

Inscription moodle: par groupe

Clé d'auto-inscription

BDR_2016_Gx

Remplacez x par votre numéro de groupe

Plan

- Opérateurs relationnels unaires: SELECT et PROJECT
- Opérations d'algèbre relationnelle venant de la théorie des ensembles
- Opérateurs relationnels binaires: JOIN et DIVISION
- Opérateurs relationnels supplémentaires
- Exemples de requêtes en algèbre relationnelle

Algèbre relationnelle

- **Algèbre relationnelle**
 - Ensemble d'opérations de base pour le modèle relationnel
- **Expression d'algèbre relationnelle**
 - Sequence d'opérations en algèbre relationnelle

Exemple

- Une société et ses employés
- Pour chaque employé, nous connaissons
 - nom
 - numéro de sécu
 - date de naissance, sexe
 - adresse personnelle, revenu
 - numéro de son département dans la société
 - son supérieur
 - qqs informations sur sa famille

Addison-Wesley
is an imprint of

PEARSON

Copyright © 2011 Ramez Elmasri and Shamkant Navathe

Figure 3.6

One possible database state for the COMPANY relational database schema.

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
John	B	Smith	123456789	1965-01-09	731 Fondren, Houston, TX	M	30000	333445555	5
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Alicia	J	Zelaya	999887777	1968-01-19	3321 Castle, Spring, TX	F	25000	987654321	4
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5
Joyce	A	English	453453453	1972-07-31	5631 Rice, Houston, TX	F	25000	333445555	5
Ahmad	V	Jabbar	987987987	1969-03-29	980 Dallas, Houston, TX	M	25000	987654321	4
James	E	Borg	888665555	1937-11-10	450 Stone, Houston, TX	M	55000	NULL	1

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
Research	5	333445555	1988-05-22
Administration	4	987654321	1995-01-01
Headquarters	1	888665555	1981-06-19

DEPT_LOCATIONS

Dnumber	Dlocation
1	Houston
4	Stafford
5	Bellaire
5	Sugarland
5	Houston

Addison-Wesley
is an imprint of

PEARSON

Continued next page...

Copyright © 2011 Pearson Education, Inc. Publishing as Pearson Addison-Wesley

EMPLOYEE

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
-------	-------	-------	-----	-------	---------	-----	--------	-----------	-----

DEPARTMENT

Dname	Dnumber	Mgr_ssn	Mgr_start_date
-------	---------	---------	----------------

DEPT_LOCATIONS

Dnumber	Dlocation
---------	-----------

PROJECT

Pname	Pnumber	Plocation	Dnum
-------	---------	-----------	------

WORKS_ON

Essn	Pno	Hours
------	-----	-------

DEPENDENT

Essn	Dependent_name	Sex	Bdate	Relationship
------	----------------	-----	-------	--------------

Figure 3.5

Schema diagram for the COMPANY relational database schema.

Addison-Wesley
is an imprint of

PEARSON

Copyright © 2011 Pearson Education, Inc. Publishing as Pearson Addison-Wesley

Opérations relationnelles unaires: SELECT et PROJECT

• L'opération SELECT

Sous-ensemble des tuples d'une relation satisfaisant une condition de selection:

$$\sigma_{\langle \text{selection condition} \rangle}(R)$$

- Expression booléenne contenant des clauses de la forme

– <nom d'attribut> <op comparaison> <valeur constante>

ou

– <nom d'attribut> <op comparaison> <nom d'attribut>

Addison-Wesley
is an imprint of

PEARSON

Addison-Wesley
is an imprint of

PEARSON

Copyright © 2011 Ramez Elmasri and Shamkant Navathe

Opération SELECT (cont'd.)

Exemple:

$\sigma_{(Dno=4 \text{ AND } Salary > 25000) \text{ OR } (Dno=5 \text{ AND } Salary > 30000)}(EMPLOYEE)$

- <condition de selection> appliquée indépendamment à chaque tuple t de R
 - Si l'évaluation rend TRUE, le tuple est sélectionné
- Conditions booléennes **AND**, **OR**, and **NOT**

Opération SELECT (cont'd.)

Selectivité

– Proportion des tuples sélectionnés par une condition de sélection

- Cascade** d'opérations SELECT dans une seule opération avec une condition **AND**
- L'opération SELECT est commutative

L'opération PROJECT

- Selectionne certaines colonnes d'une table, et jète les autres colonnes:

$\pi_{\langle \text{attribute list} \rangle}(R)$

Elimination de doublons

- Le résultat d'une opération PROJECT est un ensemble de tuples distincts l'un de l'autre

Figure 6.1

Results of SELECT and PROJECT operations. (a) $\sigma_{(Dno=4 \text{ AND } Salary > 25000) \text{ OR } (Dno=5 \text{ AND } Salary > 30000)}(EMPLOYEE)$. (b) $\pi_{\langle Lname, Fname, Salary \rangle}(EMPLOYEE)$. (c) $\pi_{\langle Sex, Salary \rangle}(EMPLOYEE)$.

