

COURS de BASES de DONNEES

Céline Robardet

INSA-Lyon

Point de départ

- ◆ Une base de données est une collection de données ayant une origine commune
- ◆ Un Système de Gestion de Base de Données (SGBD) est un logiciel conçu pour remplir deux buts :
 - ▶ Ajouter, effacer et mettre à jour les données dans la base
 - ▶ Fournir différentes manières de visualiser tout ou partie des données de la base

Exemple

ISBN	Titre	IdAuteur	NomAuteur	TelAuteur	IdEditeur	NomEditeur	TelEditeur	Prix
0-103-4578-9	Iliad	3	Homer	333-333-33	1	BigHouse	123-456-78	25,00 F
0-11-3456-3	Moby Dick	2	Melville	222-222-22	3	Small Hous	714-000-00	49,00 F
0-12-33345-3	On liberty	8	Mill	888-888-88	1	BigHouse	123-456-78	25,00 F
0-12-34678-9	Jane Eyre	1	Austen	111-111-11	3	Small Hous	714-000-00	49,00 F
0-123-52395	Ulysses	6	Joyce	666-666-66	2	Alpha Press	999-999-99	34,00 F
0-312-3143-3	Balloon	13	Sleepy	312-321-11	3	Small Hous	714-000-00	34,00 F
0-312-3143-3	Balloon	11	Snoopy	321-321-22	3	Small Hous	714-000-00	34,00 F
0-312-3143-3	Balloon	12	Grumpy	312-321-00	3	Small Hous	714-000-00	34,00 F
0-55-12554	Main Street	10	Jones	123-333-33	3	Small Hous	714-000-00	22,95 F
0-55-12554	Main Street	9	Smith	123-222-22	3	Small Hous	714-000-00	22,95 F
0-555-555-9	Macbeth	5	Shakespeare	555-555-55	2	Alpha Press	999-999-99	12,00 F
0-91-04578-5	Hamlet	5	Shakespeare	555-555-55	2	Alpha Press	999-999-99	20,00 F
0-91-35678-7	Faerie Queen	7	Spenser	777-777-77	1	BigHouse	123-456-78	15,00 F
0-99-77777-7	King Lear	5	Shakespeare	555-555-55	2	Alpha Press	999-999-99	49,00 F
0-99-9999-9	Emma	1	Austen	111-111-11	1	BigHouse	123-456-78	20,00 F
1-1111-111-1	c++	4	Roman	444-444-44	1	BigHouse	123-456-78	29,95 F
1-22-23370-0	Visual Basic	4	Roman	444-444-44	1	BigHouse	123-456-78	25,00 F

Table unique

- ◆ Gestion des données dans une simple table
 - un traitement de texte
- ◆ Bibliothèque du congrès 16 millions de livres
 - Redondances
 - Problèmes pour visualiser les données
- ⇒ Ventiler les données dans plusieurs tables

ISBN	Titre	IdEditeur	Prix
0-312-3143-3	Balloon	3	34,00 F
1-1111-111-1	c++	1	29,95 F
0-99-9999-9	Emma	1	20,00 F
0-91-35678-7	Faerie Queer	1	15,00 F
0-91-04578-5	Hamlet	2	20,00 F
0-103-4578-9	Iliad	1	25,00 F
0-12-34678-9	Jane Eyre	3	49,00 F
0-99-77777-7	King Lear	2	49,00 F
0-555-555-9	Macbeth	2	12,00 F
0-55-12554	Main Street	3	22,95 F
0-11-3456-3	Moby Dick	3	49,00 F
0-12-33345-3	On liberty	1	25,00 F
0-123-52395	Ulysses	2	34,00 F
1-22-23370-0	Visual Basic	1	25,00 F

IdAuteur	NomAuteur	TelAuteur
1	Austen	111-111-1111
12	Grumpy	312-321-0000
3	Homer	333-333-3333
10	Jones	123-333-3333
6	Joyce	666-666-6666
2	Melville	222-222-2222
8	Mill	888-888-8888
4	Roman	444-444-4444
5	Shakespeare	555-555-5555
13	Sleepy	312-321-1111
9	Smith	123-222-2222
11	Snoopy	321-321-2222
7	Spenser	777-777-7777

