Info710 : compléments de bases de données
Pierre Hyvernat

complément de cours

On rappelle quelques définitions :

Def. *Si X et Y sont des ensembles d'attributs pour une table T, on dit que Y dépend fonctionnellement de X, noté $X \to Y$ si la propriété suivante est vérifiée pour toutes les instances de la table*

$$u, v \in T \text{ et } \prod_X(u) = \prod_X(v) \quad \Rightarrow \quad \prod_Y(u) = \prod_Y(v)$$

Axiomes de Armstrong.
1) si $Y \subseteq X$ alors $X \to Y$
2) si $X \to Y$ alors $X, Z \to Y, Z$ (où Z est un ensemble d'attributs)
3) si $X \to Y$ et $Y \to Z$ alors $X \to Z$

Def. *Si F est un ensemble (fini) de dépendances fonctionnelles, alors on note F^+ l'ensemble des dépendances qui peuvent être déduites de F par utilisation des axiomes de Armstrong.*

Par exemple, si $A_1, A_2 \to B$ et $B \to C_1, C_2$ sont deux éléments de F, alors $A_1, A_2 \to C_1, C_2$ sera nécessairement un élément de F^+ (application de l'axiome 3).

Def. *Soit F un ensemble de dépendances ; si X est un ensemble (fini) d'attributs, alors on note X^+ pour l'ensemble des attributs A vérifiant $X{\to}A \in F^+$.*

On a trivialement que $X{\to}Y \in F^+$ ssi $Y \subseteq X^+$.

Étant donné un ensemble F de dépendances (fixé) et X un ensemble d'attributs, on définit récursivement les ensembles X_n de la manière suivante :
- $X_0 = X$
- $X_{n+1} = X_n \cup \{A \mid Y{\to}A, Z \in F \text{ pour un } Y \subseteq X_n\}$.

Dans la deuxième ligne, A est un attribut et Z un ensemble d'attributs fini.

Lemme. *Pour un ensemble de dépendances fixé, on a :*

$$X \subseteq X' \quad \Rightarrow \quad \forall n \; X_n \subseteq X'_n$$

Il faut faire la preuve par récurrence sur n. (Faites la en détails...)

Proposition. *Pour tout ensemble de dépendances et pour tout ensemble X d'attributs, on a*

$$X^+ = \bigcup_{n \geq 0} X_n \ .$$

Preuve : on commence par montrer que $X_n \subseteq X^+$ pour tout n. Ceci permettra de conclure que $\bigcup_{n\geq 0} X_n \subseteq X^+$. On la preuve par récurrence sur n :
- si $n = 0$, alors $X_0 = X$ et on a trivialement que $X \subseteq X^+$;
- pour $n+1$, supposons (HR) que $X_n \subseteq X^+$, montrons que $X_{n+1} \subseteq X^+$. Soit $A \in X_{n+1}$; par la définition de X_{n+1}, il y a deux cas possibles :
  . soit $A \in X_n$, et on a directement que $A \in X_{n+1}$ par l'hypothèse de récurrence,
  . soit F contient une dépendance de la forme $Y \to A, Z$ pour un $Y \subseteq X_n$. Pour montrer que $A \in X^+$, il faut montrer que la dépendance $X \to A$ peut être obtenue à partir de F et des axiomes de Armstrong. Par HR, on sait que $Y \subseteq X^+$, c.à.d. $X{\to}Y \in F^+$. Ceci signifie que l'on peut obtenir $X \to Y$ à partir de F et des axiomes de Armstrong. Comme $Y \to A, Z$, on peut obtenir (en utilisant l'axiome 3) la dépendance $X \to A, Z$. Comme $A \subseteq A, Z$, on peut (axiomes 1 et 3) obtenir la dépendance $X \to A$ à partir de F et des axiomes. Ceci termine la preuve que $A \in X^+$.

  On a donc bien que $A \in X_{n+1} \;\Rightarrow\; A \in X^+$, ce qui termine la preuve que $X_{n+1} \subseteq X^+$.

