Database Management Systems

Queries and Relational Algebra

Topics

- Definitions
  - Queries and Query Types
  - Query Languages
  - Relational Algebra
- Relational Algebra Procedures
  - Union, Intersection, Difference
  - Renaming
  - Selection
  - Projection
  - Join (natural join, outer join, equi join)

Operations on Databases

- Operations on databases:
  - Read queries: read data from the database
  - Update queries: change the content of the database
- Mathematical foundations of these operations can be studied with ***"Relational Algebra"***

Queries and Query Languages

- Query is a set of operations that take one or some relations (database) as input and return one or some relations (database) as output.

- Query languages are used for writing queries for databases

Relational Algebra

- A collection of operators that
  - are defined on relations
  - produce relations as results
  - and therefore can be combined to form complex expressions
- Operators
  - union, intersection, difference
  - renaming
  - selection
  - projection
  - join (natural join, outer join, equi join)

Union, Intersection, and Difference

- Relations are sets, so we can apply set operators
- However, we want the results to be relations
- Therefore: it is meaningful to apply union, intersection, difference only to relations defined over the same attributes (relations that have the same schema)

Union

Graduates

Number	Surname	Age
7274	Robinson	37
7432	O’Malley	39
9824	Darkes	38

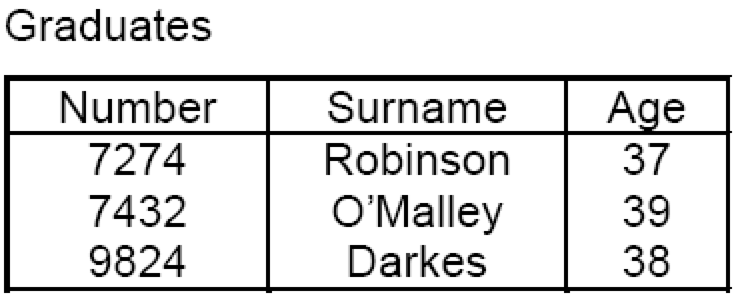
Managers

Number	Surname	Age
9297	O’Malley	56
7432	O’Malley	39
9824	Darkes	38

Graduates ∪ Managers

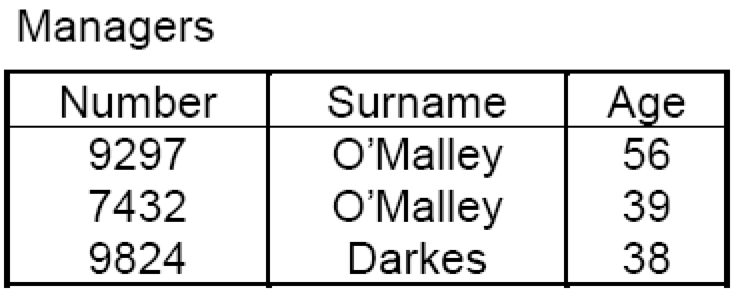
Number	Surname	Age
7274	Robinson	37
7432	O’Malley	39
9824	Darkes	38
9297	O’Malley	56

Intersection



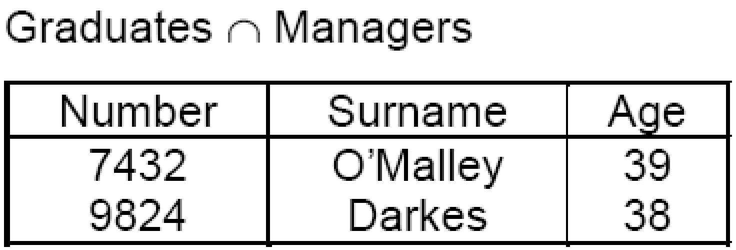
Graduates

Number	Surname	Age
7274	Robinson	37
7432	O’Malley	39
9824	Darkes	38



Managers

Number	Surname	Age
9297	O’Malley	56
7432	O’Malley	39
9824	Darkes	38



Graduates ∩ Managers

Number	Surname	Age
7432	O’Malley	39
9824	Darkes	38

Difference

Graduates

Number	Surname	Age
7274	Robinson	37
7432	O'Malley	39
9824	Darkes	38

Managers

Number	Surname	Age
9297	O'Malley	56
7432	O'Malley	39
9824	Darkes	38

Graduates - Managers

Number	Surname	Age
7274	Robinson	37

A meaningful but Impossible Union

Paternity

Father	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael

Maternity

Mother	Child
Eve	Cain
Eve	Seth
Sarah	Isaac
Hagar	Ishmael

$$\text{Paternity} \cup \text{Maternity} = ??$$

- the problem: Father and Mother are different attribute names, but both represent a "Parent"
- the solution: rename the attributes

Renaming

- unary operator
- changes attribute names without changing values
- notation: $\rho_{y/x}(r)$

  - *Example:* : $\rho_{parent/father}(Paternity)$
- Two or more attributes can be renamed with one expression
  - *Example*: $\rho_{Location, Pay/ Branch, Salary}(Employee)$