ISBN	IdAuteur
1-1111-111-1	4
0-99-9999-9	1
0-91-35678-7	7
0-91-04578-5	5
0-103-4578-9	3
0-12-34678-9	1
0-99-77777-7	5
0-555-555-9	5
0-11-3456-3	2
0-12-33345-3	8
0-312-3143-3	13
0-312-3143-3	11
0-312-3143-3	12
0-55-12554	10
0-55-12554	9
0-123-52395	6
1-22-23370-0	4

Quatre tables : les avantages

- ◆ Réduction de la taille de la base
- ◆ Certaines données sont imbriquées i.e. elles apparaissent à plusieurs endroits dans les tables

Problèmes liés à l'utilisation d'une table unique

- ◆ Problèmes relatifs aux valeurs multiples
- ◆ Anomalie de mise à jour
- ◆ Anomalie d'insertion
- ◆ Anomalie de suppression

Problèmes liés aux BDD relationnelles

- ◆ Eviter la perte de données
 - Comment placer les données dans les différentes tables sans perte d'information ?
- ◆ Maintien de l'intégrité référentielle
- ◆ Création de vues
 - Comment obtenir la liste des éditeurs dont les livres publiés coûtent moins de 10 euros ?

Base de Données relationnelle

- ◆ Ensemble organisé de données non indépendantes interrogable par le contenu
- ◆ Collection de fichiers reliés par des pointeurs multiples organisés pour répondre à des questions

Système de Gestion de Base de Données

- ◆ Ensemble de logiciels systèmes gérant les opérations
 - ▶ d'insertion
 - ▶ de modification
 - ▶ de recherche
 - ▶ d'accès concurrents
 - ▶ de confidentialité

Elements d'algèbre relationnelle

Chapitre 2

Le modèle relationnel

- ◆ Inventé par E; F; Codd en 1970
- ◆ Fondé sur la théorie mathématique des ensembles et sur la notion de base qui lui est rattachée : la relation.

Objectifs d'un modèle relationnel: Rappel

- ◆ Proposer une indépendance des données et des traitements.
- ◆ Permettre de gérer les problèmes de cohérence et de redondance des données.
- ◆ Utiliser des structures de données simples et issues de la vie courante (tables).
- ◆ Proposer des langages de manipulation de données non procéduraux
- ◆ Permettre des vues utilisateurs différentes des relations implantées

Le modèle relationnel de CODD: Les 10 concepts clés (1/2)

- ◆ La définition du domaine d'un attribut
- ◆ Le concept de relation
- ◆ Les attributs
- ◆ Les tuples
- ◆ Le schéma d'une relation

Le modèle relationnel de CODD: Les 10 concepts clés (2/2)

- ◆ La clé
- ◆ La contrainte référentielle
- ◆ Les valeurs nulles
- ◆ La contrainte d'entité
- ◆ La contrainte de domaine

Les structures de données : Domaine d'un attribut

- ◆ Définition : ensemble des valeurs prises par un attribut
- ◆ Un domaine se définit soit :
 - en extension
 - ▶ ex. : couleur_yeux = {bleu, vert, marron, noir}
 - en compréhension
 - ▶ ex. : âge_enfants = {x dans N tel que x < 15}

Définition du produit cartésien: Rappel

Soient deux ensembles O et P, on appelle *Produit cartésien* de O et P, l'ensemble noté $O \times P$, des couples (x,y) où $x \in O$ et $y \in P$.

$$\text{Étudiants} = \{ \text{Anne}, \text{Tom}, \text{Maud} \}$$

$$\text{Professeurs} = \{ \text{Lydie}, \text{Yvan} \}$$

$$\text{Étudiants} \times \text{Professeurs} = \{ (\text{Anne}, \text{Lydie}), (\text{Anne}, \text{Yvan}), (\text{Tom}, \text{Lydie}), (\text{T}, \text{Y}), (\text{M}, \text{L}), (\text{M}, \text{Y}) \}$$

