CSE 560

Fall 2024

Databases and Query Languages Homework -2

Name: Sahithya Arveti Nagaraju

Person number: 50559752

UBITname: sarvetin

Problem 1: Consider a database that includes the entity sets student, course, and section from the university schema and that additionally records the marks that students receive in different exams of different sections.

a. Construct an E-R diagram that models exams as entities and uses a ternary relationship as part of the design.

Solution:

To create an E-R diagram for the described university schema, I will consider the attributes below for Student, exam, section and course entities as per the textbook,

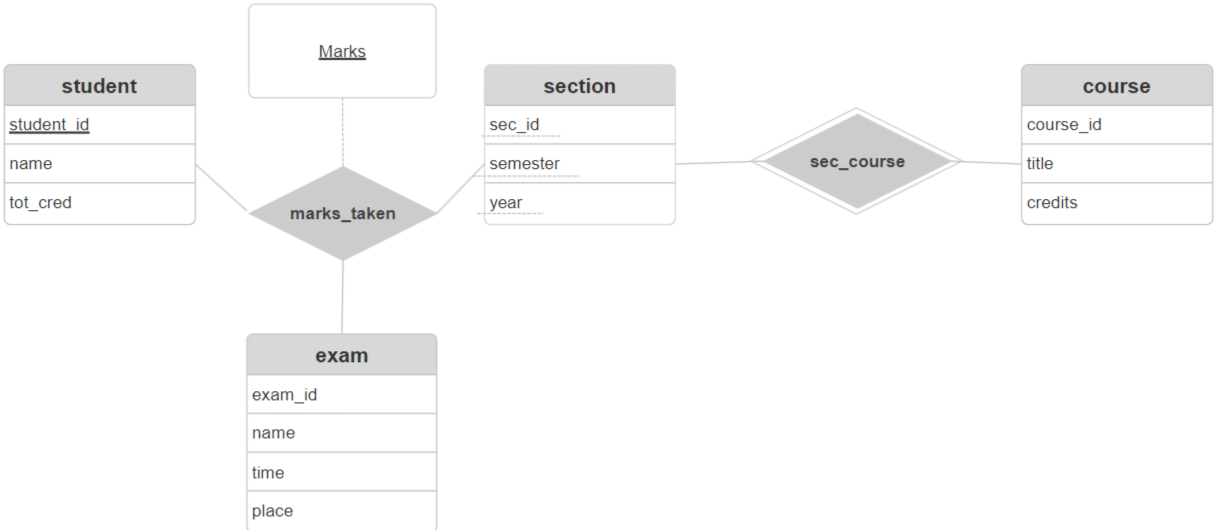
Entities

1. **student:** studentID (Primary Key), name, tot_cred
2. **course**: course_id (Primary Key), title, credits
3. **section**: section_id (Primary Key), semester, year
4. **exam**: exam_id (Primary Key), exam_name, place, time

Ternary Relationship

- **marks_taken**
  - This relationship connects the three entities: **student**, **section**, and **exam**.
  - Attributes of this relationship:
    - Marks (the score the student received in that exam for that section).

E-R Diagram Representation



Problem 2: Consider the schema R = (A, B, C, D, E, G,H) and the set F of functional dependencies: AB →CD,D → C,DE → B,DEH → AB, AC → DC Compute

a. A list of all candidate keys

b. A canonical cover or minimal basis for F, Fc

Solution:

a.

Closures:

$AB^+ = ABCD$

$D^+ = CD$

$DE^+ = BCDE$

$DEH^+ = ABCDEH$

$AC^+ = ACD$

$DEGH^+ = ABCDEGH$

$ABEGH^+ = ABCDEGH$

$ACEGH^+ = ABCDEGH$

$ADEGH^+ = ABCDEGH$

G is not there on right side of any functional dependency and so it must be part of candidate key. The keys that have all attributes in their closures are: DEGH, ABEGH, ADEGH and ACEGH. Among these, DEGH has a minimal number of attributes therefore it is considered as candidate key.

b.

