

Introduction to databases

ER Diagram; Relational Model; mapping ER-Relational; Normalization

Week 2 – Lab

Relational schema

- The relational schema is the primary element of the relational database, which refers to the meta-data that describes the **structure of data within a certain domain.**
- It is the blueprint of a database that outlines the way its structure organizes data into tables.

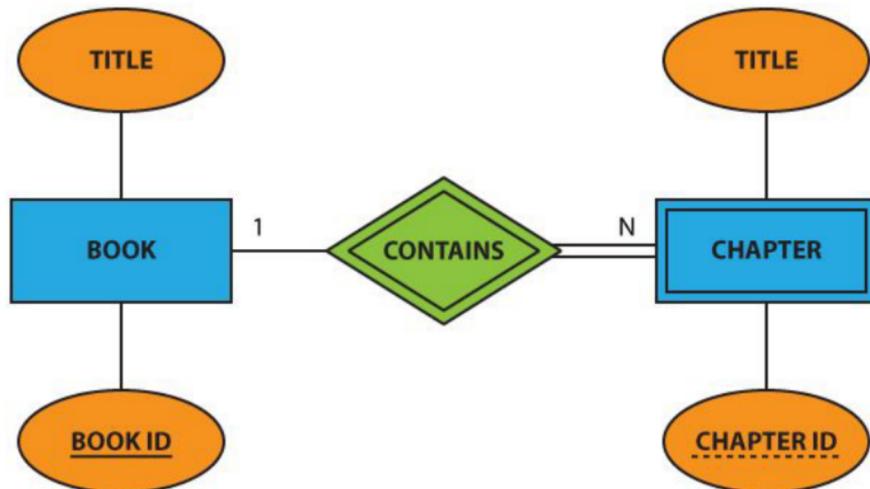
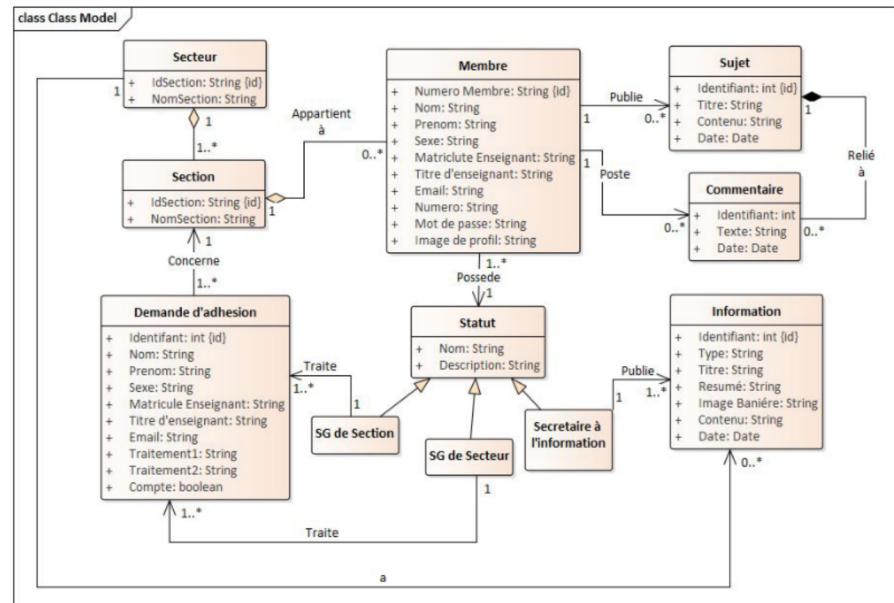
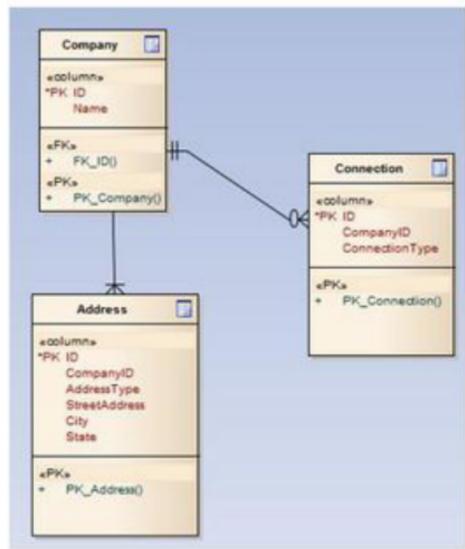