Les structures de données : la relation

Définition : Sous-ensemble du produit cartésien d'un ensemble d'attributs

Nom de la relation

PRODUIT

*Occurrences ,
Tuples,
Enregistrements*

Attributs

Num_prod	Nom_prod	Qté_stock
24	Chaise	63
141	Table	27
67	Lit	12

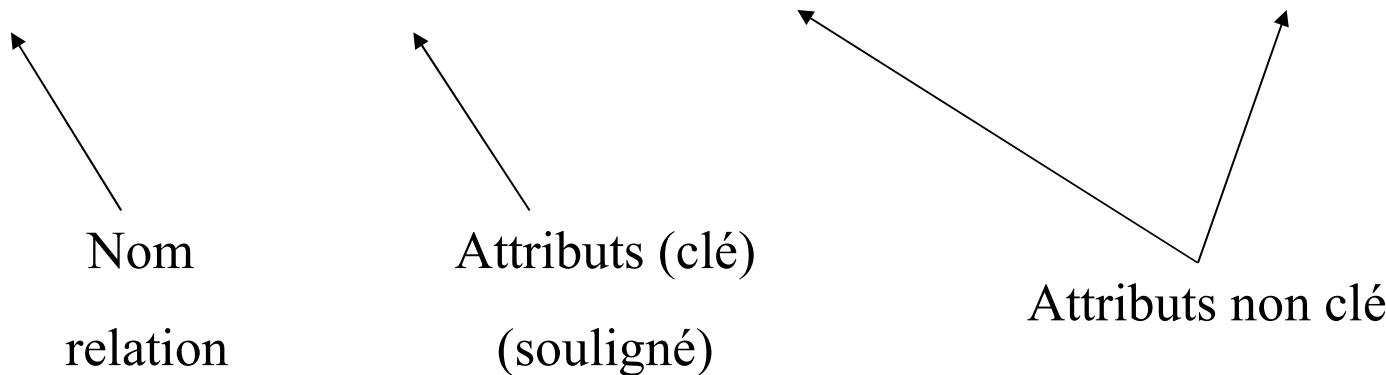
Caractérisé par :

- ◆ Cardinalité de la relation (nombre de lignes)
- ◆ Unicité du nom des attributs
- ◆ Indifférence de l'ordre des lignes et des colonnes¹⁸
- ◆ Nombre de colonnes = degré ou arité

Les structures de données : Schéma de relations

- ▶ Le schéma de relation représente la structure invariante d'une relation.
- ▶ Il est constitué du nom de la relation suivie de la liste des attributs et de leurs domaines associés

PRODUIT (N°PRODUIT : entier, **NOM** : chaîne, **QTE EN STOCK** : entier >0)



Règles d'intégrité structurelle

- ◆ Assertions qui doivent être vérifiées par les données contenues dans la base.
 - Règles inhérentes au modèle relationnel
 - Règles inhérentes au schéma particulier d'une application

Règles d'intégrité structurelle du modèle relationnel

- ◆ Unicité de clé
- ◆ Contrainte de référence
- ◆ Contrainte d'entité
- ◆ Contrainte de domaine

Clé d'une relation

- ◆ Une relation étant un ensemble de tuples, il ne peut y avoir deux tuples identiques dans une relation.
- ◆ Clé : ensemble minimal d'attributs dont la connaissance des valeurs permet d'identifier un unique tuple

Contrainte référentielle(1/2)

- ◆ Contrainte d'intégrité portant sur une relation R1, consistant à imposer que la valeur d'un groupe d'attributs apparaisse comme valeur de clé dans une autre relation R2.
 - ▶ Elles définissent des liens obligatoires entre relations.
 - ▶ Garantissent le succès des opérations de mise à jour.

Contrainte référentielle(2/2)

- ◆ Les contraintes de référence sont liées à la notion de clés étrangères :
 - ▶ Lors de l'insertion d'un tuple dans une relation soumise à une contrainte référentielle (« *relations référençantes* »), il faut vérifier que les valeurs de clés étrangères existent dans les relations référencées.
 - ▶ Lors de la suppression d'un tuple dans une relation référencée, il faut vérifier qu'aucun tuple de la relation référençante ne fasse référence au tuple que l'on souhaite supprimer.
 - soit le SGBD refuse la suppression
 - soit il la répercute en cascade.

Contrainte d'entité

- ◆ Lorsque la valeur d'un attribut est inconnue, on utilise une valeur conventionnelle appelée valeur nulle
- ◆ La contrainte d'entité impose que toute relation possède une clé primaire et que tout attribut participant à cette clé primaire soit non nul.

Contrainte de domaine

- ◆ Les valeurs d'un attribut doivent vérifier une assertion logique
 - Permet de restreindre le domaine de valeur

Algèbre relationnelle : Définition

- ◆ Collection d'opérations formelles qui agissent sur des relations et produisent une relation en résultat.
- ◆ Dans la plupart des systèmes relationnels, la réponse à une requête s'obtient par l'utilisation d'un ou plusieurs opérateurs relationnels.