We have the functional dependencies F={AB→CD,D→C,DE→B,DEH→AB,AC→DC}

First, we can remove C from the right side of AC→DC because it's unnecessary. This gives us F_c={AB→CD,D→C,DE→B,DEH→AB,AC→D}.

Next, C is also extraneous in AB→CD. We can derive AB→C from AB→D and D→C. So, we update it to F_c ={AB→D,D→C,DE→B,DEH→AB,AC→D}.

Additionally, B is unnecessary in DEH→AB because we can infer DEH→B from DE→B. This leads to F_c ={AB→D,D→C,DE→B,DEH→A,AC→D}.

After checking for any further extraneous attributes, the canonical cover of F is: F_c={AB→D,D→C,DE→B,DEH→A,AC→D}.

Problem 3: For each of the following relation schemas and sets of FD's:

a) R (A , B ,C ,D) with FD's AB→C, C→D and D→A.

b) R(A, B, C, D) with FD's A B →C ,B C→D, CD → A, and AD →B. do the following:

i) Indicate all the BCNF violations. Do not forget to consider FD's that are not in the given set, but follow from them. However, it is not necessary to give violations that have more than one attribute on the right side.

ii) Decompose the relations, as necessary, into collections of relations that are in BCNF.

Solution:

a) R (A ,B ,C ,D) with FD's AB→C, C→D and D→A.

i) Candidate keys are AB (AB^+ = ABCD), BC (BC^+=ABCD) and BD (BD^+=ABCD)

BCNF Violations:

1. C → D: C is not a superkey (it does not uniquely identify all attributes of R). Violation.
2. D → A: D is not a superkey. Violation.

ii) BCNF Decomposition:

C→D: This will be decomposed into R1(C, D) and R2(A, B, C)

C^+=ACD

1. Relation R1: (A,C,D)

    - FDs: C→D, D→A

    - D→A violates BCNF since D is not a superkey. Furthermore, decompose into R3(A, D) and R4(C,D).

    - R3(A,D): FD D→A Doesn't violate BCNF (because D is superkey).

    - R4(C,D): FD C→D Doesn't violate BCNF (Because C is superkey).

    - No further violation.

2. Relation R2: (B,C)

    - No BCNF violation.

Final decomposition:

- R2(B,C)

- R3(A,D)

- R4(C,D)

b) R(A, B, C, D) with FD's AB → C , BC→D, CD → A, and AD →B

i) Candidate keys are AB, BC, CD and AD because AB^+= BC^+ = CD^+ = AD^+ = ABCD

BCNF Violations:

1. AB → C: AB is a super key. No Violation
2. BC → D: BC is a super key. No Violation
3. CD → A: CD is a super key. No Violation.
4. AD → B: AD is a super key. No Violation.

ii) BCNF Decomposition:

Since there are no BCNF violations there is no need of BCNF decomposition.

Problem 4: For the given relation schemas and sets of FD's R (A ,B ,C ,D) with FD's AB→C, C→D and D→A.

i) Indicate all the 3NF violations.

ii) Decompose the relations, as necessary, into collections of relations that are in 3NF.

Solution:

Candidate key: AB (AB^+ = ABCD), BC (BC^+=ABCD) and BD (BD^+=ABCD)

i) 3NF Violations:

1. $C \rightarrow D$: C is not a super key but D is a prime attribute of candidate key BD. No violation.
2. $AB \rightarrow C$: AB is the super key. No violation.
3. $D \rightarrow A$: D is not a super key but A is a prime attribute of candidate key AB. No violation

ii) 3NF decomposition:

Since there are no 3NF violations. It is in 3NF.

Problem 5: Suppose we have relation R (A ,B ,C ,D ,E), with some set of FD's, and we wish to project those FD's onto relation S(A, B, C). Give the FD's that hold in S if the FD's for R are:

a) $AB \rightarrow DE$, $C \rightarrow E$, $D \rightarrow C$, and $E \rightarrow A$.

b) $A \rightarrow D$, $BD \rightarrow E$, A $C \rightarrow E$ and $DE \rightarrow B$.

Solution:
a) $AB \rightarrow DE$, $C \rightarrow E$, $D \rightarrow C$, and $E \rightarrow A$.

