CSE-3421M Test #1

Design

Sur / Last Name:
Given / First Name:
Student ID:

- **Instructor:** Parke Godfrey
- **Exam Duration:** 75 minutes
- **Term:** Winter 2013

Answer the following questions to the best of your knowledge. Your answers may be brief, but be precise and be careful. The exam is closed-book and closed-notes. Calculators, etc., are fine to use. Write any assumptions you need to make along with your answers, whenever necessary.

There are four major questions, each with parts. Points for each question and sub-question are as indicated. In total, the test is out of 50 points.

If you need additional space for an answer, just indicate clearly where you are continuing.

Marking Box	
1.	/15
2.	/10
3.	/15
4.	/10
Total	/50

1. (15 points) **Relational Schema to E-R Diagram.** *I table that motion!* [Exercise]

Person(p#,name∗,birthdate,nationality,gender)
Actor(p#,aguild#∗)
FK (p#) refs **Person**
Director(p#,dguild#∗)
FK (p#) refs **Person**
Writer(p#, wguild#∗)
FK (p#) refs **Person**
Studio(name)
ScreenPlay(title,year)
Authored(title,year,writer)
FK (title, year) refs **ScreenPlay**
FK (writer) refs **Writer** (p#)
Movie(title,studio,year,genre∗,director∗,length)
FK (studio) refs **Studio** (name)
FK (title, year) refs **ScreenPlay**
FK (director) refs **Director** (p#)
Cast(title,studio,year,role actor,minutes)
FK (title, studio, year) refs **Movie**
FK (actor) refs **Actor** (p#)
Affiliated(director,studio)
FK (director) refs **Director** (p#)
FK (studio) refs **Studio** (name)

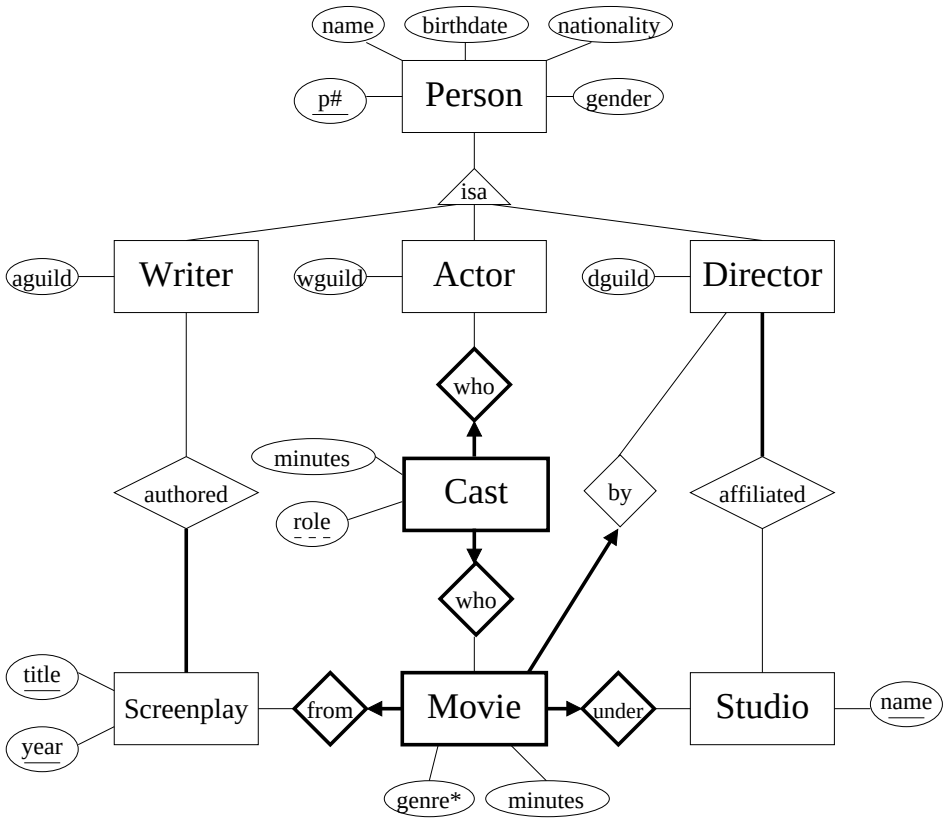
Figure 1: Movie Schema.

The basic schema of a database for tracking movies, actors, directors, and screenplay writers is shown in Figure 1. The underlined attributes indicate a table's primary key (and are hence not nullable). Additionally, attributes marked with a "∗" are not nullable. Foreign keys are indicated by FK.

