LNCT GROUP OF COLLEGES

Name of Faculty: Shubha Mishra & Abhishek Dwivedi

Designation: Assistant Professor

Department: Information Technology

Subject: Database Management System

Unit: IV

Topic: Candidate Key

Finding Candidate Key (s) of a relation R using Functional Dependency Set F

A candidate key of a relation schema R is a subset X of the attributes of R with the following two properties:

Every attribute is functionally dependent on X, i.e., X+= all attributes of R (also denoted as X+= R).

No proper subset of X has the property (1), i.e., X is minimal with respect to the property (or No proper subset of X has contain any key of R

A sub-key of R is a subset of a candidate key.

A super-key is a set of attributes containing a candidate key.

For example, Let R(ABCDE) be a relation schema and consider the following functional dependencies F = {AB → E, AD → B, B → C, C → D}.

result := α;
repeat
    for each functional dependency $\beta \rightarrow \gamma$ **in** F **do**
      begin
        if $\beta \subseteq result$ **then** $result := result \cup \gamma$;
      end
until (*result* does not change)

Figure 8.8 An algorithm to compute α^+, the closure of α under F.

Here you will write the procedure to find out closure of attribute

Since $(AC)^+$ =ACDBE, $(AB)^+$ =ABECD, $(AD)^+$ =ADBCE

we know that AC is a candidate key, both A and C are sub-keys, AB is a candidate key, both A and B are sub-keys,

AD is a candidate key, both A and D are sub-keys,

So sub keys of relation R={A, B, C, D}.

ABC is a super-key because A, B, C are the set of attributes containing a candidate key of relation R.

Note that since nothing determines A, A is in every candidate key.

Please note that ABC, ABD, ACD,& ABCD are not a candidate key bacause it proper subset contain a key i.e AB, AC & AD

Computing All Candidate Keys of a Relation R.

Necessary attributes: An attribute A is said to be a necessary attribute if

(a) A occurs only in the L.H.S. (left hand side) of the fd's in F;

or

(b) A is an attribute in relation R, but A does not occur in either L.H.S. or R.H.S. of any fd in F.

In other words, necessary attributes NEVER occur in the R.H.S. of any fd in F.

Useless attributes: An attribute A is a useless attribute if A occurs ONLY in the R.H.S. of fd's in F.

Middle-ground attributes: An attribute A in relation R is a middle-ground attribute if A is neither necessary nor useless.

Example. Consider the relation R(ABCDEG) with set of fd'sF = {AB → C, C → D, AD → E}

Necessary attributes	Useless attributes	Middle-ground attributes
A, B, G	E	C, D

The algorithm for computing all candidate keys of R.

Input:A relation R={ A1, A2, ..., An}, and F, a set of functional dependencies.

Output: K={ K1,..., Kt}, the set of all candidate keys of R.

Step1. Set F′ to a minimal cover of F(This is needed because otherwise we may not detect all useless attributes).

Step2. Partition all attributes in R into necessary, useless and middle-ground attribute sets according to F′. Let X={ C1 ,...,Cl} be the necessary attribute set, Y = {B1 , ..., Bk} be the useless attribute set, and M = {A1 , ..., An} − (X ground attribute set. If X={}, then go to step4.

Step3. Compute $\mathbf{X}^+$. If $\mathbf{X}^+$ =R, then set K= {X}, terminate.

Step4. Let L = ⟨Z1 , Z2 ,..., Zm⟩ be the list of all non-empty subsets of M (the middle- ground attributes) such that L is arranged in ascending order of the size of Zi . Add all attributes in X (necessary attributes) to each Zi in L.

Set K = {}.

i ← 0.

WHILE (L ≠ empty)

do BEGIN

i ← i+1.

Remove the first element Z from L.

Compute Z^+. If Z^+ = R,

Then

Begin

set K ← K ∪ {Z};

for any Z j ∈ L, if Z ⊂ Z j

then L ← L − {Z j}.

end

END

Example. (Computing all candidate keys of R.)

Let R = R(ABCDEG) and F = {AB → CD, A → B, B → C, C → E, BD → A}. The process to compute all candidate keys ofRis as follows:

(1) The minimal cover ofFis {A → B, A → D, B → C, C → E, BD → A}.

(2) Since attribute G never appears in any fd's in the set of functional dependencies, G must be included in a candidate key of R. The attribute E appears only in the right hand side of fd's and hence E is not in any key of R. No attribute of R appears only in the left hand side of the set of fd's. Therefore X=G at the end of step 2.

(3) Compute G^+ = G, so Gis not a candidate key.

(4) The following table shows the L, K, Z and Z^+ at the very beginning of each.

i	Z	Z^+	L	K
0	-	-	⟨AG, BG, CG, DG, ABG, ACG, ADG, BCG, BDG, CDG, ABCG, ABDG, ACDG, BCDG, ABCDG⟩	{}
1	AG	ABCDEG = R	⟨BG, CG, DG, BCG, BDG, CDG,BCDG⟩	{AG}
2	BG	BCEG ≠ R	⟨CG, DG, BCG, BDG, CDG, BCDG⟩	{AG}
3	CG	CEG ≠ R	⟨DG, BCG, BDG, CDG, BCDG⟩	{AG}
4	DG	DG ≠ R	⟨BCG, BDG, CDG, BCDG⟩	{AG}
5	BCG	BCEG ≠ R	⟨BDG, CDG, BCDG⟩	{AG}
6	BDG	ABCDEG = R	⟨CDG⟩	{AG, BDG}
7	CDG	CEDG ≠ R	⟨ ⟩	{AG, BDG}