

Le langage algébrique

HLIN511

Pascal Poncelet

Pascal.Poncelet@umontpellier.fr

<http://www.lirmm.fr/~poncelet>



Introduction

- Introduit par CODD en 1970
- Traitement de requêtes de lecture et écriture
- Deux types d'opérateurs
 - Opérateurs ensemblistes : UNION, INTERSECTION, DIFFERENCE, PRODUIT CARTESIEN
 - Opérateurs relationnels : SELECTION, PROJECTION, JOINTURE et DIVISION
- Notation : t un tuple d'une relation et $t(A)$ dans R , le sous tuple de R relatif à A



Opérateurs ensemblistes

- Pour l'union, l'intersection, la différence les relations doivent être unicompatibles :
 - les relations doivent avoir même degré
 - les attributs associés deux à deux doivent être du même type syntaxique



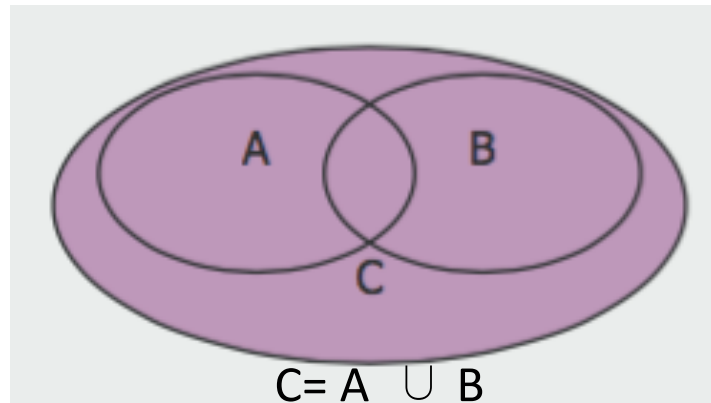
UNION

- UNION

$R \cup S$: ensemble des tuples qui appartiennent soit à R , soit à S , soit à R et S :

$$R \cup S = \{t/t \in R \text{ OU } t \in S\}$$

- Opérateur commutatif ($R \cup S = S \cup R$)



Exemple

Pilote1 : ensemble des pilotes habitant PARIS

PILOTE1	PLNUM	ADR
	100	PARIS
	101	PARIS
	120	PARIS
	110	PARIS

Pilote2 : ensemble des pilotes assurant un vol au départ de PARIS de TOULOUSE

PILOTE2	PLNUM	VD
	130	TOUL
	140	TOUL
	150	TOUL
	100	TOUL
	120	TOUL
	130	PARIS
	101	PARIS
	140	PARIS
	110	PARIS



Pilote1 UNION pilote 2

PILOTE1 UNION
PILOTE2

Ensemble des pilotes
habitant PARIS ou
assurant
un vol au départ de
PARIS ou TOULOUSE

PLNUM	VILLE
100	PARIS
101	PARIS
120	PARIS
110	PARIS
130	TOUL
140	TOUL
150	TOUL
100	TOUL
120	TOUL
130	TOUL
130	PARIS
140	PARIS

Pilote1

Pilote2

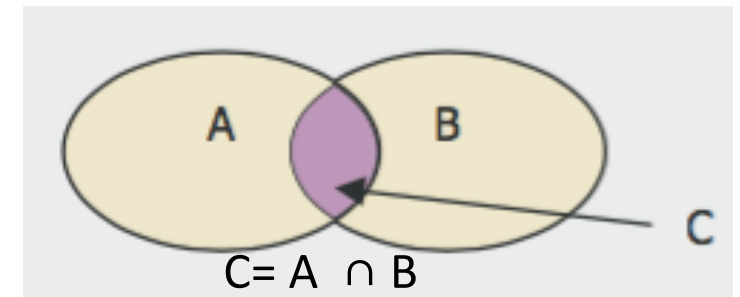
Les duplicats sont éliminés

L'Union permet de faire de l'ajout de tuples



INTERSECTION

- $R \cap S$: ensemble des tuples qui appartiennent à R et à S :
$$R \cap S = \{t / t \in R \text{ et } t \in S\}$$
- Opérateur commutatif : $R \cap S = S \cap R$



