Database Systems

Chapter # 5

The Relational Data Model and Relational Database Constraints

eman.shahid@nu.edu.pk

Chapter Outlines

- **Relational Model Concepts**
- **Relational Model Constraints**
- **Relational Database Schemas**
- **Update Operations**
- **Transactions**
- **Dealing with Constraint Violations**

Relational Model Concepts

- Relational model: represents the database as a collection of relations.
  - Relation: Table of values
  - Tuple represent facts: Row
  - Column Header: Attribute
  - Domain: Data type describing the types of values that can appear in each column.
- The model was first proposed by Dr. E.F. Codd of IBM Research in 1970 in the following paper:“ A Relational Model for Large Shared Data Banks," Communications of the ACM, June 1970

STUDENT

Name	Student_number	Class	Major
Smith	17	1	CS
Brown	8	2	CS

COURSE

Course_name	Course_number	Credit_hours	Department
Intro to Computer Science	CS1310	4	CS
Data Structures	CS3320	4	CS
Discrete Mathematics	MATH2410	3	MATH
Database	CS3380	3	CS

SECTION

Section_identifier	Course_number	Semester	Year	Instructor
85	MATH2410	Fall	07	King
92	CS1310	Fall	07	Anderson
102	CS3320	Spring	08	Knuth
112	MATH2410	Fall	08	Chang
119	CS1310	Fall	08	Anderson
135	CS3380	Fall	08	Stone

GRADE_REPORT

Student_number	Section_identifier	Grade
17	112	B
17	119	C
8	85	A
8	92	A
8	102	B
8	135	A

PREREQUISITE

Course_number	Prerequisite_number
CS3380	CS3320
CS3380	MATH2410
CS3320	CS1310

Figure 1.2
A database that stores student and course information.

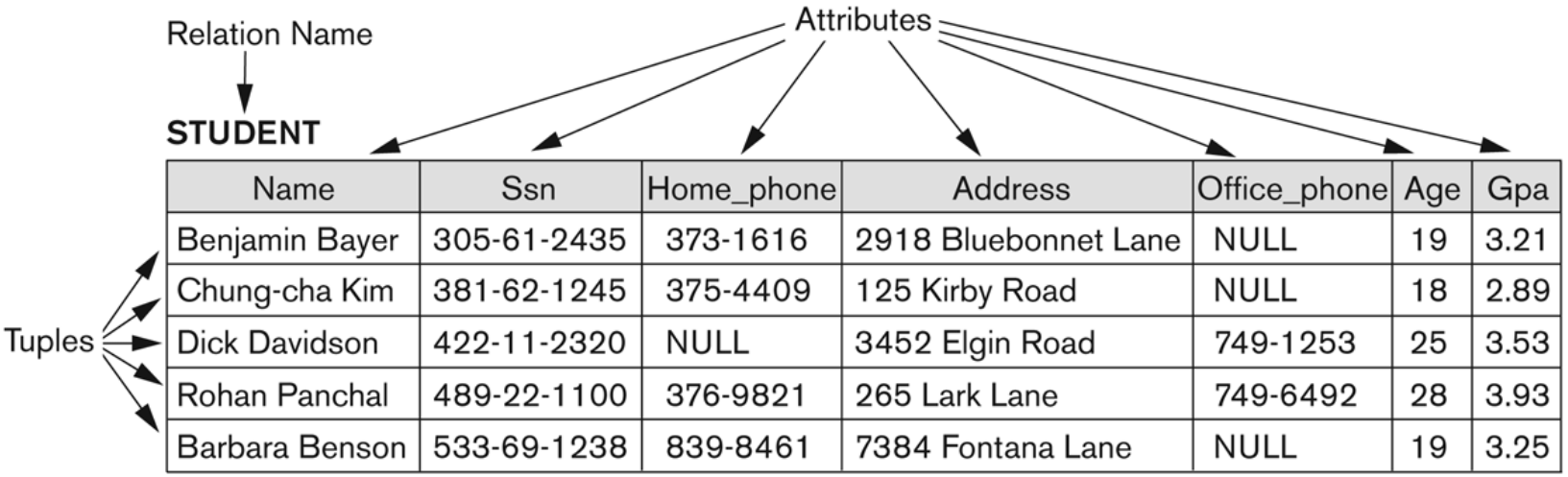
Relational Model Concepts

- **Domain:**
  - Set of atomic values
- **Represents data type, and format.**
  - For example, the PAK_phone_numbers may have a format: (ddd)-ddddddddd where each d is a decimal digit.
  - Dates have various formats such as month, date, year or yyyy-mm-dd, or dd mm,yyyy etc.
  - Names: set of character strings that represent names of persons.
  - Grade_point_averages: floating-point number between 0 and 4.
  - Employee_ages: integer value between 15 and 80.