Renaming Example

Paternity

Father	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael

$\rho_{\text{Parent} \leftarrow \text{Father.}}(\text{Paternity})$

Parent	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael

Renaming and Union

Paternity

Father	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael

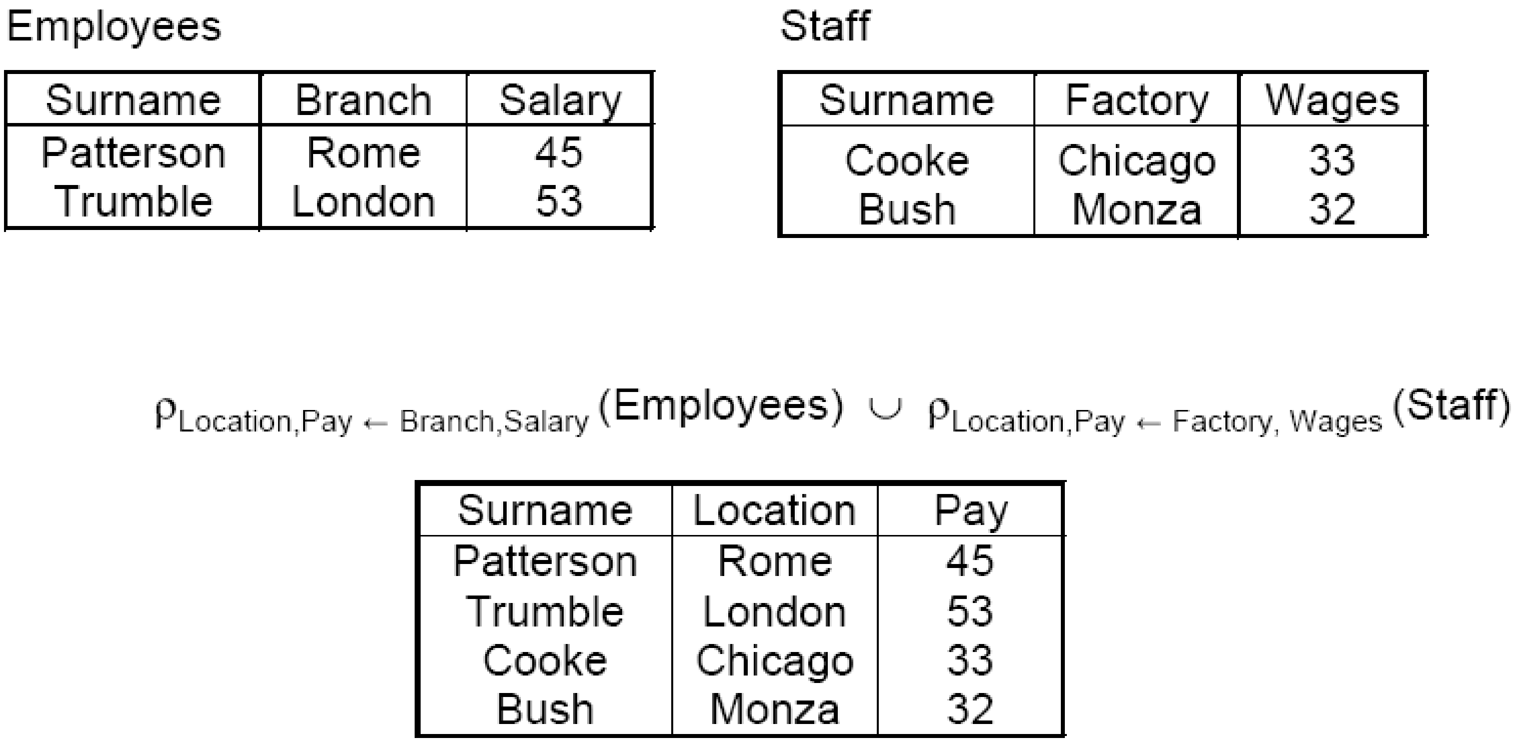
Maternity

Mother	Child
Eve	Cain
Eve	Seth
Sarah	Isaac
Hagar	Ishmael

$\rho_{\text{Parent} \leftarrow \text{Father.}}(\text{Paternity}) \cup \rho_{\text{Parent} \leftarrow \text{Mother.}}(\text{Maternity})$

Parent	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael
Eve	Cain
Eve	Seth
Sarah	Isaac
Hagar	Ishmael

Renaming and Union with More Attributes



Employees

Surname	Branch	Salary
Patterson	Rome	45
Trumble	London	53

Staff

Surname	Factory	Wages
Cooke	Chicago	33
Bush	Monza	32

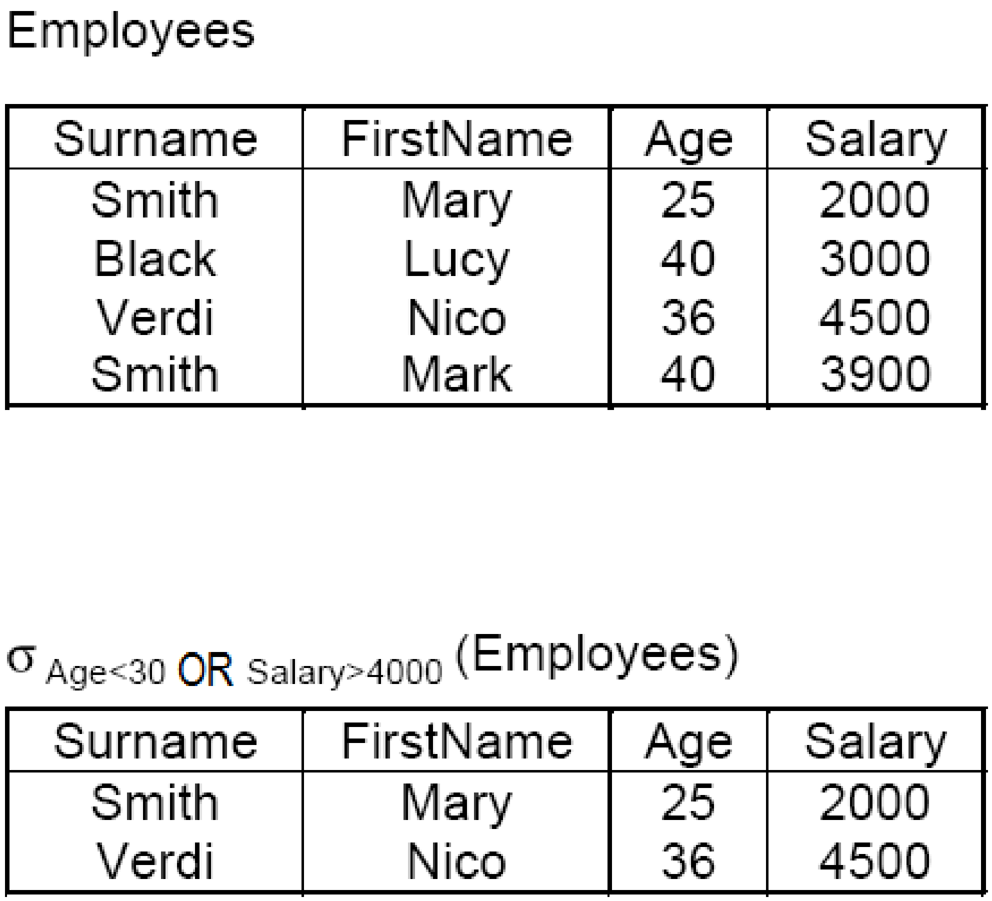
$$\rho_{\text{Location,Pay} \leftarrow \text{Branch,Salary}}(\text{Employees}) \cup \rho_{\text{Location,Pay} \leftarrow \text{Factory, Wages}}(\text{Staff})$$