Avec Pilote1 INTERSECTION Pilote2

**PILOTE1
INTERSECTION
PILOTE2**

PLNUM	VILLE
101	PARIS
110	PARIS

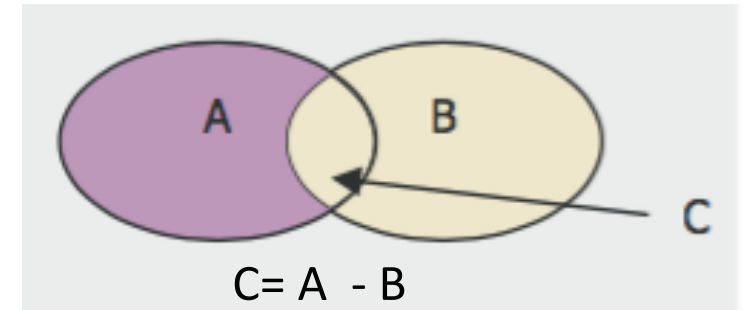


L'intersection permet de traduire le ET logique

DIFFERENCE

- $R - S$: ensemble des tuples qui appartiennent à R sans appartenir à S . Complémentaire de l'intersection :

$$R - S = \{t / t \in R \text{ ET } t \notin S\}$$



- Opérateur non commutatif : $R - S \neq S - R$

PILOTE1 – PILOTE2	PLNUM	ADR
	100	PARIS
	120	PARIS

PILOTE 1 – PILOTE 2 : ensemble des pilotes habitant PARIS et n'assurant pas de vol au départ de PARIS ou TOULOUSE

PRODUIT CARTESIEN

- $R \otimes S$: ensemble de tous les tuples obtenus par concaténation des tuples de R et de S . Ensemble de paires ordonnées :

$$R \otimes S = \{(t(r), t(s)) \text{ avec } t(r) \in R \text{ et } t(s) \in S\}$$

- Opérateur commutatif : $R \otimes S = S \otimes R$

PILOTE3

PLNUM	ADR
103	NICE
106	NICE

AVION

AVNOM	CAP
AIRBUS	350
CARAV	250

Produit
Cartésien

PLNUM	ADR	AVNOM	CAP
103	NICE	AIRBUS	350
103	NICE	CARAV	250
106	NICE	AIRBUS	350
106	NICE	CARAV	250



Opérateurs relationnels

- Deux catégories
 - unaires de restriction
 - binaires d'extension
- Les opérateurs de restrictions permettent :
 - soit un découpage horizontal d'une relation (SELECTION)
 - soit un découpage vertical d'une relation (PROJECTION)



SELECTION

- Soit θ un comparateur binaire $\theta = \{<, <=, >, >=, <>\}$ applicable à l'attribut (ou à l'ensemble d'attributs) A et au tuple c (de la relation R)
- La sélection $R(A \theta c)$ est l'ensemble des tuples de R pour lesquels θ est vérifié entre la (les) composante(s) A et le tuple c :

$$R(A \theta c) = \{t / t \in R \text{ ET } tA \theta c\}$$



SELECTION

PILOTE4

PLNUM	PLNOM	ADR
100	JEAN	PARIS
101	PIERRE	PARIS
120	PAUL	PARIS
130	SERGE	TOUL
140	MICHEL	TOUL

R

PLNUM	PLNOM	ADR
130	SERGE	TOUL
140	MICHEL	TOUL

R = PILOTE4 (ADR = « TOUL »)



SELECTION - Notations

$R = \text{PILOTE4 (ADR = « TOUL »)}$

$R = \text{SELECTION(PILOTE4/ADR = « TOUL »)}$

Notation sigma : $\sigma_Q(R)$ où Q est le critère de la forme :
Ai θ Valeur

$R = \sigma_{\text{ADR=« TOUL »}}(\text{PILOTE4})$

- Toutes les notations sont équivalentes mais il ne faut pas les mélanger



SELECTION - remarques

- Il est possible d'utiliser des opérateurs logiques :
ET, OU, NON

$R = \text{PILOTE4}(\text{ADR}=\text{« TOUL » ET NOM}=\text{« MICHEL »})$

$R = \sigma_{(\text{ADR}=\text{« TOUL » ET NOM}=\text{« MICHEL »})}(\text{PILOTE4})$