Formal Definition of Relation

- A relation schema.
  - denoted by R(A1, A2, … , An),
  - made up of a relation name R
  - Contains a list of attributes, A1, A2, … , An.
  - Each attribute Ai is the name of a role played by some domain D in the relation schema R. D is called the domain of Ai and is denoted by dom(Ai).
- Degree (or arity) of a relation: number of attributes/columns
- Tuple t is a row:
  - Ordered/ unordered list of n values t =<v1, v2, … , vn>
  - each value vi, $1 \leq i \leq n$, is an element of dom (Ai) or is a special NULL value.

Example of a Relation



Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

Figure 5.1
The attributes and tuples of a relation STUDENT.

Degree of Relation?

Relation Definition Using set theory

- A relation (or relation state) r(R) is a mathematical relation of degree n on the domains dom(A1), dom(A2), … , dom(An), which is a subset of the Cartesian product (denoted by ×) of the domains that define
  - R: r(R) ⊆ (dom(A1) × dom(A2) × . . . × (dom(An))
- **Cartesian product:** All possible combinations of values from the underlying domains.
- **Total number of tuples in the Cartesian product:** |dom(A1)| × |dom(A2)| × . . . × |dom(An)|

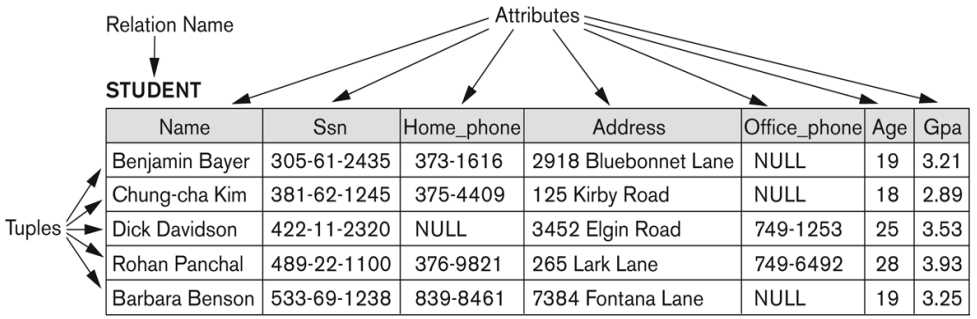
Relational Model Concepts

- Relation State: The state of the whole database will correspond to the states of all its relations at a particular point in time.

Informal Terms	Formal Terms
Table	Relation
Column	Attribute/Domain
Row	Tuple
Values in a column	Domain
Table Definition	Schema of a Relation
Populated Table	Extension

Characteristics of Relations

- Relation: set of tuples.
  - Ordering of Tuples in a Relation: tuples in a relation do not have any particular order.
- Hence, the relation displayed in Figure 5.2 is considered identical to the one shown in Figure 5.1



Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

Figure 5.1
The attributes and tuples of a relation STUDENT.

Figure 5.2
The relation STUDENT from Figure 5.1 with a different order of tuples.

STUDENT

Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	(817)749-1253	25	3.53
Barbara Benson	533-69-1238	(817)839-8461	7384 Fontana Lane	NULL	19	3.25
Rohan Panchal	489-22-1100	(817)376-9821	265 Lark Lane	(817)749-6492	28	3.93
Chung-cha Kim	381-62-1245	(817)375-4409	125 Kirby Road	NULL	18	2.89
Benjamin Bayer	305-61-2435	(817)373-1616	2918 Bluebonnet Lane	NULL	19	3.21

Characteristics of Relations

- Ordering of Attributes is not important, because the attribute name appears with its value.

