Introduction to Relational Databases

- Licence Informatique, Université Lille 1
- Sept 7, 2011 (lecture 2/12)
- Today: C. Kuttler

Evaluation

- **NOTE=(TP + sup(CC+ 2*DS, 3*DS)) /4**
- **TP:** une note sur 20 de Travaux Pratiques: participation, TPs rendus et **contrôle en fin de semestre**
- **CC:** une note sur 20 de Travaux Dirigés, attribuée par l'enseignant de Travaux Dirigés: interrogations écrites, devoirs maison, participation, …
- **DS** : une note sur 20 pour l'examen de fin de semestre en décembre.

Yesterday's summary

- What is a relation
- What is an attribute
- What is a domain for an attribute
- What is NULL
- What is a key and a superkey, in a relations

Relational Algebra

Classification of query languages

a formal languages

- **Relational algebra**
- **Relational calculus**
- **Logic programming**

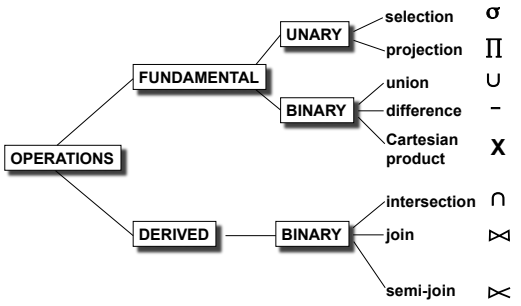
B programing languages

- **SQL: Structured Query Language**
- **QBE: Query By Example**

Relational algebra

- **defined by Codd (70)**
- **useful to learn to write database queries**
- **useful to undertand optimization of SQL queries**
- **minimal set of 5 operators provide expressive power of this language**
- **same expressive power as SQL**

Overview



7

Example: university exams

STUDENT

SID	NAME	CITY	MAJOR
123	Pierre	Lyon	Inf
415	Celine	Lille	Inf
702	Estelle	Paris	Log

Note: Celine moved from Bordeaux to Lille last night!

EXAM

SID	CLASS	DATE	GRADE
123	1	7-9-03	10
123	2	8-1-03	8
702	2	7-9-03	5

CLASS

CID	TITLE	TEACHER
1	maths	Leguichet
2	CS	Duchat

8

Selection

$\sigma_{NAME='Celine'}$ STUDENT

is a (nameless) table with

- schema :
  same schema as STUDENT
- instance :
  those tuples of STUDENT satisfying the selection predicate

SID	NAME	CITY	MAJOR
415	Celine	Lille	Inf

9

Syntax of the selection predicate P in σ_P R

Boolean expression of simple predicates

Boolean operators:

- AND (P1 AND P2)
- OR (P1 OR P2)
- NOT (P1)

Simple predicates:

- TRUE, FALSE
- term comparator term

comparators :

- =, !=, <, <=, >, >=

terms :

- constant, attribute
- arithmetic expression with constants and attributes

10

Example of selection

σ STUDENT
(City='Lille') OR
((City='Paris') AND NOT (MAJOR='Log'))

SID	NAME	CITY	MAJOR
123	Pierre	Lyon	Inf
415	Celine	Lille	Inf
702	Estelle	Paris	Log

11

Projection

$\Pi_{NAME,MAJOR}$ STUDENT

is a (nameless) table of

- schema :
  attributes NAME and MAJOR
- instance :
  restriction of tuples on the attributes NAME and MAJOR

NAME	MAJOR
Pierre	Inf
Celine	Inf
Estelle	Log

12

Question de révision pour les DUTs

- **Expliquez la différence entre**
  - **l'operateur σ de sélection, en algèbre**
  - **et le mot-clé select en SQL?**

13

Projection and duplicates

- **in the formal model, projection eliminates duplicates**

Π_{MAJOR} STUDENT

MAJOR
Inf
Log

- **in informal models (and systems), duplicate elimination must be explicitly requested**

14

Assignment

- **allows to give name to the result of an algebraic expression**
- **is not itself an algebraic expression**

GEEKS = $\sigma_{MAJOR='Inf'}$ STUDENT

CHTIS= $\sigma_{CITY='Lille'}$ STUDENT

15

Union

TABLE1 ∪ TABLE2

Only works if TABLE1 and TABLE2 are compatible

→ **With the same grade, or (in systems) with domains of same type, in same order**

16

Union

GEEKS ∪ CHTIS

Is a (nameless) table of
- **schema :**
  Schema of GEEKS
- **instance :**
  union of the tuples of GEEKS and CHTIS

ID	NAME	CITY	MAJOR
123	Pierre	Lyon	Inf
415	Celine	Lille	Inf

Commutative: order doesn't matter.