Surname	Location	Pay
Patterson	Rome	45
Trumble	London	53
Cooke	Chicago	33
Bush	Monza	32

Selection

- Selection operation produces a relation with the same schema as its operand.
- The resulted (output) relation is a subset of the operand (input) relation.
- Notation: σ_F (r)
  - r is the relation
  - F is the condition

Selection Example 1



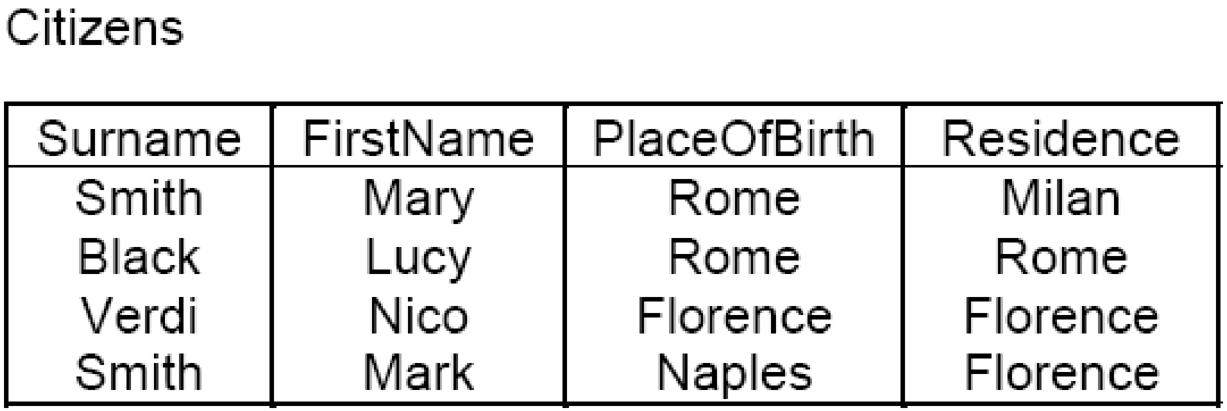
Employees

Surname	FirstName	Age	Salary
Smith	Mary	25	2000
Black	Lucy	40	3000
Verdi	Nico	36	4500
Smith	Mark	40	3900

$\sigma_{Age<30 \text{ OR } Salary>4000}$ (Employees)

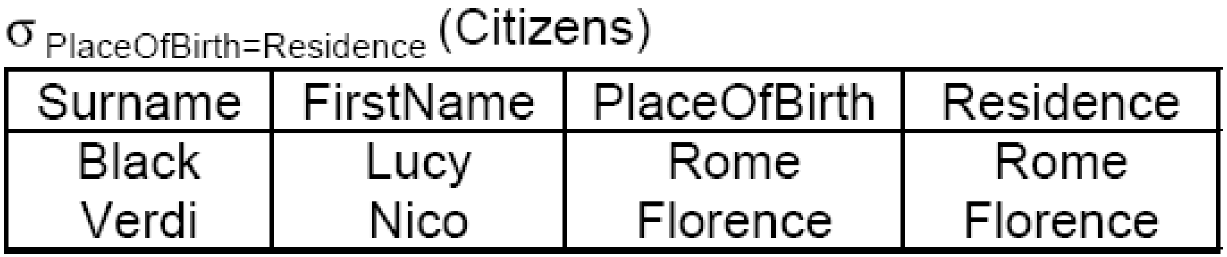
Surname	FirstName	Age	Salary
Smith	Mary	25	2000
Verdi	Nico	36	4500

Selection Example 2



Citizens

Surname	FirstName	PlaceOfBirth	Residence
Smith	Mary	Rome	Milan
Black	Lucy	Rome	Rome
Verdi	Nico	Florence	Florence
Smith	Mark	Naples	Florence



$\sigma_{\text{PlaceOfBirth=Residence}}$ (Citizens)