Additional implicit constraints on the database are that a director should be affiliated with at least one studio, and a screenplay ought to have at least one author.

a. (12 points) *Reverse engineer* the relational schema in Figure 1 to provide an *accurate* entity-relationship diagram for it.

Do *not* make anything an entity that does not need to be.



Grading criteria:

- *2 ISA*
- *2 Cast (weak, additional key)*
- *2 Movie (weak, no extra key)*
- *2 M-M rel-ships (authored, affiliated)*
- *1 M-M mandatories*
- *2 proper keys*
- *no missing attr's*

b. (3 points) We want to ensure that the studio that produces a movie is directed by a director who is affiliated with that same studio.

Explain how the current schema ensures this already, if it does.

Otherwise, if it does not ensure this already, suggest a change to the relational schema so that it would.

Replace

Movie

FK (studio) refs **Studio** *(name)*
FK (director) refs **Director** *(p#)*

by

Movie

FK (director, studio) refs **Affiliated**

2. (10 points) **General.** *So much choice!* [Multiple Choice]

a. A relational database management system (RDBMS) does *not* do which of the following?

A. Provide mechanisms to help protect the integrity of the data.
B. Allow for concurrent transactions against the database.
C. Facilitate crash recovery of the database in case of hardware failure.
D. Optimize query evaluation for arbitrary SQL queries.
E. Ensure that relational schemas are in at least 3NF.

b. In an E-R diagram, *aggregation* is

A. a means to allow any relationship to be related as an entity to another entity via a relationship.
B. a means to allow a many-many relationship to be related as an entity to another entity via a relationship.
C. a means to treat any portion of an E-R diagram as if it were an entity.
D. a way to express aggregate operators like `min` and `max` from `SQL` in E-R.
E. another way to express a weak entity.

c. In an E-R diagram, if one sees a bold line with an arrow between an entity set and relationship set, this means

A. every entity in the entity set must participate in the relationship set.
B. every entity in the entity set must appear exactly once in the relationship set.
C. a relationship in the relationship set need not involve an entity from the entity set.
D. the entity set is weak.
E. the entity set *is an* instance of the relationship set.

d. Consider a relation with *three* attributes. How many candidate keys could it have at most at the same time?

A. 1
B. 3
C. 7
D. 8
E. *There is no limit.*

e. We know that table **Q** has only one candidate key. From this, we know the following.

A. **Q** is in 2NF.
B. **Q** is in 2NF but is not in 3NF.
C. If **Q** is in 3NF, it is also in BCNF.
D. **Q** cannot be in BCNF.
E. *None of the above.*

Consider the following schema.

```
create table T (
    c integer primary key,
    d integer);
create table S (
    b integer primary key,
    c integer references T(c) on delete cascade);
create table R (
    a integer primary key,
    b integer references S(b) on delete set null);
```

Assume that each of the tables **R**, **S**, and **T** each contain some tuples.

Consider the `SQL` statement "`DELETE FROM T;`" for Questions 2f–2h.

---

f. The `DELETE` statement

A. is certain to remove *all* the tuples from **T**.
B. is certain to remove *no* tuples from **T**.
C. may remove *some* tuples from **T**.
D. is syntactically incorrect because it is missing a `WHERE` clause.
E. *Not enough information to determine.*

---

g. The `DELETE` statement

A. is certain to remove *all* the tuples from **S**.
B. is certain to remove *no* tuples from **S**.
C. may remove *some* tuples from **S**.
D. is syntactically incorrect because it is missing a `WHERE` clause.
E. *Not enough information to determine.*

---

h. The `DELETE` statement

A. is certain to remove *all* the tuples from **R**.
B. is certain to remove *no* tuples from **R**.
C. may remove *some* tuples from **R**.
D. is syntactically incorrect because it is missing a `WHERE` clause.
E. *Not enough information to determine.*

The statement "check (one between 1 and 10)" means that the database system will only allow tuples for which one has a value between 1 and 10.

i. Consider the following table.

```
create table test (
    one integer,
    two integer,
    primary key(one, two),
    check (one between 1 and 10),
    check (two between 1 and 5)
);
```

How many tuples at most can this table contain?

A. 5
B. 10
C. 15
D. 50
E. *There is no limit.*

j. Consider the following table.

```
create table test (
    one integer,
    two integer,
    primary key (one),
    unique (two),
    check (one between 1 and 10),
    check (two between 1 and 5)
);
```

How many tuples at most can this table contain?

A. 5
B. 10
C. 15
D. 50
E. *There is no limit.*

3. (15 points) **Modelling.** *Strutting the E-R catwalk.* [Exercise]

   For the following, you may write it in the abbreviated style of the relational schema as presented in Figure 1 in Question 1. Do not make a table except if it is absolutely necessary. Take an attribute-minimalist point of view.