Algèbre relationnelle : Opérateurs relationnels

- **Opérateurs unaires :**

- ▶ Sélection
- ▶ Projection
- ▶ Complément

- **Opérateurs binaires :**

- ▶ Union
- ▶ Intersection
- ▶ Différence
- ▶ Les produits et Jointures



Théorie des ensembles

Motivation : interroger des données

- Qu'est-ce qu'une requête ?
 - une expression dans une algèbre dite algèbre relationnelle.
 - i.e., des opérations sur des tables
- ◆ Filtrer :
 - Sélection (filtre / critères)
 - Projection (choix de colonnes)
 - Différence (suppression de lignes)
- ◆ Combiner :
 - Produit cartésien (toutes les possibilités)
 - Jointure (complément d'information)
 - Union (union d'ensembles de lignes)

Quelques définitions

◆ Ligne

- Une ligne L est une séquence de valeurs V_i

$$L = \langle V_1, \dots, V_i, \dots, V_n \rangle$$

◆ Extraction

- La ligne $\langle V_i \rangle$ construite à partir de L est notée $L(i)$

◆ Concaténation

- Soient deux lignes

$$L_1 = \langle V_1, \dots, V_n \rangle \text{ et } L_2 = \langle W_1, \dots, W_m \rangle,$$

la concaténation de L_1 et de L_2 , notée $L_1 \bullet L_2$ est :

$$L_1 \bullet L_2 = \langle V_1, \dots, V_n, W_1, \dots, W_m \rangle$$

◆ Table :

- un ensemble de lignes

Algèbre relationnelle : Sélection (ou restriction)

- Suppression des occurrences de la relation qui ne satisfont pas à une condition donnée.
- Opération sur une relation $R1$ produisant une relation $R2$ de même schéma, mais comportant les seuls tuples qui vérifient la condition précisée en argument.
- Notation : $\sigma_{\text{condition}}(R1)$

Sélection (restriction)

T =

c1 (nom)	c2 (âge)	c3 (adresse)	c4 (néA)
Bob	13	Lyon	Nice
Sam	7	Nice	Nice
Cathy	13	Brest	Brest
Julie	20	Lyon	Brest

- ◆ Sélection / constante : Personnes habitant Lyon

$\sigma_{c3="Lyon"}(T)$

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice
Julie	20	Lyon	Brest

- ◆ Sélection / critère inter-colonnes : Personnes nées dans la ville d'habitation

$\sigma_{c3=c4}(T)$

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Sam	7	Nice	Nice
Cathy	13	Brest	Brest

Sélection (restriction)

- ◆ Autre comparateur : Personnes n'ayant pas plus de 14 ans

 $\sigma_{c2 \leq 14}(T)$

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice
Sam	7	Nice	Nice
Cathy	13	Brest	Brest

- ◆ Composition de sélections : Habitants de Lyon n'ayant pas plus de 14 ans

 $\sigma_{c3="Lyon"}(\sigma_{c2 \leq 14}(T))$

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice

- ◆ Conjonction de critères :

 $\sigma_{c3="Lyon"}(\sigma_{c2 \leq 14}(T)) \Leftrightarrow \sigma_{c3="Lyon" \text{ and } c2 \leq 14}(T)$

Une formalisation de la sélection (restriction)

- ◆ Sélection / constante :

L'ensemble des lignes L de la table T telles que $L(i)=a$ est noté $\sigma_{i=a}(T)$

$$\sigma_{i=a}(T) = \{ L \mid L \text{ dans } T \text{ et } L(i) = a \}$$

- ◆ Sélection / inter-colonnes :

L'ensemble des lignes L de la table T telles que $L(i) = L(j)$ est noté $\sigma_{i=j}(T)$

$$\sigma_{i=j}(T) = \{ L \mid L \text{ dans } T \text{ et } L(i) = L(j) \}$$

- ◆ Autres comparateurs : $\neq < > \leq \geq$ peuvent être employés à la place de $=$
- ◆ Conjonction de comparaisons : Pour alléger les notations des conjonctions de critères de sélection peuvent être employées

Algèbre relationnelle : Sélection (ou restriction)

- ◆ Opérateurs de comparaison
 $=, \neq, <, >, \leq, \geq$, BETWEEN, IN, LIKE
- ◆ Combinaison de plusieurs conditions reliées par des opérateurs logiques :
and, or, not
- ◆ Valeurs NULL

Algèbre relationnelle : Projection

- ◆ Opération sur une relation $R1$ consistant à composer une relation $R2$ en enlevant à la relation initiale tous les attributs non mentionnés en opérandes.
- ◆ Notation $\pi_{A_1, A_2, \dots, A_n}(R1)$