- Pas indispensable pour le moment



PROJECTION

- Soit $R(A)$ une relation et un ensemble d'attributs A_1, \dots, A_n de R tels que $(A_1, A_2, \dots, A_n) \subseteq A$
- La projection $R' (A_1, A_2, \dots, A_n)$ est la relation obtenue à partir de $R(A)$ en éliminant de $R(A)$ les attributs autres que ceux spécifiés par A_1, A_2, \dots, A_n

$$R (A_1, \dots, A_n) = \{t(a_1, \dots, a_n)\}$$

- Suppression des tuples dupliqués



PROJECTION

Avion1

AVNUM	AVNOM	CAP	LOC
100	AIRBUS	350	TOUL
101	AIRBUS	350	TOUL
104	AIRBUS	150	PARIS
105	CARAV	250	PARIS

R

AVNOM	CAP
AIRBUS	350
AIRBUS	350
AIRBUS	150
CARAV	250

$R = \text{AVION1}(\text{AVNOM}, \text{CAP})$



PROJECTION - Notations

$R = \text{AVION1 (AVNOM, CAP)}$

$R = \text{PROJECTION(AVION1/AVNOM, CAP)}$

Notation sigma : $\pi_{A1,A2,\dots,Ap}(R)$

$R = : \pi_{\text{AVNOM,CAP}}(\text{AVION1})$

- Toutes les notations sont équivalentes mais il ne faut pas les mélanger



Opérateurs binaires d'extension

- L'opérateur JOIN et DIVISION
- JOIN : permettre de pouvoir relier des relations entre elles – Attention à la sémantique des requêtes
- DIVISION : Opérateur qui permet de sélectionner les tuples d'une relation (dividende) qui satisfont un critère de couverture énoncé via le contenu d'une autre relation (diviseur). Le résultat est une troisième relation, appelée le quotient. Exprime le « tous les »



JOIN

- Soient les relations $R(A, B_1)$ et $S(B_2, C)$ avec B_1 et B_2 attributs définis sur le même domaine, soit $\theta = \{=, >, >=, <, <=, <>\}$ applicables aux valeurs des attributs B_1 et B_2
- Le JOIN de R sur B_1 avec S sur B_2 est la relation dont les tuples sont ceux obtenus par concaténation des tuples de R avec ceux de S pour lesquels la relation θ entre les composantes B_1 et B_2 est vérifiée :

$$R \Join_{\theta} S = \{t / t \in R \otimes S \text{ ET } t(B_1) \theta t(B_2)\}$$

- L'opérateur JOIN est équivalent à un produit cartésien suivi d'une sélection



JOIN - Vocabulaire

$$R \theta (B_1 \theta B_2) S = \{t / t \in R \otimes S \text{ ET } t(B_1) \theta (B_2)\}$$

- Lorsque $\theta = \{=\}$ on parle d'équijointure autrement de thétajointure

$$R (B_1 = B_2) S = \{t / t \in R \otimes S \text{ ET } t(B_1) \theta (B_2)\}$$

- Il est possible d'avoir des autojointures

$$R \theta (B_1 \theta B_2) R = \{t / t \in R \otimes R \text{ ET } t(B_1) \theta (B_2)\}$$



JOIN - Exemple

PILOTE1

PLNUM	PLNOM	ADR
100	JEAN	PARIS
101	PIERRE	PARIS
120	PAUL	PARIS

VOL1

VOLNUM	AVNUM	PLNUM
IT500	110	100
IT501	130	100
IT503	110	100
IT504	110	120
IT506	120	120
IT507	130	110

JOIN

PILOTE1 (PLNUM=PLNUM) VOL1

- Ensemble des pilotes habitant PARIS en service avec les numéros des vols et des avions correspondants