---

a. (5 points) Write a relational schema for the E-R in Figure 2.

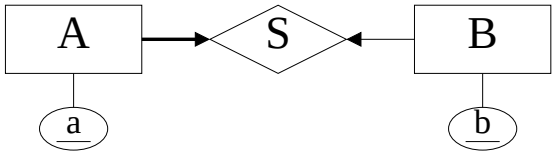


Figure 2: One to one.

A *(a, b∗)*
FK (b) refs **B** *% (for* **S***)*
unique (b)

B *(b)*

Careful that you only model the one *relationship.*

---

b. (5 points) Write a relational schema for the E-R in Figure 3.

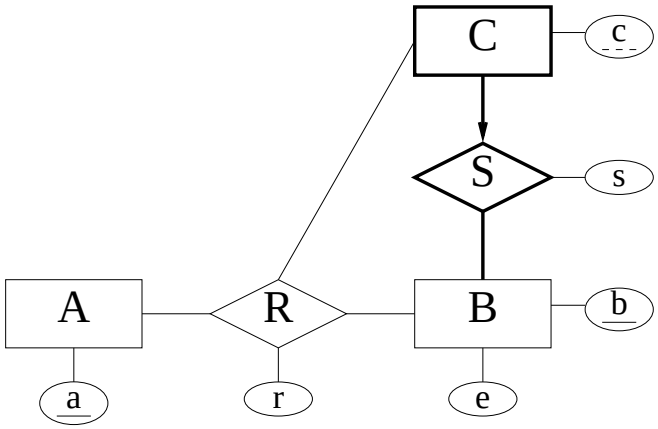


Figure 3: Ternary.

A *(a)*
B *(b, e)*
R *(a, b, c, r)*
FK (a) refs **A**
FK (b, c) refs **C**

C *(b, c, s)*
FK (b) refs **B** *% (for* **S***)*

c. (5 points) Draw an E-R diagram that replaces *relationship* **R** with an *entity* instead, but has the same *meaning* as the E-R diagram in Figure 4. (Do *not* add a new attribute to be the key of **R**.)

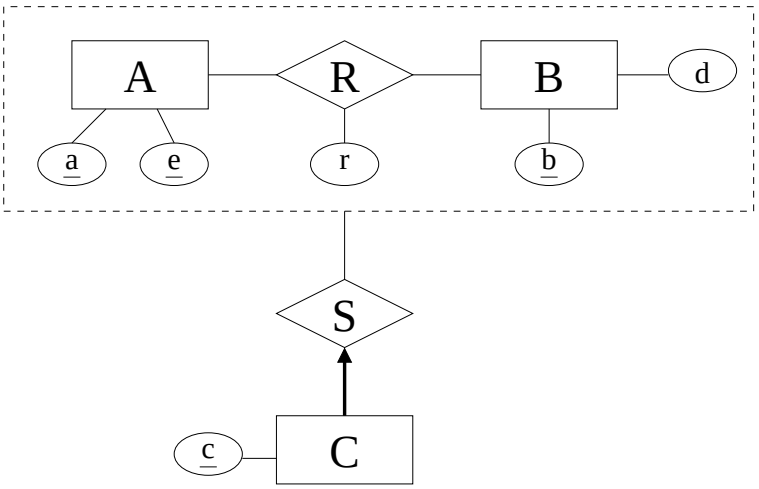
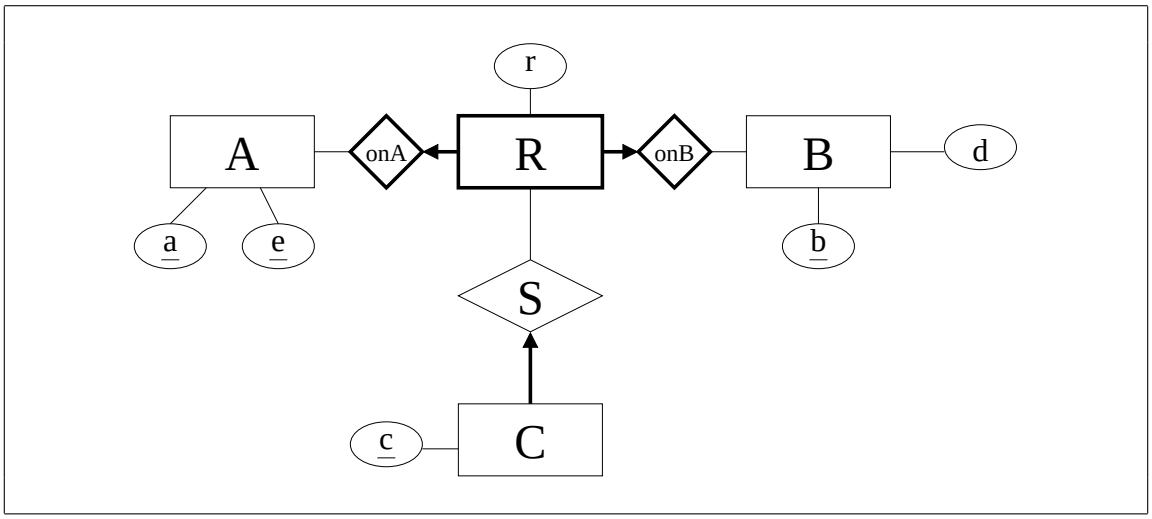