Projection

$T =$

	c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice	
Sam	7	Nice	Nice	
Cathy	13	Brest	Brest	
Julie	20	Lyon	Brest	

◆ Projection sur les colonnes 1 et 2 :

Nom et age de toutes les personnes $\pi_{c1,c2}(T)$

	c1 (nom)	c2 (âge)
Bob	13	
Sam	7	
Cathy	13	
Julie	20	

◆ Projection sur les colonnes 1 et 3

Nom et adresse de toutes les personnes

(renumérotation des colonnes)

$\pi_{c1,c3}(T)$

	c1 (nom)	c2 (adr)
Bob	Lyon	
Sam	Nice	
Cathy	Brest	
Julie	Lyon	

Algèbre relationnelle : Projection

Consiste à supprimer des attributs d'une relation

Soit la relation ETUDIANT :

Num_étu	Nom_étu	Nom_départ	Adr_départ
521	Loulou	Informatique	Lyon
632	Babette	Mathématique	Marseille
569	Fifi	Informatique	Lille
451	Loulou	Informatique	Lille

La projection sur *nom_étu*,
nom_départ donne :

Nom_étu	Nom_départ
Loulou	Informatique
Babette	Mathématique
Fifi	Informatique

La projection sur *nom_départ*,
adr_départ donne :

Nom_départ	Adr_départ
Informatique	Lyon
Mathématique	Marseille
Informatique	Lille

Composition (Combinaison)

T =

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice
Sam	7	Nice	Nice
Cathy	13	Brest	Brest
Julie	20	Lyon	Brest

- ◆ Quels sont les noms des personnes habitant Lyon ?

$\sigma_{c3="Lyon"}(T)$

c1 (nom)	c2 (âge)	c3 (adr)	c4 (néA)
Bob	13	Lyon	Nice
Julie	20	Lyon	Brest

1) Sélection :

2) ... suivie d'une projection :

$\pi_{c1}(\sigma_{c3="Lyon"}(T))$

c1 (nom)
Bob
Julie

Une formalisation de la projection

◆ Projection sur colonnes :

L'ensemble des lignes de la table T obtenues en ne conservant que les colonnes de numéros i_1, i_2, \dots, i_k est noté $\pi_{i_1, i_2, \dots, i_k}(T)$

$$\pi_{i_1, i_2, \dots, i_k}(T) = \{ L(i_1)L(i_2) \dots L(i_k) \mid L \text{ dans } T \}$$

◆ Retour sur la composition : le *principe de clôture*

Le résultat d'une opération portant sur des tables
est aussi une table

⇒ composition possible de différentes opérations

NB. Propriété fondamentale pour expliquer le confort des interrogations

Algèbre relationnelle : Produit cartésien

Le produit cartésien se construit en combinant toutes les possibilités.

Soit la relation LIVRE

Titre	Auteur
X	Toto
Y	Loulou

Soit la relation EDITION

Couleur	Edition
Rouge	Luxe
Blanc	Broché
Vert	Cartonné

Le produit cartésien permet de construire la relation suivante

Titre	Auteur	Couleur	Edition
X	Toto	Rouge	Luxe
X	Toto	Blanc	Broché
X	Toto	Vert	Cartonné
Y	Loulou	Rouge	Luxe
Y	Loulou	Blanc	Broché
Y	Loulou	Vert	Cartonné

Algèbre relationnelle : Produit cartésien

- ◆ Opération portant sur deux relations $R1$ et $R2$, consistant à construire une relation $R3$ ayant pour schéma la concaténation de ceux des relations opérandes et pour tuples toutes les combinaisons des tuples des relations opérandes
- ◆ Notation $R1 \times R2$

Produit cartésien

homme =

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

femme =

c1 (nom)	c2 (adr)
Cathy	Brest
Julie	Lyon
Linda	Lyon

- ◆ Quels sont tous les couples homme-femme ?

homme \otimes femme

à partir de 'Bob'

à partir de 'Sam'

c1 (nom)	c2 (adr)	c3 (nom)	c4 (adr)
Bob	Lyon	Cathy	Brest
Bob	Lyon	Julie	Lyon
Bob	Lyon	Linda	Lyon
Sam	Nice	Cathy	Brest
Sam	Nice	Julie	Lyon
Sam	Nice	Linda	Lyon

- ◆ Quels sont les couples homme-femme d'une même ville ?

