

BASES DE DONNÉES : ALGÈBRE RELATIONNELLE

I. Opérateurs ensemblistes

Soient R_1 et R_2 deux relations de même schéma $S = ((A_1, D_1), \dots, (A_n, D_n))$. R_1 et R_2 sont en fait des parties de $D_1 \times D_2 \times \dots \times D_n$.

On peut donc leur appliquer un opérateur ensembliste : l'union $\cup$, l'intersection $\cap$, la différence $-$. On obtient alors une partie de $D_1 \times D_2 \times \dots \times D_n$, donc une relation de schéma S .

1) Union

L'union de deux relations $R_1(S)$ et $R_2(S)$ est l'ensemble des valeurs appartenant à R_1 ou R_2 . On la note $R_1 \cup R_2$.

Exemple

On considère deux relations *livres_alice* et *livres_bob* ayant pour schéma (titre, auteur) :

titre	auteur
Un mari idéal	Oscar Wilde
Le Mariage forcé	Molière
Le Rouge et le Noir	Stendhal
Les Châtiments	Victor Hugo
Zadig	Voltaire

titre	auteur
La Mare au diable	George Sand
Le Bourgeois gentilhomme	Molière
Le Rouge et le Noir	Stendhal
Zadig	Voltaire

titre	auteur
Un mari idéal	Oscar Wilde
Le Mariage forcé	Molière
Le Rouge et le Noir	Stendhal
Les Châtiments	Victor Hugo
Zadig	Voltaire
La Mare au diable	George Sand
Le Bourgeois gentilhomme	Molière

2) Intersection

L'intersection de deux relations $R_1(S)$ et $R_2(S)$ est l'ensemble des valeurs appartenant à R_1 et R_2 . On la note $R_1 \cap R_2$.

Exemple

titre	auteur
Le Rouge et le Noir	Stendhal
Zadig	Voltaire

3) Différence

La différence entre deux relations $R_1(S)$ et $R_2(S)$ est l'ensemble des valeurs appartenant à R_1 mais pas à R_2 . On la note $R_1 - R_2$.

Exemple

titre	auteur
Un mari idéal	Oscar Wilde
Le Mariage forcé	Molière
Les Châtiments	Victor Hugo

II. Opérateurs de l'algèbre relationnelle

1) Projection

Soit $R(S)$ une relation de schéma S . Pour tout élément e de R et tout attribut A , on note $e.A$ la composante de e associée à l'attribut A .

Plus généralement, si $X = (A_1, \dots, A_p) \subset S$, notons $e.X = (e.A_1, \dots, e.A_p)$.

On appelle alors projection de R selon X et on note $\pi_X(R)$ la relation

$$\pi_X(R) = \{e(X), e \in R\}$$

Le schéma de $\pi_X(R)$ est donc X .

Exemple

Reprenons la relation *eleves* du chapitre précédent. La relation $\pi_{nom, dateNaissance}(eleves)$ est :

nom	dateNaissance
ALCAZAR	1998-02-01
ALCAZAR	1999-07-06
CASTAFIORE	2000-05-04
HADDOCK	1999-03-02
LAMPION	2000-07-14
RASTAPOPOULOS	1998-12-25
TOURNESOL	2000-04-03

On ne garde en fait que les colonnes *nom* et *dateNaissance*.

2) Sélection

Soit $R(S)$ une relation de schéma S et C une condition portant sur un ou plusieurs attributs de S .

On appelle sélection de R selon le critère C et on note $\sigma_C(R)$ la relation obtenue en sélectionnant dans R uniquement les valeurs qui vérifient le critère C .

Par exemple, $\sigma_{A=1}(R)$ sélectionne les éléments e de R tels que $e.A = 1$; $\sigma_{A \geq 2}(R)$ sélectionne les éléments e de R tels que $e.A \geq 2$.

On peut aussi utiliser des conditions portant sur plusieurs attributs : $\sigma_{A_1=A_2}(R)$ sélectionne les éléments e de R tels que $e.A_1 = e.A_2$.

Enfin, il est possible d'utiliser les opérateurs logiques « ou », « et » dans le critère de sélection.

