Movies Schema Design

We're going to design a movies database. Each movie has a title and year, one (and only one) director, and some number of actors. Actors can star in multiple movies. Directors can direct multiple movies. Some movies have the same title such as Ocean's Eleven (the 2001 version directed by Steven Sodenbergh had George Clooney, Brad Pitt, Matt Damon, Julia Roberts, and many others, but the 1960 version was directed by Lewis Milestone and starred Frank Sinatra, Dean Martin and Sammy Davis Jr).

The schema should be normalized enough to avoid duplicating strings too much, and also to be able to efficiently answer questions like these two:

- Who acted in Fight Club (1999)?
- What are the 10 most recent movies that George Clooney starred in?

Solution

Movies Schema Design

Tables :

Movies :

- **mov_id (PK) int ,**
- Title Varchar(255) ,
- Years int,,
- **Dir_id (FK) int**

Actors :

- **Actor_id (PK) int ,**
- Act_Name Varchar(150)

Directors

- **Dir_id (PK) int ,**
- Dir_Name Varchar(150)

Movies_Actors

- **mov_id(FK) int**
- **Actor_id(FK) int ,**

Primary Keys :

- **Movies : mov_id**
- **Actors : Actor_id**
- **Directors : Dir_id**
- **Movies_Actors : mov_id , Actor_id (Composite Key)**

Foreign Keys :

- **Movies(Dir_id)** refers to **Directors(Dir_id)**
- **Movies_Actors(mov_id)** refers **Movies(mov_id)**
- **Movies_Actors(Actor_id)** refers **Actors(Actor_id)**

Indexes :

- **Movies : mov_id(pk) , Dir_id(FK)**
- **Actors : Actor_id (PK)**
- **Directors : Dir_id (Pk)**
- **Movies_Actors : mov_id , Actor_id (Composite Key index)**

Cardinality and Relationships :

1. **Movies to Director : one - to - many (1:m) -->**
    - Each movies has only one director but one director have multiple movies
2. **Movies to Actor : Many - to - many -->**
    - Each actor may have multiple movies and each movie have muliple actors
    - Managed by Movies_Actors table

Normalization - Avoiding Duplicates :

- **Directors Table:** Stores each director’s information only once.
- **Actors Table:** Stores each actor’s information only once.
- **Movies Table:** Stores each movie’s title and year only once

Queries :

- Who acted in Fight Club (1999)?

```
SELECT a.Act_Name
FROM Actors a
JOIN Movies_Actors ma
ON a.Actor_id = ma.Actor_id
JOIN Movies m
ON ma.mov_id = m.mov_id
WHERE m.Title = 'Fight Club' AND m.Years = 1999;
```

- What are the 10 most recent movies that George Clooney starred in?

```
SELECT m.Title, m.Years
FROM Movies m
JOIN Movies_Actors ma ON m.mov_id = ma.mov_id
JOIN Actors a ON ma.Actor_id = a.Actor_id
WHERE a.Act_Name = 'George Clooney'
ORDER BY m.Years DESC
LIMIT 10;
```