Database Systems (CS F212)

BITS Pilani
Pilani Campus

Introduction to DBMS

Data: Facts that have some implicit meaning

Database: Collection of related data

DBMS: A system used for managing database

Components of DBMS:

- Database: A collection of inter-related data
- Managing: A tool or software used for *defining*, *constructing*, *manipulating*, and *sharing* databases among various users and applications.

Problems in File Processing Systems

1. Data redundancy and inconsistency
2. Difficulty in accessing data
3. Data isolation
4. Security problems
5. Concurrent access anomalies
6. Integrity problem

Purpose of DBMS

A DBMS attempts to resolve the following problems:

- Avoid data redundancy and inconsistency
- Data isolation (multiple files and formats)
- Integrity problems by enforcing constraints
- Atomicity of updates
- Concurrent access by multiple users
- Security problems

Data Abstraction

- The primary causes of all the problems in FPS is
  - Different people writing different applications programs independently
  - Absence *of* *program-data independence*

- DBMS provides users with an abstract view of data to hide the complex data structure used for the representation of data in the databases.

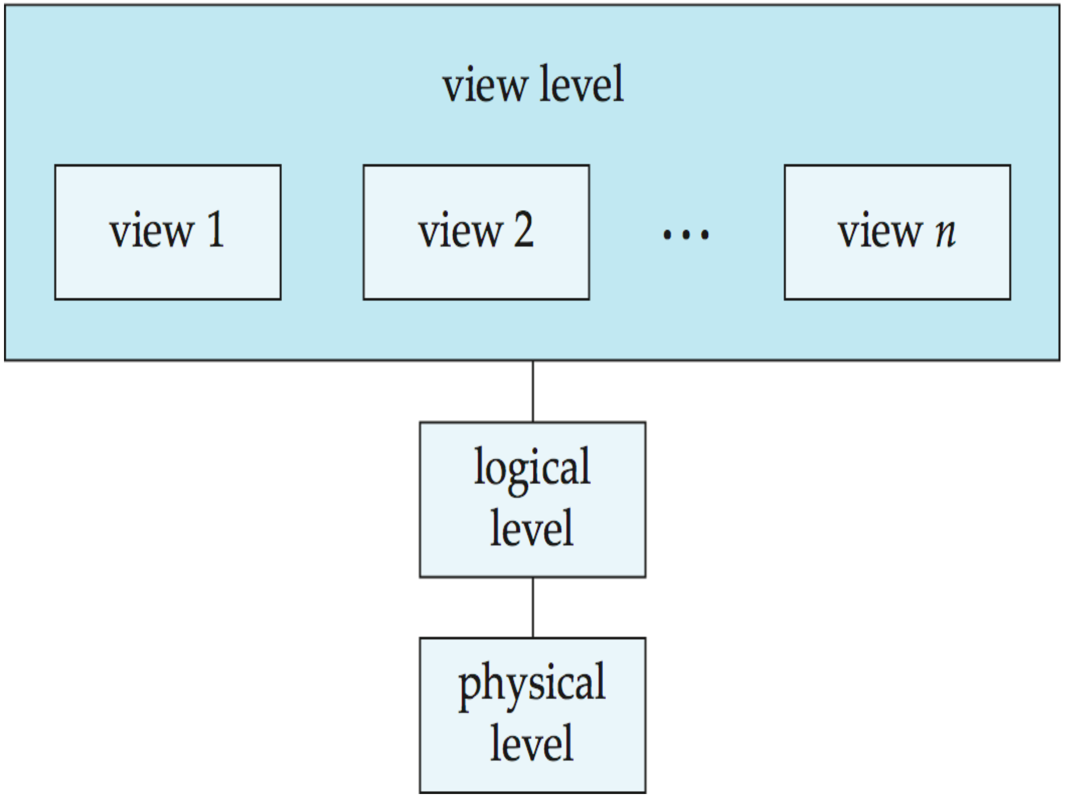
Data Abstraction

Levels of Data Abstraction:

1. **Physical : *How* the data are actually stored**
   - How a tables are stored in the disk? How each and every record of a table in a database are kept together? etc.
2. **Conceptual/Logical: *What* data are actually stored**
   - What are the attributes of a table? What index is used to access a table? etc.
3. **View: Describes only a part of the database**
   - The data stored in a table that could be shown to a particular type of user

Data Abstraction

Modification in any level does not require modification in the above levels



An architecture for a database system

Data-Independence due to Data Abstraction

Two types of data independence is achieve:

bottom to top independence

1. Physical-data independence : Modification in physical level does not require modification in the conceptual level

2. Logical-data independence: Modification in conceptual level does not require modification in the view level

Instances and Schemas

Schema: The overall design of the database

Instance/State: The collection of information stored in the database at a particular moment