$t = <$ (Name, Dick Davidson),(Ssn, 422-11-2320),(Home_phone, NULL),(Address, 3452 Elgin Road), (Office_phone, (817)749-1253),(Age, 25),(Gpa, 3.53)>

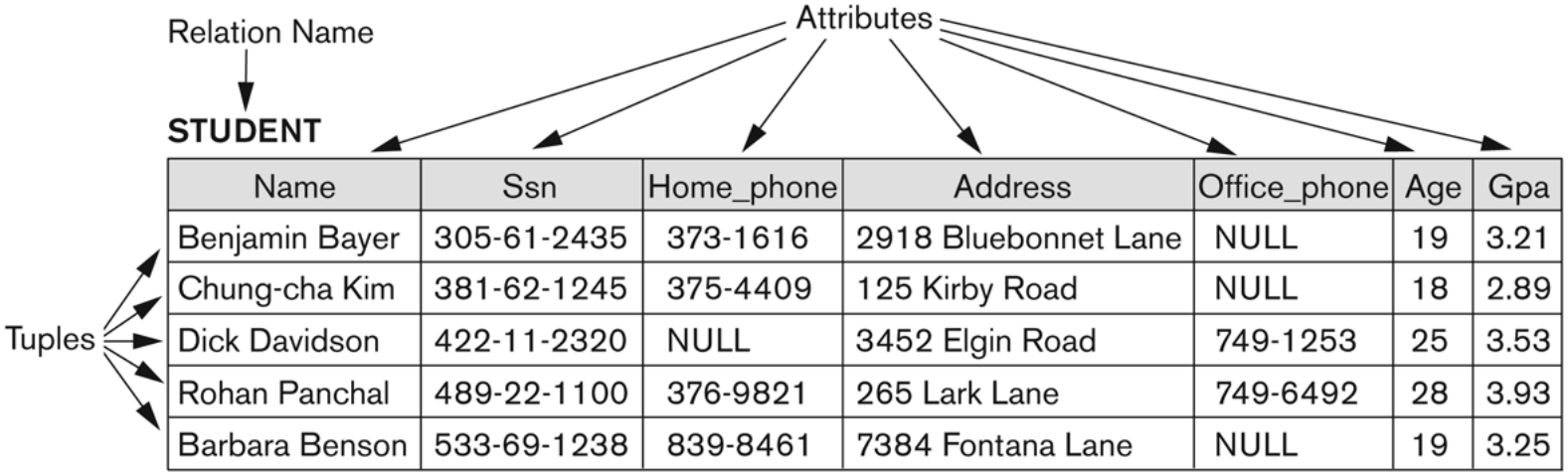
$t = <$ (Address, 3452 Elgin Road),(Name, Dick Davidson),(Ssn, 422-11-2320),(Age, 25), (Office_phone, (817)749-1253),(Gpa, 3.53),(Home_phone, NULL)>

Figure 5.3
Two identical tuples when the order of attributes and values is not part of relation definition.

- When the attribute name and value are included together in a tuple, it is known as self-describing data.

Characteristics of Relations

- **Values and NULLs in the Tuples:**
  - Each value in a tuple is an atomic value.
  - composite and multivalued attributes are not allowed
- NULL values: values of attributes that may be unknown or may not apply to a tuple.



Name	Ssn	Home_phone	Address	Office_phone	Age	Gpa
Benjamin Bayer	305-61-2435	373-1616	2918 Bluebonnet Lane	NULL	19	3.21
Chung-cha Kim	381-62-1245	375-4409	125 Kirby Road	NULL	18	2.89
Dick Davidson	422-11-2320	NULL	3452 Elgin Road	749-1253	25	3.53
Rohan Panchal	489-22-1100	376-9821	265 Lark Lane	749-6492	28	3.93
Barbara Benson	533-69-1238	839-8461	7384 Fontana Lane	NULL	19	3.25

Figure 5.1
The attributes and tuples of a relation STUDENT.

Relational Model Constraints

- **Constraints:** set of rules or restrictions. They are used to maintain the quality of data.
- Constraints ensure that the data insertion, updating, and other processes have to be performed in such a way that data integrity is not affected.
- Constraints are used to guard against accidental damage to the database.