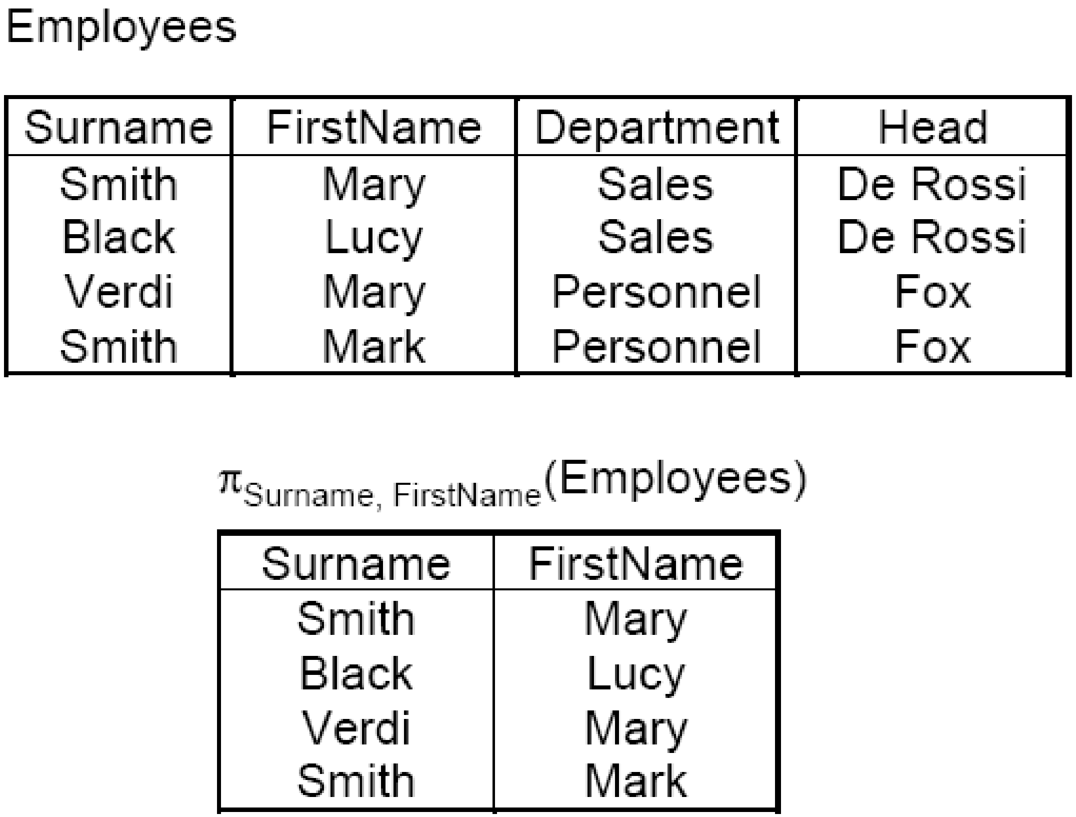
Surname	FirstName	PlaceOfBirth	Residence
Black	Lucy	Rome	Rome
Verdi	Nico	Florence	Florence

Projection

- Projection operation produces a relation with a sub-set of the attributes of its operand
- All tuples from the operand will be in the resulted relation

- Notation: $\pi_Y(r)$
  - r is a relation
  - Y is the list of selected attributes

Projection Example 1



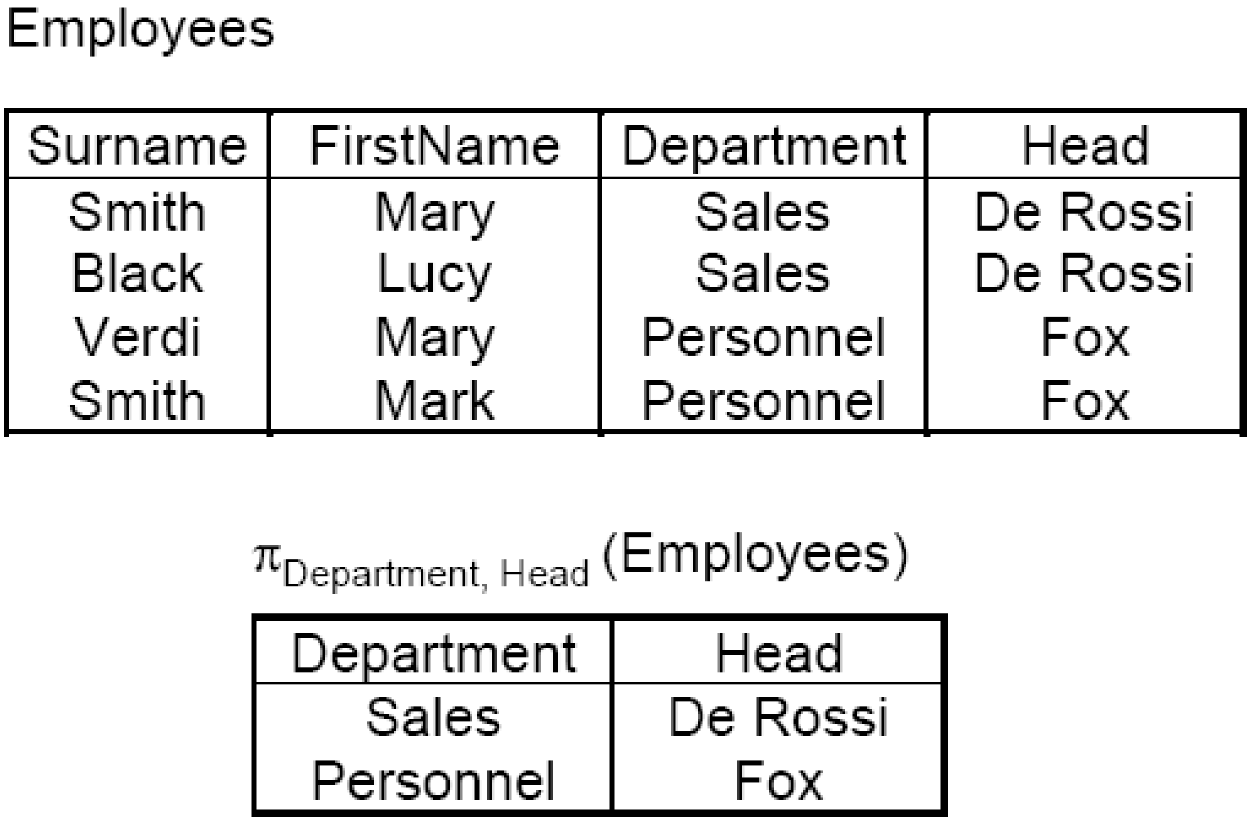
Employees

Surname	FirstName	Department	Head
Smith	Mary	Sales	De Rossi
Black	Lucy	Sales	De Rossi
Verdi	Mary	Personnel	Fox
Smith	Mark	Personnel	Fox

$\pi_{\text{Surname, FirstName}}(\text{Employees})$

Surname	FirstName
Smith	Mary
Black	Lucy
Verdi	Mary
Smith	Mark

Projection Example 2



Employees

Surname	FirstName	Department	Head
Smith	Mary	Sales	De Rossi
Black	Lucy	Sales	De Rossi
Verdi	Mary	Personnel	Fox
Smith	Mark	Personnel	Fox