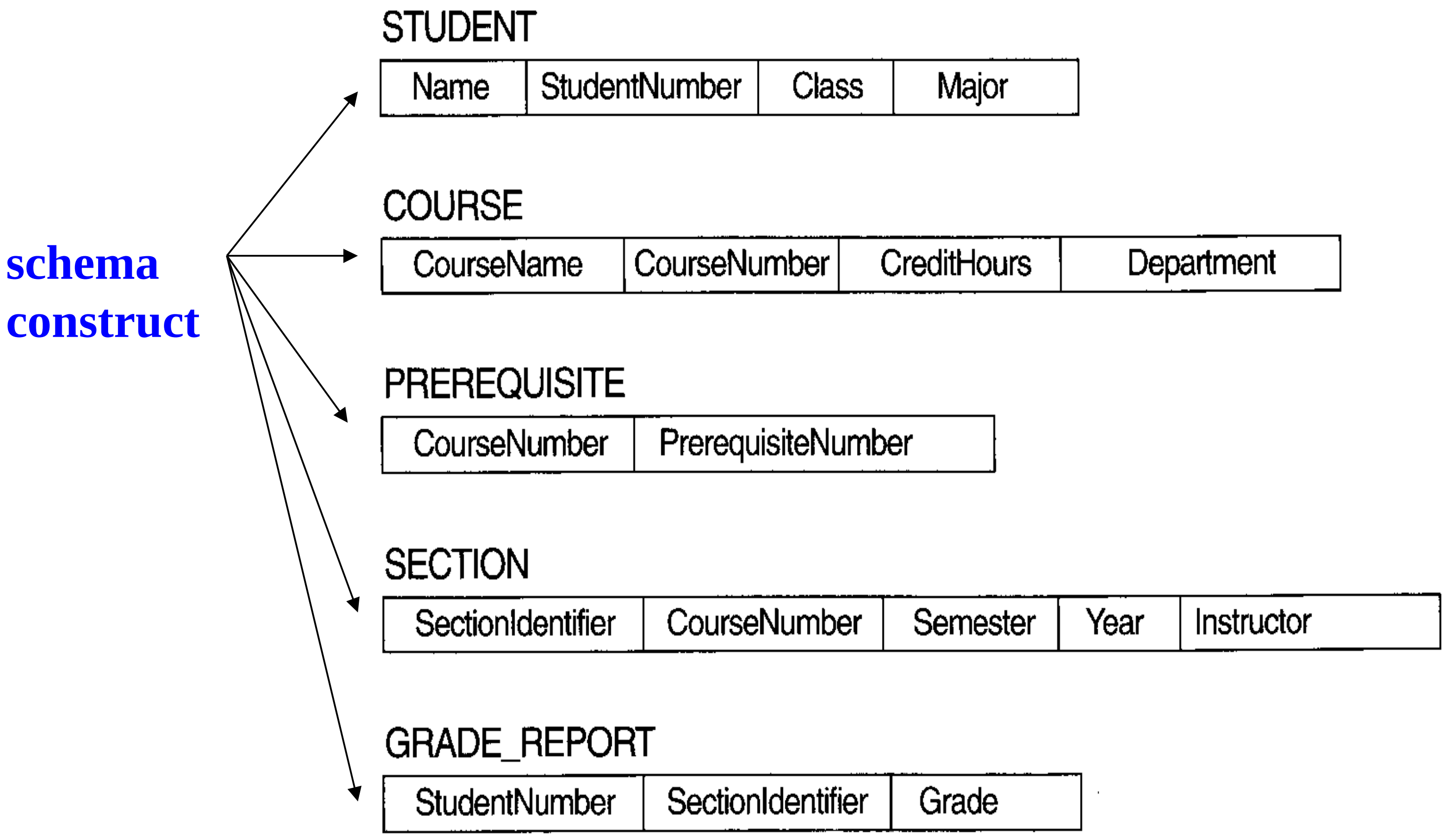
Physical schema:

- database design at the physical level
- remains hidden beneath the logical schema, and can usually be changed easily without affecting application programs

Logical schema:

- database design at the logical level
- programmers construct applications by using the logical schema

