数据建模2

1. **找实体**

这是一个有关电影数据库的设计。

actors or actress name (first, middle, last)

director name (first, middle, last)

release date of the film

title of the film

the studio which releases the film

the parent studio (if any)

the name of the characters in each film

who played each role in each film

根据上述信息，该数据库可以表述为：演员（actor）出演电影（film），电影

（film）被导演（director）执导，电影（movie）由工作室（studio）制作。

确定实体为actor、film、director、studio。

  

1. **找关系**

“Each ACTOR may be a cast member of one or more FILMs.”

“Each FILM may be staffed by one or more ACTORs.”

“Each FILM must be produced by one and only one STUDIO.”

“Each STUDIO may be the producer of one or more FILMs.”

“Each FILM must be directed by one or only one DIRECTOR.”

“Each DIRECTOR may be the director of one or more FILMs.”

对演员而言，一个演员出演多部电影。对电影而言，一部电影由多个演员出

演。因此演员与电影间关系为多对多。

对电影而言，一部电影必须仅由一部工作室制作。对工作室而言，一部工作

室可以制作多部电影。因此电影与工作室间关系为多对一。

对电影而言，一部电影必须由一位导演执导。对电影而言，一部导演可以执

导多部电影。因此电影与导演关系为多对一。



1. **确定实体属性**

Film number, title, release date, actor last name, actor first name, actor middle name, director last name, director first name, director middle name, actor number, director number, studio name, studio’s parent name, character.

电影属性：film\_number、title、release time、film type以及actor\_information

director \_information、studio\_information

导演属性：director\_number、first\_name、last\_name、middle\_name

演员属性：actor\_number、first\_name、last\_name、middle\_name

工作室属性：studio\_number、studio name以及parent\_studio information

1. **确定独特标识符（主键及外键）**







1. **规范化数据模型**

考虑到业务逻辑，因为电影发布时间不唯一，美国发行时间或中国发行时间

或印度发行时间。release\_date可以再分，不能作为属性，所以需要单独建表date。



1. **解决多对多关系**

Film表和actor表关系为多对多，因此另建一张表actor\_film表，即采用联合

主键film\_id和actor\_id的方式，这种方式可唯一标识一个演员饰演一个角色。



若出现一人饰演多个角色，这种方式无法标识。可以在actor\_film表中添加

主键role\_number，而将原本联合主键actor\_number和film\_number设为外键。



1. **数据库模型的递归关系**

“Each STUDIO may be a subsidiary of one and only one STUDIO.”

“Each STUDIO may be the parent of one or more STUDIOs.”

每个工作室有可能是一家工作室的子工作室，每个工作室能成为一家或多家

工作室的母工作室。即对子工作室而言，有可能没有母工作室，若有母工作室，只能有一个。对母工作室而言，能有多个子工作室。因此母工作室与子工作室间关系为0…1对N。

在这里，studio表中，有studio\_number和studio\_pnumber，如：有一条记

录，1号工作室没有母工作室。则studio\_number为1，studio\_pnumber为0。另有一条记录，2号工作室是1号工作室的子工作室，则studio\_number为2，studio\_pnumber为1。

在数据库查询中，若要找出某工作室所有的子工作室。

select studio\_number from studio where studio\_pnumber = ‘3’;

1. **识别子类型和超类型的实体（supertype and subtype）**

Many of the “actors” in the documentaries are animals, wild or trained. For the musical films,

we will need to track the specific music group rather than an individual actor.

we need to expand the concept of the ACTOR entity into a supertype of PERFORMER, and

make ACTOR, MUSICAL GROUP, and ANIMAL subtypes.

Every PERFORMER will have a performer number. For each musical group, track the group

name and their musical speciality (e.g. rock, jazz, country). For each animal performer, track an optional name, whether the animal is trained or wild, and type of animal (e.g. dog, lion).

大多数纪录片中actor为动物animal（野生或驯服的）。对音乐电影而言，我

们需要追踪某个独特音乐团体musical group，而不是某个actor个人。

因此我们需要将actor实体扩展为超类型实体performer，同时让actor、

musical group、animal subtypes成为子类型。

每个performer会有一个performer number。对每个musical group而言，能

够追踪其不同于performer的属性musical group\_name、musical speciality(eg. rock,jazz,country)。对每个animal performer而言，也能追踪其不同于performer的属性name、trained\_wild、type。



1. **完整的数据库模型**

