TD 9 : Normalisation et dépendances

Dépendances fonctionnelles

2024-11-22

- **L3 MIASHS/Ingémath**
- **Université Paris Cité**
- Année 2024-2025
- Course Homepage
- Moodle

Définitions

Une *dépendance fonctionnelle* est une expression de la forme

$$A_1, A_2, \ldots, A_k \rightarrow A_{k+1}, \ldots, A_n$$

où $A_1, A_2, \ldots, A_k, A_{k+1}, \ldots, A_n$ sont des attributs (colonnes) d'une base de données.

Elle signifie que deux tuples ayant la même valeur sur $A_1, \ldots, A_k$ doivent avoir la même valeur sur chaque colonnes $A_{k+1}, \ldots, A_n$ (en français : $A_1, \ldots, A_k$ *déterminent* $A_{k+1}, \ldots, A_n$. On dit que les attributs $A_{k+1}, \ldots, A_n$ *dépendent fonctionnellement* de $A_1, A_2, \ldots, A_k$.

La notion de dépendance est transitive : si $A \rightarrow B$ et $B \rightarrow C$ alors $A \rightarrow C$.

Un ensemble de dépendances fonctionnelles $\mathcal{F}$ est *irréductible* si aucune dépendance ne peut être déduite des autres en utilisant les règles suivantes :

- trivialité : si $Y \subseteq X$ alors $X \rightarrow Y$
- augmentation : si $X \rightarrow Y$ alors $X, Z \rightarrow Y, Z$ pour toute suite d'attributs Z.
- transitivité : si $X \rightarrow Y$ et $Y \rightarrow Z$ alors $X \rightarrow Z$
- union : si $X \rightarrow Y$ et $X \rightarrow Z$ alors $X \rightarrow Y, Z$
- décomposition/séparation si $X \rightarrow Y$ et $Z \subseteq Y$ alors $X \rightarrow Z$

La *clôture transitive* des attributs $A_1, \ldots, A_k$ pour un ensemble de dépendances fonctionnelles $\mathcal{F}$ est l'ensemble des attributs $B_1, \ldots, B_\ell$ qui dépendent fonctionnellement de $A_1, \ldots, A_k$.

On la note

$$[A_1, \ldots, A_k]^+_{\mathcal{F}}$$

en oubliant $\mathcal{F}$ si le contexte est clair.

Un ensemble d'attributs $A_1, \ldots, A_k$ est une *super-clé pour une relation* $R(B_1, \ldots, B_\ell)$ si ce sont des attributs de R et si sa clôture transitive contient $B_1, \ldots, B_\ell$. C'est une clé si elle est minimale, c'est-à-dire, aucun sous-ensemble strict de cette super-clé n'est une clé.

Un schéma est en :

- FN_3 si pour toute dépendance fonctionnelle non triviale, le membre de gauche contient une clef ou tout attribut du membre de droit appartient à une clef.
- FNBC si pour toute dépendance fonctionnelle non triviale, le membre de gauche contient une clef.

Un schéma et un ensemble de dépendances fonctionnelles peut se *décomposer* en une collection de schémas, dans le sens où chaque relation R peut se décomposer en $R_1, \ldots, R_k$ tels que $R_i = \pi_i(R)$ pour une certaine projection π_i.

On dit cette décomposition *sans perte d'information* si toute relation R du schéma d'origine peut être retrouvée à partir des relations $R_1, \dots, R_k$: $R = \pi_1(R) \bowtie \dots \bowtie \pi_k(R)$.

On dit que cette décomposition *respecte les dépendances fonctionnelles* si celles-ci sont toujours satisfaites par la nouvelle décomposition.

Exercice

Soit une relation concernant des personnes en France avec les attributs suivants~ :

`Nom, Numéro de sécurité sociale, Commune, Département, Code postal, Numéro de téléphone`

Quelles sont les dépendances fonctionnelles censées être satisfaites ?

Exercice

Soit un schéma d'attributs $A_1, A_2, \dots A_n$ et un ensemble de dépendances fonctionnelles. Calculer le nombre de super-clefs (en fonction de n) dans les cas suivants~ :

- La seule clef est $\{A_1\}$.
- Les seules clefs sont $\{A_1\}$ et $\{A_2\}$.
- Les seules clefs sont $\{A_1, A_2\}$ et $\{A_3, A_4\}$.
- Les seules clefs sont $\{A_1, A_2\}$ et $\{A_1, A_3\}$.

Exercice

Soit le schéma $\mathcal{A} = \{A, B, C, D\}$ et l'ensemble de dépendances fonctionnelles

$$\Sigma = \{A \longrightarrow B, B \longrightarrow C\}$$

- Quelle est la fermeture $\{A\}^+$ de $\{A\}$?
- Quelles sont les super-clés ? Les clés ?

Exercice

Soit le schéma $\mathcal{A} = \{A, B, C, D, E, F\}$ et l'ensemble de dépendances fonctionnelles

$$\Sigma = \Big\{\{A, B\} \to C, \{B, C\} \to \{A, D\}, D \to E, \{C, F\} \to B\Big\}$$

- Calculer la fermeture $\{A, B\}^+$ de $\{A, B\}$.
- Est-ce que Σ implique la dépendance fonctionnelle $\{A, B\} \to D$~ ?
- Est-ce que Σ implique la dépendance fonctionnelle $D \to A$~ ?

