Linnæus University
Sweden

Report

Assignment 2
2DV514

Author: marwan laroussi
Supervisor: Ilir Jusufi
Semester: HT17
Course code: 2dv513

Task 1

1. What are the names of students enrolled in 1dv513 ?

$\hookrightarrow$ $R_1 := \pi_{name} \left(\sigma_{code = "1dv513"} (student \bowtie enrolledIn) \right)$

2. What are the names of students in both 1dv513 and 2dv513 ?

$\hookrightarrow$ $R_2 := \pi_{name} \left(\sigma_{code = "1dv513"} (student \bowtie enrolledIn) \right)$
$\cap \; \pi_{name} \left(\sigma_{code = "2dv513"} (student \bowtie enrolledIn) \right)$

3. Who teaches 2dv610 ?

$\hookrightarrow$ $R_3 := \pi_{lecturer} \left(\sigma_{code = "2dv610"} (subject) \right)$

4. Who teaches 1dv513 and 2dv513 ?

$\hookrightarrow$ $R_4 := \pi_{lecturer} \left(\sigma_{code = "1dv513"} (subject) \right)$
$\cup \; \pi_{lecturer} \left(\sigma_{code = "2dv513"} (subject) \right)$

5. What are the names of students who are taking a subject not taught by Ilir ?

$\hookrightarrow$ $R_5 := \pi_{name} \left(\sigma_{lecturer \neq "Ilir"} (student \bowtie subject \bowtie enrolledIn) \right)$

TASK 2

1. Find functional dependencies.

Each manager uses the same room for all interviews during the day so:

Manager day -> room

A room may, however, be used by different managers during a day, as long as the interviews are not at the same time. In other words, one room is used by one manager (and applicant) for a certain time.

Day time room-> Manager Applicant

Since one applicant can only do one interview a day with one specific manager in a specific room, we can also assume that a functional dependency would be:

Applicant day -> Manager room time

2. Find the keys of the relation.

Regarding the result of question 1, {Applicant day} would be a key. {day time room} would be another key.

3. Show that the relation is in 3NF but not in BCNF.

{Applicant day} is a key of the relation Interviews.

Manager day -> room is a FD

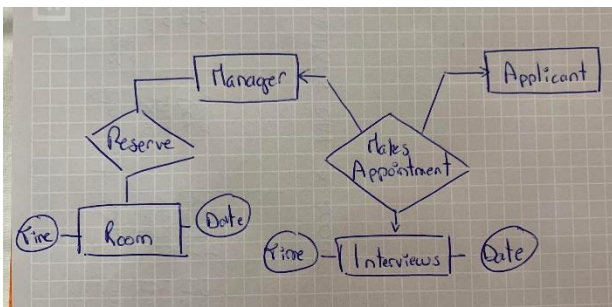But left side of this FD is not a superkey as it should contain {Aplicant, day}

On the other side, day is a member of some key ({Applicant day}) so the relation is a 3NF.

4. Decompose the relation in relations that are in BCNF.

Interview1(Manager day room)

Interview2(Manager day Applicant)

5. Draw an E/R diagram that describes the system. Try to incorporate all dependencies.

TASK 3

After importing the JSON script from task 4 I had to modify a bit the tables to make them as simple as possible otherwise it took ages to import the file in its entirety. I decided to create only two tables, as Subreddit is FD of Subreddit_id (Subreddit_id -> Subreddit) so creating a table for Subreddit was needed. I finally have two tables in my database, one with all the information of comments and one with all the information of a subreddit.
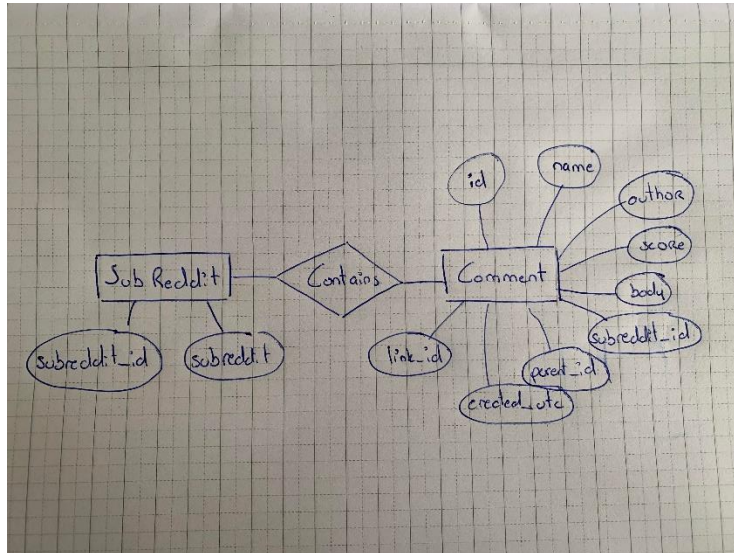

E/R diagram of the design

Table1: Comment

Column	Type	Default Value	Nullable	Character Set	Collation	Privileges
id	varchar(7)		NO	utf8	utf8_general_ci	select,insert,update,references
name	varchar(20)		NO	utf8	utf8_general_ci	select,insert,update,references
author	varchar(20)		NO	utf8	utf8_general_ci	select,insert,update,references
score	int(11)		NO			select,insert,update,references
body	varchar(9900)		NO	utf8	utf8_general_ci	select,insert,update,references
subreddit_id	varchar(8)		NO	utf8	utf8_general_ci	select,insert,update,references
parent_id	varchar(10)		NO	utf8	utf8_general_ci	select,insert,update,references
created_utc	timestamp	CURRENT_TIMESTA...	NO			select,insert,update,references
link_id	varchar(8)		NO	utf8	utf8_general_ci	select,insert,update,references

Table2 : SubReddit

Column	Type	Default Value	Nullable	Character Set	Collation	Privileges
id	varchar(8)		NO	utf8	utf8_general_ci	select,insert,update,references
name	varchar(20)		NO	utf8	utf8_general_ci	select,insert,update,references

TASK 4

I made two-timing tests, one without any constraints (even without keys) and one adding all the constraints to see the difference between the two results. I got a JavaScript error after 10 minutes:" JavaScript heap out of memory" so I had to find on the internet how to add memory to import the whole file into my database.

```
Timer:   582553.5168999992

<--- Last few GCs --->

[8124:000001AD4EF941E0]    582757 ms: Mark-sweep (reduce) 4059.6 (4143.6) -> 4058.8 (4143.6) MB, 8668.2 / 0.0 ms  (average
 mu = 0.156, current mu = 0.006) allocation failure scavenge might not succeed
[8124:000001AD4EF941E0]    591703 ms: Mark-sweep (reduce) 4059.9 (4143.6) -> 4059.2 (4144.1) MB, 8895.3 / 0.0 ms  (average
 mu = 0.086, current mu = 0.006) allocation failure scavenge might not succeed


<--- JS stacktrace --->

FATAL ERROR: Reached heap limit Allocation failed - JavaScript heap out of memory
 1: 00007FF72E4D30AF v8::internal::CodeObjectRegistry::~CodeObjectRegistry+112511
```

Screen of first try importing rows into database

I had to increase the amount of memory allocated to the command by running the following command in the PowerShell:

```
$env:NODE_OPTIONS="--max-old-space-size=8192"
```

Even after increasing the amount of memory I still had a huge performance problem (which delayed my project release as I really didn't expect it to take that much time) .

I'v never managed to fix this problem even after the deadline. I posted all of my work in my Github which is accessible here :