Deuxième partie II

L'algèbre relationnelle

Que vient faire l'algèbre dans les bases de données?

Jusqu'à présent nous nous sommes concentrés sur les concepts mais...

Comment interroger une base de données pour obtenir les informations qui y sont stockées?

Solution

Définir un langage à base d'opérateurs algébriques pour manipuler les tables et les relations. Ce langage est l'algèbre relationnelle.

Plan du cours

- Partie I : Introduction aux bases de données relationnelles
 - Cours 1 : Concepts des bases de données relationnelles
 - Cours 2 : L'algèbre relationnelle
- Partie II : Utilisation des bases de données relationnelles
 - Cours 3: Le langage SQL DML (1)
 - Cours 4 : Le langage SQL DML (2)
 - Cours 5 : Le langage SQL DDL
- Partie III : Developpement des bases de données relationnelles
 - Cours 6 : Le modèle entité-association
 - Cours 7 : Élaboration d'un schéma conceptuel
 - Cours 8 : Production du schéma de la base de données

Un exemple : la gestion des étudiants (1/2)

Etudiants

- nom, prénom, numéro, date de naissance, date de première inscription, régime de sécurité sociale, etc.
- Enseignants
 - numéro de sécurité sociale, nom, prénom, date de naissance, emploi occupé, salaire, catégorie, date de titularisation, nombre d'enfants, adresse, téléphone, etc.
- UFR (Unité de Formation et de Recherche)
 - nom, responsable, bâtiment, téléphone, secrétariat, diplômes préparés, modules, cours, etc.
- Inscriptions
 - étudiants, UFR, cours et diplômes, etc.
- Affectation des enseignants
 - calcul des services d'enseignement, congés, modification de salaires, etc.

Un exemple : la gestion des étudiants (2/2)

• Questions :

- Q1 : Qui suit le cours de BD et qui le donne?
- Q2 : A quels cours est inscrit Lulu?
- Q3 : Qui donne le cours de droit?
- Q4 : Quels sont les professeurs de Toto et de Zoé?
- Q5 : Quels sont les étudiants inscrits à au moins un cours?
- Q6 : Quels sont les étudiants qui suivent tous les cours?
- Q7 : Quel est le nombre d'étudiants inscrits au cours de BD?
- Q8 : Quels sont les étudiants de Michel et de Roland?

L'algèbre relationnelle

- L'algèbre relationnelle a été proposée par Codd en 1970
- Elle est composée d'une collection d'opérateurs algébriques unaires ou binaires
- Possibilité de composition des opérateurs (propriétés de fermeture de l'algèbre relationnelle)
- Une requête relationnelle est la composition d'un nombre fini d'opérateurs algébriques
 - L'ordre d'évaluation des opérateurs a un impact sur le temps de réponse du SGRD
 - Principe de l'optimisation des requêtes

Pour ceux qui ne sauraient pas?

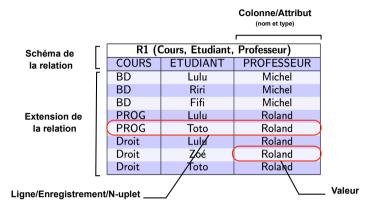
- Une algèbre est un ensemble d'opérateurs de base formellement définis, qui peuvent être combinés à souhait pour construire des expressions algébriques
 - Les opérateurs arithémtiques : +; -, ×, / etc.
 - les opérateurs logique : ∧, ∨, ¬
- Chaque opérateur à une arité
 - Exemples d'opérateurs binaires : +, -, ×, /
 - Exemple d'opérateurs unaires : −, ¬
- Règles de construction des opérateurs ou axiomes
 - Axiomes pour +, × (associativité, commutativité, distributivité, etc.)
 - Axiomes (loi de Morgan) pour ∧ et ∨
- Une algèbre est dite fermée si le résultat de tout opérateur est du même type que les opérandes (ce qui est indispensable pour construire des expressions)