$\sigma c2 = c4$ (homme \otimes femme)

c1 (nom)	c2 (adr)	c3 (nom)	c4 (adr)
Bob	Lyon	Julie	Lyon
Bob	Lyon	Linda	Lyon

Produit cartésien

homme=

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

ville=

c1 (nomV)	c2 (nbHab)	c3 (départ)
Nice	340.000	Alp-Mar
Brest	160.000	Finistère
Lyon	420.000	Rhône

- ◆ Dans quels départements habitent les personnes de la table homme ?

1) Combinaison par produit cartésien

(homme \otimes ville)

à partir de 'Bob' {

à partir de 'Sam' {

c1 (nom)	c2 (adr)	c3 (nomV)	c4 (nbHab)	c5 (départ)
Bob	Lyon	Nice	340.000	Alp-Mar
Bob	Lyon	Brest	160.000	Finistère
Bob	Lyon	Lyon	420.000	Rhône
Sam	Nice	Nice	340.000	Alp-Mar
Sam	Nice	Brest	160.000	Finistère
Sam	Nice	Lyon	420.000	Rhône

Produit cartésien

2) ... suivi d'une sélection :

$\sigma \text{ c2} = \text{c3} \text{ (homme } \otimes \text{ ville)}$

c1 (nom)	c2 (adr)	c3 (nomV)	c4 (nbHab)	c5 (départ)
Bob	Lyon	Lyon	420.000	Rhône
Sam	Nice	Nice	340.000	Alp-Mar

Une formalisation du produit cartésien

L'ensemble des lignes pouvant être obtenues par concaténation d'une ligne d'une table T1 avec une ligne d'une table T2 est noté $T1 \otimes T2$

$$T1 \otimes T2 = \{ L1.L2 \mid L1 \text{ dans } T1 \text{ et } L2 \text{ dans } T2 \}$$

Le produit $T1 \otimes T2$ est souvent suivi d'une sélection dans laquelle intervient une colonne issue de T1 et une colonne issue de T2

⇒ définition d'un nouvel opérateur : **la jointure**

Algèbre relationnelle : Thêta-produit (thêta-jointure)

Le thêta-produit consiste en un produit cartésien doublé d'une sélection.

On ne retient que les occurrences qui vérifient une condition logique.

Thêta prend les valeurs : $<$, \leq , $>$, \geq , \neq ou \neq

Soit la relation EMPLOYE

Nom_emp	Salaire_emp
E1	20000
E2	10000
E3	6000

Soit la relation CHEF

Nom_chef	Salaire_chef
Toto	25000
Loulou	12000

Le thêta-produit permet de répondre à la question : *Donnez le nom des employés qui gagnent plus qu'un chef de service*

On effectue d'abord un produit cartésien puis une sélection dont la condition est $\text{salaire_emp} > \text{salaire_chef}$

Algèbre relationnelle : Thêta-produit (thêta-jointure)

Produit cartésien

Nom_emp	Salaire_emp	Nom_chef	Salaire_chef
E1	20000	Toto	25000
E1	20000	Loulou	12000
E2	10000	Toto	25000
E2	10000	Loulou	12000
E3	6000	Toto	25000
E3	6000	Loulou	12000

Sélection

Nom_emp	Salaire_emp	Nom_chef	Salaire_chef
E1	20000	Loulou	12000

Algèbre relationnelle : Thêta-produit (thêta-jointure)

- ◆ Opérande consistant à rapprocher selon une condition les tuples de deux relations $R1$ et $R2$ afin de former une relation $R3$ qui contient l'ensemble de tous les tuples obtenus en concaténant un tuple de $R1$ et un de $R2$ vérifiant la condition de rapprochement.
- ◆ Notation $R1 \theta$ condition $R2$

Algèbre relationnelle : Jointure naturelle (ou jointure)

La jointure naturelle permet de réaliser une liaison logique entre deux tables. La condition de sélection est l'égalité entre les deux clés des deux relations (l'une primaire, l'autre secondaire)

C'est un thêta-produit qui prend la valeur ‘ = ’ entre des attributs identiques.