Exemples

— Les étudiants de PCSI :

$$\sigma_{classe="PCSI"}(eleves)$$

nom	prenom	dateNaissance	classe	etabOrigine
ALCAZAR	Ramon	1999-07-06	PCSI	Ottokar
CASTAFIORE	Bianca	2000-05-04	PCSI	Moulinsart
LAMPION	Séraphin	2000-07-14	PCSI	Ottokar
RASTAPOPOULOS	Roberto	1998-12-25	PCSI	Rackham le Rouge

— Les noms et prénoms des étudiants de MPSI ayant effectué leur terminale au lycée Moulinsart :

$$\pi_{nom, prenom}(\sigma_{classe="MPSI" \text{ et } etabOrigine="Moulinsart"}(eleves))$$

nom	prenom
HADDOCK	Archibald
TOURNESOL	Tryphon

Ici, on a appliqué une projection à notre sélection pour ne conserver que les attributs *nom* et *prenom*.

3) Renommage

Soit $S = (A_1, \dots, A_n)$ un schéma et $R(S)$ une relation de schéma S . Soit B un attribut de même domaine que A_i où $i \in \llbracket 1, n \rrbracket$.

On note $\rho_{A_i \leftarrow B}(S)$ le schéma $(A_1, \dots, A_{i-1}, B, A_{i+1}, \dots, A_n)$ et $\rho_{A_i \leftarrow B}(R(S))$ la relation contenant exactement les mêmes valeurs que $R(S)$ et dont le schéma est $\rho_{A_i \leftarrow B}(S)$.

Cela revient à renommer l'attribut A_i en B .

Par extension, si $i \neq j$, $\rho_{A_i, A_j \leftarrow B, D}(R(S))$ est la relation $\rho_{A_j \leftarrow D}(\rho_{A_i \leftarrow B}(S))$ (renommage de A_i en B et de A_j en D). Cette notation se généralise à un nombre quelconque de renommages.

Cette opération sert généralement à lever une ambiguïté lorsqu'on manipule plusieurs tables.

$\rho_{\text{titre, auteur} \leftarrow \text{title, author}}(\text{livres_alice})$

title	author
Un mari idéal	Oscar Wilde
Le Mariage forcé	Molière
Le Rouge et le Noir	Stendhal
Les Châtiments	Victor Hugo
Zadig	Voltaire

Exemple

4) Produit cartésien

Si $R(S)$ et $R'(S')$ sont deux relations de schémas disjoints $S = (A_1, \dots, A_n)$ et $S' = (B_1, \dots, B_p)$, leur produit cartésien est la relation :

$$R \times R' = \{(x_1, \dots, x_n, y_1, \dots, y_p) \mid (x_1, \dots, x_n) \in R \text{ et } (y_1, \dots, y_p) \in R'\}$$

de schéma $(A_1, \dots, A_n, B_1, \dots, B_p)$.

Pour se ramener à deux schémas disjoints, on peut toujours utiliser des renommages.

Exemple

<i>fromages</i>
nom
Cantal
Bleu de Gex

<i>vins</i>
nom
Chablis
Volnay
Fitou

$\rho_{\text{nom} \leftarrow \text{fromage}}(\text{fromage}) \times \rho_{\text{nom} \leftarrow \text{vin}}(\text{vins})$

fromage	vin
Cantal	Chablis
Cantal	Volnay
Cantal	Fitou
Bleu de Gex	Chablis
Bleu de Gex	Volnay
Bleu de Gex	Fitou

5) Division cartésienne

Soient $R(S)$ et $R'(S')$ deux relations. On suppose que $S' \subset S$.

La division cartésienne de R par R' , notée $R \div R'$, est la plus grande relation (pour l'inclusion) telle que $(R \div R') \times R' \subset R$, autrement dit telle qu'il existe R'' vérifiant

$$((R \div R') \times R') \cup R'' = R \text{ et } ((R \div R') \times R') \cap R'' = \emptyset$$

Exemple

On utilise la relation *eleves* du premier exemple du chapitre précédent. On définit deux relations R et R' de la manière suivante :

$$R = \pi_{classe, etabOrigin}(eleves)$$

classe	etabOrigin
MPSI	Ottokar
PCSI	Ottokar
PCSI	Moulinsart
MPSI	Moulinsart
PCSI	Ottokar
PCSI	Rackham le Rouge
MPSI	Moulinsart

$$R' = \pi_{classe}(eleves)$$

classe
MPSI
PCSI

$R \div R'$ correspond alors aux établissements ayant des anciens élèves dans toutes les classe :

$$R \div R'$$

etabOrigin
Moulinsart
Ottokar

6) Jointure

Soient $R(S)$ et $R'(S')$ sont deux relations de schémas disjoints, A un attribut de S et A' un attribut de S' de même domaine.

