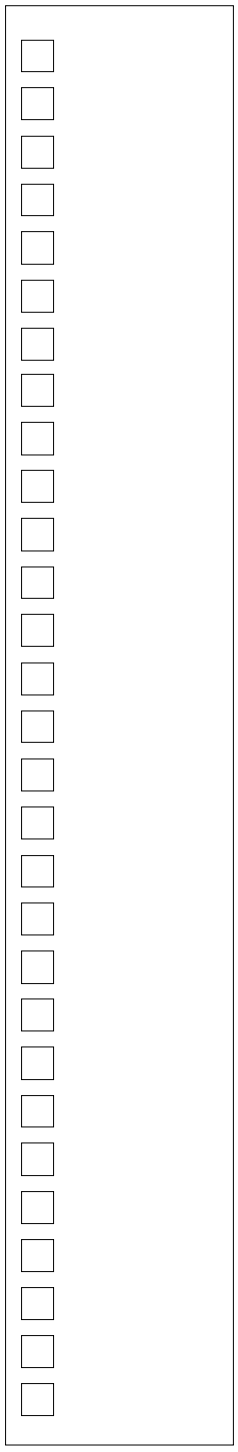

CSC 742
Database Management Systems

Topic #7:
Relational Algebra

Relational Algebra

Offers operations defined at a high level

- input >=1 relations; output 1 relation
- independent of implementation
- procedural
- nesting of an expression defines the *essential* order of operations independent of implementation
- intuitively, like a virtual machine for databases

Achtung!

- Operations are
  - written in terms of R, S, etc. referring to schema names
  - applied to specific instances

Select Operation

- Denoted $\sigma_{<\text{selection condition c}>}(R)$
  - Filters out tuples from r that don't match c
  - c is a boolean combination of comparison clauses
  - Terms in c are attributes or constants
  - c makes sense only for one tuple at a time

Salary

SSN	Lname	Fname	Salary
111-22-3333	Smith	John	30000
121-23-3333	Wong	Frank	45000
153-32-1342	Wallace	Jennifer	43000
154-33-3333	Borg	James	56000
555-44-5555	English	Joyce	53000

$$\sigma_{salary>50000}(Salary)$$

Achtung!

- σ is not the same as SELECT in SQL
- A series of σ's can be reordered or joined into one
  - $\sigma_{<cond1>}(\sigma_{<cond2>}(\ldots(\sigma_{<condn>}(R))\ldots)) =$ $\sigma_{<cond1> \text{ AND } <cond2> \text{ AND } <condn>}(R)$.

Project

- Denoted $\pi_{<\text{attribute list}>}(R)$
- Returns a relation that has
  - only the attributes in the list
  - all the tuples in r restricted to these attributes
  - except that duplicates are removed

Salary

SSN	Lname	Fname	Salary
111-22-3333	Smith	John	30000
121-23-3333	Wong	Frank	45000
153-32-1342	Wallace	Jennifer	43000
154-33-3333	Borg	James	56000
555-44-5555	English	Joyce	53000

$$\pi_{\text{Lname, Fname, Salary}}(\text{Salary})$$

$$\pi_{\text{Lname, Fname}}\ (\pi_{\text{Lname, Fname, Salary}}(\text{Salary}))$$

$$\pi_{\text{Lname, Fname}}\ (\sigma_{\text{salary}>50000}(\text{Salary}))$$

Conventions

- Relations may be
  - temporary (named or anonymous)
  - predefined in the schema
- Attributes may be renamed
- Define a new relation
  - $A \leftarrow B$

Rename Operation

- Change the name of a relation, the names of attributes, or all of the above.
- $\rho_S(R)$
- $\rho_{(B1, B2, \ldots, Bn)}(R)$
- $\rho_{S(B1, B2, \ldots, Bn)}(R)$

Salary

SSN	Lname	Fname	Salary
111-22-3333	Smith	John	30000
121-23-3333	Wong	Frank	45000
153-32-1342	Wallace	Jennifer	43000
154-33-3333	Borg	James	56000
555-44-5555	English	Joyce	53000

- $\rho_{CurrentSalary}(Salary)$
- $\rho_{(B1, B2, B3, B4)}(Salary)$
- $\rho_{CurrentSalary(B1, B2, B3, B4)}(Salary)$

Set Operations

- Union
- Intersection
- Difference
- Apply only on "(union) compatible" arguments
  - same degree
  - same domains for each matching pair of attributes
- Attribute names come from first argument

