Feuille de note normalisation

Réflexivité

Soit X et Y deux ensembles d'attributs, alors

$$\text{Si } Y \subseteq X \text{ alors } X \to Y$$

Exemples :

$$A \to A \qquad A, B \to A, B$$
$$A, B \to A \qquad A, B \to B$$

Augmentation

Soit X, Y, Z trois ensembles d'attributs, alors

$$X \to Y \Rightarrow XZ \to YZ$$

Exemples :

$$A \to B \Rightarrow A, C \to B, C$$
$$B, C \to A \Rightarrow B, C, E \to A, E$$

Transitivité

Soit X, Y, Z trois ensembles d'attributs, alors

$$X \to Y \text{ et } Y \to Z \Rightarrow X \to Z$$

Exemples :

$$A \to B \text{ et } B \to C \Rightarrow A \to C$$
$$B, C \to C, E \text{ et } C, E \to D \Rightarrow B, C \to D$$

Décomposition

Soit X, Y, Z trois ensembles d'attributs, alors

$$X \to YZ \Rightarrow X \to Y \text{ et } X \to Z$$

Exemple :

$$A, B \to C, D \Rightarrow A, B \to C \text{ et } A, B \to D$$

Composition

Soit X, Y, A, B quatre ensembles d'attributs, alors

$$X \to Y \text{ et } A \to B \Rightarrow XA \to YB$$

Exemple :

$$A \to B \text{ et } C \to D \Rightarrow AC \to BD$$

Union

Soit X, Y, Z trois ensembles d'attributs, alors

$$X \to Y \text{ et } X \to Z \Rightarrow X \to YZ$$

Pseudo-transitivité

Soit X, Y, Z, W quatre ensembles d'attributs, alors

$$X \to Y \text{ et } YZ \to W \Rightarrow XZ \to W$$

Exemple :

$$A \to B \text{ et } BC \to D \Rightarrow AC \to D$$

Détermination de soi (Self-determination)

$$I \to I \quad \forall I$$

Extensivité

Soit X, Y, Z trois ensembles d'attributs, alors

$$X \to Y \Rightarrow X \to XY$$

Fermeture de F

Soit $\mathscr{F}$ un ensemble de dépendances fonctionnelles, alors

$$\mathscr{F}^+ = \{f \text{ tel que } \mathscr{F} \models f\}$$

Autrement dit, $\mathscr{F}^+$ est l'ensemble de toutes les dépendances fonctionnelles qui sont logiquement impliquées par $\mathscr{F}$ en utilisant les axiomes d'Armstrong.

Fermeture de X

Soit X un ensemble d'attributs et $\mathscr{F}$ un ensemble de dépendances fonctionnelles, alors

$$X^+ = \{A \text{ tel que } X \to A \text{ est dérivable de } \mathscr{F}\}$$

Clés et superclés

Soit X un ensemble d'attributs et $\mathscr{F}$ un ensemble de dépendances fonctionnelles

Superclé :
X est une superclé ssi

$$X^+ = R$$

Clé :
X est une clé ssi

$$X^+ = R \text{ et } \nexists Y \subset X \text{ tel que } Y^+ = R$$

DF triviale

Une dépendance fonctionnelle est triviale si elle est de la formes suivante :

$$X \to Y \text{ où } Y \subseteq X$$

Dépendance fonctionnelle élémentaire

Soit X un ensemble d'attributs et A un attribut, alors $X \to A$ est élémentaire ssi

1. $A \notin X$,
2. $\nexists X' \subset X$ tel que $X' \to Y$, c-à-d, que on ne peut pas trouver un sous-ensemble de X qui détermine Y.

Note : La partie de droite ne peut pas être un groupe d'attributs.

Couverure minimale

Soit $\mathscr{F}$ un ensemble de dépendances fonctionnelles, alors $\mathscr{F}$ est minimale ssi on ne peut pas enlever une dépendance fonctionnelle sans perdre de l'information.

Calculer la couverture minimale

1. Décomposer chaque DF de type $X \to Y, Z$ en $X \to Y$ et $X \to Z$.
2. Enlever les DF non-élémentaires.
3. Enlever les DF redondantes. (Les DF qui peuvent être déduites des autres DF)

Forme normale de Boyce-Codd (BCNF)

Définition 1 :
Un schéma relationnel $R(U)$ avec ensemble F de DF est en forme BCNF ssi pour chaque dépendance élémentaire $X \to A$ dans $\mathscr{F}^+$, on a que

X est une clé (candidate)

Définition 2 :
Un schéma relationnel $R(U)$ avec ensemble $\mathscr{F}$ de DF est en forme BCNF ssi pour chaque DF non triviale $X \to A$ de $\mathscr{F}$, on a que :

X est une superclé de R

1FN,2FN,3FN

1FN :
Chaque attribut est atomique, c-à-d, chaque attribut ne contient qu'une seule valeur.

