Databases Project – Spring 2017

Prof. Anastasia Ailamaki

Team No: 15

Names: Colin Branca, Jules Courtois, Yoan Martin

Contents

Contents 1

Deliverable 1 2

Assumptions 2

Entity Relationship Schema 2

Schema 2

Description 2

Relational Schema 2

ER schema to Relational schema 2

DDL 3

General Comments 3

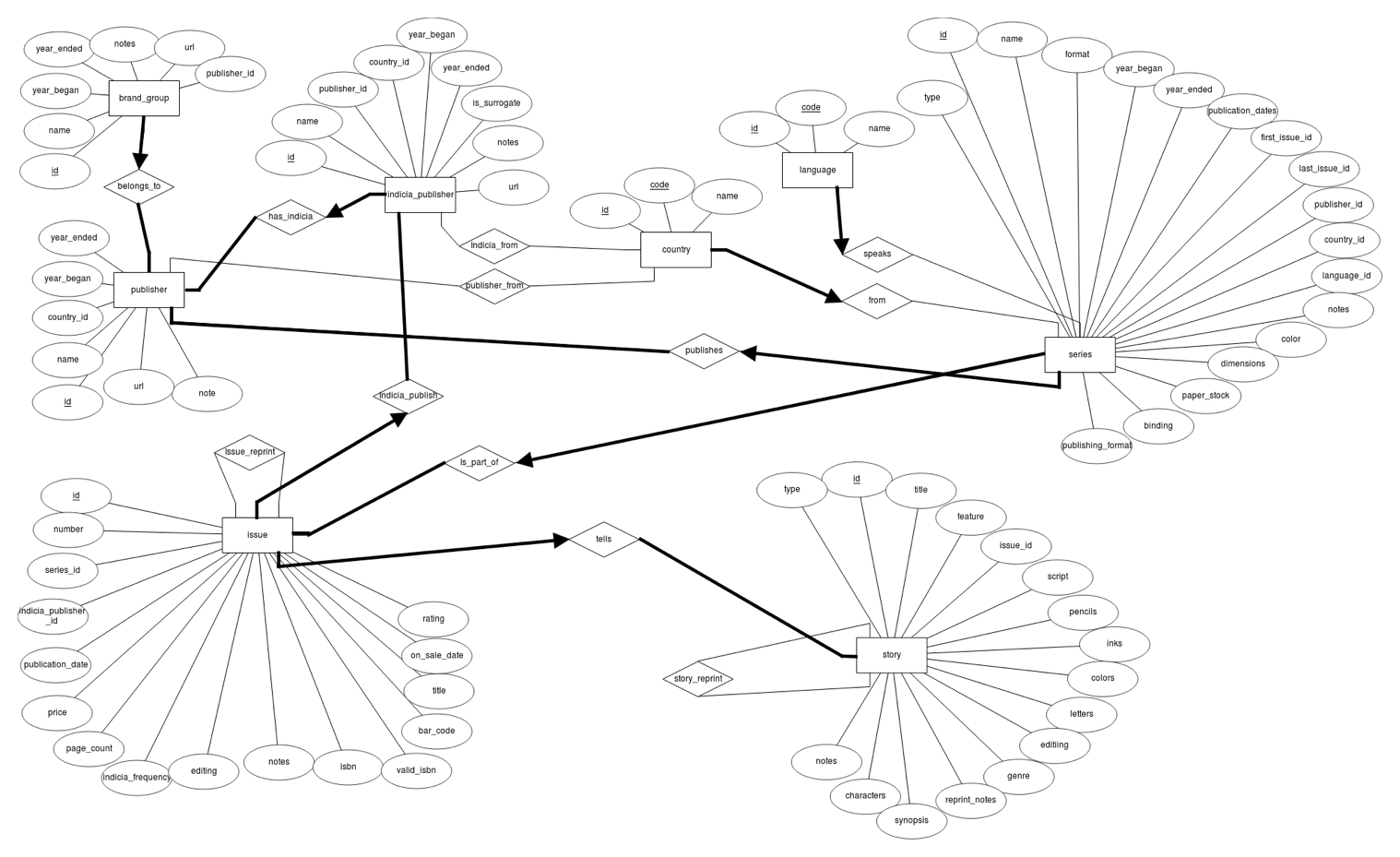
# Deliverable 1

## Assumptions

When we create the tables, we assume that each CHAR has a max size. No longer elements will be added later.

## Entity Relationship Schema

### Schema



### Description

There are two main parts in the schema: Books with Story, Issue and Series and Company with Publisher, Indicia Publisher and Brand Group. Let’s first describe each part before explaining the connections between them.

Books: This part is describing a book in the more general sense. The physical book is an issue. It tells a story and it is part of a series. For example, Harry Potter is a famous series. Harry Potter and the Philosopher’s Stone is a story from this series. Finally, the book with ISBN X is a physical book telling this story. Since these three components are highly connected. There are a lot of constraints between them.

1. An issue tells at least a story.
2. A story must be told by exactly one issue.
3. Each series has at least one issue.
4. An issue is part of exactly one series.

Company: This part is describing the company which publish some books. The main company is the publisher. It has smallest companies, the indicia publishers and it holds some brand group. For example, Marvel is the general company. It holds a brand group called Thor and the physical books are published by an indicia publisher, Thor Entertaining Group, which is part of Marvel. As for the book part, the company part has also a lot of constraints.

1. Each publisher has at least one brand group.
2. Each brand group belongs to exactly one publisher.
3. Each publisher has at least one indicia publisher.
4. Each indicia publisher is part of exactly one publisher.