PLNUM	PLNOM	ADR	VOLNUM	AVNUM	PLNUM
100	JEAN	PARIS	IT500	110	100
100	JEAN	PARIS	IT501	130	100
100	JEAN	PARIS	IT503	110	100
120	PAUL	PARIS	IT504	110	120
120	PAUL	PARIS	IT506	120	120



JOIN

PILOTE1 (PLNUM>PLNUM) VOL1

- Quels sont les pilotes les pilotes dont le numéro est supérieur à au moins un numéro de pilote dans vol (qui effectue un vol)

PLNUM	PLNOM	ADR	VOLNUM	AVNUM	PLNUM
101	PIERRE	PARIS	IT500	110	100
101	PIERRE	PARIS	IT501	130	100
101	PIERRE	PARIS	IT503	110	100
120	PAUL	PARIS	IT500	110	100
120	PAUL	PARIS	IT501	130	100
120	PAUL	PARIS	IT503	110	100
120	PAUL	PARIS	IT507	130	110

} 101>100

} 120>100

} 120 > 110

JOIN - Notations

RES = PILOTE1 (PLNUM=PLNUM) VOL1

RES = JOINTURE (PILOTE1, VOL1 / PLNUM = PLNUM)

Notation sigma : RES = R \bowtie_{θ} S

RES = PILOTE1 \bowtie VOL1
PLNUM=PLNUM



JOINTURE LEFT - RIGHT

- Cette partie sera vue lors du cours sur SQL
- Ce qu'il faut retenir surtout c'est que la jointure est l'un des opérateurs les plus importants
- Il faut bien comprendre ce qu'est la jointure



DIVISION

- Utiliser souvent pour exprimer le « tous les »
- Division d'une relation binaire par une relation unaire

$$R(A_1 \div A_2) S = \{t/t \in R[B] \text{ ET } (\{t\} \otimes S) \subseteq R\}$$

avec $R(B, A_1)$ et $S(A_2)$

- La division de R par S est le sous-ensemble des éléments de $R(B)$ dont le produit cartésien avec S est inclus dans R



DIVISION

- Avions conduits par tous les pilotes :

VOL1 (PLNUM ÷ PLNUM) PIL ?

VOL1

Dividende

AVNUM	PLNUM
30	100
30	101
30	102
30	103
31	100
31	102
32	102
32	103
33	102

Diviseur PIL1

PLNUM
100



AVNUM
30
31

Quotient

Diviseur PIL2

PLNUM
102
103



AVNUM
30
32

Quotient



DIVISION

- Chaque fois que l'on aura « pour tous les x », il suffira de mettre l'attribut x dans le diviseur. Le dividende binaire doit contenir alors le même attribut (sur lequel porte la division) et l'attribut du résultat recherché

DIVISION

- Quels sont les noms des pilotes qui conduisent tous les avions de la compagnie ?

$VOL1 = VOL (AVNUM, PLNUM) <projection>$

$AV1 = AVION (AVNUM) <projection>$

$PILOTE1 = PILOTE (PLNUM, PLNOM) <projection>$

$PILOTE2 = VOL1 (AVNUM \div AVNUM) AV1 <division \text{ pour avoir les numéros des pilotes qui conduisent tous les avions}>$

$PILOTE3 = PILOTE1 (PLNUM=PLNUM)PILOTE2 <join \text{ pour avoir les noms}>$

$RES = PILOTE3(PLNOM) <projection>$



CE QU'IL FAUT RETENIR

- L'ALGEBRE RELATIONNELLE EST COMPLETE
 - Les cinq (sept) opérations de base permettent de formaliser sous forme d'expressions toutes les questions que l'on peut poser avec la logique du premier ordre (sans fonction)
 - Le résultat de l'application d'un opérateur donne une relation
 - Possibilité d'impliquer les opérateurs directement
- NOM ET PRENOM DES BUVEURS DE VOLNAY 1988

