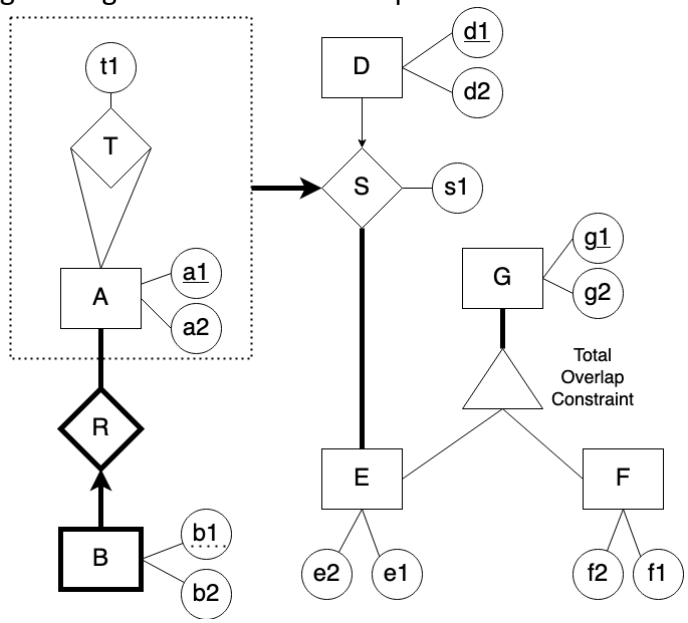
CPSC 368: Databases in Data Science
2021W2 Midterm Exam

Examining ER Diagrams

Consider the following ER diagram and answer the questions below.



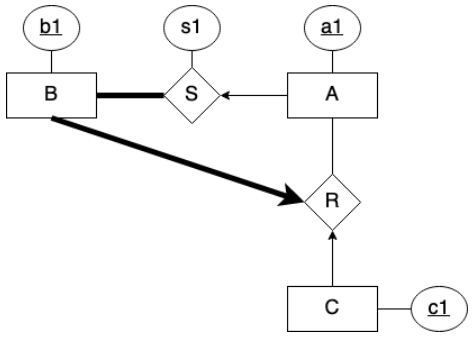
Indicate whether the following statement is True or False. Answers that are unreadable or unclear will be considered as incorrect.

Statement	True or False
It is possible for all A entities to participate in relationship set T (i.e., the ER diagram does not forbid that from happening).	
Some relationships in relationship set T may have a null value for t1.	
All A entities must participate in the relationship set T.	
There may be some A entities that do not participate in relationship set T.	
We can uniquely identify a given relationship in relationship set T if we know the value of the t1 attribute.	
When examining the A entities involved in a particular relationship T, the A entities cannot be the same. I.e., the A entities involved in a relationship T cannot have the same primary key.	
Multiple relationships in relationship set T may have the same t1 value.	
Multiple B entities can be associated with the same A entity.	

Statement	True or False
Multiple A entities can be associated with the same B entity.	
Given a particular A entity, all of the associated B entities must have unique b1 values.	
Given a particular A entity, all of the associated B entities must have unique b2 values.	
Each individual B entity must have a unique b1 value.	
Multiple B entities can have the same b1 value.	
All A entities must be associated with a B entity.	
Given a particular A entity, all of the associated B entities must have unique b1 AND b2 values.	
A G entity can either participate in entity set E or entity set F (but not both).	
Some G entities **do not** participate in entity set E or entity set F.	
All G entities participate in relationship set S.	
Some G entities participate in both entity set E or entity set F.	
Multiple E entities can have the same e1 and e2 values.	

Translate the ER Diagram

Consider the ER diagram below.



Translate the ER diagram to relations and fill in the table below.

If you create one relation to represent both an entity and a relationship, put the combined relation in the row for the entity and put 'N/A' in each box for the row for the relation. If a relation does not have a foreign key and/or a non-null attribute, put "N/A".

The relation containing…	**Attributes in the Relation**	**Primary key attribute(s)**	**Foreign key attribute(s)**	**Attributes that have NOT NULL constraints that must be stated explicitly**
The Entity Set A				
The Entity Set B				
The Entity Set C				
The Relationship Set R				
The Relationship Set S				

Writing SQL

Consider the following relations for a library:

1. Book(<u>isbn</u>, <u>edition</u>, title, author, year, category)
    - The primary key of this relation is {isbn, edition}. For the purposes of the exam, you can assume that ISBN is an integer value with no punctuation (e.g., dashes) in the value.
    - edition is the edition number for the book. Books like textbooks have different editions as the years go by.
    - title is the title of the book.
    - author is the primary author who wrote the book. In the case where a book is written by multiple authors, we only store the author who is listed first.
    - year is the year the book was first published.
    - category indicates whether a book is fiction, non-fiction, or a magazine.
2. User(<u>id</u>, firstName, lastName, age)
    - The primary key for this relation is {id}. This is an integer value that is uniquely generated for each user.
    - firstName and lastName make up the name of the user
    - age is the age when the user was first registered in our database. It is not a value that is updated as time goes on.
3. Borrows(<u>uid</u>, <u>isbn</u>, <u>edition</u>, <u>borrowDate</u>, returnDate)
    - The primary key for this relation is {uid, isbn, edition, borrowDate}.
    - uid refers to the id attribute of the User relation.
    - isbn and edition refer to the same named values in the Book relation.
    - borrowDate and returnDate refer to the did attribute in the Date relation.
4. Date(<u>did</u>, year, month, date)
    - The primary key of this relation is {did} which is an integer value.
    - This is a relation that stores multiple dates so that the Borrows relation can refer to it. In reality, we may design our database differently but we have made certain design choices to make this simpler for our exam.
    - year, month, and date are all integer values.

Categories Not Popular with 10 to 15 Year Olds

Find the category/categories of book that have never been borrowed by someone between the ages of 10 and 15 (inclusive). Do not use DISTINCT unless you need to.

If you want to, you can use views.

Most Well Read Users

Find the user (uid, firstName, and lastName) who has borrowed the most number of distinct books. I.e., if a user borrows the same book twice, that counts as one book, not two. Do not use the LIMIT or the FETCH FIRST x ROWS clauses.
If you want to, you can use views.

Equivalent?

Are the SQL queries below equivalent? By equivalent, we mean that given a relational instance, these queries would produce the same result every time. If the queries are equivalent, state "Yes". If they are not equivalent, provide a relational instance and then state what each of the queries would produce.

You can also omit columns that don't help you prove your point (much like what we do in the in-class exercise solutions).

Query 1:

```
SELECT age, COUNT(*)
FROM User
GROUP BY age
HAVING age > (SELECT AVG(age) FROM User)
```

Query 2:

```
SELECT age, COUNT(*)
FROM User
WHERE age NOT IN (SELECT AVG(age) FROM User)
GROUP BY age
```

Read All Fiction Books Between 2000 and 2015

Find the users who have read every fiction book published between 2000 and 2015. For the purposes of this question, the edition does not matter. For example, if you have read edition 10 of "Robots have a heart too!", you are considered as having read the book - even if this book has multiple editions.

If you want to, you can use views.