Les opérateurs de l'algèbre relationnelle

- Les opérateurs ensemblistes
 - 1. Union : $R \cup S$
 - 2. Intersection : $R \cap S$
 - 3. Différence : R S
 - 4. Complément : -R
- Les autres opérateurs
 - 1. La projection : R[A]
 - 2. La sélection : R : C
 - 3. Le produit cartésien : ×
 - 4. La jointure : $R(C) \times S$
 - 5. Division : R/S

Rappel de vocabulaire

- Une TABLE = une RELATION
- On pourra faire référence à un attribut d'une relation en utilisant la notation : Relation.Attribut



La projection

- Soients les relations suivantes :
 - EMPLOYE(NOM, SALAIRE, SERVICE, ADRESSE)
 - RESPONSABLE(NOM, SERVICE)
- Questions
 - Donner la liste de tous les employés
 - Pour chaque employé, donner son service et son adresse
 - Pour chaque employé, donner son salaire
 - Donner la liste des responsables de service

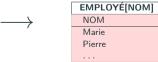
La projection

- Opérandes :
 - Une relation R de schéma X
 - A une sous-liste de X
- Résultat : une relation R' réduite aux colonnes listées dans A
- Notation : R[A]

La projection

• Donner la liste de tous les employés

EMPLOYÉ				
NOM SALAIRE SERVICE ADRESSE				
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	



La projection

• Pour chaque employé, donner son service et son adresse

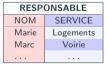
EMPLOYÉ				
NOM SALAIRE SERVICE ADRESSE				
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	



EMPLOYÉS[NOM, SERVICE, ADRESSE]		
NOM	SERVICE	ADRESSE
Marie	Logements	St-Egreve
Pierre	Voirie	Grenoble

La projection

• Donner la liste de tous les responsables de service





RESPONSABLE[NOM]
NOM
Marie
Pierre

La projection

• Pour chaque employé, donner son salaire

EMPLOYÉ				
NOM SALAIRE SERVICE ADRESSE				
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	



EMPLOYÉ[NOM, SALAIRE]		
NOM	SALAIRE	
Marie	1900	
Pierre	1300	

La sélection

- Opérandes :
 - Une relation R de schéma X
 - Une condition C
- Résultat : une relation R' de schéma X ne contenant que les lignes telles que la condition C est vraie
- Notation : R : C
- Condition : la condition est aussi appelée critère de sélection. C'est une expression composée de :
 - de valeur, e.g., "Marie"
 - de nom de colonnes ou d'attributs, e.g., "SALAIRE"
 - d'opérateurs arithmétiques $(<,>,=,\neq,\leq,\geq)$;
 - $\bullet\,$ d'opérateurs logiques $\wedge,\,\vee,\,\neg$

La sélection

• Donner les informations relatives aux employés habitant Grenoble

EMPLOYÉ				
NOM SALAIRE SERVICE ADRESSE				
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	



EMPLOYÉ : (ADRESSE = "Grenoble")					
NOM SALAIRE SERVICE ADRESSE					
Pierre	1300	Voirie	Grenoble		

La sélection

• Que produit la requête : EMPLOYÉ : (NOM = "Marie")?

EMPLOYÉ			
NOM	SALAIRE	SERVICE	ADRESSE
Marie	1900	Logements	St-Egreve
Pierre	1300	Voirie	Grenoble



EMPLOYÉ : (NOM = "Marie")			
NOM	SALAIRE	SERVICE	ADRESSE
Marie	1900	Logement	St-Ègreve

La sélection

• Donner les informations relatives aux employés dont le salaire est supérieur à 1500€

EMPLOYÉ					
NOM	NOM SALAIRE SERVICE ADRESSE				
Marie	1900	Logements	St-Egreve		
Pierre	1300	Voirie	Grenoble		



$EMPLOYE: (SALAIRE \geq 1500)$				
NOM	SALAIRE	SERVICE	ADRESSE	
Marie	1900	Logement	St-Ègreve	

La sélection

• Que produit la requête : EMPLOYÉ : (NOM = "Marie" \land ADRESSE \ne "St-Égrève" \land SALAIRE < 3000) ?