$\pi_{\text{Department, Head}}$ (Employees)

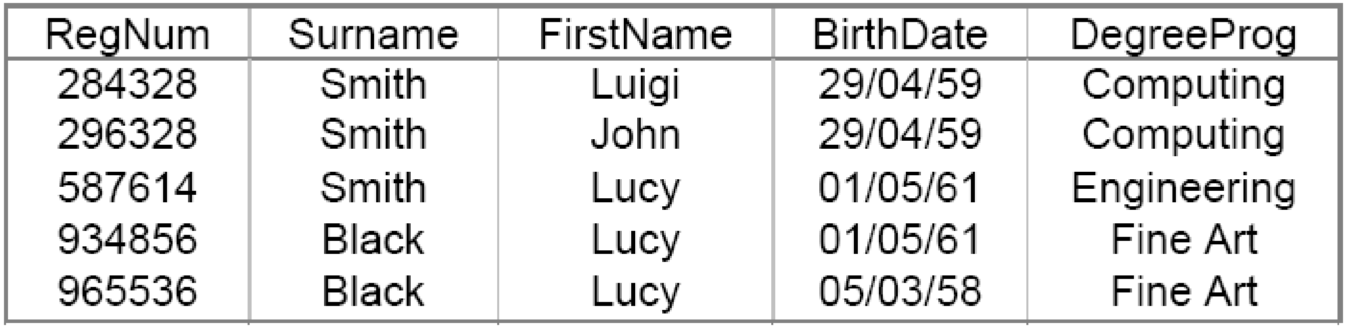
Department	Head
Sales	De Rossi
Personnel	Fox

Cardinality of Projection

- The result of a projection contains at most as many tuples as the operand
- It can contain fewer tuples, if some tuples are the same.
- $\pi_Y(r)$ contains as many tuples as ***r*** if and only if Y is a key for ***r***

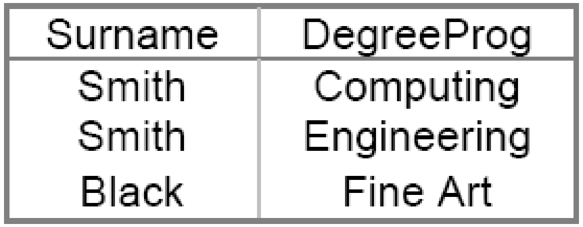
Cardinality of Projection: Example

Students



RegNum	Surname	FirstName	BirthDate	DegreeProg
284328	Smith	Luigi	29/04/59	Computing
296328	Smith	John	29/04/59	Computing
587614	Smith	Lucy	01/05/61	Engineering
934856	Black	Lucy	01/05/61	Fine Art
965536	Black	Lucy	05/03/58	Fine Art

$\pi_{\text{Surname, DegreeProg}}$ (Students)



Surname	DegreeProg
Smith	Computing
Smith	Engineering
Black	Fine Art

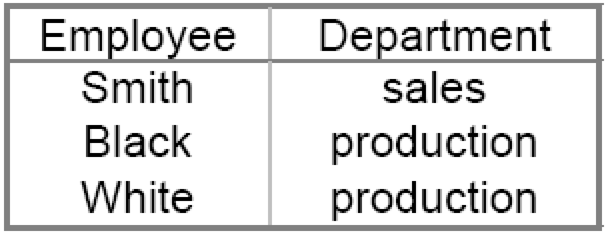
Join

- Join operation uses the connection between the relations to combine (join) them.
- Notation $\bowtie$

Natural Join

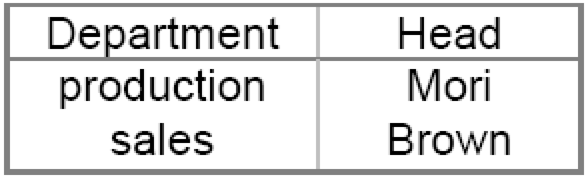
- Joins two tuples of relations r1 and r2 if the values of attributes with the same name are equal.

r_1



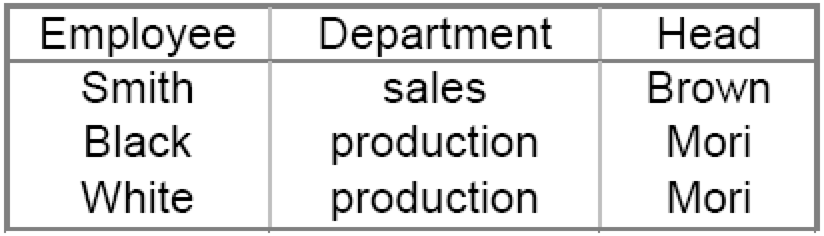
Employee	Department
Smith	sales
Black	production
White	production

