NEEDS_SKILL (SKILL-TYPE, JOB)
JOB_ASSIGNED (PERSON, JOB))

Now by applying the definition of 5NF, the join dependency is given as:

*((PERSON, SKILL-TYPE), (SKILL-TYPE, JOB), (PERSON, JOB))

The above statement is true because a join relation of these three relations is equal to the original relation PERSONS_ON_JOB_SKILLS. The consequence of these join dependencies is that the SKILL-TYPE, JOB or PERSON, is not relation key, and hence the relation is not in 5NF. Now suppose, the second tuple (row 2) is removed form relation PERSONS_ON_JOB_SKILLS, a new relation is created that no longer has any join dependencies. Thus the new relation will be in 5NF.

Review Questions

1. What do you understand by the term normalization? Describe the data normalization process. What does it accomplish?
2. Describe the purpose of normalising data.
3. What are different normal forms?
4. Define 1NF, 2NF and 3NF.
5. Describe the characteristics of a relation in un-normalised form and how is such a relation converted to a first normal form (1NF).
6. What undesirable dependencies are avoided when a relation is in 3NF?
7. Given a relation $R(A, B, C, D, E)$ and $F = (A \rightarrow B,\ BC \rightarrow D,\ D \rightarrow BC,\ DE \rightarrow \phi)$, synthesise a set of 3NF relation schemes.
8. Define Boyce-Codd normal form (BCNF). How does it differ from 3NF? Why is it considered a stronger from 3NF? Provide an example to illustrate.
9. Why is 4NF preferred to BCNF?
10. A relation $R(A, B, C)$ has FDs $AB \rightarrow C$ and $C \rightarrow A$. Is R is in 3NF or in BCNF? Justify your answer.
11. A relation $R(\underline{A}, B, C, D)$ has FD $C \rightarrow B$. Is R is in 3NF? Justify your answer.
12. A relation $R(\underline{A}, \underline{B}, C)$ has FDs $A. \rightarrow C$. Is R is in 3NF? Does $AB \rightarrow C$? Justify your answer.
13. Given the relation $R(A, B, C, D, E)$ with the FDs $(A \rightarrow BCDE,\ B \rightarrow ACDE,\ C \rightarrow ABDE)$, what are the join dependencies of R? Give the lossless decomposition of R.
14. Given the relation $R(A, B, C, D, E, F)$ with the set $X = (A \rightarrow CE,\ B \rightarrow D,\ C \rightarrow ADE,\ BD \rightarrow\!\!\rightarrow F)$, find the dependency basis of BCD.
15. Explain the following:
    (a) Why R_1 is in 1NF but not in but not 2NF, where
    $R_1 = (\{A, B, C, D\}, \{B \rightarrow D,\ AB \rightarrow C\})$
    (b) Why R_2 is in 2NF but not 3NF, where
    $R_2 = (\{A, B, C, D, E\}, \{AB \rightarrow CE,\ E \rightarrow AB,\ C \rightarrow D\})$

(c) Why R_3 is in 3NF but not BCNF, where

$R_3 = (\{A, B, C, D\}, \{A \rightarrow C, D \rightarrow B\})$

(d) What is the highest form of each of the following relations?

$R_1 = (\{A, B, C\}, \{A \leftrightarrow B, A \rightarrow C\})$

$R_2 = (\{A, B, C\}, \{A \leftrightarrow B, C \rightarrow A\})$

$R_3 = (\{A, B, C, D\}, \{A \rightarrow C, D \rightarrow B\})$

$R_4 = (\{A, B, C, D\}, \{A \rightarrow C, CD \rightarrow B\})$

16. Consider the functional dependency diagram as shown in Fig. 10.12. Following relations are given:

(a) SALE1 (SALE-NO, SALE-ITEM, QTY-SOLD)

(b) SALE2 (SALE-NO, SALE-ITEM, QTY-SOLD, ITEM-PRICE)