Then, there are some connections between these two parts. The publisher publishes a series. It means that the publisher creates a series but the issues are printed by the indicia publisher. Here are the constraints related to these connections.

1. A publisher publishes at least one series.
2. A series is published by exactly one publisher.
3. An indicia publisher publishes at least one issue.
4. An issue is published by exactly one indicia publisher.

Finally, there are some smallest relationships. Firstly, the publisher, the indicia publisher and the series come from a country and the series have a language. Secondly, there is a reprint relationship between two stories or two issues. We consider only two constraints.

1. A publisher, indicia publisher or series comes from exactly one country.
2. A series has exactly one language.

N.B. We do not create a story\_type and series\_publication\_type entity since these data contain only one attribute if we exclude the id. To simplify the schema, we add an attribute type containing the name of the type in the Story/Series entity.

## Relational Schema

### ER schema to Relational schema

To create the table, we simply took each entity or each relationship and we translate them into a table. We took care to respect the constraints using NOT NULL or PRIMARY KEY to state the “at least” constraint and we permit each “at most” relation to appear in only one column in the connected table.

### DDL

CREATE TABLE StoryReprint (

id INT,

origin\_id INT NOT NULL,

target\_id INT NOT NULL,

PRIMARY KEY (id)

);

CREATE TABLE IssueReprint (

id INT,

origin\_issue\_id INT NOT NULL,

target\_issue\_id INT NOT NULL,

PRIMARY KEY (id)

);

CREATE TABLE Country (

id INT,

code CHAR(4) NOT NULL,

name CHAR(36) NOT NULL,

PRIMARY KEY (id),

UNIQUE (code),

UNIQUE (name)

);

CREATE TABLE Story (

id INT,

type VARCHAR(50) NOT NULL,

title VARCHAR(50) NOT NULL,

feature VARCHAR(50),

issue\_id INT NOT NULL,

script VARCHAR(50),

pencils VARCHAR(50),

inks VARCHAR(50),

color VARCHAR(50),

letters VARCHAR(50),

editing VARCHAR(50),

genre VARCHAR(50),

characters VARCHAR(50),

synopsis VARCHAR(150),

reprint\_notes VARCHAR(50),

notes VARCHAR(50),

PRIMARY KEY (id),

FOREIGN KEY (issue\_id) REFERENCES issue (id)

);

CREATE TABLE Issue (

id INT,

issue\_number INT,

series\_id INT,

indicia\_publisher-id INT,

publication\_date CHAR(15), --je ne sais pas comment définir la date car il y plusieurs format dans le fichier csv

price CHAR(10),

page\_count INT,

indicia\_frequency CHAR(10),

editing VARCHAR(20),

notes VARCHAR(1000),

isbn VARCHAR(20),

valid\_isbn VARCHAR(20),

barcode VARCHAR(10),

title VARCHAR(30),

on\_sale\_date INT,

rating INT,

PRIMARY KEY (id),

FOREIGN KEY (series\_id) REFERENCES Series (id),

FOREIGN KEY (indicia\_publisher) REFERENCES IndiciaPublisher(id)

);

CREATE TABLE BrandGroup (

id INT,

name VARCHAR2(128) NOT NULL,

year\_began SMALLINT,

year\_ended SMALLINT,

notes VARCHAR2(1023),

url VARCHAR(128),

publisher\_id INT NOT NULL,

PRIMARY KEY (id),

UNIQUE (name)

);

CREATE TABLE Publisher (

id INT,

name VARCHAR2(128) NOT NULL,

year\_began SMALLINT,

year\_ended SMALLINT,

notes VARCHAR2(3500),

url VARCHAR(128),

country\_id INT,

PRIMARY KEY (id),

UNIQUE (name)

);

CREATE TABLE IndiciaPublisher (

id INT,

name VARCHAR2(128) NOT NULL,

year\_began SMALLINT,

year\_ended SMALLINT,

notes VARCHAR2(3500),

url VARCHAR(128),

country\_id INT,

publisher\_id INT,

is\_surrogate NUMBER(3) NOT NULL,

PRIMARY KEY (id),

UNIQUE (name)

);

CREATE TABLE Language (

id INT,

code CHAR(3) NOT NULL,

name CHAR(36) NOT NULL,

PRIMARY KEY (id),

UNIQUE (code),

UNIQUE (name)

);

CREATE TABLE Series (

id INT,

name VARCHAR2(1000) NOT NULL,

type VARCHAR(16),

format VARCHAR(30),

publication\_dates VARCHAR(200),

year\_began SMALLINT,

year\_ended SMALLINT,

first\_issue\_id INT,

last\_issue\_id INT,

publisher\_id INT,

country\_id INT,

language\_id INT,

notes VARCHAR2(4000),

color VARCHAR(200),

dimensions VARCHAR(200),

paper\_stock VARCHAR(200),

binding VARCHAR(200),

publishing\_format VARCHAR(200),

PRIMARY KEY (id),

UNIQUE (name),

FOREIGN KEY (first\_issue\_id) REFERENCES Issue (id),

FOREIGN KEY (last\_issue\_id) REFERENCES Issue (id),

FOREIGN KEY (publisher\_id) REFERENCES Publisher (id),

FOREIGN KEY (country\_id) REFERENCES Country (id),

FOREIGN KEY (language\_id) REFERENCES Language (id)

);

## General Comments

For this deliverable, we design the schema together. Then we split the work individually. Colin Branca and Jules Courtois wrote the SQL commands to create the tables and Yoan Martin wrote this pdf document.