UNIT 2

Lecture 33

Normalization

3NF & BCNF

Third Normal Form (3NF)

- Third normal form (3NF) is based on the concept of ***transitive dependency***.
- A relation schema R is in 3NF if it is in 2NF and no nonprime attribute of R is transitively dependent on the primary key.
- A relation schema R is in 3NF if it is in 2NF and no nonprime attribute is dependent on other non prime attribute.
- **General Def. Of 3NF** – A relation schema R is in third normal form (3NF) if, whenever a nontrivial functional dependency X→A holds in R, either
  - X is a super key of R, or
  - A is a prime attribute of R.

Third Normal Form (3NF)

E.g. 1 Consider a relation R (A, B, C, D) with a set of FDs

F = { A → B,

B → C,

A → C,

C → D }

is not in 3NF, because it has 3 transitive dependencies B → C, C → D and B → D, with candidate key as A.

3NF Decomposition Algorithm

1. Compute the canonical cover F_C. (Remove Redundant FDs from F)
2. For every dependency $X \rightarrow Y$ in F_C, create a relation Ri (X, Y).
3. If none of the relations preserve the candidate key, then create a relation to preserve the candidate key. (If the key K of R does not occur in any relation Ri, create one more relation Ri(K)).
4. If possible merge 2 or more relations on the basis of common candidate keys.

Third Normal Form (3NF)

E.g. 1 Consider a relation R (A, B, C, D) with a set of FDs

F = { $A \rightarrow B$,

$B \rightarrow C$,

$A \rightarrow C$,

$C \rightarrow D$ }

Sol : First we have to find F_C.

F_C = { $A \rightarrow B$,

$B \rightarrow C$,

$C \rightarrow D$ }

So, the required 3NF decomposition of R is R1(<u>A</u>, B), R2(<u>B</u>, C) and R3(<u>C</u>, D).

Third Normal Form (3NF)

E.g. 2 Consider a relation R (A, B, C, D, E) with a set of FDs

F = { $A \rightarrow B$,

$B \rightarrow E$,

$A \rightarrow C$,

$C \rightarrow D$,

$A \rightarrow E$ }

Sol : First we have to find F_C.

F_C = { $A \rightarrow BC$,

$B \rightarrow E$,

$C \rightarrow D$ }

So, the required 3NF decomposition of R is R1(<u>A</u>, B, C), R2(<u>B</u>, E) and R3(<u>C</u>, D).

Third Normal Form (3NF) (Approach – I)

E.g.3 Consider a relation R (A, B, C, D, E) with a set of FDs

F = { AB → C,

A → D,

D → E }

Sol : First we compute F^+.

So. F^+ = { AB → CDE,

A → DE,

D → E }

From F^+, we have 2 partial dependencies A → D, and A → E.

R (A, B, C, D, E)

A → DE

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

So, relations R1 (A, D, E) with F1 = {A → DE, D → E} and R2 (A, B, C) with F2 = {AB → C} are the required 2NF decompositions of R.

Now, relation R2 is already in 3NF but R1 is not in 3NF because it contains a transitive dependency D → E. So again we decompose R1(A, D, E) into R11(A, D) with F11 = {A → D} and R12(D, E) with F12 = {D → E}. So, R11 (A, D), R12 (D, E) and R2 (A, B, C) are the required 3NF decomposition of R.

Third Normal Form (3NF) (Approach – II)

E.g.3 Consider a relation R (A, B, C, D, E) with a set of FDs

F = { $AB \rightarrow C$,

$A \rightarrow D$,

$D \rightarrow E$ }

Sol : First we compute F_C.

So. F_C = { $AB \rightarrow C$,

$A \rightarrow D$,

$D \rightarrow E$ }

From F_C , we have 3 dependencies $AB \rightarrow C$, $A \rightarrow D$ and $D \rightarrow E$.

So, relations R1 (<u>A, B</u>, C), R2 (<u>A</u>, D), and R3 (<u>D</u>, E) are the required 3NF decompositions of R.

Third Normal Form (3NF) (Approach – I)

E.g.4 Consider a relation R (A, B, C, D) with a set of FDs

F = { A → B,

C → D }

Sol : First we compute F^+.

So. F^+ = { A → B,

C → D }

From F^+, we have 2 partial dependencies A → B, and C → D.