Example Schema diagram of an UNIVERSITY database



schema construct

STUDENT

Name	StudentNumber	Class	Major

COURSE

CourseName	CourseNumber	CreditHours	Department

PREREQUISITE

CourseNumber	PrerequisiteNumber

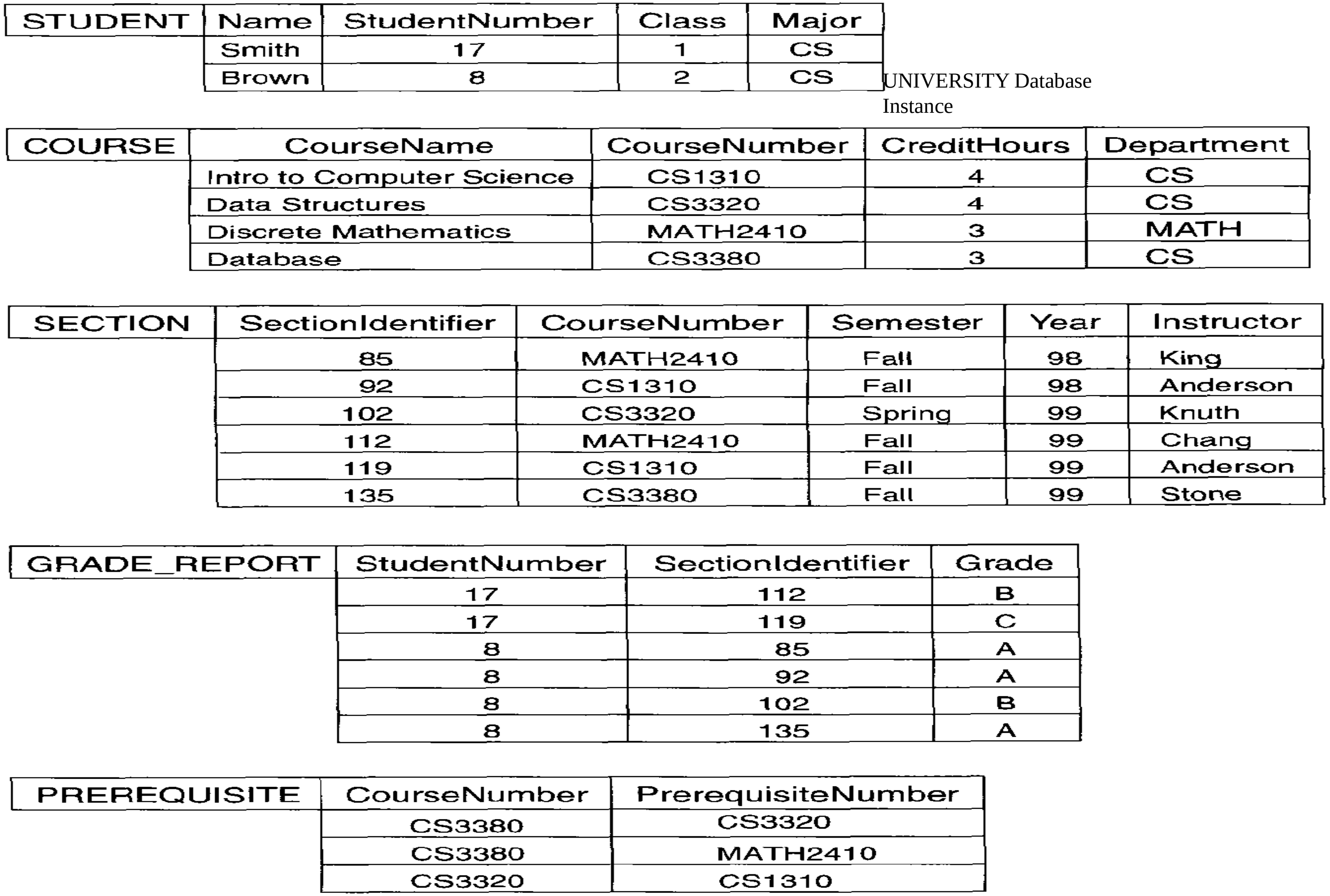
SECTION

SectionIdentifier	CourseNumber	Semester	Year	Instructor

GRADE_REPORT

StudentNumber	SectionIdentifier	Grade

Instance of UNIVERSITY database



STUDENT	Name	StudentNumber	Class	Major
	Smith	17	1	CS
	Brown	8	2	CS

UNIVERSITY Database Instance

COURSE	CourseName	CourseNumber	CreditHours	Department
	Intro to Computer Science	CS1310	4	CS
	Data Structures	CS3320	4	CS
	Discrete Mathematics	MATH2410	3	MATH
	Database	CS3380	3	CS

SECTION	SectionIdentifier	CourseNumber	Semester	Year	Instructor
	85	MATH2410	Fall	98	King
	92	CS1310	Fall	98	Anderson
	102	CS3320	Spring	99	Knuth
	112	MATH2410	Fall	99	Chang
	119	CS1310	Fall	99	Anderson
	135	CS3380	Fall	99	Stone

GRADE_REPORT	StudentNumber	SectionIdentifier	Grade
	17	112	B
	17	119	C
	8	85	A
	8	92	A
	8	102	B
	8	135	A

PREREQUISITE	CourseNumber	PrerequisiteNumber
	CS3380	CS3320
	CS3380	MATH2410
	CS3320	CS1310

Levels of University Database Schema

Conceptual / Logical Schema:

- *Students(sid: string, name: string, login: string,*
  age: integer, gpa:real)
- *Courses(cid: string, cname:string, credits:integer)*
- *Enrolled(sid:string, cid:string, grade:string)*

Physical Schema:

- Relations stored as unordered files.
- Index on first column of Students.

View Level Schema:

- *Course_info(cid:string,enrollment:integer)*