2FN :
Pour être en 2FN, il faut que :

1. Il faut être en 1FN
2. Chaque attribut non-clé ne dépend d'une partie de la clé.

3FN :
Pour être en 3FN, il faut que :

1. Il faut être en 2FN
2. Chaque attribut non-clé ne dépend pas d'un attribut non-clé.

Algo-BCNF

Entrée :
Un schéma $R(U, \mathscr{F})$ où U est l'ensemble des attributs et $\mathscr{F}$ est l'ensemble des dépendances fonctionnelles.
Sortie :
Une décomposition $D = \{R_1(U_1, \mathscr{F}_1), \ldots, R_n(U_n, \mathscr{F}_n)\}$ sans perte d'information avec $R_i(U_i, \mathscr{F}_i)$ en BCNF pour tout i (mais perte possible de dépendances).

1. Initialiser $D \leftarrow \{R(U, \mathscr{F})\}$
2. Tant qu'il existe $R'(U', \mathscr{F}') \in D$ qui n'est pas en BCNF
   (a) Trouver une DF non triviale $X \to Y \in \mathscr{F}'$ tq $X \neq X^+$ et $X^+ \neq U'$
   (b) Poser $U_1' = X^+$ et $U_2' = X \cup (U' - X^+)$, $\mathscr{F}_1'$ les dépendances sur U_1' et $\mathscr{F}_2'$ celles sur U_2'
   (c) Poser $U_1' = X^+$, $U_2' = X \cup (U' - X^+)$, $\mathscr{F}_1'$ les dépendances sur U_1' et $\mathscr{F}_2'$ celles sur U_2'
   (d) Remplacer $R'(U')$ par ses projections sur U_1' et U_2' : $R_1'(U_1', \mathscr{F}_1')$ et $R_2'(U_2', \mathscr{F}_2')$ ($R' = R_1' \bowtie R_2'$ par Heath avec X déterminant, $X^+ - X$ déterminé et $U' - X^+$ résidu)

Le calcul de $\mathscr{F}_1'$ et $\mathscr{F}_2'$ à partir de $\mathscr{F}'$ doit se faire avec l'algorithme de projection des dépendances

Projection des dépendances fonctionnelles

Entrée : Un schéma $R(U, \mathscr{F})$ et $U_1 \subset U$
Sortie : Les dépendances $\mathscr{F}_1$ induites sur $R_1(U_1) = \pi_{U_1}(R)$

1. $\mathscr{F}_1 \leftarrow \emptyset$ (initialisation)
2. Pour chaque $X \subseteq U_1$
   (a) Calculer X^+ (par rapport à $\mathscr{F}$)
   (b) Ajouter à $\mathscr{F}_1$ toutes les $X \to \{A\}$ t.q. $A \in (X^+ - X) \cap U_1$
   (optionnel : et t.q. on n'a pas déjà $Y \to \{A\} \in \mathscr{F}_1$ avec $Y \subset X$)
3. (optionnel) enlever de $\mathscr{F}_1$ les dépendances redondantes
4. retourner $\mathscr{F}_1$

Algo de Bernstein

Entrée : Un schéma R (arbitraire) et un ensemble $\mathscr{F}$ de DF qui soit en 1NF.
Sortie : Projections donnant des schémas R_i en 3-NF pour tout i sans perte d'information et préservant les dépendances.

1. Initialiser P (ensemble de projections) à l'ensemble vide (et $i = 1$)
2. Fixer $\mathscr{G}$ une **couverture minimale** de $\mathscr{F}$
3. Pour chaque X distinct d'une partie gauche d'une DF de $\mathscr{G}$
   (a) Faire la réunion Y de tous les $\{A\}$ tels que $X \to \{A\} \in \mathscr{G}$
   (b) Ajouter à P la projection de R sur X, Y (donne $R_i(\underline{X}, Y) = \pi_{XY}(R)$, avec clé primaire X)
   (c) $i = i + 1$
4. Si aucune des projections dans P ne contient une clé candidate de R, ajouter à P la projection de R sur une clé candidate. ($R_i(K) = \pi_K(R)$ pour K clé candidate).

R est alors la jointure naturelle des projections obtenues. La décomposition est sans perte.