Soit la relation EMPLOYE

Num_emp	Nom_emp	Num_service
02	Toto	S1
10	Loulou	S8
72	Babette	S6
62	Riri	S1

Soit la relation SERVICE

Num_service	Nom_service
S1	Informatique
S6	Mathématiques
S8	Sociologie
S4	Anglais

Jointure

homme =

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

femme =

c1 (nom)	c2 (adr)
Cathy	Brest
Julie	Lyon
Linda	Lyon

◆ Couples homme-femme d'une même ville ?

homme $\bigtriangleup_{c2=c2}$ femme $\iff \Sigma_{c2=c4} (\text{homme} \otimes \text{femme})$

c1 (nom)	c2 (adr)	c3 (nom)	c4 (adr)
Bob	Lyon	Julie	Lyon
Bob	Lyon	Linda	Lyon

Jointure

homme =

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

ville =

c1 (nomV)	c2 (nbHab)	c3 (départ)
Nice	340.000	Alp-Mar
Brest	160.000	Finistère
Lyon	420.000	Rhône

- ◆ Dans quels départements habitent les hommes ?

$$\text{homme} \bowtie_{c2=c1} \text{ville} \Leftrightarrow \sigma_{c2 = c3} (\text{homme} \otimes \text{ville})$$

c1 (nom)	c2 (adr)	c3 (nomV)	c4 (nbHab)	c5 (départ)
Bob	Lyon	Lyon	420.000	Rhône
Sam	Nice	Nice	340.000	Alp-Mar

Jointure

{combinaison par produit cartésien et sélection}

L'ensemble des lignes L pouvant être obtenues par concaténation d'une ligne L1 d'une table T1 avec une ligne L2 d'une table T2 telle que L1(i) = L2(j) est noté :

$$T1 \underset{i=j}{\bowtie} T2 = \{ L1 \bullet L2 \mid L1 \text{ dans } T1 \text{ et } L2 \text{ dans } T2 \\ \text{et } L1(i) = L2(j) \}$$

Cette opération est appelée **jointure de T1 avec T2**

Une formalisation de la jointure

◆ Lien avec le produit cartésien et la sélection

Soit m le nombre de colonnes de T_1 ,

$$T_1 \bowtie_{i=j} T_2 \quad \text{est "identique" à } \sigma_{i=(j+m)} (T_1 \otimes T_2)$$

Remarque : comme pour la sélection, les comparateurs
 $\neq < > \leq \geq$ ainsi que les conjonctions de critères peuvent
être employés

Algèbre relationnelle : Union

Permet de fusionner deux relations en une seule.

Cette opération n'est possible que sur des relations ayant les mêmes attributs.

Soit la relation OUVRIER

Num_empl	Nom_empl
15	Loulou
17	Fifi
56	Babette

Soit la relation CADRE

Num_empl	Nom_empl
3	Jojo
21	Sophie

L'union permet de construire la relation EMPLOYE

Num_empl	Nom_empl
3	Jojo
21	Sophie
15	Loulou
17	Fifi
56	Babette

Algèbre relationnelle : Union

- ◆ Opération portant sur deux relations de même schéma $R1$ et $R2$, consistant à construire une relation de même schéma $R3$ ayant pour tuples ceux appartenant à $R1$ ou (inclusif) $R2$.
- ◆ Notation $R1 \cup R2$

Union

homme =

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

femme =

c1 (nom)	c2 (adr)
Cathy	Brest
Julie	Lyon
Linda	Lyon

- Quel est l'ensemble des personnes ?

$$\text{homme} \cup \text{femme}$$

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice
Cathy	Brest
Julie	Lyon
Linda	Lyon

- Quel est l'ensemble des personnes habitant Lyon ?

$$\sigma_{c2="Lyon"} (\text{homme} \cup \text{femme})$$

c1 (nom)	c2 (adr)
Bob	Lyon
Julie	Lyon
Linda	Lyon

Algèbre relationnelle : Différence

Permet d'obtenir les occurrences de la relation 1 qui n'appartiennent pas à la relation 2. Les deux relations doivent avoir les mêmes attributs. Cette opération n'est pas commutative.

