Midterm

Due Jul 3 at 11:59pm **Points** 250 **Questions** 25 **Available** Jul 1 at 6am - Jul 3 at 11:59pm 3 days
Time Limit 90 Minutes

Instructions

This is the final exam. Once you get started, you will have 75 minutes + 15 minutes (just in case of technical issues) to finish the exam. If you have special accommodation through school, your exam time will be set as the form indicates.

This quiz was locked Jul 3 at 11:59pm.

Attempt History

	Attempt	Time	Score
LATEST	Attempt 1	88 minutes	254 out of 250

Score for this quiz: **254** out of 250
Submitted Jul 1 at 6:44pm
This attempt took 88 minutes.

Question 1

4 / 4 pts

Attributes in the ER model can only have atomic values.

Correct!

- (•) True
- ( ) False

Question 2

4 / 4 pts

In the ER model, A participation constraint defines how every member of the entity set participates in the relationship

Correct!

- (•) True
- ( ) False

Question 3

0 / 4 pts

In an ER diagram, given two entity sets E with attributes ($\underline{e_1}$,e_2, e_3) in which e_1 is a key; and F with atttributes ($\underline{f_1}$,f_2,f_3) with f_1 is a key; and a relationship R between them. When translating it to relational model, is it OK to combine the table of R with the table of E, if relationship R is many-one from entity set E to entity set F? (This means that we have only two relations: E(e_1,e_2,e_3,f_1) and F(f_1,f_2,f_3))

Correct Answer

- ( ) True

You Answered

- (•) False

Question 4

4 / 4 pts

If an is-a hierarchy has a total of three entity sets, then translating the entire is-a hierarchy into relational tables using the ER approach would result in

- One table
- Two tables

Correct!

- Three tables
- Four tables

Question 5

4 / 4 pts

Consider relation 'Grades' and the query given below:

Grade

Student	CourseGrade	SeminarGrade
A	45	NULL
B	NULL	90
C	100	90

SELECT Student

FROM Grades

WHERE (CourseGrade> SeminarGrade AND SeminarGrade >75 AND CourseGrade >90) OR (CourseGrade<50)

Which students' tuples are returned?

- B and C only

Correct!

- A and C only
- A only
- A, B, and C

Question 6

4 / 4 pts

The difference between a candidate key and the primary key for a given relation is

Correct!

- The primary key is the key selected by the DBA or database designer from among the group of candidate keys, all of which uniquely identify a tuple
- There is no differences between them
- The candidate key is a superkey while primary key is not
- None of the above.

Question 7

4 / 4 pts

Which of the followings is a trival functional dependency?

- A -> AB
- AB -> C
- B -> A
- AB->A

Correct!

Question 8

4 / 4 pts

Which of the followings is not a ***basic*** relational algebra operator?

- Union
- Projection
- Join
- Selection

Correct!

Question 9

4 / 4 pts

In a relational model, if a relation instance has the cardinality 2000 and degree 200, how many tuples will be present in that relation?

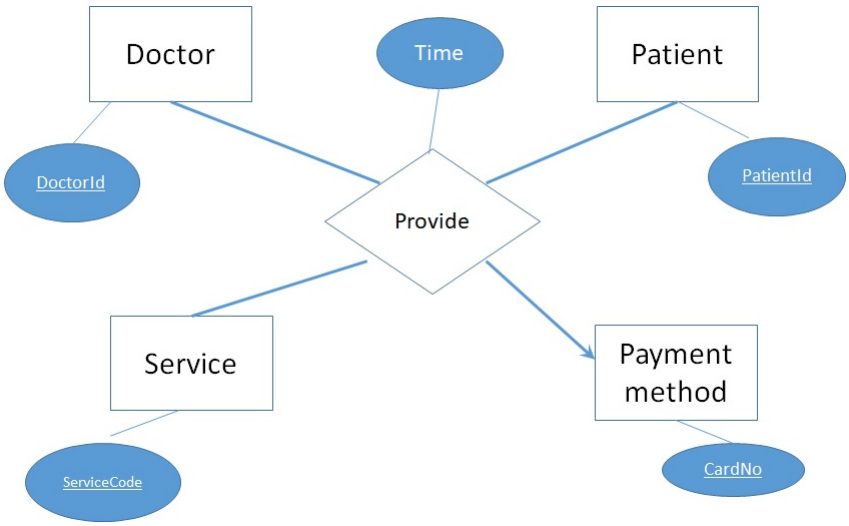
- 200
- 2000
- 400000
- None of the above

Correct!