Crow's

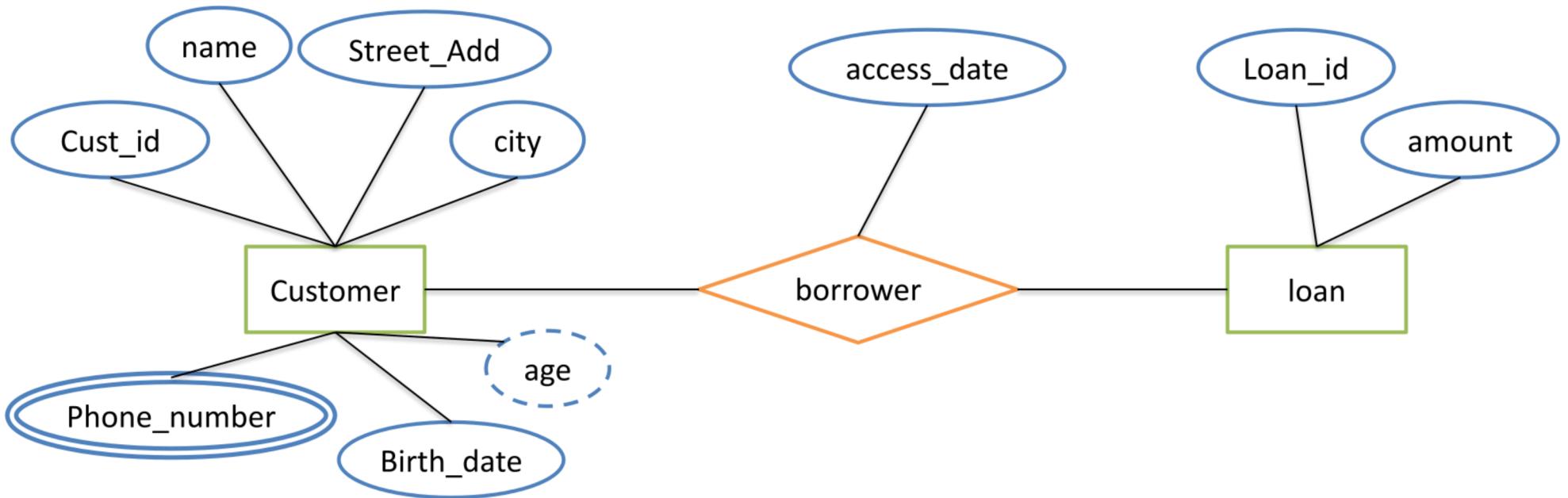


Foot

Notation



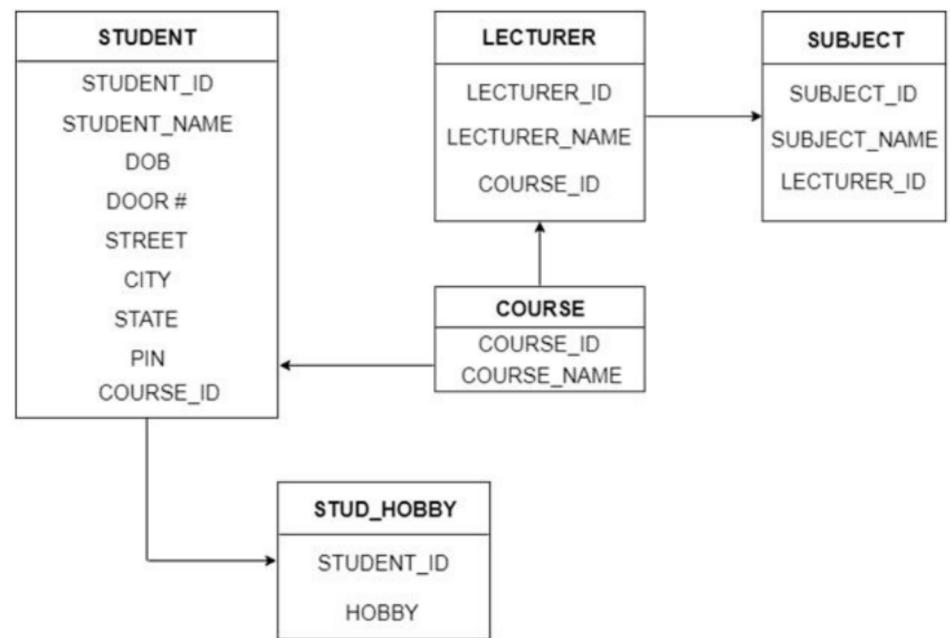
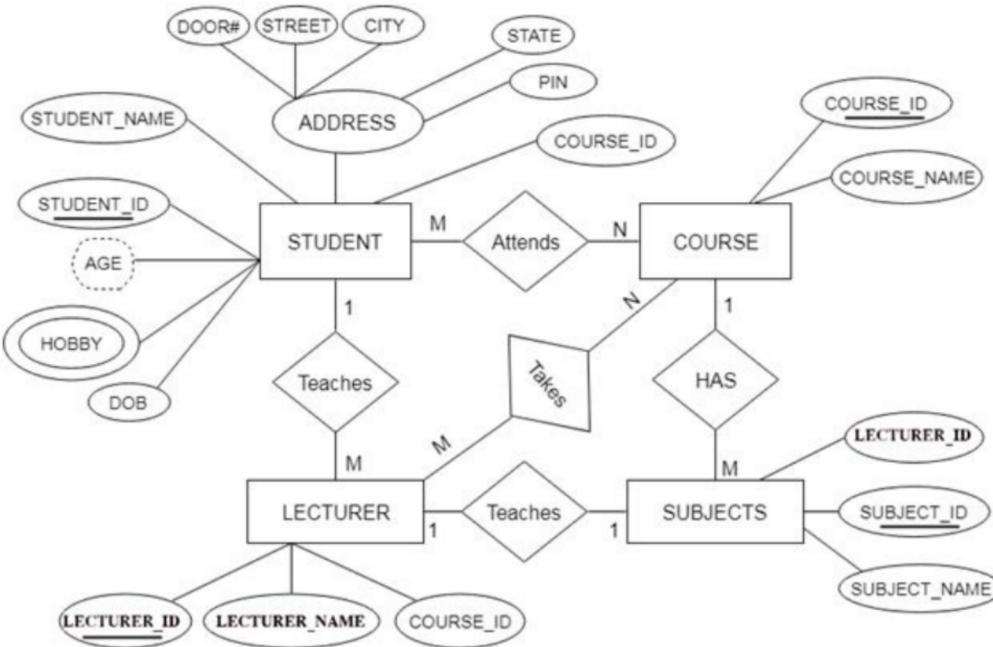
Example