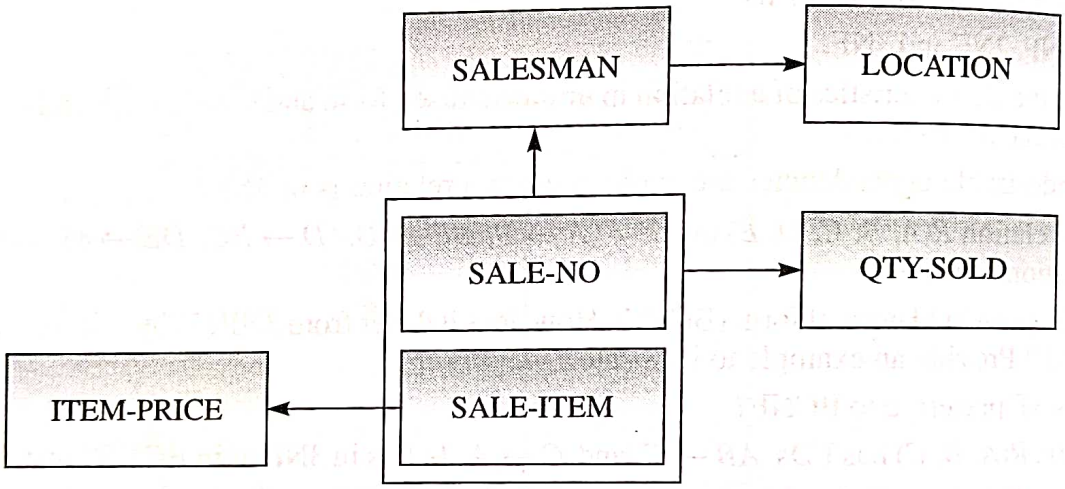
(c) SALE3 (SALE-NO, SALE-ITEM, QTY-SOLD, LOCATION)

(d) SALE4 (SALE-NO, QTY-SOLD)

(e) SALE5 (SALESMAN, SALE-ITEM, QTY-SOLD)

(f) SALE6 (SALE-NO, SALESMAN, LOCATION)

Fig. 10.12 Functional dependency diagram



(i) What are the relation keys of these relations?

(ii) What is the highest normal form of the relations?

17. Consider the following FDs:

PROJ-NO → PROJ-NAME

PROJ-NO → START-DATE

PROJ-NO, MACHINE-NO → TIME-SPENT-ON-PROJ

MACHINE-NO, PERSON-NO → TIME-SPENT-BY-PERSON

State whether the following relations are in BCNF?

R_1 = (PROJ-NO, MACHINE-NO, PROJ-NAME, TIME-SPENT-ON-PROJ)

R_2 = (PROJ-NO, PERSON-NO, MACHINE-NO, TIME-SPENT-ON-PROJ)

R_3 = (PROJ-NO, PERSON-NO, MACHINE-NO).

18. Define the concept of multi-valued dependency (MVD) and describe how this concept relates to 4NF. Provide an example to illustrate your answer.

19. Following relation is given:
STUDENT (COURSE, STUDENT, FACULTY, TERM, GRADE)
Each student receives only one grade in a course during a terminal examination. A student can take many courses and each course can have more than one faculty in a terminal.
(a) Define the FDs and MVDs in this relation.
(b) Is the relation in 4NF? If not, decompose the relation.

20. Following relation is given:
ACTING (PLAY, ACTOR, PERF-TIME)
This relation stores the actors in each play and the performance times of each play. It is assumed that each actor takes part in every performance. .
(a) What are MVDs in this relation?
(b) Is the relation in 4NF? If not, decompose the relation.
(c) If actors in a play take part in some but not all performances of the play, what will the MVDs?
(d) Is the relation of (c) is in 4NF? If not, decompose it.

21. A role of the actor is added in the relation of exercise 20, which now becomes
ACTING (PLAY, ACTOR, ROLE, PERF-TIME)
(a) Assuming that each actor has one role in each play, find the MVDs for the following cases:
(i) Each actor takes part in every performance of the play.
(ii) An actor takes part in only some performances of the play.
(b) In each case determine whether the relation is in 4NF and decompose it if it is not.

22. For exercise 6 of Chapter 9, design relational schemas for the database that are each in 3NF or BCNF.

23. Consider the universal relation R (A, B, C, D, E, F, G, H, I, J) and the set of FDs
$F = (\{A, B\} \rightarrow \{A\} \rightarrow \{D, E\}, \{B\} \rightarrow \{F\}, \{F\} \rightarrow \{G, H\}, \{D\} \rightarrow \{I, J\})$.
(a) What is the key of R?
(b) Decompose R into 2NF, then 3NF relations.