r_2



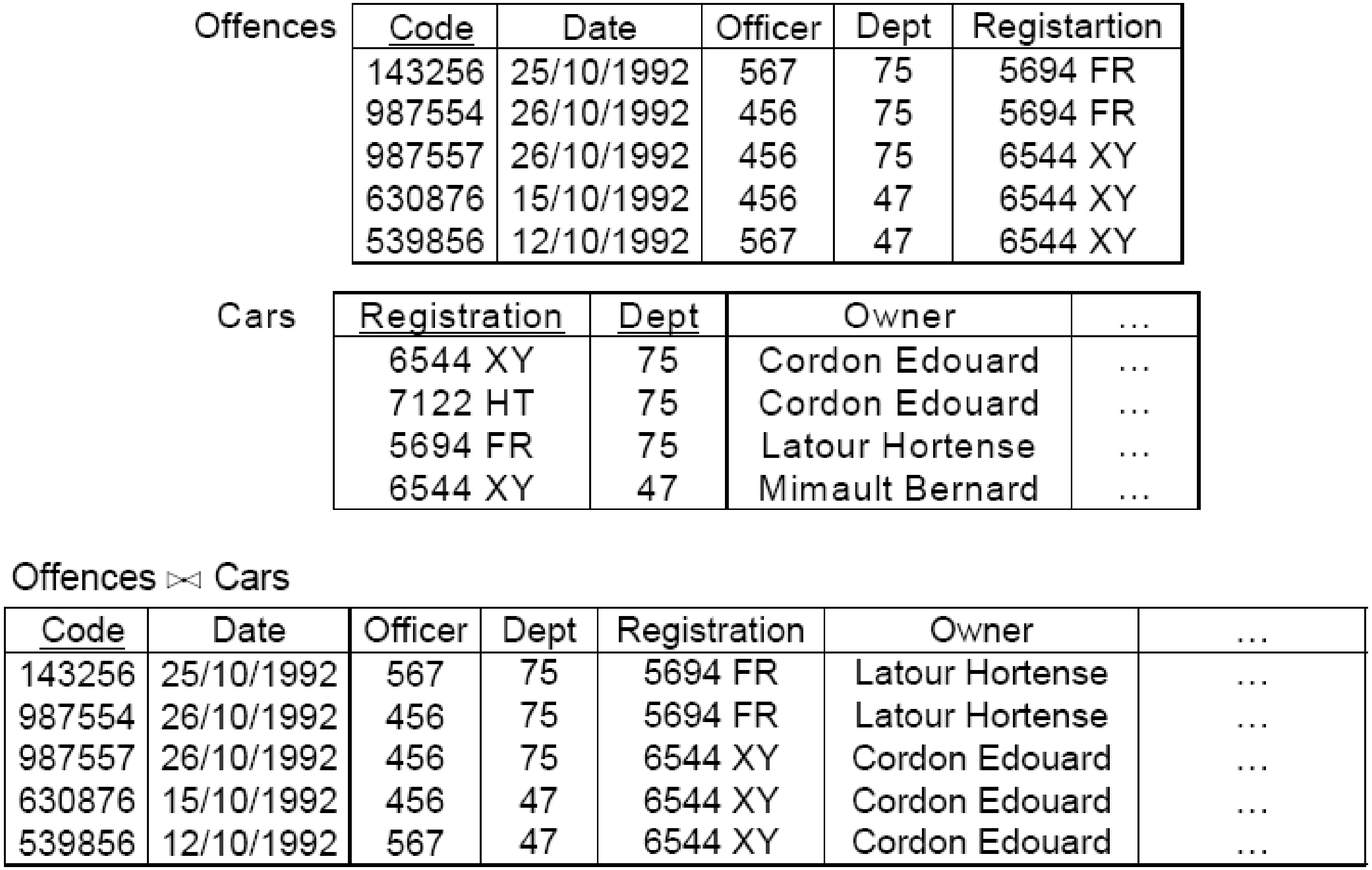
Department	Head
production	Mori
sales	Brown

$r_1 \bowtie r_2$



Employee	Department	Head
Smith	sales	Brown
Black	production	Mori
White	production	Mori

Natural Join Example 1



Offences

Code	Date	Officer	Dept	Registartion
143256	25/10/1992	567	75	5694 FR
987554	26/10/1992	456	75	5694 FR
987557	26/10/1992	456	75	6544 XY
630876	15/10/1992	456	47	6544 XY
539856	12/10/1992	567	47	6544 XY

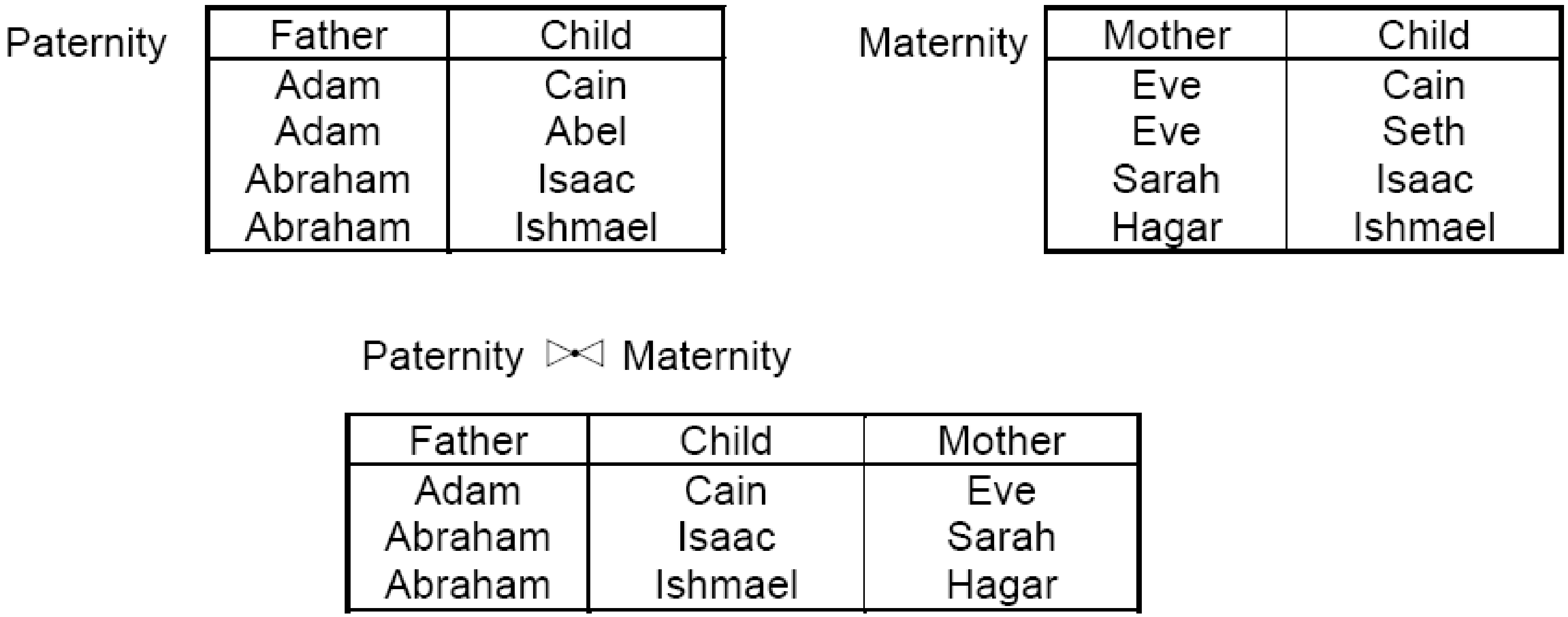
Cars

Registration	Dept	Owner	...
6544 XY	75	Cordon Edouard	...
7122 HT	75	Cordon Edouard	...
5694 FR	75	Latour Hortense	...
6544 XY	47	Mimault Bernard	...