	EMPLOYÉ				
NOM	SALAIRE SERVICE ADRESS				
Marie	1900	Logements	St-Egreve		
Pierre	1300	Voirie	Grenoble		



RESULTAT				
NOM	SALAIRE	SERVICE	ADRESSE	

Le produit cartésien

• Opérandes :

- Une relation R de schéma X
- ullet Une relation S de schéma Y
- Résultat : une relation T dont le schéma est la concaténation de X et de Y contenant tous les couples d'éléments de R et S
- Notation : $R \times C$

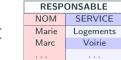
La jointure

- Opérandes :
 - Une relation R de schéma X
 - Une relation S de schéma Y
 - Une condition C
- Résultat : une relation T dont le schéma est la concaténation de X et de Y contenant tous les couples d'éléments de R et S tels que C est vraie
- Notation : $R(C) \times S$
- Condition : la condition est aussi appelée critère de jointure. C'est une expression composée de :
 - de valeur, e.g., "Marie"
 - de nom de colonnes ou d'attributs, e.g., "SALAIRE"
 - d'opérateurs arithmétiques $(<, >, =, \neq, \leq, \geq)$;
 - d'opérateurs logiques ∧, ∨, ¬
- Remarques :
 - 1. Propriété : $R(C) \times S = (R \times S) : C$
 - 2. On parle de jointure naturelle lorsque ${\it C}$ est un test d'égalité et de thêta-produit dans le cadre général

Le produit cartésien

• Quel est le produit cartésien des tables EMPLOYÉ et RESPONSABLE?

EMPLOYÉ				
NOM	SALAIRE SERVICE ADRESSE			
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	



$EMPLOYE \times RESPONSABLE$					
EMP.NOM	EMP.SALAIRE	EMP.SERVICE	EMP.ADRESSE	RESP.NOM	RESP.SERVICE
Marie	1900	Logements	St-Egreve	Marie	Logement
Marie	1900	Logements	St-Egreve	Marc	Voirie
Pierre	1300	Voirie	Grenoble	Marie	Logement
Pierre	1300	Voirie	Grenoble	Marc	Voirie

La jointure

• Pour chaque employé, donner son responsable

EMPLOYÉ				
NOM	SALAIRE SERVICE ADRESSE			
Marie	1900	Logements	St-Egreve	
Pierre	1300	Voirie	Grenoble	





EMPLOYÉ (EMPLOYÉ.SERVICE = RESPONSABLE.SERVICE) × RESPONSABLE		
EMPLOYÉ.NOM	RESPONSABLE.NOM	
Marie	Marie	
Pierre	Marc	
Marc	Marc	

La jointure

• La jointure = mise en correspondance de 2 tables selon un critère

R				
Υ	Z			
2	3			
5	6			
8	9			
	5			





$R (Y < V) \times S$				
X	Υ	Z	V	W
1	2	3	3	1
1	2	3	6	2
4	5	6	6	2

La division

- Opérandes :
 - Une relation R de schéma X
 - Une relation S de schéma Y
- Résultat : une relation T de schéma X-Y ne contenant aucune des lignes (v_1, v_2, \ldots, v_n) telles que pour chaque ligne (w_1, w_2, \ldots, w_n) de S, R contient les lignes $(v_1, v_2, \ldots, v_n, w_1, w_2, \ldots, w_n)$
- Notation : R/S

La jointure

•
$$R(C) \times S = (R \times S) : C$$

• Exemple :

R		
Υ	Z	
2	3	
5	6	
8	9	
	Y 2 5	

		$R \times S : (Y < V)$				
/	X	Υ	Z	V	W	
	1	2	3	3	1	
: -	1	2	3	6	2	
	4	5	6	3	1	
	4	5	6	6	2	
	7	8	9	3	1	
	7	8	9	6	2	