24. In a relation R (A, B, C, D, E, F, G, H, I, J), different set of FDs are given as
$G = (\{A, B\} \rightarrow \{C\} \rightarrow \{B, D\} \rightarrow \{E, F\}, \{A, D\} \rightarrow \{G, H\}, \{A\} \rightarrow \{I\}, \{H\}, \rightarrow \{J\})$.
(c) What is the key of R?
(d) Decompose R into 2NF, then 3NF relations.

25. Following relations for an order-processing application database of M/s KLY Ltd. are given:
ORDER (ORD-NO, ORD-DATE, CUST-NO, TOT-AMNT)
ORDER_ITEM (ORD-NO, ITEM-NO, QTY-ORDRD, TOT-PRICE, DISCT%)
Assume that each item has a different discount. The TOT-PRICE refers to the price of one item. ORD-DATE is the date on which the order was placed. TOT-AMNT is the amount of the order.
(a) If natural join is applied on the relations ORDER and ORDER_ITEM in the above database, what does the resulting relation schema look like?
(b) What will be its key?
(c) Show the FDs in this resulting relation.
(d) State why or why not is it in 2NF.
(e) State why or why not is it in 3NF.

26. Following relation for published books is given

BOOK (BOOK-TITLE, AUTH-NAME, BOOK-TYPE, LIST-PRICE, AUTH-AFFL, PUBLISHER)

AUTH-AFFL refers to the affiliation of author. Suppose that the following FDs exist:

BOOK-TITLE → PUBLISHER, BOOK-TYPE

BOOK-TYPE → LIST-PRICE

AUTH-NAME → AUTH-AFFL

(a) What normal form is the relation in? Explain your answer.

(b) Apply normalization until the relations cannot be decomposed any further. State the reason beh each decomposition.

27. Set of FDs given are $A \rightarrow BCDEF$, $AB \rightarrow CDEF$, $ABC \rightarrow DEF$, $ABCD \rightarrow EF$, $ABCDE$ $B \rightarrow DG$, $BC \rightarrow DEF$, $BD \rightarrow EF$ and $E \rightarrow BF$.

(a) Find the minimum set of 3NF relations.

(b) Designate the candidate key attributes of these relations.

(c) Is the set of relations that has been derived also BCNF?

28. A relation $R(\underline{A}, B, C, D)$ has FD $AB \rightarrow C$.

(a) Is R is in 3NF?

(b) Is R in BCNF?

(c) Does the MVD $AB \rightarrow\!\!\!\rightarrow C$ hold?

(d) Does the set $\{R_1(A, B, C), R_2(A, B, D)\}$ satisfy the lossless join property?

29. A relation $R(\underline{A}, B, \underline{C})$ and the set $\{R_1(A, B), R_2(B, C)\}$ satisfies the lossless decomposition property

(a) Is R in 4NF?

(b) Is B a candidate key?

(c) Does the MVD $B \rightarrow\!\!\!\rightarrow C$ hold?

30. Following relations are given:

(a) EMPLOYEE(E-ID, E-NAME, E-ADDRESS, E-PHONE, E-SKILL)
FD: E-ADDRESS → E-PHONE

(b) STUDENT(S-ID, S-NAME, S-BLDG, S-FLOOR, S-RESIDENT)
FD: S-BLDG, S-FLOOR → S-RESIDENT

(c) WORKER(W-ID, W-NAME, W-SPOUSE-ID, W-SPOUSE-NAME)
FD: W-SPOUSE-ID → W-SPOUSE-NAME

For each of the above relations,

(i) Indicate which normal forms the relations confirm to, if any.

(ii) Show how the relation can be decomposed into multiple relations each of which confirms to highest normal forms.

31. A life insurance company has a large number of policies. For each policy, the company wants to k the policy holder's social security number, name, address, date of birth, policy number, annual prem and death benefit amount. The company also wants to keep track of agent number, name, and cit residence of the agent who made the policy. A policy can have many policies and an agent can h many policies.

Create a relational database schema for the above life insurance company with all relation 4NF.

2e

Database Systems

Concepts, Designs and Application

Shio Kumar Singh

ALWAYS LEARNING

PEARSON