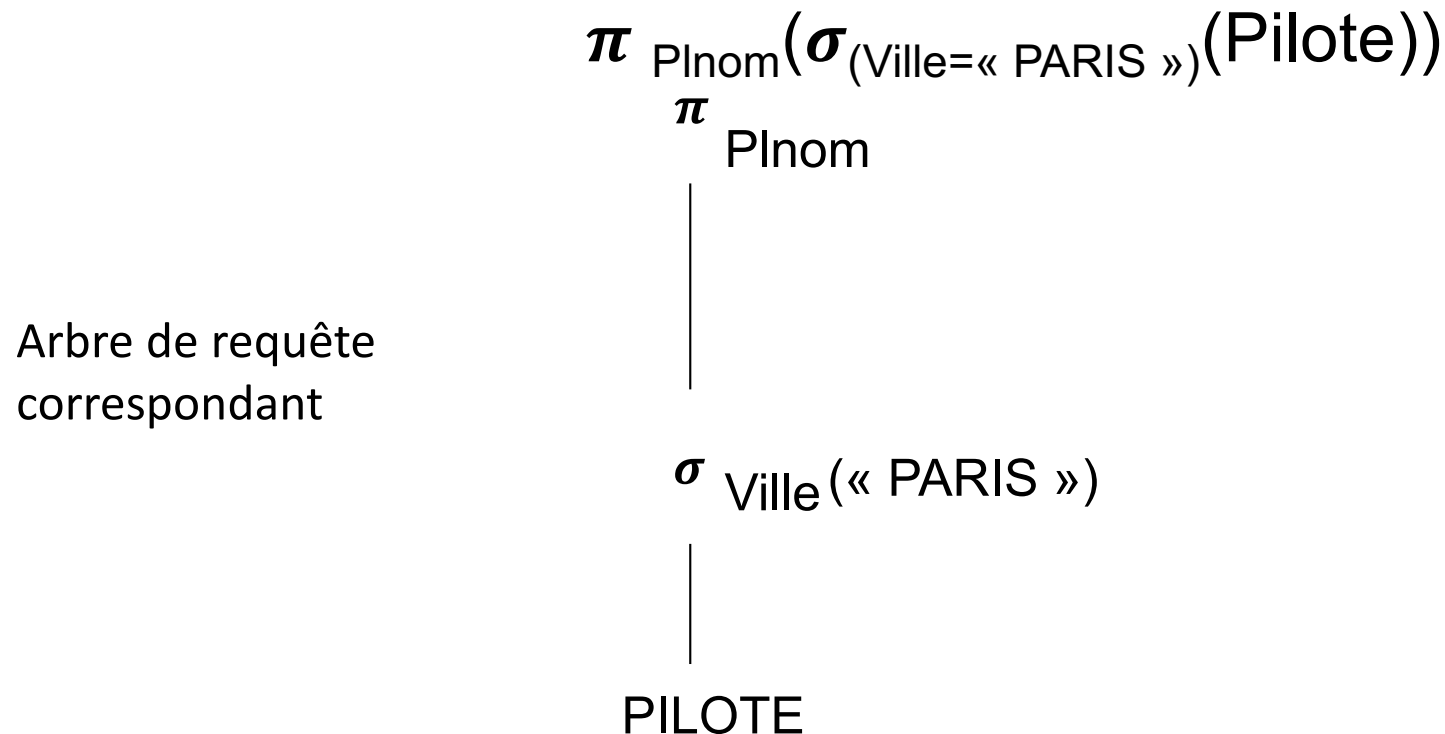
PROJECTION (NOM,PRENOM, SELECTION(CRU="VOLNAY" ET
MILL =1988, JOIN(VINS,ABUS,BUVEURS)))



ARBRE DE REQUETES

- Pour chaque requête un arbre de requête est créé. Il permet notamment de faire de l'optimisation de requêtes

Liste des noms des pilotes Parisiens :



ARBRE DE REQUETES

- Intuitivement une requête est coûteuse : remonter les opérations de sélection et de sélection
- Une jointure est un produit cartésien suivi d'une sélection !



-
- Des questions ?