Offences ⋈ Cars

Code	Date	Officer	Dept	Registration	Owner	...
143256	25/10/1992	567	75	5694 FR	Latour Hortense	...
987554	26/10/1992	456	75	5694 FR	Latour Hortense	...
987557	26/10/1992	456	75	6544 XY	Cordon Edouard	...
630876	15/10/1992	456	47	6544 XY	Cordon Edouard	...
539856	12/10/1992	567	47	6544 XY	Cordon Edouard	...

Natural Join Example 2



Paternity

Father	Child
Adam	Cain
Adam	Abel
Abraham	Isaac
Abraham	Ishmael

Maternity

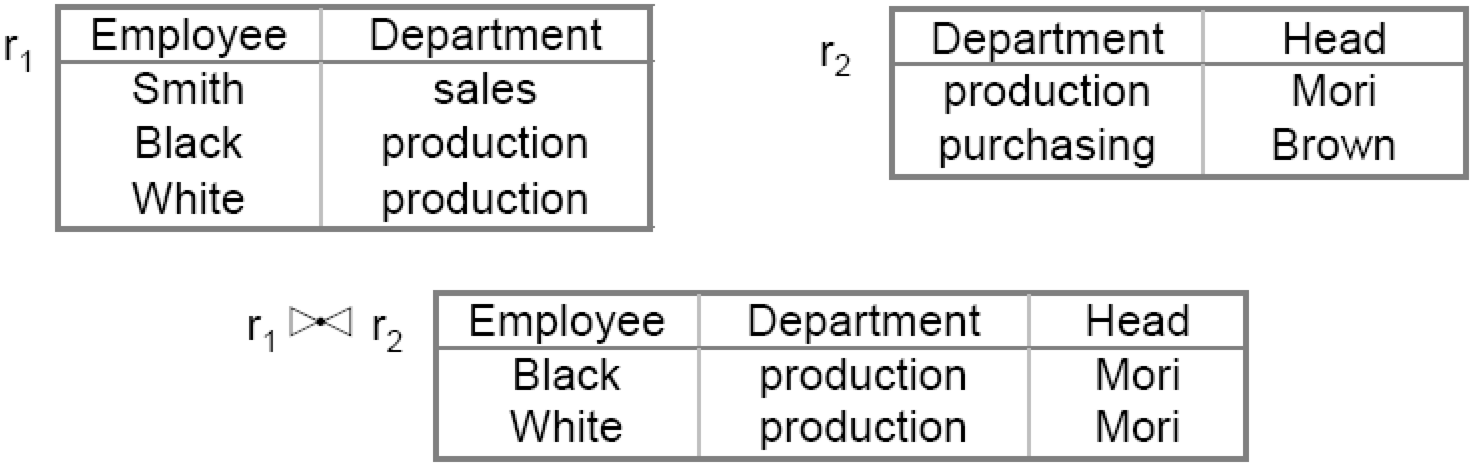
Mother	Child
Eve	Cain
Eve	Seth
Sarah	Isaac
Hagar	Ishmael

Paternity $\bowtie$ Maternity

Father	Child	Mother
Adam	Cain	Eve
Abraham	Isaac	Sarah
Abraham	Ishmael	Hagar

Incomplete Natural Join

- Some tuples from the first relation may have no match in the second relation. Natural join will be incomplete in this case.



r_1

Employee	Department
Smith	sales
Black	production
White	production

r_2

Department	Head
production	Mori
purchasing	Brown

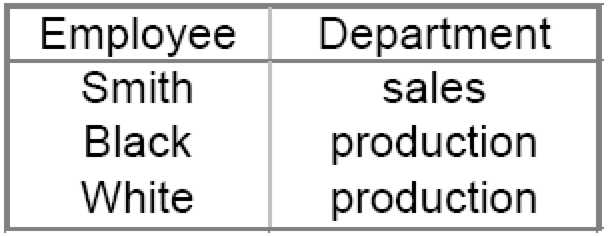
$r_1 \bowtie r_2$

Employee	Department	Head
Black	production	Mori
White	production	Mori

Empty Natural Join

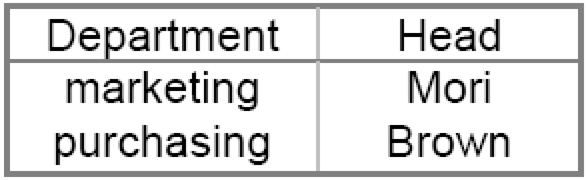
- If no tuple of the first relation matches a tuple from the second relation, the natural join will be empty.

r_1



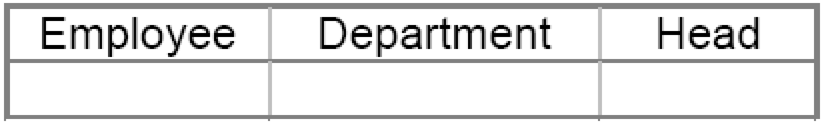
Employee	Department
Smith	sales
Black	production
White	production

r_2



Department	Head
marketing	Mori
purchasing	Brown

$r_1 \bowtie r_2$



Employee	Department	Head

Outer Join

- If no match is found for a tuple, we can join it with a ***null tuple.*** This is called outer join.
- Left outer join: the ***null tuple*** is from the second relation
- Right outer join: the ***null tuple*** is from the first relation
- Full outer join: the ***null tuple*** can be from either the first or the second relation