Figure 4: Many-many relationship.



4. (10 points) **Normalization.** *Normal, Alaska.* [Short Answer]

Consider the relation **R** with attributes A, B, C, D, E, and F, and with the functional dependencies

$$\begin{array}{ll} \mathsf{AB} \mapsto \mathsf{C} & \mathsf{B} \mapsto \mathsf{D} \\ \mathsf{AC} \mapsto \mathsf{B} & \mathsf{BC} \mapsto \mathsf{A} \\ \mathsf{AD} \mapsto \mathsf{E} & \mathsf{E} \mapsto \mathsf{F} \end{array}$$

a. (2 points) What are the *candidate keys* of **R**?

$$\begin{array}{c} AB* \mapsto CDEF \\ AC* \mapsto BDEF \\ BC* \mapsto ADEF \end{array}$$

So, AB, AC, BC are each candidate keys. (All other possibilities that determine all *the attr's are a superset of one of these, and so are super-keys.)*

b. (2 points) Is **R** in `3NF`? Justify your answer.

$B \mapsto D$ *breaks 3NF:* D *is not prime (part of some candidate key), nor is* $\{B\}$ *a key or superkey. (By the way, this breaks 2NF too.)*

c. (3 points) Consider the relation **R** with attributes A, B, C, D, and E, and with the following functional dependencies (FDs):

$$\mathsf{AB} \mapsto \mathsf{C} \qquad \mathsf{C} \mapsto \mathsf{A}$$
$$\mathsf{AD} \mapsto \mathsf{E} \qquad \mathsf{C} \mapsto \mathsf{D}$$

Dr. Mark Dogfurry claims that the decomposition step of **R** into ABC and CDE is a lossless join decomposition step.

Explain convincingly either that he is correct or that he is wrong.

$\{C\} = \{A, B, C\} \cap \{C, D, E\}$. So, we ask whether $C \mapsto AB$ or $C \mapsto DE$. $C* \mapsto ADE$, so $C \mapsto DE$!
This means Dr. Dogfurry is right; The two parts can be joined back losslessly.

d. (3 points) Consider a relation **R**. You are told that all of its candidate keys are singleton—that is, each candidate key is a set of just *one* attribute—and that **R** is in 3NF. Can you determine whether **R** is in BCNF?

Either show by an example that such a table need not be in BCNF, or prove that it is always the case that any such table is.

The only way that **R** *might not be in BCNF but be in 3NF is if there exists an FD $\mathcal{X} \mapsto A$ such that $\mathcal{X}$ is* not *a key or superkey but A is prime.*
However, since we know all *candidate keys of* **R** *are* singleton*, A must then be a candidate key itself! Since $\mathcal{X} \mapsto A$, so must $\mathcal{X}$ be. Contradiction.*

Extra space.

Extra space.

1NF: Domain of each attribute is an *elementary* type; that is, not a *set* or a *record structure*.

2NF: Whenever $\mathcal{X} \mapsto A$ is a functional dependency that holds in relation **R** and $A \notin \mathcal{X}$, then either

- A is *prime*, or
- $\mathcal{X}$ is not a proper subset of any key for **R**.

3NF: Whenever $\mathcal{X} \mapsto A$ is a functional dependency that holds in relation **R** and $A \notin \mathcal{X}$, then either

- A is *prime*, or
- $\mathcal{X}$ is a key or a super-key for **R**.

BCNF: Whenever $\mathcal{X} \mapsto A$ is a functional dependency that holds in relation **R** and $A \notin \mathcal{X}$, then

- $\mathcal{X}$ is a key or a super-key for **R**.

An attribute A is called *prime* if A is in any of the candidate keys.

Extra space.

Relax. Turn in your exam. Go home.