(a)

Fname	Minit	Lname	Ssn	Bdate	Address	Sex	Salary	Super_ssn	Dno
Franklin	T	Wong	333445555	1955-12-08	638 Voss, Houston, TX	M	40000	888665555	5
Jennifer	S	Wallace	987654321	1941-06-20	291 Berry, Bellaire, TX	F	43000	888665555	4
Ramesh	K	Narayan	666884444	1962-09-15	975 Fire Oak, Humble, TX	M	38000	333445555	5

(b)

Lname	Fname	Salary
Smith	John	30000
Wong	Franklin	40000
Zelaya	Alicia	25000
Wallace	Jennifer	43000
Narayan	Ramesh	38000
English	Joyce	25000
Jabbar	Ahmad	25000
Borg	James	55000

(c)

Sex	Salary
M	30000
M	40000
F	25000
F	43000
M	38000
M	25000
M	55000

Séquences d'opérations et l'opération RENAME

- Expression imbriquée:

$\pi_{\text{Fname, Lname, Salary}}(\sigma_{\text{Dno}=5}(\text{EMPLOYEE}))$

- Séquence d'opérations:

$\text{DEP5_EMPS} \leftarrow \sigma_{\text{Dno}=5}(\text{EMPLOYEE})$
 $\text{RESULT} \leftarrow \pi_{\text{Fname, Lname, Salary}}(\text{DEP5_EMPS})$

- **Renommer** des attributs de résultats intermédiaires

$\rho_{S(B_1, B_2, \dots, B_n)}(R)$ or $\rho_S(R)$ or $\rho_{(B_1, B_2, \dots, B_n)}(R)$

Opérations ensemblistes en RA

- **UNION, INTERSECTION et MINUS**

- Opérations pour combiner les éléments de deux ensembles de différentes manières
- Les relations doivent avoir le même type de tuples (union compatibilité)

UNION

- $R \cup S$
- Comprend tous les tuples qui sont soit dans R, soit dans S, ou dans aussi bien R que S
- Les doublons sont éliminés

Opérations ensemblistes en RA (2/2)

- **INTERSECTION**

- $R \cap S$
- Comprend les tuples qui sont aussi bien dans R que dans S

- **DIFFERENCE d'ENSEMBLES (ou MINUS)**

- $R - S$
- Les tuples qui sont dans R, mais pas dans S

L'opération de produit cartésien

- **CARTESIAN PRODUCT**

- **CROSS PRODUCT** or **CROSS JOIN**
- Denoté par $R \times S$
- Opération ensembliste binaire
- Les relations n'ont pas besoin d'être union compatible
- Schéma du resultat: concaténation des schémas de R et S
- Utile quand suivi d'une selection qui compare des valeurs d'attributs

Opérations binaires: JOIN and DIVISION

▪ L'opération **JOIN**

- Dénotee $R \bowtie S$
- Combine des tuples de deux relations ayant un certain rapport, en un seul tuple "plus long"
- Condition de jointure générale de la forme $\langle \text{condition} \rangle$ **AND** $\langle \text{condition} \rangle$ **AND...AND** $\langle \text{condition} \rangle$
- Exemple:

$\text{DEPT_MGR} \leftarrow \text{DEPARTMENT} \bowtie_{\text{Mgr_ssn}=\text{Ssn}} \text{EMPLOYEE}$
 $\text{RESULT} \leftarrow \pi_{\text{Dname, Lname, Fname}}(\text{DEPT_MGR})$

JOIN (cont'd.)

▪ **THETA JOIN**

- Chaque $\langle \text{condition} \rangle$ est de la forme $A_i \theta B_j$
- A_i est un attribut de R
- B_j est un attribut de S
- A_i ont B_j le même domaine
- θ (theta) est un opérateur de comparaison:
 - $\{=, <, \leq, >, \geq, \neq\}$

Variantes de JOIN: EQUIJOIN et NATURAL JOIN

▪ **EQUIJOIN**

- Uniquement l'opérateur de comparaison =
- Dans chaque tuple du résultat, on a une (ou n) paire(s) d'attributs avec la même valeur

▪ **NATURAL JOIN**

- Dénote $R * S$
- Equijoin des *attributs communs* de R et S
- Supprime le second attribut, qui est superflu puisque répété, du résultat d'un EQUIJOIN

Variantes de JOIN: EQUIJOIN et NATURAL JOIN (cont'd.)

