

M1104 : Introduction aux bases de données

Dut Alternance
2015-2016





Chapitre 3

L'algèbre relationnelle

Modèle relationnel de données

Une Base de Données Relationnelle (BDR) peut être vue par l'utilisateur comme un ensemble de **tableaux** ou de **tables**.

- Une table est un ensemble de lignes et de colonnes

Exemple de BDR :

- fournisseur est une table contenant le numéro (nof), le nom (nomf) et la ville (ville) de chaque fournisseur
- pièce est une table contenant le numéro (nop), le nom (nomp) et le prix (prix) de chaque pièce
- vente est une table indiquant qu'une pièce (nop) est vendue par un fournisseur (nof)

fournisseur

nof	nomf	ville
1	Girard	Lyon
2	Blanc	Paris
3	Merlin	Nancy

vente

nop	nof
1	1
1	2
2	2
2	3
3	1
3	2
3	3

pièce

nop	nomp	prix
1	vis	1.5
2	écrou	2
3	boulon	2.5

Une base de données relationnelle

Formalisation du modèle relationnel de données (*théorie des ensembles*)

- Un **domaine** est un ensemble de valeurs
 - ex : L'ensemble des nombres entiers (\mathbb{Z}) ; l'ensemble des chaînes de caractères de longueur 50 ; {jaune, vert, bleu} ; {x,y,z}
- Le **produit cartésien** des domaines D_1, D_2, \dots, D_n noté $D_1 \times D_2 \times \dots \times D_n$ est l'ensemble des tuples (v_1, v_2, \dots, v_n) tel que
 - $v_i \in D_i, 1 \leq i \leq n$
 - ex : $n=2, D_1 = \{1,2\}, D_2 = \{x,y\}, D_1 \times D_2 = \{(1,x), (1,y), (2,x), (2,y)\}$
- Une **relation** ou **table** est un sous-ensemble du produit cartésien d'un ou plusieurs domaines : $R \subseteq D_1 \times D_2 \times \dots \times D_n$
 - ex : $R = \{(1,x), (1,y), (2,y)\}$

Formalisation du modèle relationnel de données

Une relation $R \subseteq D_1 \times D_2 \times \dots \times D_n$ est caractérisée par :

- Son **nom** R
- Son **arité** n
- Un ensemble d'**attributs** A_1, A_2, \dots, A_n correspondant aux composantes d'un tuple
- D_i est le domaine de l'attribut A_i c'est à dire l'ensemble des valeurs possibles de A_i ($1 \leq i \leq n$)
- Le **schéma** de la relation : $\mathbf{R} (A_1, A_2, \dots, A_n)$

Formalisation du modèle relationnel de données

- $n=2$, $D_1 = \{1,2\}$, $D_2 = \{x,y\}$, $R = \{(1,x), (1,y), (2,y)\}$
 - ✓ attribut chiffre pour la première composante
 - ✓ attribut lettre pour la seconde composante
 - ✓ schéma : $R(\text{chiffre}, \text{lettre})$
domaine(chiffre)= D_1 et domaine(lettre)= D_2
- Autre exemple de **schéma** de relation :

Pièce (nop, nomp, prix)

domaine (nop) : entiers positifs

domaine (nomp) : chaînes de caractères

domaine (prix) : réels positifs

nop	nomp	prix
1	vis	1.5
2	écrou	2
3	boulon	2.5

Formalisation du modèle relationnel de données

L'**extension** d'une relation :

- est l'ensemble de ses tuples
- est représentée par un **tableau** à deux dimensions
 - chaque **colonne** correspond à un **attribut**
 - chaque **ligne** correspond à un **tuple**

Extension possible
de la table **pièce**

nop	nomp	prix
1	vis	1.5
2	écrou	2
3	boulon	2.5

Extension d'une table (ou d'une relation)

Nom de la relation

Pièce

Nom d'attribut

nop	nomp	prix
1	Vis10	1.5
2	écrou	2
3	boulon	2.5
4	vis15	2

N-uplets, extensions

tuple

Notion de valeur **NULL**

La valeur d'un attribut dans un tuple peut être absente

- Cette absence de valeur est notée **NULL**
 - Cas de l'attribut no-téléphone pour personne n'ayant pas de téléphone
 - Cas de l'attribut prix pour une pièce dont le prix n'est pas encore fixé

Pièce

nop	nomp	prix
1	Vis10	1.5
2	écrou	2
3	boulon	2.5
4	vis15	NULL

Autres notions dans un schéma de relation :

Clé candidate et Clé primaire

■ Une **clé candidate d'une relation** : Groupe d'attributs qui identifie de manière **unique** tout tuple dans une relation

ex : {nop} dans la relation pièce

{nof} dans la relation fournisseur.