17

Difference

TABLE1 - TABLE2

Works if TABLE1 e TABLE2 are compatible

18

Difference

GEEKS - CHTIS

is a (nameless) TABLE with

- schema :
schema of GEEKS
- instance:
tuples in GEEKS, but not in and CHTIS

ID	NAME	CITY	MAJOR
123	Pierre	Lyon	Inf

Warning: not commutative

19

Cartesian product

$R \times S$

is a (nameless) TABLE with

- schema :
attributes of R and S
(arity(RxS)= arity(R)+arity(S))
- instance :
all possible combinations from tuples in R and S
- cardinality:
(card(RxS)=card(R)*card(S))

20

Example

R1(A,B)

A	B
a	1
b	3

R2(C,D)

C	D
c	1
b	3
a	2

R1xR2 (A,B,C,D)

A	B	C	D
a	1	c	1
a	1	b	3
a	1	a	2
b	3	c	1
b	3	b	3
b	3	a	2

21

Intersection

TABLE1 $\cap$ TABLE2

As the other set operations, works if TABLE1 and TABLE2 are compatible

Can be derived by following formula
$R \cap S = R - (R - S)$

22

Intersection

GEEKS $\cap$ CHTIS

is a (nameless) TABLE with

- schema :
schema of GEEKS
- instance :
tuples in both GEEKS and CHTIS

ID	NAME	CITY	MAJOR
415	Celine	Lille	Inf

23

Join

STUDENT $|\triangleright\triangleleft|_{\text{STUDENT.SID=EXAM.SID}}$ EXAM

is equivalent to the following expression (derived operator):

$\sigma_{\text{STUDENT.SID=EXAM.SID}}$ STUDENT $\times$ EXAM

When two relations have attributes of the same name, we use the dotted notation:
EXAM.SID, STUDENT.SID

24

Join

STUDENT $|\triangleright\triangleleft|_{\text{STUDENT.SID=EXAM.SID}}$ EXAM

Produces a (nameless) TABLE

- schema :
  Concatenation of schemes of STUDENT and EXAM
- instance:
  the tuples obtained by concatenating those lines of STUDENT and EXAM that have the same SID

STUDENT. SID	NAME	CITY	MAJOR	EXAM. SID	CLASS	DATE	GRADE
123	Pierre	Lyon	Inf	123	1	7-9-03	10
123	Pierre	Lyon	Inf	123	2	8-1-03	8
702	Estelle	Paris	Log	702	2	7-9-03	5

25

Syntax of join predicate

conjunctive expression with simple predicates

ATTR1 comp ATTR2

where ATTR1 is from TAB1 and
ATTR2 from TAB2
comp: =, !=, <, <=, >, >=

EQUI-JOIN :
only comparisons of equality

26

Natural join

equi-join of all homonymous attributes (repeated columns with common attributes are eliminated)

STUDENT $|\triangleright\triangleleft|$ EXAM

SID	NAME	CITY	MAJOR	CLASS	DATE	GRADE
123	Pierre	Lyon	Inf	1	7-9-03	10
123	Pierre	Lyon	Inf	2	8-1-03	8
702	Estelle	Paris	Log	2	7-9-03	5

27

Natural join with 3 tables

STUDENT $|\triangleright\triangleleft|$ EXAM $|\triangleright\triangleleft|$ CLASS

ID	NAME	CITY	MAJOR	CLASS	DATE	GRADE	TITLE	TEACHER
123	Pierre	Lyon	Inf	1	7-9-03	10	maths	Leguichet
123	Pierre	Lyon	Inf	2	8-1-03	8	infor	Duchat
702	Estelle	Paris	Log	3	7-9-03	5	infor	Duchat

28

Equivalence of expressions

- which students have obtained a 10 in mathematics?

Π_{NAME}
(STUDENT $|\triangleright\triangleleft|$
($\sigma_{\text{GRADE=10}}$ EXAM $|\triangleright\triangleleft|$
($\sigma_{\text{TITLE='maths'}}$ CLASS)))

- equivalent to:

$\Pi_{\text{NAME}}\ \sigma_{\text{GRADE=10 AND TITLE='maths'}}$
(STUDENT $|\triangleright\triangleleft|$ EXAM $|\triangleright\triangleleft|$ CLASS)

31

Equivalence of expressions

- which TEACHERs have examined Estelle?

Π_{TEACHER}
(CLASS $|\triangleright\triangleleft|$
(EXAM $|\triangleright\triangleleft|$
$\sigma_{\text{NAME = 'Estelle'}}$ STUDENT))

- equivalent to:

$\Pi_{\text{TEACHER}}\ \sigma_{\text{NAME = 'Estelle'}}$
(STUDENT $|\triangleright\triangleleft|$ EXAM $|\triangleright\triangleleft|$ CLASS)

32

Complex expressions

• extract NAME of STUDENT that have never had a grade under 8

Π_{NAME} **STUDENT** $|\triangleright\!\triangleleft|$
(Π_{ID} **EXAM**
-
$\Pi_{ID}\ \sigma_{GRADE<8}$ **EXAM)**

• explanation: first find IDs of al students that have taken exams, minus those that had some grade unter 8. Then, find their names.

33

Summary

- We've seen the relation algebra operators
  - projection, selection
  - Cartesian product, join
  - set operations: union, difference, intersection,

36

No lecture this Friday!

- Postponed to later
- Enough to practice after our first 2 lectures
- Next week's tutorial and lab
  - Relational model
  - Algebraic queries: on paper and computer

37

Réunion ERASMUS

- Lundi 19 septembre, pause déjeuner
- Condition pour partir:
  - valider ce semestre en 1ère session
- Programme:
  - aspects administratifs (bourses,…)
  - témoignages d'étudiants
  - universités partenaires
  - parrainage entre étudiants
- Mobilité enseignants: cours en anglais?

38