Connection ————— Entity ————— Attribute ————— Multivalued attr ————— Derived attr ————— Relationship —————

Databases. Innopolis University. spring 2021.

ERD to Tables



Databases. Innopolis University. spring 2021.

Normalization(1/3)

- Database normalization is the process of structuring a relational database in accordance with a series of so-called normal forms in order to reduce **data redundancy and improve data integrity**. It was first proposed by *Edgar F. Codd* as part of his relational model.
- Normalization entails organizing the columns (attributes) and tables (relations) of a database to ensure that their dependencies are properly enforced by database integrity constraints. It is accomplished by applying some formal rules either by a process of **synthesis (creating a new database design) or decomposition (improving an existing database design)**.

Normalization(2/3)

- 1NF criteria
 - Primary key (no duplicate tuples)
 - No repeating groups
 - Atomic columns (cells have single value)
- 2NF criteria
 - Every non-trivial functional dependency either does not begin with a proper subset of a candidate key or ends with a prime attribute (no partial functional dependencies of non-prime attributes on candidate keys)

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean, Clash of the Titans	Ms.
Robert Phil	3 rd Street 34	Forgetting Sarah Marshal, Daddy's Little Girls	Mr.
Robert Phil	5 th Avenue	Clash of the Titans	Mr.

1NF

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3 rd Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3 rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5 th Avenue	Clash of the Titans	Mr.

Databases. Innopolis University. spring 2021.

FULL NAMES	PHYSICAL ADDRESS	MOVIES RENTED	SALUTATION
Janet Jones	First Street Plot No 4	Pirates of the Caribbean	Ms.
Janet Jones	First Street Plot No 4	Clash of the Titans	Ms.
Robert Phil	3 rd Street 34	Forgetting Sarah Marshal	Mr.
Robert Phil	3 rd Street 34	Daddy's Little Girls	Mr.
Robert Phil	5 th Avenue	Clash of the Titans	Mr.

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3 rd Street 34	Mr.
3	Robert Phil	5 th Avenue	Mr.

2NF

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

Databases. Innopolis University. spring 2021.

Normalization(3/3)

- 3NF Criteria
 - Every non-trivial functional dependency either begins with a super key or ends with a prime attribute (no transitive functional dependencies of non-prime attributes on candidate keys)

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION
1	Janet Jones	First Street Plot No 4	Ms.
2	Robert Phil	3 rd Street 34	Mr.
3	Robert Phil	5 th Avenue	Mr.

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

MEMBERSHIP ID	FULL NAMES	PHYSICAL ADDRESS	SALUTATION ID
1	Janet Jones	First Street Plot No 4	2
2	Robert Phil	3 rd Street 34	1
3	Robert Phil	5 th Avenue	1

MEMBERSHIP ID	MOVIES RENTED
1	Pirates of the Caribbean
1	Clash of the Titans
2	Forgetting Sarah Marshal
2	Daddy's Little Girls
3	Clash of the Titans

SALUTATION ID	SALUTATION
1	Mr.
2	Ms.
3	Mrs.
4	Dr.

