

Hand in a hardcopy of your answers (printed or handwritten). Include your name, and be sure to clearly identify which problems you are answering and which homework the problems are from since problems from multiple homeworks will be collected at the same time. These problems will be collected on Monday 9/28 and Monday 10/5.

You can hand-draw the ER diagrams, though you are encouraged to use PlantUML to draw the diagrams instead — PlantUML does the layout so you don't have to spend time formatting, and it is easy to edit your diagrams. Instructions for setting up and using PlantUML can be found on Canvas.

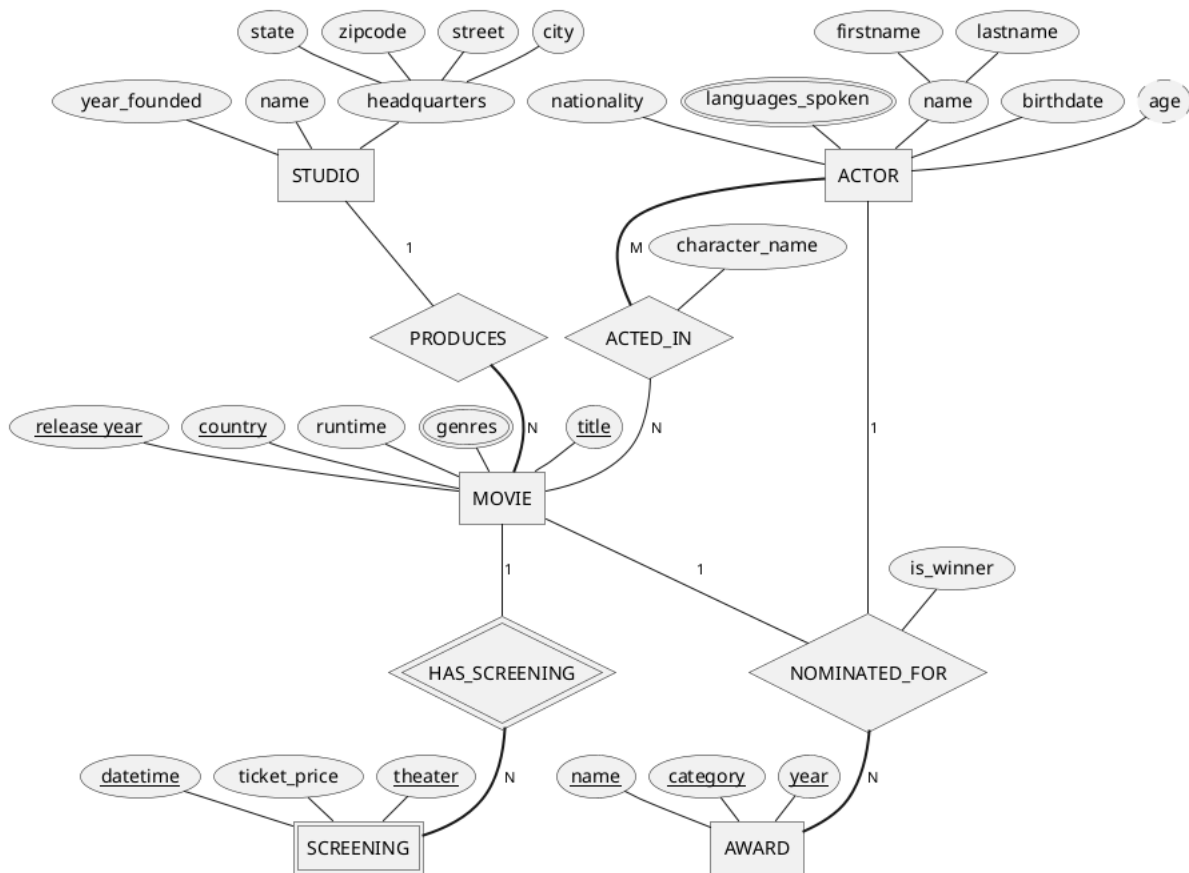


Figure 1: ER diagram for questions 1–2. Bold lines are substituted for a double lines.

1. True or false: You can get a list of actors ordered by last name.
Also give a brief justification for your answer, referencing the specific element(s) in the diagram that it depends on.

2. For each of the following statements, determine whether it is
- *allowed* — true in at least one valid instance of the data but not in every valid instance of the data,
 - *required* — true in every valid instance of the data, or
 - *forbidden* — not true in any valid instance of the data

according to the ER diagram above. Also give a brief justification for your answer, referencing the specific element(s) in the diagram that it depends on.

- (a) A movie has more than one genre.
- (b) An actor's age is consistent with their birth year.
- (c) Two actors have the same name (first name and last name).
- (d) There's a studio that hasn't produced any movies.
- (e) Every movie has a producing studio.
- (f) A movie is produced by two different studios.
- (g) Every actor has acted in more than one movie.
- (h) An actor plays multiple characters in the same movie.
- (i) An actor plays the same character in different movies.
- (j) There is a screening that isn't linked to a movie.
- (k) The same screening event involves two different movies (e.g., a double feature).
- (l) There's a movie with no screenings recorded.
- (m) There are multiple screenings at the same theater at the same date and time.
- (n) A movie has multiple screenings at the same theater.
- (o) Two different actors in a movie are nominated for the same award.
- (p) An actor is nominated for two different awards in the same movie.
- (q) An actor is nominated for the same award twice based on performances in two different movies.
- (r) There's an actor who has never been nominated for an award.
- (s) Every award has at least one nomination recorded.
- (t) There are two movies with the same title.
- (u) There are two movies with the same title, release year, and country but different running times.