Soit la relation INSCRITS

Nom_étu	Nom_UV
Toto	Maths
Jojo	Maths
Toto	Physique
Babette	Chimie
Jojo	Chimie

Soit la relation RECUS

Nom_étu	Nom_UV
Toto	Maths
Jojo	Maths
Babette	Chimie

La différence permet de construire la relation suivante

Num_empl	Nom_empl
Toto	Physique
Jojo	Chimie

Algèbre relationnelle : Différence

- ◆ Opération portant sur deux relations de même schéma $R1$ et $R2$, consistant à construire une relation de même schéma $R3$ ayant pour tuples ceux appartenant à $R1$ et n'appartenant pas à $R2$.
- ◆ Notation $R1 - R2$

Différence

homme=

c1 (nom)	c2 (adr)
Bob	Lyon
Sam	Nice

femme=

c1 (nom)	c2 (adr)
Cathy	Brest
Julie	Lyon
Linda	Lyon

- Quelles sont les villes pour lesquelles on connaît au moins un homme et aucune femme ?

$$\pi_{c_2}(\text{homme}) - \pi_{c_2}(\text{femme})$$

c1 (adr)
Nice

Formalisation des deux opérations

◆ Une formalisation de l'union (ensembliste)

Soient T_1 et T_2 deux tables ayant même nombre de colonnes, l'ensemble contenant les lignes de T_1 ainsi que celles de T_2 est noté $T_1 \cup T_2$

$$T_1 \cup T_2 = \{ L \mid L \text{ dans } T_1 \text{ ou } L \text{ dans } T_2 \}$$

◆ Une formalisation de la différence (ensembliste)

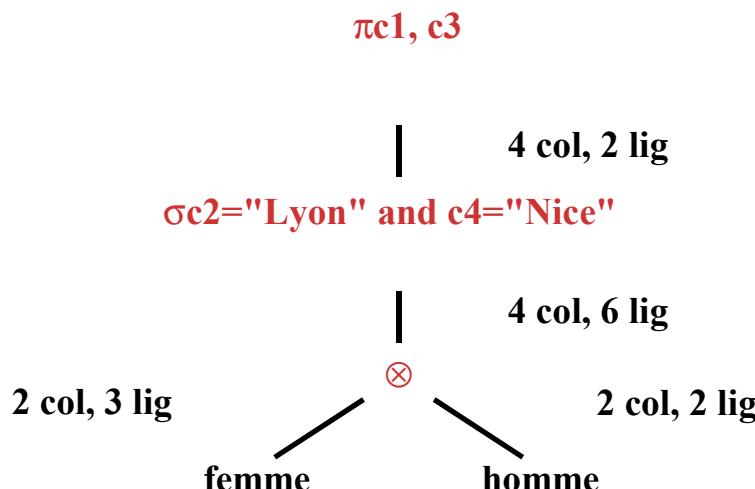
Soient T_1 et T_2 deux tables ayant même nombre de colonnes, l'ensemble contenant les lignes de T_1 ne se trouvant pas dans T_2 est noté $T_1 - T_2$

$$T_1 - T_2 = \{ L \mid L \text{ dans } T_1 \text{ et } L \text{ n'est pas dans } T_2 \}$$

Arbre et transformation de requêtes

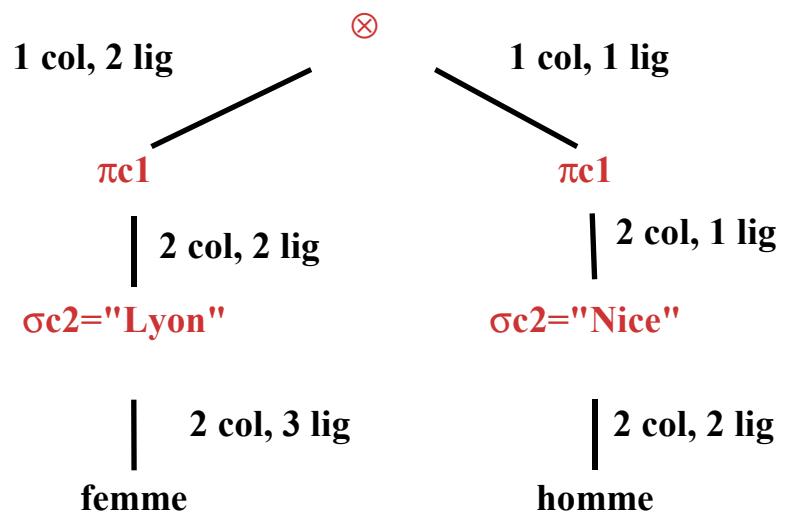
c1	c2
Julie	Sam
Linda	Sam

- Quels sont les couples possibles avec une femme de Lyon et un homme de Nice ?



$$2 \cdot 3 + 2 \cdot 2 + 4 \cdot 6 + 4 \cdot 2 =$$

42 cases manipulées



$$2 \cdot 3 + 2 \cdot 2 + 2 \cdot 2 + 2 \cdot 1 + 1 \cdot 2 + 1 \cdot 1 =$$

19 cases manipulées