Question 10

4 / 4 pts

Given the following ER diagram. We purposely omit non-key attributes in all participating entities to focus on the main idea of this question.



The relationship Provide will be translated into the following relation in a relational model:

Correct!

- Provide(**DoctorId , PatientId, ServiceCode,CardNo,time**)
- Provide(**DoctorId , PatientId, ServiceCode,CardNo,time**)
- Provide(**DoctorId , PatientId, ServiceCode,CardNo,time**)
- Provide(**DoctorId , PatientId, ServiceCode,CardNo,time**)

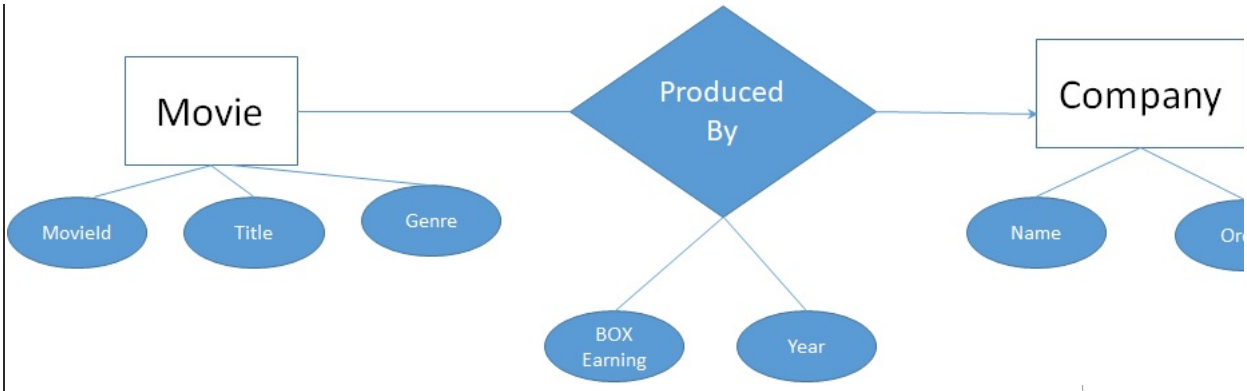
This text description is used for question 11 - question 18 regarding ER model and relational model:

We want to store information about movies. Each movie has a title, release year, genre and a movie identification (movieID). The movieID uniquely identifies each movie. Some movies are TV shows. Each TV show also specifies the starting episode, the ending episode, and the TV network that carries it (e.g HBO, Lifetime, USA…). Some movies are instructional training videos/DVDs for coaches and athletes. Each sport training video also specifies the specific sport, and trainer. Each sport training video is associated with one or multiple exercise apps that users could install in their phones or mobile devices. Each app has a name and platform. We can only uniquely identify each app using its name and the information of the sport training video that the app is associated with. Each movie is directed by one or more directors. Each director has director ID, a name, and multiple phone numbers. Each movie is also produced by exactly one motion picture company (such as Warner Bros. Entertainment, Sony Pictures, Walt Disney …) Each motion picture company has a name (which uniquely identifies the company) and a single address. When a motion picture company release a movie, we want to record the year of release and the total box office earning for that movie.