□ Clé primaire (est une clé candidate) : groupe d'attributs **minimum** permettant d'identifier de manière unique les tuples d'une relation

■ Convention : clé primaire **soulignée** dans le schéma

ex : Pièce (nop, nomp, prix)

Autres notions dans un schéma de relation :

Clé étrangère

- **Clé étrangère d'une relation** : l'attribut ou les attributs qui constituent la clé primaire d'une autre relation.
- Convention : clé étrangère précédé par # dans le schéma (suivi par #, ou * aussi)

Soit la base de données :

pièce (nop, nomp, prix)
fournisseur (nof, nomf, ville)
vente (#nop, #nof)

vente.nop est une clé étrangère et fait **référence à** pièce.nop

vente.nof est une clé étrangère et fait **référence à** fournisseur.nof

Algèbre relationnelle (A.R.)

A.R. : Langage de manipulation de données relationnelles

A.R. : huit **opérateurs** s'appliquant à une ou deux **relations** et donnant une **relation** comme résultat

- Projection
- Restriction (ou sélection)
- Produit cartésien
- Jointure
- Union
- Intersection
- Différence
- Division

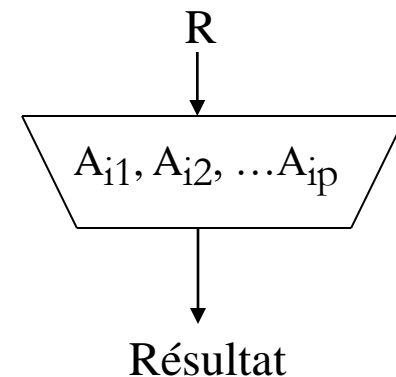
Opérateur de projection

■ Définition

La **projection** d'une relation R de schéma $R(A_1, A_2, \dots, A_n)$ sur les attributs $\{A_{i1}, A_{i2}, \dots, A_{ip}\}$ est une relation « Résultat » de schéma Résultat $(A_{i1}, A_{i2}, \dots, A_{ip})$ dont les tuples sont obtenus par élimination des attributs de R n'appartenant pas à T et par **suppression** des tuples en double.

Notation

Résultat = $\pi (R / \{A_{i1}, A_{i2}, \dots, A_{ip}\})$



Exemple de projection de relation

VINS

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12
200	Sancerre	1977	12

Resultat = π (Millésime, Degré) VINS

Millésime	Degré
1974	12
1978	13
1977	12

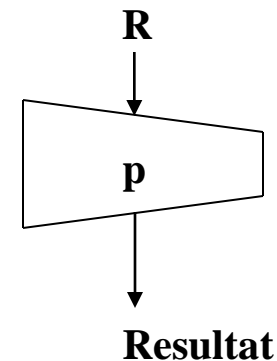
Opérateur de restriction

Définition

La restriction d'une relation R à l'aide d'un prédicat p est une relation « Resultat » de même schéma dont les n -uplets sont ceux de R satisfaisant le prédicat p .

Notation

$$\text{Resultat} = \sigma (p) R$$



Exemple de restriction de relation

VINS

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12
200	Sancerre	1977	12

VINS-8 = σ (Degré = 12) VINS

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
120	Mâcon	1977	12
200	Sancerre	1977	12

Prédicat de restriction

Le prédicat d'une restriction est une formule logique quelconque avec des connecteurs logiques **ET** (\wedge) et **OU** (\vee) entre prédicats simples de la forme $A_i \theta a$ où :

- A_i est un nom d'attribut
- a est un élément du domaine de A_i (constante)
- θ est un opérateur de comparaison ($=, <, >, \neq, \geq, \leq$)

ex. $(\text{Cru}=\text{"Chablis"} \vee \text{Cru}=\text{"Mâcon"}) \wedge \text{Millésime} < 1988$

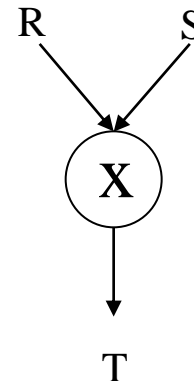
Opérateur de produit cartésien

Définition

Le produit cartésien de deux relations R et S (de schémas quelconques) est une relation T ayant pour attributs la concaténation de ceux de R et de S et dont les tuples sont toutes les combinaisons possibles d'un tuple de R et d'un tuple de S.