Outer Join Example

r_1

Employee	Department
Smith	sales
Black	production
White	production

r_2

Department	Head
production	Mori
purchasing	Brown

$r_1 \bowtie_{\text{LEFT}} r_2$

Employee	Department	Head
Smith	Sales	NULL
Black	production	Mori
White	production	Mori

$r_1 \bowtie_{\text{RIGHT}} r_2$

Employee	Department	Head
Black	production	Mori
White	production	Mori
NULL	purchasing	Brown

$r_1 \bowtie_{\text{FULL}} r_2$

Employee	Department	Head
Smith	Sales	NULL
Black	production	Mori
White	production	Mori
NULL	purchasing	Brown

Equi Join

- Equi join, combines tuples using attributes with different names if they have equal values

- Notation: $r1 \bowtie_{x=y} r2$

  (x is an attribute from r1 and y an attribute from r2)

Equi Join Example

Employees

Employee	Project
Smith	A
Black	A
Black	B

Projects

Code	Name
A	Venus
B	Mars

Employes $\bowtie_{\text{Project=Code}}$ Projects

Employee	Project	Code	Name
Smith	A	A	Venus
Black	A	A	Venus
Black	B	B	Mars

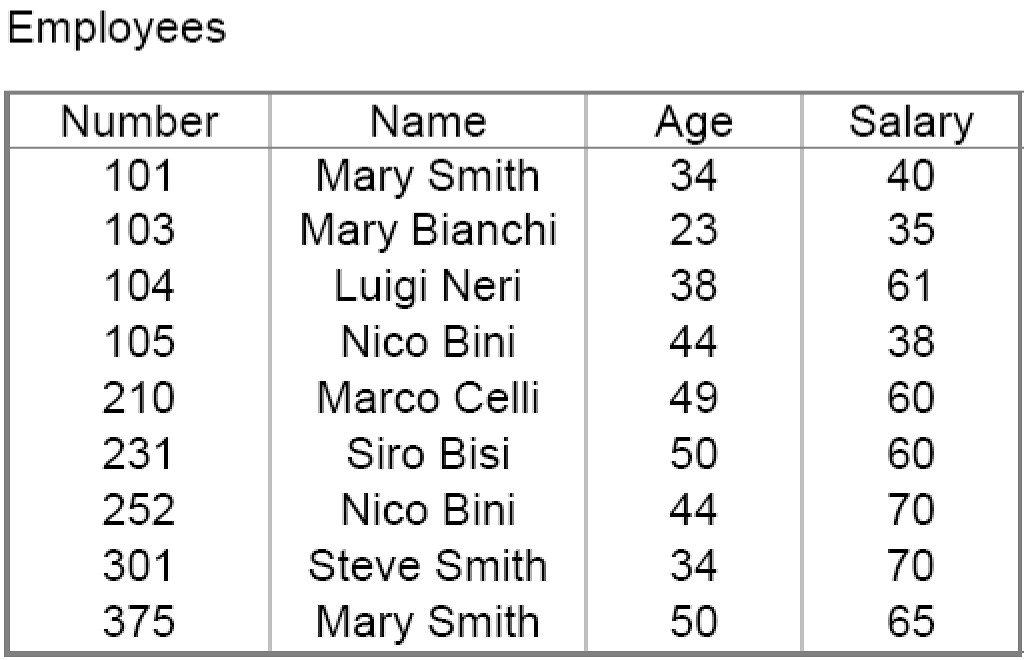
Queries using Relational Algebra

- Queries can be written using relational algebra operations

- Example: Find the numbers, names and ages of employees earning more than 40 thousand

Query Example 1

- Data Table:

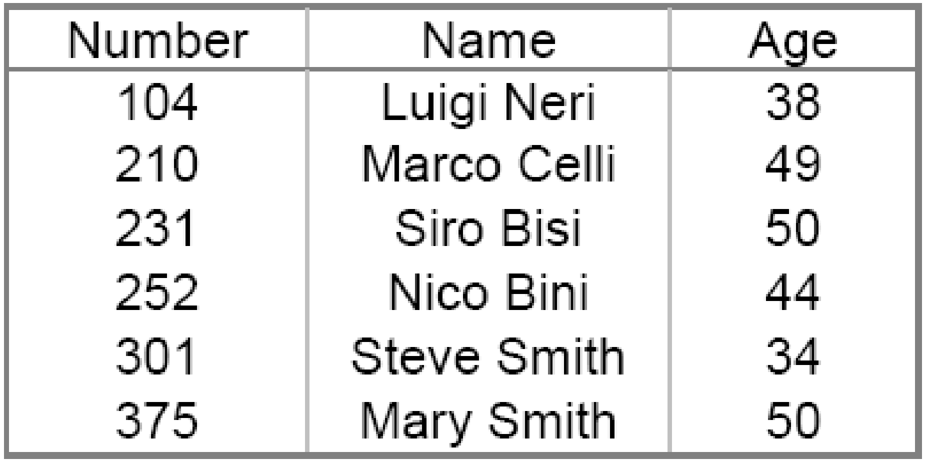


Employees

Number	Name	Age	Salary
101	Mary Smith	34	40
103	Mary Bianchi	23	35
104	Luigi Neri	38	61
105	Nico Bini	44	38
210	Marco Celli	49	60
231	Siro Bisi	50	60
252	Nico Bini	44	70
301	Steve Smith	34	70
375	Mary Smith	50	65

Query Example

$$\pi_{\text{Number,Name,Age}}(\sigma_{\text{Salary}>40}(\text{EMPLOYEES}))$$



Number	Name	Age
104	Luigi Neri	38
210	Marco Celli	49
231	Siro Bisi	50
252	Nico Bini	44
301	Steve Smith	34
375	Mary Smith	50

Summary

- Queries are used to read or update data in a database
- Queries can be expressed using Relational Algebra Procedures
- The most important operations in relational algebra are
  - Union, Intersection, Difference, Renaming
  - Selection, Projection, Join (natural join, outer join, equi join)

Questions?