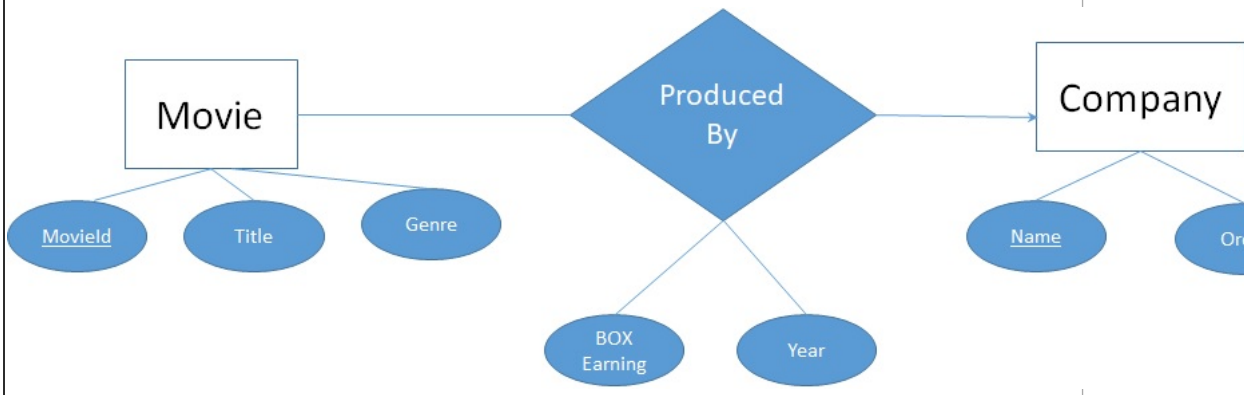
Question 11

5 / 5 pts

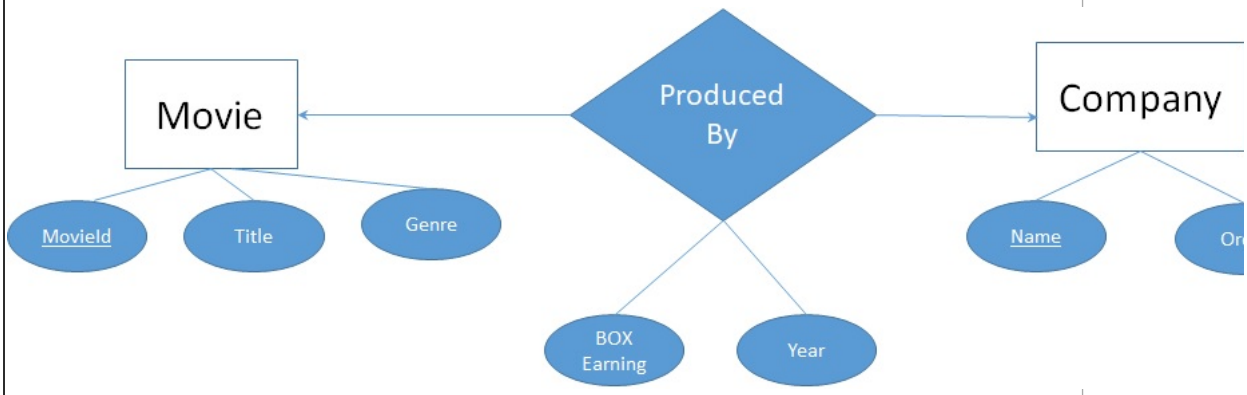
Which one of the following options correctly models the relationship between Movie and Company? Please choose only one option and justify your answer.



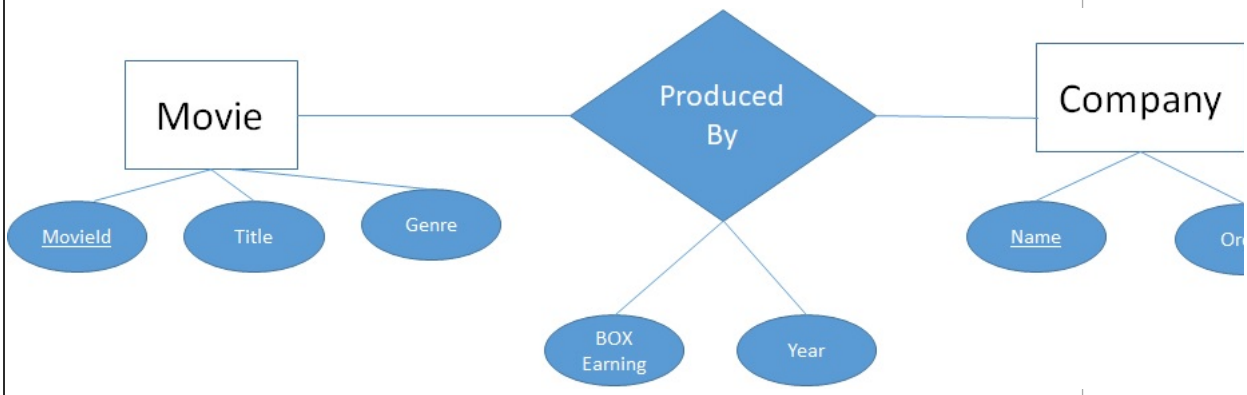
(a)



(b)



(c)



(d)

Your Answer:

B is the correct answer.

Each movie has a title, release year, genre and a movie identification (movieID).

The attributes in the Movie entity set are MovieId, Title and Genre. The primary key is MovieId because it uniquely identifies each movie

The attributes in the Company entity set are Name and Order and the name is the primary key as it uniquely identifiesthe company

The descriptive attributes for the relationship set between Movie and Company are Box earning and Year by the statement 'When a motion picture company release a movie, we want to record the year of release and the total box office earning for that movie.'