R (A, B, C, D)

A → B

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

C → D

R21 (C, D) R22 (A, C)

So, relations R1 ($\underline{A}$, B), R21 ($\underline{C}$, D) and R22 ($\underline{A, C}$) are the required 2NF decompositions of R.

Now, relations R1 ($\underline{A}$, B), R21 ($\underline{C}$, D) and R22 ($\underline{A, C}$) are also in 3NF, because they do not contain any transitive dependencies.

Third Normal Form (3NF) (Approach – II)

E.g.4 Consider a relation R (A, B, C, D) with a set of FDs

F = { $A \rightarrow B$,

$C \rightarrow D$ }

Sol : First we compute F_C.

So. F_C = { $A \rightarrow B$,

$C \rightarrow D$ }

From F_C , we have 2 dependencies $A \rightarrow B$, and $C \rightarrow D$.

So, we create 2 relations R1 (<u>A</u>, B), and R2 (<u>C</u>, D).

Now, none of the relations preserve the candidate key AC, so we create a relation R3 (<u>A, C</u>) to preserve the candidate key. So, R1 (<u>A</u>, B), R2 (<u>C</u>, D) and R3 (<u>A, C</u>) are the required 3NF decompositions of R.

Boyce Codd Normal Form (BCNF)

- **General Def. Of BCNF** – A relation schema R is in BCNF if whenever a nontrivial functional dependency X→ A holds in R, then X is a super key of R.
- A relation schema R is in BCNF if it is in 3NF and every determinant is a candidate key of R.
- Boyce-Codd normal form (BCNF) was proposed as a simpler form of 3NF, but it was found to be stricter than 3NF. That is, every relation in BCNF is also in 3NF; however, a relation in 3NF is not necessarily in BCNF.

Boyce Codd Normal Form (BCNF)

E.g. 1 Consider a relation R (A, B, C) with a set of FDs

F = { AB → C,

C → A }

is not in BCNF, because it has 2 candidate keys AB and BC, and the determinant of C → A is not an super key of R.

BCNF Decomposition Algorithm

1. Compute the closure of set of FDs F^+.
2. For every dependency $X \rightarrow Y$ in F^+ , where X is not a super key of R, create 2 relations Ri (X ,Y) and Rj(R – Y).
3. If possible merge 2 or more relations on the basis of common candidate keys.

Boyce Codd Normal Form (BCNF)

E.g.1 Consider a relation R (A, B, C) with a set of FDs

F = { $AB \rightarrow C$,

$C \rightarrow A$ }

Sol : First we compute F^+.

So. F^+ = { $AB \rightarrow C$,

$C \rightarrow A$ }

From F^+, we have 1

dependency $C \rightarrow A$, which violates BCNF property.

R (A, B, C)

$C \rightarrow A$

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

So, relations R1 (A, C) with F1 = {$C \rightarrow A$}, and R2 (B, C) with F2 = { ø } are the required BCNF decompositions of R.

Boyce Codd Normal Form (BCNF)

E.g.2 Consider a relation R (A, B, C, D) with a set of FDs

F = { ABC → D,

D → B,

D → C }

Sol : First we compute F^+.

So. F^+ = { ABC → D,

D → BC }

From F^+, we have 1 dependency D → BC which violates property.

R (A, B, C, D)

D → BC

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

So, relations R1 (B, C, $\underline{D}$) with F1 = {D → BC}, and R2 ($\underline{A, D}$) with F2 = { ø } are the required BCNF decompositions of R.

Objective of Normalization

- *"to create relations where every dependency is on the key, the whole key, and nothing but the key".*

Levels of Normalization

- 1NF
  - 2NF
    - 3NF
      - BCNF (Boyce Codd Normal Form)
        - 4NF
          - 5NF (PJNF (Project Join Normal Form))
            - 6NF (DKNF (Domain Key Normal Form))

Practice Question

Consider a relation schema $R = (A, B, C, D)$ with a set of functional dependencies $F = \{C \rightarrow D, C \rightarrow A, B \rightarrow C\}$.

1. Identify all candidate keys for R.
2. Identify the best normal form that R satisfies.
3. Decompose R into a set of BCNF relations.
4. Decompose R into a set of 3NF relations.

For Video lecture on this topic please subscribe to my youtube channel.

The link for my youtube channel is

https://www.youtube.com/channel/UCRWGtE76JlTp1iim6aOTRuw?sub_confirmation=1