

Normalization

- It's a technique/process that reduces or removes redundancy, insertion anomaly, updation anomaly, deletion anomaly by decomposing a table into multiple tables.

Functional Dependency

- It's a constraint of a given table that determines whether a set of attributes/columns can be accurately determined from another set of attributes or not.

Example:

Let, a relation/table is given,

$R = (A, B, C, G, H, I)$ [Relation/Table]

And the valid Functional Dependencies of the attributes of R are given,

$F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$

Super Key: 1 or more attributes/columns that are used to uniquely identify each row/tuple/instance.

Candidate Key: SK of minimal size.

Primary Key: The selected CK.

Attribute Closures

- It's a method to find all the attributes that can be determined or are gettable from a given set of attributes.
- It helps us find the SK or CK of a given table using the given FDs of that table.

Example:

Let, $R = (A, B, C, G, H, I)$ [Relation/Table]

$F = \{A \rightarrow B, A \rightarrow C, CG \rightarrow H, CG \rightarrow I, B \rightarrow H\}$

i) Find the closures of AG.

Here, AG is a set that contains the attributes A and G i.e., $\{A, G\}$.

Solution with detailed steps:

$(AG)^+ = AG$	[AG determines AG, rule of thumb]
$= AG B$	[$A \rightarrow B$]
$= AG B C$	[$A \rightarrow C$]
$= AG B C H$	[$CG \rightarrow H$]
$= AG B C H I$	[$CG \rightarrow I$]

As **all the columns can be determined by AG**, so, **AG is an SK**.

And, **if this SK is of minimum size**, then **AG is also a CK**.

ii) Find the closures of AH.

Single line Solution:

$(AH)^+ = AH B C$

As **all the columns can NOT be determined by AG**, so, **AG is NOT an SK**.

Finding all the CKs of a given relation/table

- The closure method is used.

Example:

Given,

$R = (A, B, C, D, E)$
 $F = \{A \rightarrow C, CD \rightarrow B\}$

Find all the CKs of relation R.

Solution:

$\alpha = \{A, D, E\}$ // Not in the right side of any FD
- Must be present in the CK

$\beta = \{B\}$ // Present in the right side, but not in the Left side
- Can never be considered to be a part of CK

Step-1:

$(ADE)^+ = ADE C B$ [all columns found | SK of minimum size | CK]

Candidate Key Set, CK = { ADE } [ANS]

Practice 1:

Given,

$$R = (A, B, C, D, E)$$
$$F = \{A \rightarrow BC, CD \rightarrow E, B \rightarrow D, E \rightarrow A\}$$

Find all the CKs of the relation R.

$\alpha = \{ \}$ // Not present in the right side of any FD

$\beta = \{ \}$ // In right side, but not in left side

Step-1:

$$\{ \}^+ = \{ \}$$

Step-2:

$$(A)^+ = A B C D E \text{ [all columns found | SK of minimum size | CK]}$$

$$(B)^+ = B D \quad [X]$$

$$(C)^+ = C \quad [X]$$

$$(D)^+ = D \quad [X]$$

$$(E)^+ = E A B C D \text{ [all columns found | SK of minimum size | CK]}$$

Step-3:

$$(BA)^+ = [\text{SK but not of minimum size | B is extraneous}]$$

$$(BC)^+ = B C D E A \text{ [all columns found | SK of minimum size | CK]}$$

$$(BD)^+ = B D \quad [X]$$

$$(BE)^+ = [\text{SK but not of minimum size | B is extraneous}]$$

$$(CD)^+ = C D E A B \text{ [all columns found | SK of minimum size | CK]}$$

Step-4:

$$(BDC)^+ = [\text{SK but not of minimum size | B is extraneous}]$$

Candidate Key Set, CK = { A, E, BC, CD }

[ANS]

