Subject : DBMS
Topic:Closure properties

NISHA C.D,Assistant professor
Department of Computer Science

Closure of an Attribute Set-

Closure of an Attribute Set-

- The set of all those attributes which can be functionally determined from an attribute set is called as a closure of that attribute set.
- Closure of attribute set {X} is denoted as $\{X\}^+$.

Steps to Find Closure of an Attribute Set-

Following steps are followed to find the closure of an attribute set-

Step-01:

Add the attributes contained in the attribute set for which closure is being calculated to the result set.

Step-02:

Recursively add the attributes to the result set which can be functionally determined from the attributes already contained in the result set.

Example-

Consider a relation R (A , B , C , D , E , F , G) with the functional dependencies-

$$A \rightarrow BC$$

$$BC \rightarrow DE$$

$$D \rightarrow F$$

$$CF \rightarrow G$$

Now, let us find the closure of some attributes and attribute sets-

Closure of attribute A-

$A^+ = \{ A \}$

$= \{ A , B , C \}$ (Using $A \to BC$)

$= \{ A , B , C , D , E \}$ (Using $BC \to DE$)

$= \{ A , B , C , D , E , F \}$ (Using $D \to F$)

$= \{ A , B , C , D , E , F , G \}$ (Using $CF \to G$)

Thus,

$$\mathbf{A^+ = \{ A , B , C , D , E , F , G \}}$$

Closure of attribute D-

$D^+ = \{ D \}$

$= \{ D , F \}$ (Using $D \rightarrow F$)

We can not determine any other attribute using attributes D and F contained in the result set.

Thus,

$$\mathbf{D^+ = \{ D , F \}}$$

Closure of attribute set {B, C}-

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

$= \{ B , C , D , E \}$ (Using BC → DE)

$= \{ B , C , D , E , F \}$ (Using D → F)

$= \{ B , C , D , E , F , G \}$ (Using CF → G)

Thus,

$$\mathbf{\{ B , C \}^+ = \{ B , C , D , E , F , G \}}$$

Finding the Keys Using Closure-

Super Key-

- If the closure result of an attribute set contains all the attributes of the relation, then that attribute set is called as a super key of that relation.
- Thus, we can say-

"The closure of a super key is the entire relation schema."

Example-

In the above example,

- The closure of attribute A is the entire relation schema.
- Thus, attribute A is a super key for that relation.

Candidate Key-

- If there exists no subset of an attribute set whose closure contains all the attributes of the relation, then that attribute set is called as a candidate key of that relation.

Example-

In the above example,

- No subset of attribute A contains all the attributes of the relation.
- Thus, attribute A is also a candidate key for that relation.

Consider the given functional dependencies-

$$AB \rightarrow CD$$

$$AF \rightarrow D$$

$$DE \rightarrow F$$

$$C \rightarrow G$$

$$F \rightarrow E$$

$$G \rightarrow A$$

Solution-

Let us check each option one by one-

Option-(A):

$\{ CF \}^+ = \{ C , F \}$

$= \{ C , F , G \}$ (Using $C \rightarrow G$)

$= \{ C , E , F , G \}$ (Using $F \rightarrow E$)

$= \{ A , C , E , E , F \}$ (Using $G \rightarrow A$)

$= \{ A , C , D , E , F , G \}$ (Using $AF \rightarrow D$)

Candidate Key-

A set of minimal attribute(s) that can identify each tuple uniquely in the given relation is called as a candidate key.

OR

A minimal super key is called as a candidate key.

For any given relation,

- It is possible to have multiple candidate keys.
- There exists no general formula to find the total number of candidate keys of a given relation.

Example-

Consider the following Student schema-

Student (roll , name , sex , age , address , class , section)

Given below are the examples of candidate keys-

- (class , section , roll)
- (name , address)

These are candidate keys because each set consists of minimal attributes required to identify each student uniquely in the Student table.

Finding Candidate Keys-

We can determine the candidate keys of a given relation using the following steps-

Step-01:

- Determine all essential attributes of the given relation.
- Essential attributes are those attributes which are not present on RHS of any functional dependency.
- Essential attributes are always a part of every candidate key.
- This is because they can not be determined by other attributes.

<u>Example</u>

Let R(A, B, C, D, E, F) be a relation scheme with the following functional dependencies-

$$A \rightarrow B$$

$$C \rightarrow D$$

$$D \rightarrow E$$

Here, the attributes which are not present on RHS of any functional dependency are A, C and F.

So, essential attributes are- **A, C and F**.

Step-02:

- The remaining attributes of the relation are non-essential attributes.
- This is because they can be determined by using essential attributes.

Now, following two cases are possible-

Case-01:

If all essential attributes together can determine all remaining non-essential attributes, then-

- The combination of essential attributes is the candidate key.
- It is the only possible candidate key.

Case-02:

If all essential attributes together can not determine all remaining non-essential attributes, then-

- The set of essential attributes and some non-essential attributes will be the candidate key(s).
- In this case, multiple candidate keys are possible.
- To find the candidate keys, we check different combinations of essential and non-essential attributes.

Let R = (A, B, C, D, E, F) be a relation scheme with the following dependencies-

$$C \rightarrow F$$

$$E \rightarrow A$$

$$EC \rightarrow D$$

$$A \rightarrow B$$

Which of the following is a key for R?

1. CD
2. EC
3. AE
4. AC

Also, determine the total number of candidate keys and super keys.

Step-01:

- Determine all essential attributes of the given relation.
- Essential attributes of the relation are- C and E.
- So, attributes C and E will definitely be a part of every candidate key.

Step-02:

Now,

- We will check if the essential attributes together can determine all remaining non-essential attributes.
- To check, we find the closure of CE.

So, we have-

$\{ CE \}^+$

$= \{ C , E \}$

$= \{ C , E , F \}$ (Using $C \rightarrow F$)

$= \{ A , C , E , F \}$ (Using $E \rightarrow A$)

$= \{ A , C , D , E , F \}$ (Using $EC \rightarrow D$)

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

We conclude that CE can determine all the attributes of the given relation.
So, CE is the only possible candidate key of the relation.