La division

La division

- Soients les relations suivantes :
 - ÉTUDIANT(<u>NUMET</u>, NOM, PRÉNOM)
 - MATIÈRE(NUMMAT, LIBELLÉ)
 - ÉPREUVE(NUMEPR, NUMMAT, DATE)
 - INSCRIPTION(NUMET, NUMMAT)
 - RÉSULTAT(NUMET, NUMEPR, NOTE)
- Questions
 - Quels sont les étudiants inscrits dans toutes les matières?
 - Quels sont les étudiants ayant une note à chacune des épreuves?

La division

• Quels sont les étudiants ayant obtenus une note à chacune des épreuves?

RÉSULTAT			
NUMET	NUMEPR NOTE		
2	22	12	
2	53	8	
2	85	11	
3	43	14	
5	22	13	
5	43	15	
5	53	11	
5	85	18	

EPREUVE			
NUMEPR	NUMMAT	DATE	
22	12	01/12/2014	
43	12	11/01/2015	
53	14	11/06/2014	
85	13	14/09/2014	

RÉSULTAT / ÉPREUVE				
NUMET				
5				

La division

• Quels sont les étudiants inscrits dans toutes les matières?

INSCRIPTION					
NUMET	NUMMAT				
2	12				
2	14				
2	13				
3	13				
5	12				
5	14				
5	13				

MATIÈRE			
NUMMAT LIBELLE			
12	Mathématiques		
13	Droit		
14	Informatique		

INSCRIPTION / MATIÈRE			
NUMET			
2			
5			

La division

• Si l'on veut maintenant les noms et prénoms des étudiants ayant obtenus une note à chacune des épreuves :

RÉSULTAT / ÉP	REUVE
NUMET	
5	



=

ÉTUDIANT				
NUMET NOM PRÉNC				
2	Fourier	Joseph		
3	Beyle	Henri		
5	Mendès-France	Pierre		

	$R(RÉSULTAT / ÉPREUVE) (R.NUMET = S.NUMET) \times S(ÉTUDIANT)$				
	NUMET	NOM	PRÉNOM		
ĺ	5	Mendès-France	Pierre		

L'union

• Opérandes :

- Une relation R de schéma X
- Une relation S de schéma X
- Résultat : une relation ${\cal T}$ de schéma ${\cal X}$ contenant toutes lignes de ${\cal R}$ et de ${\cal S}$
- Notation : $R \cup S$
- Remarque : R et S doivent avoir le même schéma

L'intersection

- Opérandes :
 - Une relation R de schéma X
 - Une relation S de schéma X
- Résultat : une relation T de schéma X ne contenant que les lignes qui sont à la fois dans R et dans S
- Notation : $R \cap S$
- Remarque : R et S doivent avoir le même schéma

L'union





$R \cup S$				
Χ	Υ	Z		
а	b	С		
d	е	f		
k	- 1	m		
g	h	i		

• Remarques :

- La ligne [d, e, f] n'apparaît qu'une seule fois dans la relation $R \cup S$
- Une relation définit un ensemble

L'intersection

R			
X	Υ	Z	
а	b	С	
d	е	f	
g	h	i	



	S	
X	Υ	Z
k	- 1	m
d	е	f

$R\capS$		
X		
d		

La différence

• Opérandes :

- Une relation R de schéma X
- Une relation S de schéma X
- Résultat : une relation ${\mathcal T}$ de schéma ${\mathcal X}$ contenant que les lignes de ${\mathcal R}$ qui ne sont pas dans ${\mathcal S}$
- Notation : R S
- Remarque : R et S doivent avoir le même schéma

