

FONDEMENTS DES BASES DE DONNÉES

Complétude du système d'Armstrong et relations d'Armstrong

Équipe pédagogique BD



Complétude du système d'Armstrong (1/4)

Preuve formelle

Une preuve (formelle) de f à partir de Σ dans le système d'Armstrong notée $\Sigma \vdash f$ est une *séquence* $\langle f_0, \dots, f_n \rangle$ de DFs telles que $f_n = f$ et $\forall i \in [0..n]$:

- ▶ soit $f_i \in \Sigma$;
- ▶ soit f_i est la conséquence d'une règle du système d'Armstrong (réflexivité, transitivité, augmentation) dont toutes les prémisses $f_0 \dots f_p$ apparaissent avant f_i dans la séquence.

Complétude : prouver que $\Sigma \models X \rightarrow Y \Rightarrow \Sigma \vdash X \rightarrow Y$

Il faut bien faire la différence entre

- ▶ la fermeture **sémantique** : $X^+ = \{A \mid \Sigma \models X \rightarrow A\}$
- ▶ la fermeture **syntactique** : $X^* = \{A \mid \Sigma \vdash X \rightarrow A\}$

Complétude du système d'Armstrong (2/4)

Lemme : $\Sigma \vdash X \rightarrow Y \Leftrightarrow Y \subseteq X^*$

- (\Rightarrow) On suppose sans perte de généralité que $Y = A_1 \dots A_n$. Pour chaque A_i , par réflexivité on a $Y \rightarrow A_i$ et on peut prolonger la preuve de $\Sigma \vdash X \rightarrow Y$ par transitivité pour obtenir une preuve $\Sigma \vdash X \rightarrow A_i$. Ainsi $\forall A_i \in Y$ on a bien $A_i \in X^*$ et donc $Y \subseteq X^*$.
- (\Leftarrow) Supposons que $Y \subseteq X^* = \{A \mid \Sigma \vdash X \rightarrow A\}$, par définition de X^* on a une preuve de $\Sigma \vdash X \rightarrow A_i$ pour chaque A_i , dès lors, on peut concaténer toutes ces preuves et conclure par l'application de la règle de composition qu'on a prouvé valide.

Complétude du système d'Armstrong (3/4)

On va avoir le raisonnement logique suivant :

$$\begin{aligned} & \Sigma \models X \rightarrow Y \Rightarrow \Sigma \vdash X \rightarrow Y \\ \equiv & \Sigma \not\models X \rightarrow Y \Rightarrow \Sigma \not\vdash X \rightarrow Y \\ \equiv & \Sigma \not\models X \rightarrow Y \Rightarrow \exists r. (r \models \Sigma \wedge r \not\models X \rightarrow Y) \end{aligned}$$

L'astuce consiste à exhiber l'instance r ,
avec $X^* = X_1 \dots X_n$ et $Z_1 \dots Z_p = R \setminus X^*$

r	X_1	\dots	X_n	Z_1	\dots	Z_p
s	x_1	\dots	x_n	z_1	\dots	z_p
t	x_1	\dots	x_n	y_1	\dots	y_p

Cette instance vérifie bien $r \models \Sigma$ et $r \not\models X \rightarrow Y$

Complétude du système d'Armstrong (4/4)

- ▶ Prouvons $r \models \Sigma$ par l'absurde. Supposons $V \rightarrow W \in \Sigma$ et $r \not\models V \rightarrow W$ pour aboutir à une contradiction. $r \not\models V \rightarrow W$, par définition de \models on a $s[V] = t[V] \wedge s[W] \neq t[W]$, par construction de r cela implique $V \subseteq X^*$ et $W \not\subseteq X^*$. D'après le lemme $V \subseteq X^*$ implique $\Sigma \vdash X \rightarrow V$ et par transitivité avec $V \rightarrow W \in \Sigma$ on obtient $\Sigma \vdash X \rightarrow W$. Or d'après le lemme (dans l'autre sens) on déduit que $W \subseteq X^*$, une contradiction.
- ▶ Prouvons que $r \not\models X \rightarrow Y$. Pour cela, supposons, $r \models X \rightarrow Y$ par construction de r on a $s[Y] = t[Y]$ et donc $Y \subseteq X^*$. D'après le lemme on a $\Sigma \vdash X \rightarrow Y$, une contradiction.

Le théorème est prouvé : $\Sigma \models X \rightarrow Y \Leftrightarrow \Sigma \vdash X \rightarrow Y$ et $X^* = X^+$

Les relations d'Armstrong

Le corps de la preuve de complétude

- ▶ La construction de r est un exemple d'une relation qui vérifie toutes les dépendances de Σ de la forme $X \rightarrow Y$ mais ne satisfait pas les autres avec X en partie gauche.
- ▶ On va **répéter la construction** de s, t dans la preuve pour **chaque** dépendance de Σ et avoir une instance qui concerne toutes les dépendances possibles.

Objectif et intérêt des relations d'Armstrong

- ▶ Représenter sur un exemple un ensemble de contraintes et uniquement celui-ci, i.e. toutes les autres contraintes sont fausses.
- ▶ Celle permet une représentation *par l'exemple* d'ensemble de dépendances : on manipule des valeurs, visualisation plus simple des éventuels conflits, d'incohérences, de mauvaise conception. . .

Construction de la relation

- ▶ Calculer les fermés de F : $Cl(F) = \{X^+ \mid X \subseteq R\}$
- ▶ Construire la relation d'Amstrong r correspondante :
 - ▶ **Etape 0** : pour le fermé particulier R , construire le tuple $t_0 = \langle 0, \dots, 0 \rangle$
 - ▶ **Etape i** : pour chaque $X \in Cl(F)$, ajouter un tuple t_i à r :
 - ▶ tel que $t_i[A] = 0$ pour tout $A \in X^+$
 - ▶ tel que $t_i[B] = i$ pour tout $B \in (R \setminus X^+)$.

Instance canonique

Avec les mêmes arguments que dans la preuve de complétude :

- ▶ r vérifie chaque dépendance : $f \in \Sigma$ implique $r \models f$
- ▶ r ne vérifie aucune dépendance qui ne soit pas déductible : $\Sigma \not\models f$ implique $r \not\models f$

Exemple

Considérons :

- ▶ la relation $R = \{A, B, C\}$
- ▶ l'ensemble de DFs $\Sigma = \{A \rightarrow BC, B \rightarrow C\}$
- ▶ $CI(F) = \{ABC, BC, C\}$

Relation d'Armstrong pour Σ

r	A	B	C
ABC	0	0	0
BC	1	0	0
C	2	2	0

Exercice

Soit Σ l'ensemble de dépendances

$$AB \rightarrow C \quad C \rightarrow A \quad BC \rightarrow D$$

$$ACD \rightarrow B \quad D \rightarrow E \quad ABE \rightarrow C$$

$$C \rightarrow BD \quad CE \rightarrow A$$

1. Calculer l'ensemble des fermés de Σ ;
2. Construire la base à partir des fermés.