**.

- Para conectarse al servidor directamente, usted puede usar una terminal en Linux/Mac/Windows o Putty en Windows. El servidor está hosteado en `cc3201.dcc.uchile.cl` y escucha en el puerto 240. Su usuario es `cc3201` y la contraseña se publicó en u-cursos en Material Docente. El comando que deberían ejecutar en la terminal para conectarse al servidor desde Linux/Mac es:

```
ssh -p 240 cc3201@cc3201.dcc.uchile.cl
```

Para conectarse a la base de datos basta con ejecutar el comando `psql cc3201`. Para cerrar la sesión utilice el comando `\q`.

- Para conectarse al servidor programáticamente, puede crear una conexión en python como se describe a continuación:

```
conn = psycopg2.connect(host="cc3201.dcc.uchile.cl",  
    database="cc3201",  
    user="cc3201",  
    password="j'<3_cc3201", port="5440")
```

Usted deberá entregar un archivo de texto (`.txt`) y un archivo python (`.py`) con el desarrollo de los ejercicios que siguen.

- P1.** 30 PUNTOS Cree las tablas necesarias de modo que satisfagan el modelo entidad relación en el schema **superheroes**, agregando las restricciones de llave primaria y foránea según corresponda. Tenga en cuenta que el nombre de cada tabla debe tener un prefijo único para cada grupo, como por ejemplo `equiporocket_tablita`.

Hints:

- El campo id lo puede crear de tipo `serial`.
- Sugerimos traducir la herencia (isA) usando dos tablas (para `character` y `superhero`).

- P2.** 60 PUNTOS Cree y corra un archivo python para poblar la base de datos creada en el punto anterior a partir del archivo de datos csv provisto en ucursos. Este archivo tiene varios errores e inconsistencias, por lo que le sugerimos leer las siguientes consideraciones:

- Le recomendamos hacer la tarea en python. Para instalar la librería necesaria para conectarse a la base de datos utilice `pip3` ejecutando `pip3 install psycopg2-binary`
- Puede asumir por simplicidad que no hay dos personajes (`characters`) que se llamen igual. Si hay personajes sin nombre, por simplicidad puede usar el de `superhero`.
- Para validar un pariente y separar el nombre de su parentezco (e.g. en `Ashley Zolomon (wife, deceased)`), use el comando `m = re.search("([^(]+) [*]\(((^)+)\)", texto)`. Luego, si `m` esta definido, `m.group(1)` apunta al nombre (e.g. `Ashley Zolomon`), y `m.group(2)` al tipo de relación y estado (e.g. `"wife, deceased"`).
- A veces el nombre del pariente es el nombre de `superhero` y otras veces el nombre real, por lo que se sugiere procesar el archivo csv dos veces: una para ingresar los `superheroes` y otra para procesar las relaciones (y crear personajes que no sean `superheroes`). Esto significa que al buscar parientes debe buscar en ambas tablas (puede escoger el primer valor encontrado).

- Algunos strings pueden venir con comillas dobles extras (e.g. `adventurer`, `"agent"`, `mechanic`)
- Un mismo oficio pueden venir en distintas combinaciones de mayúsculas y/o minúsculas en distintas filas.

P3. 10 PUNTOS Entregue las consultas sql que responden las siguientes preguntas (entregamos las respuestas para que validen su trabajo).

- (a) **Pregunta:** Calcule los nombres de los 3 superheroes con más parientes (note que hay nombres de superheroes que se repiten, e.g. Batman (Bruce Wayne), Batman (Terry McGinnis) y Batman (Dick Grayson)).
Respuesta: Namor (37), Aquaman (36), Havok (26)
- (b) **Pregunta:** Calcule los nombres de los 3 personajes no superheroes con más parientes.
Respuesta: Jr. (7), Philip Summers (6), unnamed mother (6)
- (c) **Pregunta:** Calcule los nombres de los 5 superheroes con más parientes superheroes.
Respuesta: Havok (7), Polaris (7), Batman (6), Jean Grey (6), Batman II (5)
- (d) **Pregunta:** Calcule el nombre de relación más común
Respuesta: father (107)
- (e) **Pregunta:** Calcule los 3 trabajos más populares
Respuesta: adventurer(114), student(27), mercenary(26)