Le complément

- Opérandes :
 - Une relation R de schéma X
- Résultat : une relation R' de schéma X contenant regroupant exclusivement toutes les occurrences possibles créées à partir des valeurs d'attributs présentes dans la relation R, à l'exception des occurrences de R
- Notation : -R
- Remarques :
 - Comme l'ensemble des valeurs possibles pour chaque attribut est généralement inconnu ou infini, seules les valeurs déjà présentes dans un attribut de R sont utilisées, dans cet attribut, pour créer de nouvelles occurrences
 - 2. Si R est vide, la relation qui résulte du complément est vide
 - 3. Propriété : --R=R

La différence



Le complement

Soit la relation

	R	
Χ	Υ	Z
а	b	С
d	е	f

son complément est

-R		
X	Υ	Z
а	b	f
а	е	С
a	е	f
d	b	С
d	b	f
d	е	С

Règles permettant l'optimisation des requêtes

- L'ordre dans lequel s'effectue les opérateurs ont un impact sur les performances des SGBD
- À savoir :
 - Les sélections diminuent le nombre de lignes et donc la taille des tables
 - Les projections diminuent un peu la taille des tables
 - Les produits et les jointures augmentent considérablement la taille des tables
- Pour ces raisons, il est important d'effectuer les opérateurs dans l'ordre suivant :
 - 1. Sélections (qui diminuent beaucoup la taille des tables)
 - 2. Projections (qui diminuent un peu la taille des tables)
 - 3. jointures et produits (qui augmentent beaucoup la taille des tables)

Exercices

- Rappel du schéma :
 - CLIENT (NUMCLIENT, NOM, PRÉNOM, E-MAIL, NUMCB)
 - RESERVATION (NUMCLIENT, CODEVOYAGE, DATERES)
 - VOYAGE (CODEVOYAGE, DESTINATION, DURÉE, PRIX)
- Questions
 - 1. Nom, prénom et e-mail des clients ayant une réservation en cours



2. Nom, prénom et e-mail des clients n'ayant aucune réservation en cours



Exercices

- Soit le schéma relationnel de la base de données d'une agence de voyage :
 - CLIENT (NUMCLIENT, NOM, PRÉNOM, E-MAIL, NUMCB)
 - RESERVATION (NUMCLIENT, CODEVOYAGE, DATERES)
 - VOYAGE (CODEVOYAGE, DESTINATION, DURÉE, PRIX)
- Formuler en algèbre relationnelle les requêtes suivantes :
 - 1. Nom, prénom et e-mail des clients ayant une réservation en cours
 - Nom, prénom et e-mail des clients n'ayant aucune réservation en cours
 Destination et liste des clients ayant réservés pour un voyage de plus de 10 jours et coûtant moins de 1000€
 - 4. Numéros de tous les clients ayant réservés sur tous les voyages proposés

Exercices

- Rappel du schéma :
 - CLIENT (NUMCLIENT, NOM, PRÉNOM, E-MAIL, NUMCB)
 - RESERVATION (NUMCLIENT, CODEVOYAGE, DATERES)
 - VOYAGE (CODEVOYAGE, DESTINATION, DURÉE, PRIX)
- Questions
 - Destination et liste des clients ayant réservés pour un voyage de plus de 10 iours et coûtant moins de 1000€



4. Numéros de tous les clients ayant réservés sur tous les voyages proposés

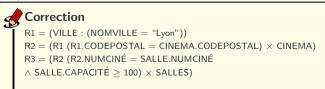


Exercices

- Soit le schéma relationnel de la base de données d'un cinéma :
 - VILLE (CODEPPOSTAL, NOMVILLE)
 - CINEMA (NUMCINÉ, NOMCINÉ, ADRESSE, CODEPOSTAL)
 - SALLE (NUMSALLE, CAPACITÉ, NUMCINÉ)
 - FILM (NUMEXP, TITRE, DUREÉ, CODEDIST)
 - PROJECTION (NUMEXP. NUMSALLE, NUMSEMAINE, NBENTRÉE)
- Formuler en algèbre relationnelle les requêtes suivantes :
 - 1. Titre des films dont la durée est supérieure ou égale à deux heures
 - 2. Nom des villes abritant un cinéma nommé "Le Capitole"
 - 3. Nom des cinémas situés à Lyon ou contenant au moins une salle de plus 100 places
 - 4. Nom, adresse et ville des cinémas dans lesquels on joue le film "Hypnose" la semaine 18
 - 5. Numéro d'exploitation des films projetés dans toutes les salles
 - 6. Titre des films qui n'ont pas été projetés