Databases. Innopolis University. spring 2021.

3NF

Data definition language (DDL)

- In the context of SQL, data definition or data description language (DDL) is a syntax for creating and modifying database objects such as tables, indices, and users.

Examples

- Create statement
 - **CREATE OBJECTTYPE** *OBJECTNAME* [(optional parameter)]
 - **CREATE TABLE** [*table name*] ([column definitions]) [table parameters]
- Drop statement
 - **DROP OBJECTTYPE** *OBJECTNAME*
 - **DROP TABLE** *employees*;

Examples

- Alter statement
 - **ALTER objecttype objectname parameters.**
 - **ALTER table customer ADD country_code varchar2.**
- Truncate statement
 - **TRUNCATE TABLE table_name;**

Primary key(PK)

- Primary key is a specific choice of a minimal set of attributes (columns) that uniquely specify a tuple (row) in a relation (table).
- Informally, a primary key is "which attributes identify a record", and in simple cases are simply a single attribute: a unique id. More formally, a primary key is a choice of candidate key (a minimal superkey); any other candidate key is an alternate key.

Primary Key(PK)

Creating a simple primary key

```
CREATE TABLE Customer (
    cust_id INTEGER PRIMARY KEY,
    name TEXT,
    Street_address TEXT,
    city TEXT,
    birthday DATE,
    phone_number TEXT
);
```

Creating a complex primary key

```
CREATE TABLE Customer (
    cust_id INTEGER,
    name TEXT,
    Street_address TEXT,
    city TEXT,
    birthday DATE,
    phone_number TEXT,
    PRIMARY KEY
    (cust_id,phone_number)
);
```

Modifying an exiting table

```
ALTER TABLE customer ADD
PRIMARY KEY (cust_id);
```

Hint: when you are creating a complex primary key, sort the attributes according to the probability that they can be repeated

Foreign Key(FK)

- Foreign key is a set of attributes in a table that refers to the primary key of another table.
- The foreign key links these two tables. Another way to put it: In the context of relational databases, a foreign key is a set of attributes subject to a certain kind of inclusion dependency constraints, specifically a constraint that the tuples consisting of the foreign key attributes in one relation, R, must also exist in some other (not necessarily distinct) relation, S, and furthermore that those attributes must also be a candidate key in S.
- *In simpler words, a foreign key is a set of attributes that references a candidate key*

Foreign Key(FK)

Basic syntax

```
[CONSTRAINT fk_name]
  FOREIGN KEY(fk_columns) REFERENCES
  parent_table(parent_key_columns)
  [ON DELETE delete_action]
  [ON UPDATE update_action]
```

Simple example

```
CREATE TABLE borrower (
  borrower_id INTEGER PRIMARY KEY,
  cust_id TEXT,
  loan_id TEXT,
  access_date DATE,
  CONSTRAINT cust_borrow_fk
    FOREIGN KEY (cust_id)
    REFERENCES customer (cust_id),
  FOREIGN KEY (loan_id) REFERENCES
  loan (loan_id)
);
```

Modifying an exiting table

- SET NULL
- SET DEFAULT
- RESTRICT
- NO ACTION
- CASCAD

Exercise 1(1/2)

- Consider a MOVIE database in which data is recorded about the movie industry, in a piece of paper create the **relational schema** based on the data requirements are summarized as follows:
 - Each movie is identified by title and year of release. Each movie has a length in minutes. Each has a production company, and each is classified under one or more genres (such as horror, action, drama, and so forth). Each movie has one or more directors, and one or more actors appear in it. Each movie also has a plot outline. Finally, each movie has zero or more quotable quotes, each of which is spoken by a particular actor appearing in the movie.

Exercise 1(2/2)

- Actors are identified by name and date of birth and appear in one or more movies. Each actor has a role in the movie.
- Directors are also identified by name and date of birth and direct one or more movies. It is possible for a director to act in a movie (including one that he or she may also direct).
- Production companies are identified by name and each has an address. A production company produces one or more movies.

Exercise 2

- Using the relational schema built in the previous exercise, write an SQL script to create the database called MOVIE_DB and its tables, including primary and foreign keys and any other functional constraints.
- Check if your database meets 1NF, 2NF and 3NF normalization criteria, if it's not, let's normalize it.
- And write a second script to create new tables or alter the existing ones.

Useful resources

- <https://www.guru99.com/database-normalization.html>
- <https://www.javatpoint.com/dbms-reduction-of-er-diagram-int-o-table>

See you next week 😊