Exercice

Montrer que les assertions suivantes sont fausses :

- $A \to B$ implique $B \to A$.
- Si $\{A, B\} \to C$ et $A \to C$ alors $B \to C$.
- Si $\{A, B\} \to C$ alors $A \to C$ ou $B \to C$.

Exercice : déductions de DF

Soit le schéma $\mathcal{A} = \{A, B, C, D, E, F, G, H\}$ et soit

$$\Sigma = \{AB \longrightarrow C;\ B \longrightarrow D;\ CD \longrightarrow E;\ CE \longrightarrow GH;\ G \longrightarrow A\}$$

Est-ce que les dépendances

- $A, B \longrightarrow E$
- $B, G \longrightarrow C$
- $A, B \longrightarrow G$

sont déductibles de Σ ?

Σ est elle irréductible/minimale ?

Rappel

Pour que Σ soit minimale, il y a 3 conditions à remplir :
- Σ est sous forme canonique, un seul attribut à droite.
- Aucune DF redondante , i.e. aucune DF ne peut être déduite des autres.
- Aucune DF redondante à gauch, .i.e. les déterminants sont minimaux

Exercice : équivalence d'ensembles de DFs

- Soit
$$\Sigma_1 = \{A \longrightarrow B; C, E \longrightarrow H; C \longrightarrow E; A \longrightarrow C, H\}$$
et
$$\Sigma_2 = \{A \longrightarrow B, C; C \longrightarrow E, H\}$$
Les deux ensembles de dépendances fonctionnelles Σ_1 et Σ_2 sont-ils *équivalents* ?

Exercice : Décomposition et perte d'information

- On considère le schéma de relation $\mathcal{A} = A, B, C$ et la dépendance fonctionnelle suivante :

$$\Sigma = \{A, B \longrightarrow C\}.$$

Déterminer si la décomposition suivante est sans perte d'information

$$\mathcal{A}_1 = \{A, B\}, \quad \mathcal{A}_2 = \{B, C\}$$

en étudiant le cas de la table suivante :

A	B	C
1	2	3
4	2	5

Exercice : poursuite

- On considère le schéma de relation $\mathcal{A} = \{A, B, C, D, E\}$ et les dépendances fonctionnelles suivantes :

$$\Sigma = \{A \longrightarrow C; B \longrightarrow C; C \longrightarrow D; D, E \longrightarrow C; C, E \longrightarrow A\}.$$

Appliquer l'algorithme de poursuite pour déterminer si la décomposition suivante est sans perte d'information :

$$\mathcal{A}_1 = \{A, D\}, \mathcal{A}_2 = \{A, B\}, \mathcal{A}_3 = \{B, E\}, \mathcal{A}_4 = \{C, D, E\}, \mathcal{A}_5 = \{A, E\}$$

Même question pour la décomposition :

$$\mathcal{A}_1 = \{A, D\}, \mathcal{A}_2 = \{A, B\}, \mathcal{A}_3 = \{B, E\}, \mathcal{A}_4 = \{C, D\}, \mathcal{A}_5 = \{D, E\}, \mathcal{A}_6 = \{A, E\}$$

Exercice

Soit $\mathcal{A} = \{A, B, C, D, E\}$ un schéma et soit la décomposition $\{\mathcal{A}_1, \mathcal{A}_2, \mathcal{A}_3\}$ où

$$\mathcal{A}_1 = A, B, C \quad \mathcal{A}_2 = B, C, D \quad \mathcal{A}_3 = A, C, E$$

Pour chaque ensemble Σ de dépendances fonctionnelles ci-dessous, appliquer l'algorithme de poursuite pour déterminer si la décomposition est sans perte d'information. Dans le cas où il y a perte d'information, donner une relation R de schéma $\mathcal{A}$ satisfaisant Σ et telle que

$$\pi_{\mathcal{A}_1}(R) \bowtie \pi_{\mathcal{A}_2}(R) \bowtie \pi_{\mathcal{A}_3}(R) \not\subseteq R$$

- $\Sigma = \{B \to E, CE \to A\}$
- $\Sigma = \{AC \to E, BC \to D\}$
- $\Sigma = \{A \to D, D \to E, B \to D\}$
- $\Sigma = \{A \to D, CD \to E, E \to D\}$

Exercice : Normalisation

On considère le schéma de relation `R(C,T,H,S,E,N)` :

`R(Cours, Enseignant, Horaire, Salle, Étudiant, Note)`

et les dépendances fonctionnelles suivantes :

$$\mathcal{F} = \{\texttt{C} \to \texttt{T}; \quad \texttt{H,S} \to \texttt{C}; \quad \texttt{H,T} \to \texttt{S}; \quad \texttt{C,E} \to \texttt{N}; \quad \texttt{H,E} \to S\}.$$

- Calculer une clé.
- Mettre en Boyce-Codd Normal Form (BCNF), donner plusieurs résultats possibles.