Exercices

- Rappel du schéma relationnel :
 - VILLE (CODEPPOSTAL, NOMVILLE)
 - CINEMA (NUMCINÉ, NOMCINÉ, ADRESSE, CODEPOSTAL)
 - SALLE (NUMSALLE, CAPACITÉ, NUMCINÉ)
 - FILM (NUMEXP. TITRE, DUREÉ, CODEDIST)
 - PROJECTION (NUMEXP, NUMSALLE, NUMSEMAINE, NBENTRÉE)
- Questions :
 - Nom des cinémas situés à Lyon ou contenant au moins une salle de plus 100 places



Exercices

• Rappel du schéma relationnel :

- VILLE (CODEPPOSTAL, NOMVILLE)
- CINEMA (NUMCINÉ, NOMCINÉ, ADRESSE, CODEPOSTAL)
- SALLE (NUMSALLE, CAPACITÉ, NUMCINÉ)
- FILM (NUMEXP, TITRE, DUREÉ, CODEDIST)
- PROJECTION (NUMEXP, NUMSALLE, NUMSEMAINE, NBENTRÉE)

• Questions :

1. Titre des films dont la durée est supérieure ou égale à deux heures



2. Nom des villes abritant un cinéma nommé "Le Capitole"



Exercices

• Rappel du schéma relationnel :

- VILLE (CODEPPOSTAL, NOMVILLE)
- CINEMA (NUMCINÉ, NOMCINÉ, ADRESSE, CODEPOSTAL)
- SALLE (NUMSALLE, CAPACITÉ, NUMCINÉ)
- FILM (NUMEXP. TITRE. DUREÉ. CODEDIST)
- PROJECTION (NUMEXP, NUMSALLE, NUMSEMAINE, NBENTRÉE)

Questions :

4. Nom, adresse et ville des cinémas dans lesquels on joue le film "Hypnose" la semaine 18

```
Correction

R1 = (PROJECTION : (SEMAINE = 18))

R2 = (FILM : (TITRE = "Hypnose"))

R3 = (R1 (R1.NUMEXP = R2.NUMEXP) × R2)

R3 = (R2 (R2.NUMSALLE = SALLE.NUMSALLE) × SALLE

R4 = (R3 (R3.NUMCINÉ = CINÉMA.NUMCINÉ) × CINÉMA

R5 = (R4 (R4.CODEPOSTAL = CILLE.CODEPOSTAL) × VILLE

R6 = R5[NOMCINÉ, ADRESSE, NOMVILLE]
```

Exercices

- Rappel du schéma relationnel :
 - VILLE (CODEPPOSTAL, NOMVILLE)
 - CINEMA (NUMCINÉ, NOMCINÉ, ADRESSE, CODEPOSTAL)
 - SALLE (NUMSALLE, CAPACITÉ, NUMCINÉ)
 - FILM (NUMEXP, TITRE, DUREÉ, CODEDIST)
 - PROJECTION (NUMEXP, NUMSALLE, NUMSEMAINE, NBENTRÉE)
- Questions :
 - 5. Numéro d'exploitation des films projetés dans toutes les salles



R = (PROJECTION[NUMEXP, NUMSALLE] / SALLE)

6. Titre des films qui n'ont pas été projetés



R1 = (FILM[NUMEXP] - PROJECTION[NUMEXP])

 $R2 = (R1 (R1.NUMEXP = FILM.NUMEXP) \times FILM)[TITRE]$