

Relational Model & Algebra

CPS 116

Introduction to Database Systems

Announcements (Thu. Aug. 28)

- ◆ Homework #1 will be assigned next Thursday
- ◆ Meeting room: LSRC 344 today and next Tuesday
 - Still looking for a permanent solution; watch your emails
- ◆ Office hours: see also course website
 - Jun: LSRC D327
 - Tue. 1.5 hours before class; Thu. 1.5 hours after
 - Ying: LSRC D125
 - Mon. & Wed. 4:05-5:05; Fri. 3-5pm
- ◆ Lecture notes
 - I will bring hardcopies of the "notes" version to lectures
 - The "complete" version will be posted after lecture, so be selective in what you copy down

Relational data model

- ◆ A database is a collection of relations (or tables)
- ◆ Each relation has a list of attributes (or columns)
- ◆ Each attribute has a domain (or type)
 - Set-valued attributes not allowed
- ◆ Each relation contains a set of tuples (or rows)
 - Each tuple has a value for each attribute of the relation
 - Duplicate tuples are not allowed
 - Two tuples are identical if they agree on all attributes

☞ Simplicity is a virtue!

Example

Student

SID	name	age	GPA
142	Bart	10	2.3
123	Milhouse	10	3.1
257	Lisa	8	3.3
356	Ralph	8	2.1
-	-	-	-

Course

CID	title
CP5116	Intro. to Database Systems
CP5120	Analysis of Algorithms
CP5114	Computer Networks
-	-

Enroll

SID	CID
142	CP5116
142	CP5114
123	CP5116
257	CP5116
257	CP5120
356	CP5114
-	-

Ordering of rows doesn't matter
(even though the output is
always in *some* order)

Schema versus instance

❖ Schema (metadata)

- Specification of how data is to be structured logically
- Defined at set-up
- Rarely changes

❖ Instance

- Content
- Changes rapidly, but always conforms to the schema

☞ Compare to type and objects of type in a programming language

Example

❖ Schema

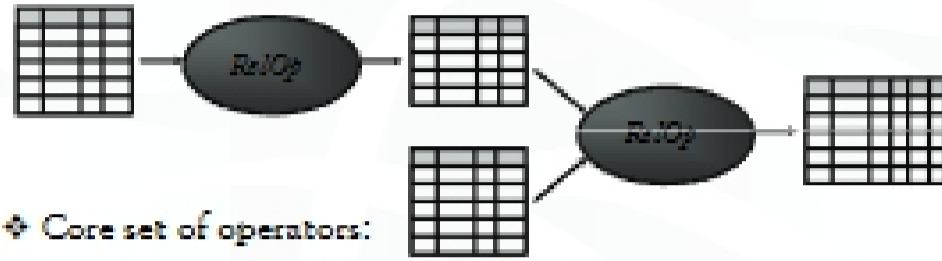
- *Student* (SID integer, name string, age integer, GPA float)
- *Courses* (CID string, title string)
- *Enroll* (SID integer, CID integer)

❖ Instance

- { (142, Bart, 10, 2.3), (123, Milhouse, 10, 3.1), ... }
- { (CP5116, Intro. to Database Systems), ... }
- { (142, CP5116), (142, CP5114), ... }

Relational algebra

A language for querying relational databases based on operators:



- ◆ Core set of operators:
 - Selection, projection, cross product, union, difference, and renaming
- ◆ Additional, derived operators:
 - Join, natural join, intersection, etc.
- ◆ Compose operators to make complex queries

Selection

- ◆ Input: a table R
- ◆ Notation: $\sigma_p R$
 - p is called a selection condition/predicate
- ◆ Purpose: filter rows according to some criteria
- ◆ Output: same columns as R , but only rows of R that satisfy p

Selection example

- ◆ Students with GPA higher than 3.0

$\sigma_{GPA > 3.0} Student$