EngineeringSalary

SSN	Lname	Fname	Salary
111-22-3333	Smith	John	30000
121-23-3333	Wong	Frank	45000
153-32-1342	Wallace	Jennifer	43000

AccountingSalary

SSN	Lname	Fname	Salary
154-33-3333	Borg	James	56000
555-44-5555	English	Joyce	53000

Cartesian Product

- Also called cross product or cross join
- Notated $R \times S$
- All possible pairings
- All the attributes, renamed if necessary
- lots of tuples: $|r \times s| = |r| \times |s|$
- Not to be used except in defining or simplifying expressions

Employee

Fname	Lnane	SSN
Alice	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333
Joyce	White	222-33-4444

Dependent

Fname	Lnane	ESSN
Eric	Zelaya	999-88-7777
Alex	Wallace	111-22-3333

Employee × Dependent

Fname	Lnane	SSN	D_FName	D_Lnane	D_ESSN
Alice	Zelaya	999-88-7777	Eric	Zelaya	999-88-7777
Alice	Zelaya	999-88-7777	Alex	Wallace	111-22-3333
Jennifer	Wallace	111-22-3333	Eric	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333	Alex	Wallace	111-22-3333
Joyce	White	222-33-4444	Eric	Zelaya	999-88-7777
Joyce	White	222-33-4444	Alex	Wallace	111-22-3333

Theta Join

- $R \mid\!><\!\mid_{condition} S$
- Equivalent to a Cartesian product nested within a select
  - $\sigma_{condition} (R \times S)$
- Tuples with null values for any attribute in condition are eliminated

Employee

Fname	Lnane	SSN
Alice	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333
Joyce	White	222-33-4444

Dependent

Fname	Lnane	ESSN
Eric	Zelaya	999-88-7777
Alex	Wallace	111-22-3333

Employee $|><|_{SSN=ESSN}$ Dependent

Fname	Lnane	SSN	D_FName	D_Lnane	D_ESSN
Alice	Zelaya	999-88-7777	Eric	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333	Alex	Wallace	111-22-3333

Equijoin

- Theta join where only = may be used as the comparison operator in the join condition

Natural Join

- Denoted R*S or R |><| S
- Basically an equijoin followed by removal of the superfluous attributes.
  - Equijoin results in pairs of attributes with identical values
- Natural join eliminates the second of each pair
  - Matching attributes must have the same name
  - Equality for all matching attributes is required

Employee

Fname	Lnane	SSN
Alice	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333
Joyce	White	222-33-4444

Dependent

Fname	Lnane	ESSN
Eric	Zelaya	999-88-7777
Alex	Wallace	111-22-3333

Employee |><| $\rho_{(DFname, DLname, SSN)}$Dependent

Fname	Lnane	SSN	DFName	DLnane
Alice	Zelaya	999-88-7777	Eric	Zelaya
Jennifer	Wallace	111-22-3333	Alex	Wallace

Join

- The heart of the relational model
- Helps combine information from multiple tables
- Essential because there are no direct pointers or references in the model

Outer Join

- Includes tuples in result even if no match is found
- The result tuples are padded with nulls
- Variants: left, right, full

Employee

Fname	Lnane	SSN
Alice	Zelaya	999-88-7777
Jennifer	Wallace	111-22-3333
Joyce	White	222-33-4444

Dependent

Fname	Lnane	ESSN
Eric	Zelaya	999-88-7777
Alex	Wallace	111-22-3333

Employee]><| $\rho_{(DFname, DLname, SSN)}$Dependent

Fname	Lnane	SSN	DFName	DLnane
Alice	Zelaya	999-88-7777	Eric	Zelaya
Jennifer	Wallace	111-22-3333	Alex	Wallace
Joyce	White	222-33-4444	NULL	NULL

Outer Union

- Applies on incompatible relations
- For attributes in only one input relation, tuples of the other are padded with nulls

Faculty

Name	SSN	Rank	Department
Alice	999-88-7777	Professor	CSC
Jennifer	111-22-3333	Assistant Prof.	ECE
Joyce	222-33-4444	Associate Prof.	CSC

Student

Name	SSN	Advisor	Department
Eric	999-88-7777	Alice	CSC
Alex	111-22-3333	Jennifer	ECE

Faculty OUTER UNION Student

Name	SSN	Rank	Advisor	Department
Alice	999-88-7777	Professor	NULL	CSC
Jennifer	111-22-3333	Assistant Prof.	NULL	ECE
Joyce	222-33-4444	Associate Prof.	NULL	CSC
Eric	999-88-7777	NULL	Alice	CSC
Alex	111-22-3333	NULL	Jennifer	ECE