Since each movie is produced by exactly one company, the relationship from movie to company many to one. So the pointing arrow should be from the relationship set to the Company entity set

Question 12

5 / 5 pts

After you have chosen one ER diagram in question 11, please convert it to relational model. You could make decision to which domain/datatype for each attribute as you see fit.

Your Answer:

Company (Name: String, Order: String)

Movie (MovieID: Integer, Title: String, Genre: String, Name: String, BoxEarnings: Integer, Year: Integer)

Assumption

1. Assume Identification is integer type

2. Assume Integer type can express the Box Earnings

3. Name is the foreign key constraint that refers to the Name of Company

Correct Answer

Question 13

5 / 5 pts

If we want to make sure that every single movie in the Movie entity participated in ProducedBy relationship, what should we do? How do we represent it in the chosen ER diagram? You don't need to draw, you just need to describe what needs to be done.

Your Answer:

We can use total participation constraint from Movie to the relationship set ProducedBy.

To represent in ER diagram, The line between the entity set Movie and the relationship set ProducedBy should be a bold line.

Question 14

5 / 5 pts

If we want to make sure every company and movie in the ProducedBy relationship valid companies and movies, what should we do? How do we represent this constraint in the chosen ER diagram? Again, you don't need to draw, you just need to describe what needs to be done

Your Answer:

We can use referential integrity constraint to represent every company and movie in the ProducedBy relationship is valid.

To represent in the ER diagram, we use one semi-circle pointing arrow from ProducedBy to Movie and one semi-circle pointing arrow from ProducedBy to Company.

Movie (- ProducedBy -) Company

Question 15

5 / 5 pts

How do we represent the relationships between TVShow to Movie and Training Video to Movie? Please describe it as best as you could with or without pictures.

Your Answer:

Relationship between TV Show to Movie and Training Video to Movie is an ISA hierarchy relation.

The superclass entity in the relation is Movie

Below the Movie, a triangle representing ISA is drawn, instead of diamond, to represent the relationship set.

Then two arrows from the bottom of triangle should point to TV Show and Training Video.

TV Show will have the starting episode, the ending episode, and the TV network as attributes but no key.

Training video will have the specific sport, and trainer as attributes but no key.

Question 16

5 / 5 pts

How do we represent the relationship between Video App and Training video? Please describe it as best as you could with or without pictures

Your Answer:

Video App is an weak entity set of Training Video entity set.

The relationship set between the two entity sets will be named as 'Associate' while it is drawn by two diamonds (Inner and Outer diamonds).

Then the pointing arrow will be placed from the relationship set to Training video to represent Many to One relationship from Video App to Training Video.

The representation of Video App should be drawn with a rectangle in a bold line.

Line from Video App to the 'Associate' should be bold to represent total participation of Video App to the relationship set.

Question 17

5 / 5 pts

How would you describe the Director entity set? Do you consider Phone as an attribute of this entity? Why/Why not?

Your Answer:

In the Director entity set, the attributes would be DirectorId, and Name

Phone cannot be represented as an attribute of Director because a director can use multiple phones.

If it is represented as an attribute of Director, a director uses at most one phone, which is not corresponding to the condition.

Therefore, Phone is represented as an entity set that contains an attribute 'Phone Number', which is a key for the entity set.

Question 18

5 / 5 pts

How would you describe the relationship between Movie and Director. Please describe it as detailed as possible, with or without pictures

Your Answer:

Statement that refers to the relationship between Movie and Director is following:

- Each movie is directed by one or more directors.

We can infer that Movie is totally participating into the relationship between Movie and Director and the multiplicity in the relationship is Many to Many.

First, we create a relationship set called DirectedBy by drawing a diamond.

Then a line should be drawn from DirectedBy to Movie and from DirectedBy to Director to represent Many to Many relationship set.

Finally, the line from DirectedBy to Movie will be drawn by a bold line to represent total participation of Movie to the relationship set.

The following text is applied for SQL and relational algebra related questions (questions 19 and 20)

Consider a database schema with the following relations. The keys are underlined.

Emp(eId: integer, eName: string, age: integer, salary: real)

Works(eId: integer, dId: integer, pctTime: integer)

Dept(dId: integer, budget: real, managerId: integer)

An employee can work in more than one department; the **pctTime** of the Works relation shows the percentage of time that a given employee works in a given department. **managerId** in the Dept relation represents the **employee ID** of that manager.