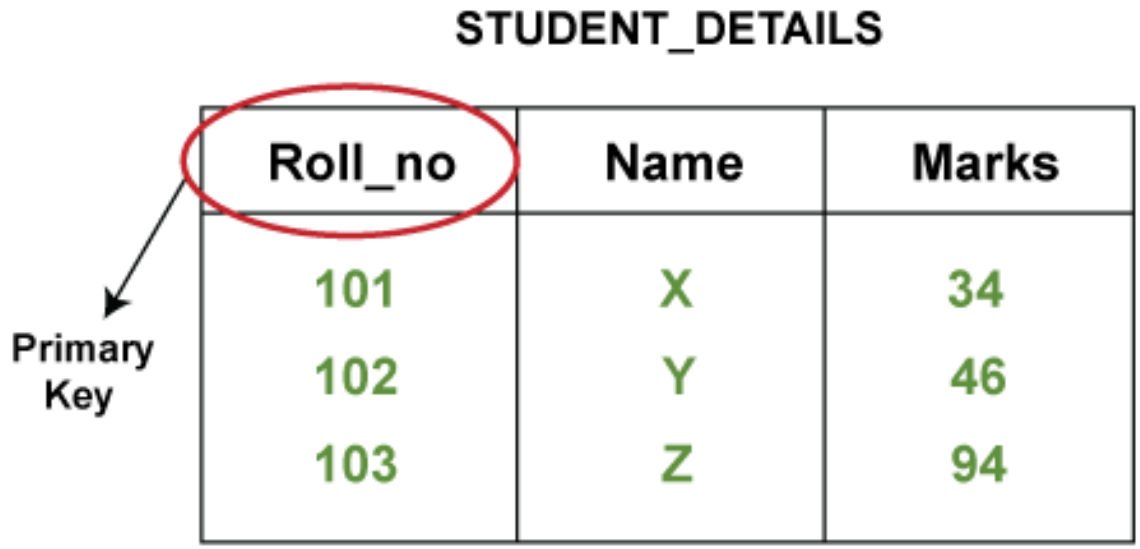
Relational Model Constraints

- Schema Based Constraints
  - Divided into three main categories.
    - **Key** constraints
    - **Entity integrity** constraints
    - **Referential integrity** constraints
- Key Constraints: used to identify an entity within its entity set uniquely.
  - Primary Key
  - Candidate Key
  - Super Key
  - Foreign Key

Key Constraints

- **Primary key:** uniquely identifies each record in a table. It must have unique values and cannot have null values.
- **FOREIGN KEY:** Foreign keys are the columns of a table that points to the primary key of another table. They act as a cross-reference between tables.
- **Candidate Key:** A super key with no redundant attribute is known as candidate key.
- **Super Key:** A super key is a set of one of more columns (attributes) to uniquely identify rows in a table.

Primary Key Examples



Candidate Key Example

Emp_Id	Emp_Number	Emp_Name
E01	2264	Steve
E22	2278	David
E23	2288	Joseph
E45	2290	Robert

The **candidate keys** we have selected are:
{Emp_Id}
{Emp_Number}

Super Key Example

Emp_Id	Emp_Number	Emp_Name
E01	2264	Steve
E22	2278	David
E23	2288	Joseph
E45	2290	Robert

How many super keys the above table can have?

1. {Emp_Id, Emp_Number}
2. {Emp_Id, Emp_Name}
3. {Emp_Id, Emp_Number, Emp_Name}
4. {Emp_Number, Emp_Name}

Foreign Key Example

Course_enrollment table:

Course_Id	Stu_Id
C01	101
C02	102
C03	101
C05	102
C06	103
C07	102

Student table:

Stu_Id	Stu_Name	Stu_Age
101	Chaitanya	22
102	Arya	26
103	Bran	25
104	Jon	21

Entity Integrity Constraint

- **Entity integrity constraint:** primary key value can't be null.
- A table can contain a null value other than the primary key field.

EMPLOYEE

EMP_ID	EMP_NAME	SALARY
123	Jack	30000
142	Harry	60000
164	John	20000
	Jackson	27000

Not allowed as primary key can’t contain a NULL value

Referential integrity constraint

- **Referential integrity constraint is** specified between two tables.
  - Foreign key in Relation 1 refers to the Primary Key of Relation 2.
- Used to specify a relationship among tuples in two relations: the referencing relation and the referenced relation.
- A referential integrity constraint can be displayed in a relational database schema as a directed arc from R1.FK to R2.PK

Relational Model Constraints

- **Domain Constraints:**
  - Value of each attribute must be an atomic value from the domain dom(A).
  - The data type of domain includes string, character, integer, time, date, currency, etc. The value of the attribute must be available in the corresponding domain.

ID	NAME	SEMENSTER	AGE
1000	Tom	1st	17
1001	Johnson	2nd	24
1002	Leonardo	5th	21
1003	Kate	3rd	19
1004	Morgan	8th	A

Not allowed. Because AGE is an integer attribute