▪ **Sélectivité d'une jointure**

- Taille de la jointure attendue, divisée par la taille maximale $n_R * n_S$

▪ **Jointure intérieures**

- Type des opérations de comparaison et combinaison
- Défini formellement par une combinaison de CARTESIAN PRODUCT et SELECTION

Ensemble complet d'opérations d'Algèbre relationnelle

- L'ensemble des opérateurs d'algèbre relationnelle $\{\sigma, \pi, \cup, \rho, -, \times\}$ est **complet**
 - N'importe quelle opération d'algèbre relationnelle peut être exprimée avec ces opérateurs

Operations d'Algèbre relationnelle

Table 6.1 Operations of Relational Algebra

OPERATION	PURPOSE	NOTATION
SELECT	Selects all tuples that satisfy the selection condition from a relation R .	$\sigma_{\langle \text{selection condition} \rangle}(R)$
PROJECT	Produces a new relation with only some of the attributes of R , and removes duplicate tuples.	$\pi_{\langle \text{attribute list} \rangle}(R)$
THETA JOIN	Produces all combinations of tuples from R_1 and R_2 that satisfy the join condition.	$R_1 \bowtie_{\langle \text{join condition} \rangle} R_2$
EQUIJOIN	Produces all the combinations of tuples from R_1 and R_2 that satisfy a join condition with only equality comparisons.	$R_1 \bowtie_{\langle \text{join condition} \rangle} R_2$, OR $R_1 \bowtie_{(\langle \text{join attributes 1} \rangle), (\langle \text{join attributes 2} \rangle)} R_2$
NATURAL JOIN	Same as EQUIJOIN except that the join attributes of R_2 are not included in the resulting relation; if the join attributes have the same names, they do not have to be specified at all.	$R_1 \star_{\langle \text{join condition} \rangle} R_2$, OR $R_1 \star_{(\langle \text{join attributes 1} \rangle), (\langle \text{join attributes 2} \rangle)} R_2$

Opérations d'Algèbre relationnelle (cont'd.)

Table 6.1 Operations of Relational Algebra

UNION	Produces a relation that includes all the tuples in R_1 or R_2 or both R_1 and R_2 ; R_1 and R_2 must be union compatible.	$R_1 \cup R_2$
INTERSECTION	Produces a relation that includes all the tuples in both R_1 and R_2 ; R_1 and R_2 must be union compatible.	$R_1 \cap R_2$
DIFFERENCE	Produces a relation that includes all the tuples in R_1 that are not in R_2 ; R_1 and R_2 must be union compatible.	$R_1 - R_2$
CARTESIAN PRODUCT	Produces a relation that has the attributes of R_1 and R_2 and includes as tuples all possible combinations of tuples from R_1 and R_2 .	$R_1 \times R_2$
DIVISION	Produces a relation $R(X)$ that includes all tuples $t[X]$ in $R_1(Z)$ that appear in R_1 in combination with every tuple from $R_2(Y)$, where $Z = X \cup Y$.	$R_1(Z) \div R_2(Y)$

Exemples de requêtes en Algèbre relationnelle

Query 1. Retrieve the name and address of all employees who work for the 'Research' department.

```
RESEARCH_DEPT ← σDname='Research'(DEPARTMENT)
RESEARCH_EMPS ← (RESEARCH_DEPT ⋈Dnumber=Dno EMPLOYEE)
RESULT ← πFname, Lname, Address(RESEARCH_EMPS)
```

As a single in-line expression, this query becomes:

```
πFname, Lname, Address(σDname='Research'(DEPARTMENT ⋈Dnumber=Dno(EMPLOYEE)))
```

Exemples de requêtes en Algèbre relationnelle (cont'd.)

Query 2. For every project located in 'Stafford', list the project number, the controlling department number, and the department manager's last name, address, and birth date.

```
STAFFORD_PROJS ←  $\sigma_{Plocation='Stafford'}$ (PROJECT)
CONTR_DEPTS ← (STAFFORD_PROJS  $\bowtie_{Dnum=Dnumber}$  DEPARTMENT)
PROJ_DEPT_MGRS ← (CONTR_DEPTS  $\bowtie_{Mgr\_ssn=Ssn}$  EMPLOYEE)
RESULT ←  $\pi_{Pnumber, Dnum, Lname, Address, Bdate}$ (PROJ_DEPT_MGRS)
```

Notation pour les arbres de requêtes

■ Arbre de requête

- Les noeuds feuilles de l'arbre représentent les relations en entrée de la requête
- Les opérations d'algèbre relationnelle sont les noeuds intérieurs de l'arbre