Question 19

30 / 30 pts

a. Please write the following SQL query: Print the names and ages of each employee who works full time (pctTime = 100) for department with ID being 1. (15 points)

b. Please convert the SQL query in (1) to relational algebra (15 points)

For relational algebra, use the following notation:

Selection: σ (Click to symbol $\sqrt{x}$ -> Greek tab) OR write it out as SELECTION

Projection: π (Click to symbol $\sqrt{x}$ -> Greek tab) OR write it out as PROJECTION

Natural join: $\bowtie$ (Click to $\sqrt{x}$ symbol -> Relationship tab) OR write it out as NATURAL JOIN

Product: $\times$ (Click to $\sqrt{x}$ symbol -> Basic tab) OR write it out as PRODUCT

Your Answer:

A)

```
SELECT   e.eName, e.age
FROM     Emp e, Works w
WHERE    e.eId = w.eId
        AND w.pctTime =100
        AND w.dId = 1;
```

B)

$\pi_{eName,\ age}(\sigma_{pctTime = 100\ AND\ did = 1}(Emp \bowtie Works))$

Question 20

30 / 30 pts

a. Find the **managerIds** of managers who manage departments with budgets greater than $1,000,000 (15 points)

b. Please convert the SQL query in (1) to relational algebra (15 points)

For relational algebra, use the following notation:

Selection: σ (Click to symbol $\sqrt{x}$ -> Greek tab) OR write it out as SELECTION

Projection: π (Click to symbol $\sqrt{x}$ -> Greek tab) OR write it out as PROJECTION

Natural join: $\bowtie$ (Click to $\sqrt{x}$ symbol -> Relationship tab) OR write it out as NATURAL JOIN

Product: $\times$ (Click to $\sqrt{x}$ symbol -> Basic tab) OR write it out as PRODUCT

Your Answer:

Emp(eId: integer, eName: string, age: integer, salary: real)

Works(eId: integer, dId: integer, pctTime: integer)

Dept(dId: integer, budget: real, managerId: integer)

A)

```
SELECT  managerId
FROM    Dept
WHERE budget >$1000000;
```

B)

$\pi_{managerId}(\sigma_{budget > 1000000}(Dept))$

Question 21

25 / 25 pts

Key question

Suppose you are given a relation R with four attributes G, H, I and J. Assuming that the following functional dependencies hold for R:

$G \rightarrow H,\ HI \rightarrow J,\ G \rightarrow I$

a. Find G^+

b. Is *G* a candidate key for R? Please explain your answer.

Your Answer:

A)

G^+ = G, H, J, I

B)

G is a candidate key for the R because G can functionally determines all attributes in the Relation and there is no subset of G that determines all attributes of R

Question 22

25 / 25 pts

Keys and Joins question

Assume we have two relations ***R(A,B)*** and ***S(B,C)***. All three attributes **(A, B, and C)** are integer attributes. Assume that Relation R contains the following tuples: (1,2), (2,3), and (3,4). Assume that Relation S contains the following tuples (2,2), (2,3) and (5,1). Assume that these given tuples have reflected all functional dependencies hold in R and S.

a. Is A a candidate key for R? Please explain.

b. Is BC a superkey for S. Please explain.

c. How many tuples are in the result of the cross-product between R and S? How many attributes are in the results of the corss-product between R and S?

d. How many tuples are in the result of the natural join between R and S? How many attributes are in the result of the natural join between R and S?

Your Answer:

A	B	B	C
1	2	2	2
2	3	2	3
3	4	5	1

A.

A is a candidate key for R. In R, FD: A->B can be found. By attribute closure, A^+ = A, B.

Then we know A can functionally determines all attributes in R and there is no subset of A that functionally determines all attributes.

Therefore, A is the candidate key for R

B.

BC is a superkey for S because S has only two attributes which are B and C. Since BC contains all attributes S, it must contain candidate key.

Since BC contains a candidate key, it is a superkey

C.

By cross-product of R and S, it will produces 9 tuples (3X3) and 4 attributes (A,B,B,C)

D.

By natural join, there would be 3 attributes (A, B, C) and 2 tuples ({1, 2, 2} and {1, 2, 3}).

