The Relational Data Model and Relational Database Constraints

Part 1

Relational Model Concepts

- When a relation is thought of as a **table** of values, each row in the table represents a collection of related data values.
- A row represents a fact that typically corresponds to a real-world entity or relationship.
- The table name and column names are used to help to interpret the meaning of the values in each row.
- All values in a column are of the same data type.

- In the formal relational model terminology, a row is called a *tuple*, a column header is called an *attribute*, and the table is called a *relation*.
- The data type describing the types of values that can appear in each column is represented by a *domain* of possible values.

Domains, Attributes, Tuples, and Relations

- A **domain** *D* is a set of atomic values.
- By **atomic** we mean that each value in the domain is indivisible as far as the formal relational model is concerned.
- A common method of specifying a domain is to specify a data type from which the data values forming the domain are drawn.
- It is also useful to specify a name for the domain, to help in interpreting its values.
- Some examples of domains follow:
  - `Usa_phone_numbers`
  - `Local_phone_numbers`
  - `Social_security_numbers`
  - `Names`
  - `Grade_point_averages`
  - `Employee_ages`
  - `Academic_department_names`
  - `Academic_department_codes`

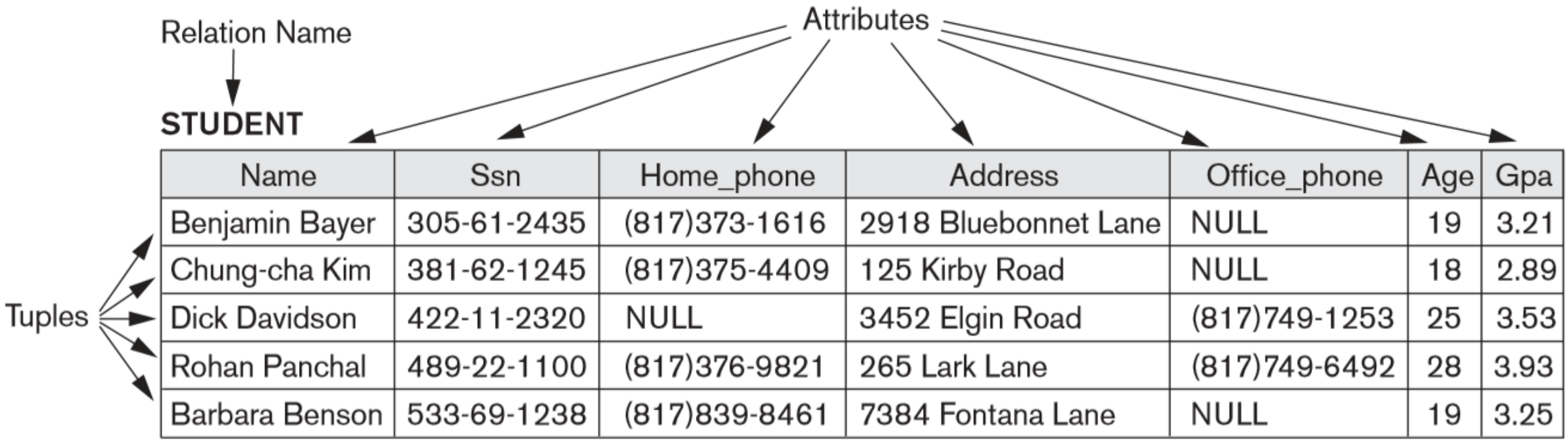
- A **relation schema** *R*, denoted by $R(A_1,A_2,...,A_n)$, is made up of a relation name *R* and a list of attributes, A_1, A_2, ..., A_n.
- Each attribute A_i is the name of a role played by some domain *D* in the relation schema *R*.
- *D* is called the domain of A_i and is denoted by **dom**($\mathbf{A_i}$).
- A relation schema is used to *describe* a relation; *R* is called the **name** of this relation.
- The **degree** (or **arity**) of a relation is the number of attributes *n* of its relation schema.

- A relation of degree seven, which stores information about university students, would contain seven attributes describing each student, as follows:
  - `STUDENT(Name, Ssn, Home_phone, Address, Office_phone, Age, Gpa)`
- Using the data type of each attribute, the definition is sometimes written as:
  - `STUDENT(Name: string, Ssn: string, Home_phone: string, Address: string, Office_phone: string, Age: integer, Gpa: real)`

- More precisely, we can specify the following previously defined domains for some of the attributes of the `STUDENT` relation:
  - dom(`Name`) = `Names`
  - dom(`Ssn`) = `Social_security_numbers`
  - dom(`HomePhone`) = `USA_phone_numbers`
  - dom(`Office_phone`) = `USA_phone_numbers`
  - dom(`Gpa`) = `Grade_point_averages`

- It is also possible to refer to attributes of a relation schema by their position within the relation; thus, the second attribute of the `STUDENT` relation is `Ssn`, whereas the fourth attribute is `Address`.

- A **relation** (or **relation state**) *r* of the relation schema $R(A_1, A_2, ..., A_n)$, also denoted by *r(R)*, is a set of *n*-tuples $r = \{t_1, t_2, ..., t_m\}$.
- Each *n*-tuple *t* is an ordered list of *n* values $t = <v_1, v_2, ..., v_n>$, where each value v_i, $1 \leq i \leq n$, is an element of dom(A_i) or is a special NULL value.
- The i^{th} value in tuple *t*, which corresponds to the attribute A_i, is referred to as $t[A_i]$ or $t.A_i$ (or $t[i]$ if we use the positional notation).
- The terms **relation intension** for the schema *R* and **relation extension** for a relation state *r(R)* are also commonly used.



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

- The earlier definition of a relation can be restated more formally using set theory concepts as follows.
- A relation (or relation state) *r(R)* is a **mathematical relation** of degree *n* on the domains dom(A_1), dom(A_2), ..., dom(A_n), which is a subset of the **Cartesian product** (denoted by ×) of the domains that define *R*:

  $r(R) \subseteq (\text{dom}(A_1) \times \text{dom}(A_2) \times \ldots \times \text{dom}(A_n))$

Characteristics of Relations

- **Ordering of Tuples in a Relation**
  - A relation is defined as a *set* of tuples.
  - Mathematically, elements of a set have *no order* among them; hence, tuples in a relation do not have any particular order.
  - In other words, a relation is not sensitive to the ordering of tuples.

- **Ordering of Values within a Tuple and an Alternative Definition of a Relation**
  - According to the preceding definition of a relation, an *n*-tuple is an *ordered list* of *n* values, so the ordering of values in a tuple—and hence of attributes in a relation schema—is important.
  - However, at a more abstract level, the order of attributes and their values is *not* that important as long as the correspondence between attributes and values is maintained.
  - A **tuple** can be considered as a **set** of (<attribute>, <value>) pairs, where each pair gives the value of the mapping from an attribute A_i to a value v_i from dom(A_i).

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

$t = <$ (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)>

- **Values and NULLs in the Tuples**
  - Each value in a tuple is an **atomic** value; that is, it is not divisible into components within the framework of the basic relational model.
  - Hence, composite and multivalued attributes are not allowed.
  - In general, we can have several meanings for NULL values, such as **value unknown**, **value** exists but is **not available**, or **attribute does not apply** to this tuple(also known as **value undefined**).
  - The exact meaning of a NULL value governs how it fares during arithmetic aggregations or comparisons with other values.
  - During database design, it is best to avoid NULL values as much as possible.

- **Interpretation (Meaning) of a Relation**
  - The relation schema can be interpreted as a declaration or a type of **assertion**.
  - An alternative interpretation of a relation schema is as a **predicate**.