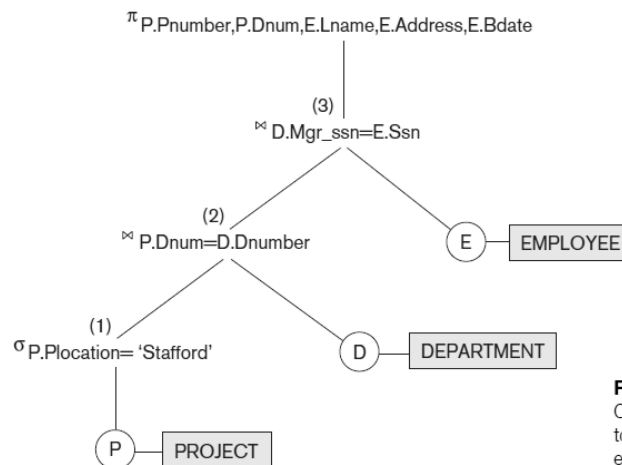


Figure 6.9
Query tree corresponding
to the relational algebra
expression for Q2.

Opérations relationnelles supplémentaires

■ Projection généralisée

- Permet d'appliquer des fonctions aux attributs de la liste de projection

$$\pi_{F_1, F_2, \dots, F_n}(R)$$

■ Fonctions d'aggrégation et groupage

- Des fonctions bien connues, appliquées à des ensembles de valeurs numériques
- SUM, AVERAGE, MAXIMUM, and MINIMUM

Opérations relationnelles supplémentaires (cont'd.)

- Grouper les tuples par la valeur d'un de leurs attributs

- Appliquer la fonction d'aggrégation indépendamment à chaque groupe

<grouping attributes> \mathcal{F} <function list> (R)

Figure 6.10

The aggregate function operation.

- $\rho_{R(Dno, No_of_employees, Average_sal)}(Dno \mathcal{F} COUNT Ssn, AVERAGE Salary(EMPLOYEE))$.
- $\rho_{Dno} \mathcal{F} COUNT Ssn, AVERAGE Salary(EMPLOYEE)$.
- $\mathcal{F} COUNT Ssn, AVERAGE Salary(EMPLOYEE)$.

R

(a)

Dno	No_of_employees	Average_sal
5	4	33250
4	3	31000
1	1	55000

(b)

Dno	Count_ssn	Average_salary
5	4	33250
4	3	31000
1	1	55000

(c)

Count_ssn	Average_salary
8	35125

⁸Note that this is an arbitrary notation we are suggesting. There is no standard notation.

Opération OUTER JOIN

- **Jointure externe**

- Garder tous les tuples de R, tous ceux de S, ou des deux relations, qu'ils aient un tuple leur correspondant dans l'autre relation, ou non

- **Types**

- **LEFT OUTER JOIN, RIGHT OUTER JOIN, FULL OUTER JOIN**

- **Example:**

TEMP \leftarrow (EMPLOYEE $\bowtie_{Ssn=Mgr_ssn}$ DEPARTMENT)

RESULT $\leftarrow \pi_{Fname, Minit, Lname, Dname}(TEMP)$

Exemples de requêtes en algèbre relationnelle (cont'd.)

Query 6. Retrieve the names of employees who have no dependents.

This is an example of the type of query that uses the MINUS (SET DIFFERENCE) operation.

ALL_EMPS $\leftarrow \pi_{Ssn}(EMPLOYEE)$
 EMPS_WITH_DEPS(Ssn) $\leftarrow \pi_{Essn}(DEPENDENT)$
 EMPS_WITHOUT_DEPS $\leftarrow (ALL_EMPS - EMPS_WITH_DEPS)$
 RESULT $\leftarrow \pi_{Lname, Fname}(EMPS_WITHOUT_DEPS * EMPLOYEE)$

Query 7. List the names of managers who have at least one dependent.

MGRS(Ssn) $\leftarrow \pi_{Mgr_ssn}(DEPARTMENT)$
 EMPS_WITH_DEPS(Ssn) $\leftarrow \pi_{Essn}(DEPENDENT)$
 MGRS_WITH_DEPS $\leftarrow (MGRS \cap EMPS_WITH_DEPS)$
 RESULT $\leftarrow \pi_{Lname, Fname}(MGRS_WITH_DEPS * EMPLOYEE)$

Résumé

- Langages formels pour le modèle relationnel des données:
 - Algèbre relationnelle: opérations, opérateurs unaires et binaires
 - Certaines requêtes ne peuvent pas être posées avec les opérations élémentaires d'algèbre
 - Mais sont importantes en pratique
- Relational calculus
 - Based predicate calculus