Nous allons maintenant montrer que $X^+ \subseteq \bigcup_{n\geq 0} X_n$. Ceci équivaut à démontrer que, pour tout attribut A tels que $X \to A$ soit dans F^+, on peut trouver un entier n tels que $A \in X_n$. Nous allons en fait démontrer le résultat un peu plus fort suivant :

$$X{\to}Y \ \in \ F^+ \quad \Rightarrow \quad (\exists n) \ Y \subseteq X_n$$

Si $X{\to}Y \ \in \ F^+$, cela veut dire que $X \to Y$ peut-être obtenue à partir de F et des axiomes de Armstrong. Nous allons faire une preuve par récurrence sur le nombre d'axiomes qu'il faut utiliser pour obtenir $X \to Y$ à partir de F.

Appelons l le nombre d'axiomes utilisés pour obtenir $X \to Y$ à partir de F :

- si $l = 0$, c'est forcément que $X{\to}Y \ \in \ F$ (on n'a pas besoin d'axiome pour obtenir $X \to Y$ à partir de F). On sait alors que $Y \subseteq X_1$ par définition de X_1. (Vérifiez le !)
- si $l > 0$, c'est qu'il faut utiliser au moins un axiome pour obtenir $X \to Y$ à partir de F. On peut supposer, par hypothèse de récurrence, que à chaque fois que l'on peut déduire une dépendance avec strictement moins de l axiomes, alors cette dépendance est dans un X_k. On regarde le dernier axiome utilisé : trois cas sont possibles.
  . si le dernier axiome utilisé était l'axiome 1, c'est que l'on peut obtenir $X \to Y$ en utilisant le fait que $Y \subseteq X$: on a alors nécessairement que $Y \subseteq X_0 = X$. (Remarque : dans ce cas, on avait $l = 1$...)
  . si le dernier axiome utilisé était l'axiome 2, alors $X \to Y$ est nécessairement de la forme $X', Z \to Y', Z$, et on peut obtenir $X' \to Y'$ à partir de F en utilisant un axiome de moins. Par hypothèse de récurrence, on sait que $Y' \subseteq X'_n$ pour un n bien choisi. Ceci implique que $Y' \cup Z \subseteq X'_n \cup Z$. Nous voulons montrer que $Y' \cup Z \subseteq (X' \cup Z)_n$. Il suffit d'utiliser le lemme précédent : comme $X' \subseteq X' \cup Z$, on a $X'_n \subseteq (X' \cup Z)_n$. Comme on sait également que $Z \subseteq X' \cup Z \subseteq (X' \cup Z)_n$, on peut conclure que $X'_n \cup Z \subseteq (X' \cup Z)_n$. Par transitivité, on obtient finalement $Y' \cup Z \subseteq (X' \cup Z)_n$.
  . si le dernier axiome utilisé était l'axiome 3, alors on peut obtenir $X \to Y$ à partir de $X \to Z$ et $Z \to Y$ pour deux dépendances dans F^+. Chacune de ces dépendances peut s'obtenir avec strictement moins d'axiomes, et donc, par hypothèse de récurrence, on sait que $Z \subseteq X_n$ et $Y \subseteq Z_m$. En utilisant le lemme précedent, on sait que $Z_i \subseteq (X_n)_i$. Il est facile de voir que $(X_n)_i = X_{n+i}$, et donc, pour $i = m$, on obtient $Z_m \subseteq X_{n+m}$. Par transitivité, on peut conclure que $Y \subseteq X_{n+m}$.

  Ceci termine la preuve de la proposition.

Remarque : le théorème de complétude nous apprend que F^+ contient en fait *toutes* les dépendances qui sont conséquences de F. Le théorème de complétude dit que si $X \to Y$ est conséquence de F (*i.e.* pour toute instance de la table qui satisfait les dépendances de F, l'instance satisfait également la dépendance $X \to Y$), alors on peut obtenir $X \to Y$ à partir de F et des axiomes de Armstrong...