Question 23

18 / 20 pts

Normalization

Suppose you are given a relation R with four attributes A,B,C,D and the following set of functional dependencies. Assuming those are the **only dependencies that hold for R**

A→ B, BC → D, A → C

a. Identify the candidate key(s) for R.

b. Identify the **best normal form** that R satisfies (2NF, 3NF, or BCNF).

c. If R is not in BCNF, decompose it into a set of BCNF relations that satisfy lossless-join and preserve the dependencies. If such decomposition doesn't exist, show your reasoning.

Your Answer:

A.

A+ = A, B, C, D

B+ = B

C+ = C

D+ = D

BC+ = B, C, D

BD+ = B, D

BCD+ = BCD

Therefore, A is the only candidate key.

B.

R is not in BCNF because for FD: BC -> D, BC is not a superkey that contains the candidate key A

R is not in 3NF because for FD: BC -> D, BC is not a superkey and D is not a prime attribute that is not part of the candidate key

C.

A ->B is in BCNF because A is the superkey.

BC -> D violates BCNF because BC is not the superkey. So we decompose into R into R1(ABC) and R2(BCD)

Then R1 is still in BCNF while R2 is also in BCNF as BC becomes the candidate key.

A->C is in BCNF.

Check lossless join

R1 intersection R2 = BC.

If BC-> R1 or R2 is in S+ of FD (S is the set of FDs in R), it is lossless join

BC->BCD is in the S+ of FD by the BC->D so it is loseless join decomposition.

Check dependency preserving

A->B and A->C can be enforced in R1(ABC) while BC-> can be enforced in R2(BCD).

Therefore, the decomposition is dependency preserving.

As a result, the final decomposition of R is (ABC)(BCD)

b. need to specify that R is 2NF

Question 24

20 / 20 pts

Normalization

Given the relation R(E,F,G,H) with the following only functional dependencies

{E-> FG, G -> H}

a. Is R in 3NF form? Please explain.

b. Is R in BCNF form? If not, please find a good decomposition into BCNF. Please show the decomposition process and justify why it is considered as a good decomposition. If you could not find such a decomposition, please explain why.

Your Answer:

A.

By attribute closure,

E+ = E, F, G, H

F+ = F

G+ = G, H

H+ = H

FG+ = F, G ,H

FH+ = F, H

FGH+ = F, G, H

Therefore, the candidate key for R is E.

E->FG does not violate 3NF but G -> H violates 3NF because G is not a superkey of R while H is not a prime attribute.

So R is not in 3NF form

B.

FD: E->FG does not violate BCNF because E is a superkey of R

But G->H violates BCNF because G is not a superkey. So we decompose R into R1(EFG) and R2(GH)

R1(EFG) is in BCNF because E is the candidate key and R2(GH) is also in BCNF as G is the candidate key

Check lossless join

R1 intersection R2 = G

G+ = {G, H} which means G -> GH

Since closure of FD: G->H contains G->GH, it is lossless join decomposition

Check dependency preserving

FD: E->FG can be enforced in R1(EFG) while FD: G->H can be enforced in R2(GH).

Therefore, the decomposition is dependency preserving.

As a result, good decomposition of R is (EFG)(GH)

Question 25

20 / 20 pts

Suppose you are given a relation R(A,B,C,D). For the following set of functional dependencies, assuming they are the **only dependencies** that hold for R

A -> *B, B*-> CD, C -> *D*

a. Identify the candidate key(s) for R.

b. A developer proposed to decompose this relation into decompose into *AB* and *ACD*. State whether or not the proposed decomposition of R into smaller relations is a lossless-join and dependency preserving and briefly explain why or why not

Your Answer:

A)

By attribute closure,

A+ = A, B, C, D

B+ = B, C, D

C+ = C, D

D+ = D

We know A can functionally determine all attribute so A is a candidate key.

Also, we know no attribute except A can functionally determine A so there is no candidate key that does not contain A.

Therefore, A is the only candidate key.

B.

(AB)(ACD)

Check lossless join

(AB) intersection (ACD) = A

A -> ACD is contained in S+ of FDs (S is the set of FD) by the closure of FD: A->B.

So it is lossless join decomposition.

Check dependency preserving

B->CD cannot be enforced in either (AB) or (ACD) because there is no relation that contains B, C, D attributes.

Since there is a FD that cannot be enforced into the new relations, the decomposition is not dependency preserving.

The final decomposition is lossless join but not dependency preserving decomposition so it is not good decomposition.

Quiz Score: **254** out of 250

This quiz score has been manually adjusted by +10.0 points.