Closure of all subsets of {A,B,C}:

$A^+ = \{A\}$

$B^+ = \{B\}$

$C^+ = \{A, C, E\}$ FD: $C \rightarrow A$

$AB^+ = \{A, B, C, D, E\}$ FD: $AB \rightarrow C$

$BC^+ = \{A, B, C, D, E\}$ FD: $BC \rightarrow A$

$AC^+ = \{A, C, E\}$

$ABC^+ = \{A, B, C, D, E\}$

Thus, the FDs that hold in S are $AB \rightarrow C$ and $C \rightarrow A$ ($BC \rightarrow A$: since A can be determined from C of $C \rightarrow A$, it is redundant).

b) $A \rightarrow D$, $BD \rightarrow E$, $AC \rightarrow E$ and $DE \rightarrow B$.

Closure of all subsets of {A,B,C}:

$A^+ = \{A, D\}$

$B^+ = \{B\}$

$C^+ = \{C\}$

$AB^+ = \{A, B, D, E\}$

$BC^+ = \{B, C\}$

$AC^+ = \{A, B, C, E, D\}$ FD: $AC \rightarrow B$

$ABC^+ = \{A, B, C, D, E\}$

FD that project on S(A,B,C) is $AC \rightarrow B$ (from the closure of AC).

Problem 6: Let R(A, B, C, D, E) be decomposed into relations with the following three sets of attributes: {A, B, C}, {B, C, D}, and {A, C, E }. For each of the following sets of FD's, use the chase test to tell whether the decomposition of R is lossless. A→D , D→E , and B→D

Solution:

Initial Tableau:

A	B	C	D	E
a	b	c	d1	e1
a2	b	c	d	e2
a	b3	c	d3	e

Let's replace d3 with d since first and third rows agree in their A components (A→D)

A	B	C	D	E
a	b	c	d3	e1
a2	b	c	d	e2
a	b3	c	d3	e

Let's replace e1 with e since first and third rows agree in their D components (D→E)

A	B	C	D	E
a	b	c	d3	e
a2	b	c	d	e2
a	b3	c	d3	e

Let's replace d3 with d since first and second rows agree in their B components (B→D)

A	B	C	D	E
a	b	c	d	e
a2	b	c	d	e2
a	b3	c	d3	e

The first row has no symbols with subscript and the decomposition is lossless.

Problem 7: For each of the following relation schemas and dependencies

b) R(A, B, C, d) with MVD's A→→B and A→→C

c) R(A, B, C, D) with MVD A→→B and FD B→→C D.

do the following:

i) Find all the 4NF violations.

ii) Decompose the relations into a collection of relation schemas in 4NF.

Solution:

b) R(A, B, C, d) with MVD's A→→B and A→→C

i) 4NF violations:

A→→B: A is not a key. This is a violation of 4NF

A→→C: A is not a key and does not determine B and D. This is a violation of 4NF.

ii) 4NF Decomposition:

A→→B violates 4NF since A is not a key. So, decompose,

R1(A,B), R2(A,C,D)

R1 is in 4NF since A is key.

R2 is not in 4NF So decomposing into R3(A,C) and R4(A,D)

1. **Relation** R1(A,B):
    - Contains A→→B. A is a key. No violation
2. **Relation** R3(A,C):
    - Contains A→→C. A is a key. No violation.
3. **Relation** R4(A,D):
    - There are no FDs with A and D. It is in 4NF.

Final decomposition:

R1(A,B), R3(A,C) and R4(A,D)

c) R(A, B, C, D) with MVD A→→B and FD B→→CD.

i) 4NF violations:

A→→B: A is not a super key. Violation

B→→CD: B is not a key. Violates 4NF

ii) 4NF Decomposition:

A→→B: R1(A,B) R2(A,C,D)

R1 is in 4NF

R2(A,C,D) is in 4NF because of MVD transitivity of A→→B and B→→CD => A→→CD

So, R2 doesn't violate 4NF.

Final decomposition:

R1(A,B) and R2(A,C,D)