Notation

$$T = R \times S$$



Exemple de produit cartésien de relations

VINS

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
200	Sancerre	1979	11

VITICULTEURS

Nom	Ville	Région
Nicolas	Pouilly	Bourgogne
Martin	Bordeaux	Bordelais

Résultat = VINS x VITICULTEURS

Numéro	Cru	Millésime	Degré	Nom	Ville	Région
100	Chablis	1974	12	Nicolas	Pouilly	Bourgogne
100	Chablis	1974	12	Martin	Bordeaux	Bordelais
200	Sancerre	1979	11	Nicolas	Pouilly	Bourgogne
200	Sancerre	1979	11	Martin	Bordeaux	Bordelais

Opérateur de jointure

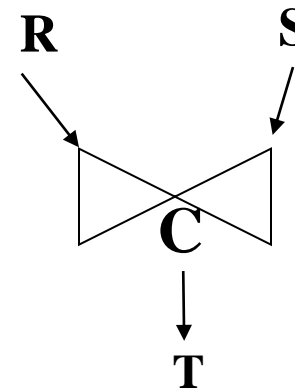
Définition

La jointure de deux relations R et S selon une condition C est l'ensemble des tuples du produit cartésien R X S satisfaisant la condition C.

Notation

$$T = R \bowtie_C S$$

Condition de jointure



Une jointure = produit cartésien + restriction (condition de jointure)

Exemple de jointure de relations

VINS

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12

VITICULTEURS

Nom	Ville	Région
Nicolas	Pouilly	Bourgogne
Félix	Mâcon	Bourgogne

Resultat = VINS  **VITICULTEURS**
Cru = Ville

Numéro	Cru	Millésime	Degré	Nom	Ville	Région
120	Mâcon	1977	12	Félix	Mâcon	Bourgogne

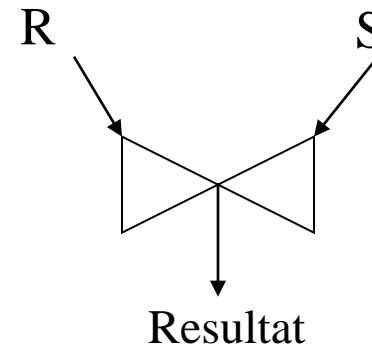
Jointure naturelle

Définition

La jointure naturelle de deux relations R et S est l'équi-jointure (opérateur d'égalité) de R et S sur tous leurs attributs communs.

Notation

$$T = R \bowtie S$$



Exemple de jointure naturelle de relations

VINS

Numéro	Cru	Millésime	Degré
150	Riesling	1984	11
110	Mecurey	1978	13
120	Mâcon	1977	12

VITIC

Nom	Numéro	Région
Nicolas	150	Alsace
Félix	120	Bourgogne

VINS-C = VINS  **VITIC**

Numéro	Cru	Millésime	Degré	Nom	Région
150	Riesling	1984	11	Nicolas	Alsace
120	Mâcon	1977	12	Félix	Bourgogne

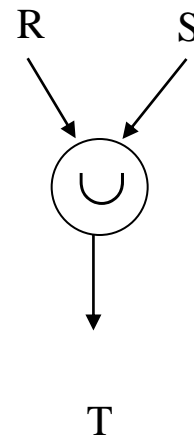
Opérateur d'union

Définition

L'**union** de deux relations R et S **de même schéma** est une relation T de même schéma contenant l'ensemble des tuples de R et S.

Notation

$$T = R \cup S$$



$$\text{VINS-3} = \text{VINS-1} \cup \text{VINS-2}$$

Exemple d'union de relations

VINS-1

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12

VINS-2

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
200	Sancerre	1979	11

VINS-3

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12
200	Sancerre	1979	11

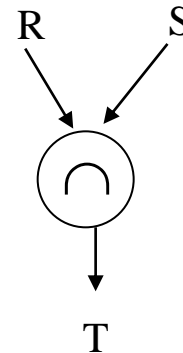
Opérateur d'intersection

Définition

L'**intersection** de deux relations R et S de **même schéma** est une relation T de même schéma contenant l'ensemble des tuples appartenant simultanément à R et à S.