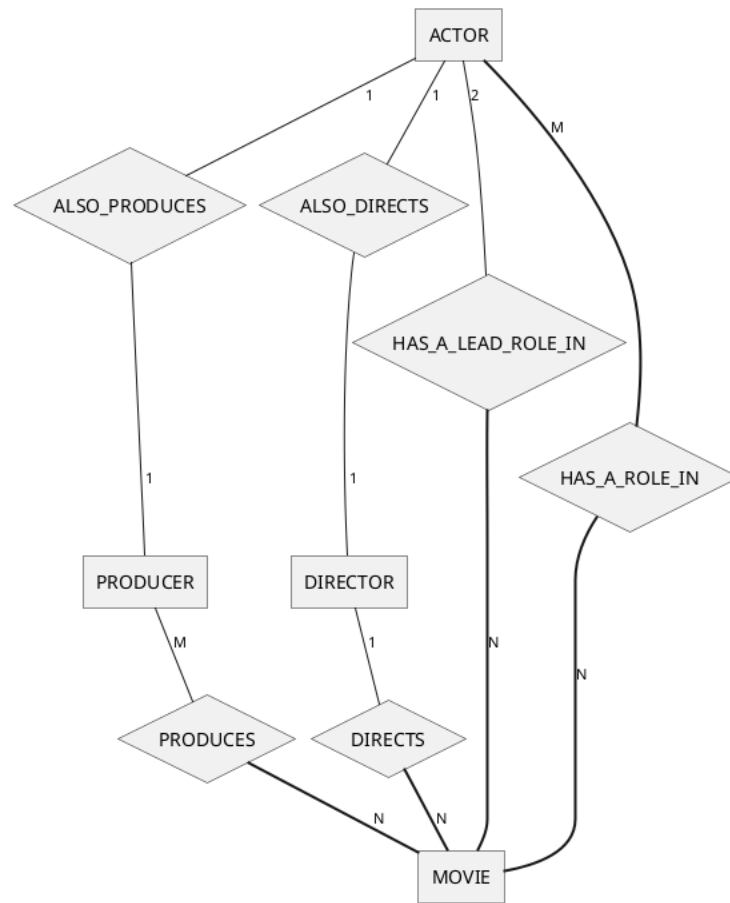


Figure 2: ER diagram for question 3. Bold lines are substituted for a double lines.

3. For each of the following statements, determine whether it is

- *allowed* — true in at least one valid instance of the data but not in every valid instance of the data,
- *required* — true in every valid instance of the data, or
- *forbidden* — not true in any valid instance of the data

according to the ER diagram above. Also give a brief justification for your answer, referencing the specific element(s) in the diagram that it depends on.

- No director has ever acted in a movie.
- A director who is also an actor has acted in at least one movie.
- No movie has a director that has also acted in that movie.
- Some producers have also been directors.
- An actor who has a lead role in a movie also has a (regular) role in the same movie.

For each of the following scenarios, draw an ER diagram to model what is described. Represent the situation as accurately as possible, including participation, cardinality, and key constraints as appropriate. Separately identify any attributes which are required (i.e. should be NOT NULL) and note any other constraints present that you did not capture in the diagram along with a reason why they were omitted. If necessary specifics are lacking in the description of the situation, make reasonable assumptions and state those assumptions.

4. Design a database for recording information about the activities (specifically votes) taken in the US House of Representatives, including:
 - Each Congress has a unique number (such as 115) and range of dates when it meets (such as January 3, 2017 to January 3, 2019).
 - Each US state has a name and region (Northeast, Midwest, Southeast, Southwest, West).
 - Each Representative is described by his or her name, represents a particular district in their state, and belongs to a political party (Republican, Democrat, Independent, Other). Note that a Representative's district and/or party affiliation may be different for different Congresses, but assume that it won't change during a single Congress.
 - A bill has a unique bill number, a title, at least one sponsor, the Congress it was introduced in, and the date on which it was introduced.
 - For each bill that has been voted on, the date of the vote, the number of "ayes" and "nos", and whether the bill passed or failed is recorded. Assume there can only be one vote on a given bill. In addition, how each Representative voted on the bill (yes, no, abstain, absent) is also recorded.
5. Design a database to manage conference papers. Researchers interested in presenting at a conference submit their research papers for consideration. Reviews by reviewers are recorded for use in the paper selection process. The database system caters primarily to reviewers who record answers to evaluation questions for each paper they review and make recommendations regarding whether to accept or reject the paper. The data requirements are summarized as follows:
 - Authors of papers are uniquely identified by email id. First and last names are also recorded.
 - Each paper is assigned a unique identifier by the system and is described by a title, abstract, and the name of the electronic file containing the paper.
 - A paper may have multiple authors, but one of the authors is designated as the contact author.
 - Reviewers of papers are uniquely identified by email address. Each re-

viewer's first name, last name, phone number, affiliation, and topics of interest are also recorded.

- Each paper is assigned between two and four reviewers. A reviewer rates each paper assigned to him or her on a scale of 1 to 10 in four categories: technical merit, readability, originality, and relevance to the conference. Finally, each reviewer provides an overall recommendation regarding each paper.
- Each review contains two types of written comments: one to be seen by the review committee only and the other as feedback to the author(s).