On appelle jointure symétrique de R et R' selon (A, A') et on note $R \bowtie_{A=A'} R'$ la relation :

$$R \bowtie_{A=A'} R' = \{e \in R \times R' \mid e.A = e.A'\} = \sigma_{A=A'}(R \times R')$$

Exemple

On définit à partir des relations *classes* et *eleves* les relations $R = \pi_{eleves}$ et $R' = \rho_{nom \leftarrow sigle}(classe)$.

On peut ensuite définir la relation $R \bowtie_{classe=sigle} R'$:

nom	pre nom	dateNaissance	classe	etabOrigine	sigle	profPrincipal
ALCAZAR	Peggy	1998-02-01	MPSI	Ottokar	MPSI	DUPONT
ALCAZAR	Ramon	1999-07-06	PCSI	Ottokar	PCSI	DUPOND
CASTAFIORE	Bianca	2000-05-04	PCSI	Moulinsart	PCSI	DUPOND
HADDOCK	Archibald	1999-03-02	MPSI	Moulinsart	MPSI	DUPONT
LAMPION	Séraphin	2000-07-14	PCSI	Ottokar	PCSI	DUPOND
RASTAPOPOULOS	Roberto	1998-12-25	PCSI	Rackham le Rouge	PCSI	DUPOND
TOURNESOL	Tryphon	2000-04-03	MPSI	Moulinsart	MPSI	DUPONT

Exercices

EXERCICE 1

On considère la relation *notes* suivante :

nom	prenom	classe	note
ALCAZAR	Peggy	MPSI	12,5
ALCAZAR	Ramon	PCSI	11
CASTAFIORE	Bianca	PCSI	8,75
HADDOCK	Archibald	MPSI	9
LAMPION	Séraphin	PCSI	13,5
RASTAPOPOULOS	Roberto	PCSI	16
TOURNESOL	Tryphon	MPSI	14

Traduire les requêtes suivantes en opérations de l'algèbre relationnelle :

- Obtenir la liste des classes.
- Obtenir toutes les informations concernant la classe de MPSI.
- Obtenir les prénoms des élèves de PCSI.
- Obtenir les noms et les notes des élèves ayant eu une note inférieure à 10.

EXERCICE 2

On dispose des relations suivantes :

R_1

cours	nom	note
Physique	Toto	15
Maths	Tutu	13
Latin	Toto	17
Maths	Toto	12
Physique	Tutu	8
Physique	Titi	18

R_2

matiere	prof
Physique	Dupont
Maths	Durand
Latin	Martin

On note R'_2 la relation $\rho_{matiere \leftarrow cours}(R_2)$.

Calculer et exprimer dans le langage courant les requêtes :

- $\sigma_{cours="Maths"}(R_1)$
- $\pi_{cours}(R'_2) - \pi_{cours}(R_1)$
- $\pi_{cours}(R_1) - \pi_{cours}(R'_2)$
- $\pi_{cours,nom}(R_1) \div \pi_{cours}(R'_2)$
- $\pi_{prof} \left(\sigma_{note \geq 15 \text{ et } nom="Toto"} \left(R_1 \bowtie_{cours=matiere} R_2 \right) \right)$

EXERCICE 3

On dispose des relations *films*, *artistes* et *roles* définies dans la partie II du chapitre précédent.

Traduire les requêtes suivantes en opérations de l'algèbre relationnelle :

- En quelle année est sorti *Psychose* ?
- Quels sont les titres des films sortis en 2016 ?
- En quelles années sont sortis des films réalisés par Gus Van Sant ?
- Quels sont les films dans lesquels a joué l'acteur dont l'identifiant est 42 ?
- Dans quels films joue Scarlett Johansson ?
- À quels films a participé Ben Affleck ?