Notation

$$T = R \cap S$$



$$\text{VINS-4} = \text{VINS-1} \cap \text{VINS-2}$$

Exemple d'intersection de relations

VINS-1

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12

VINS-2

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
200	Sancerre	1979	11

VINS-4

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12

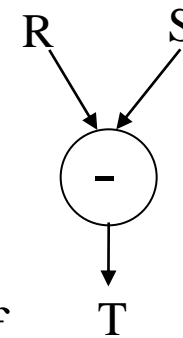
Opérateur de différence

Définition

La **différence** de deux relations R et S de **même schéma** est une relation T de même schéma contenant l'ensemble des tuples appartenant à R et n'appartenant pas à S.

Notation

$$T = R - S$$



Remarque : l'opérateur de différence n'est pas commutatif
 $R - S \neq S - R$

$$\text{VINS-5} = \text{VINS-1} - \text{VINS-2}$$

Exemple de différence de relations

VINS-1

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
110	Mecurey	1978	13
120	Mâcon	1977	12

VINS-2

Numéro	Cru	Millésime	Degré
100	Chablis	1974	12
200	Sancerre	1979	11

VINS-5

Numéro	Cru	Millésime	Degré
110	Mecurey	1978	13
120	Mâcon	1977	12

Exemples de requêtes algébriques (1/3)

Soit la base de données :

pièce (nop, nomp, prix)

fournisseur (nof, nomf, ville)

vente (#nop, #nof)

Exemple de question : Trouver le **numéro** des fournisseurs vendant le produit **boulon**.

Requête algébrique :

Exemples de requêtes algébriques (2/3)

Base de données VPC (Vente par Correspondance)

Client (numCli, nom, prénom, rue, CP, ville)

Produit (numProd, libellé, pu)

Commande (#numCli, #numProd, date, quantité)

Exemple de question : Trouver les produits commandés en quantité supérieure à 100 et dont le prix dépasse 1000€ . On affichera les numéros de produit, leur libellé et leur prix unitaire ainsi que la date de la commande.

Notons **Res** la relation résultat et R_1, R_2, R_3 des relations intermédiaires

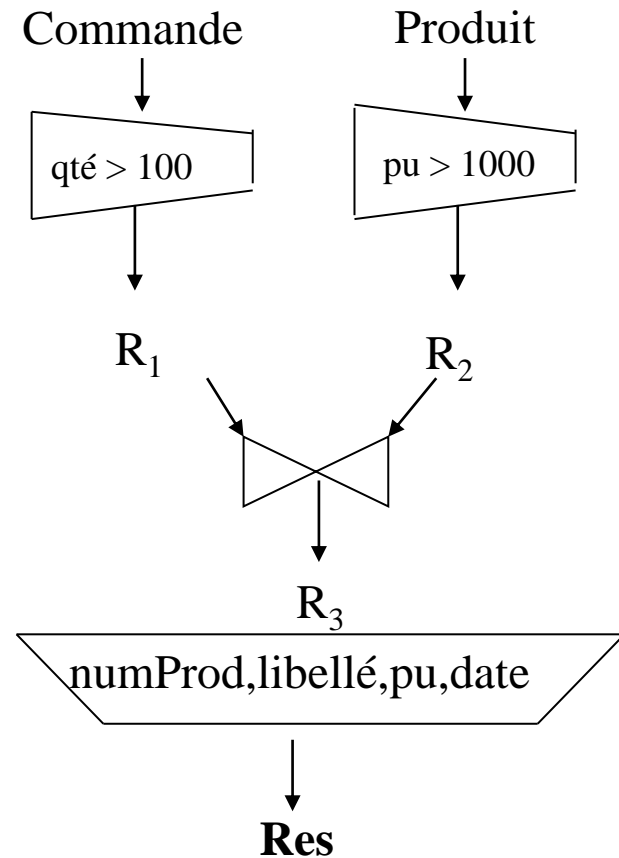
Exemple de requête algébrique (3/3)

$$R_1 = \sigma (\text{Produit} / \text{pu} > 1000)$$

$$R_2 = \sigma (\text{Commande} / \text{qté} > 100)$$

$$R_3 = R_1 \bowtie R_2$$

$$\text{Res} = \pi (\text{R3} / \{\text{numProd}, \text{libellé}, \text{pu}, \text{date}\})$$



A vous : construire les requêtes algébriques répondant aux questions (BD VPC)

- 1) Libellé et prix unitaire de tous les produits
- 2) Libellé des produits de prix inférieur à 50€
- 3) Nom et prénom des clients ayant commandé le produits
numéro 56.
- 4) Nom des clients n'ayant pas commandé le produit numéro 56.