Relational Data Model – Part 1

Schema and State

What is the Relational Data Model?

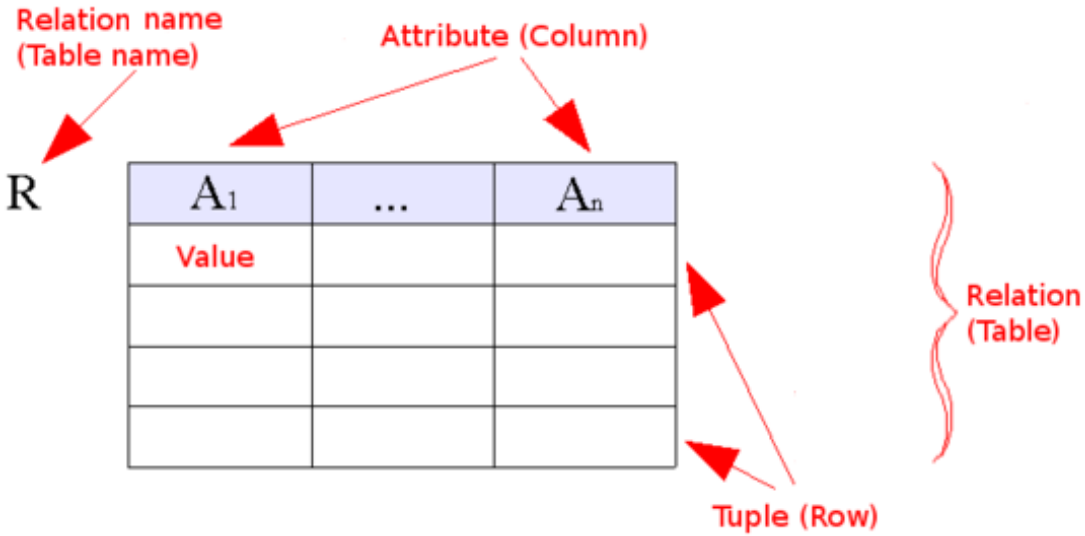
- Introduced by Edgar F. Codd of IBM Research in 1970.

  "A Relational Model for Large Shared Data Banks", Communications of the ACM.

  - A database contains tables (called relations), and each table is made up of columns and rows.
  - Humans have used tables for centuries to keep track of data.

- Used as the standard for relational DBMSs (e.g., Oracle, IBM DB2, Microsofts Access, Microsofts SQL Server, MySQL, postgreSQL, etc.).

Relation



- Correspondence of informal and formal terms:

Informal Terms	Formal Terms
Table	Relation
Column	Attribute
Data type	Domain
Row	Tuple
Table definition	Relation schema

The Basics

- **Attributes** are used to describe the properties of information. In the relational model, they usually refer to atomic data.

  Example: To capture the information of a person, we can use attributes like Name, Age, Gender, Address and PhoneNumber.

- **Domains** are the sets of all possible values for attributes.

  Example:
  - $\texttt{STRING} = \{A, B, CD, ...\}$;
  - $\texttt{DATE} = \{01/01/2005, 03/07/1978, ...\}$;
  - $\texttt{INT} = \{..., -1, 0, 1, 2, ...\}$.

- Recall that, **Cartesian product** $D_1 \times ... \times D_n$ is the set of all possible combinations of values from the sets $D_1, ..., D_n$.

  Example: Let D_1={book,pen}, D_2={1,2} and D_3={red}. Then

  - $D_1 \times D_2 \times D_3$ ={(book,1,red),(book,2,red),(pen,1,red),(pen,2,red)}

The Basics

- The attributes are StudentID, CourseNo, Semester, Status and EnrolDate.
- The domains of attributes are as follows.
  dom(StudentID)=`INT`; dom(CourseNo)=`STRING`;
  dom(Semester)=`STRING`; dom(Status)=`STRING`;
  dom(EnrolDate)=`DATE`.
- The whole table can be considered as a set {(456, COMP2400, 2016 S2, active, 25/05/2016), (458, COMP1130, 2016 S1, active, 20/02/2016), (459, COMP2400, 2016 S2, active, 11/06/2016)}.

ENROL				
StudentID	CourseNo	Semester	Status	EnrolDate
456	COMP2400	2016 S2	active	25/05/2016
458	COMP1130	2016 S1	active	20/02/2016
459	COMP2400	2016 S2	active	11/06/2016

- Is the above set a subset of

$$\texttt{INT} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{DATE}?$$

Answer: Yes.

The Basics

- A **relation schema** has a **relation name** and a list of **attributes**.
- Each attribute is associated with a **domain**.
- A relation schema can be expressed by
  - $R(A_1, ..., A_n)$, or
  - $R(A_1 : dom(A_1), ..., A_n : dom(A_n))$,

  where $A_1, .., A_n$ are attributes of R and $dom(A_i)$ is the domain of A_i.

  Example: The relation schema in the previous example is

  - ENROL(StudentID, CourseNo, Semester, Status, EnrolDate), or
  - ENROL(StudentID: `INT`, CourseNo: `STRING`, Semester: `STRING`, Status: `STRING`, EnrolData: `DATE`).

The Basics

- Let $R(A_1, ..., A_n)$ be a relation schema.
- A **tuple** in R is a list t of values, i.e., $t \in dom(A_1) \times ... \times dom(A_n)$.

  Example: The previous example has the following tuples:

  - (456, COMP2400, 2016 S2, active, 25/05/2016)$\in$ $\texttt{INT} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{DATE}$.
  - (458, COMP1130, 2016 S1, active, 20/02/2016)$\in$ $\texttt{INT} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{DATE}$.
  - (459, COMP2400, 2016 S2, active, 11/06/2016)$\in$ $\texttt{INT} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{DATE}$.

- A **relation** $r(R)$ is a set of tuples $r(R) \subseteq dom(A_1) \times ... \times dom(A_n)$.

  Example: The previous example has the following relation:

  - $r(\textsc{Enrol}) \subseteq \texttt{INT} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{STRING} \times \texttt{DATE}$.

The Basics

- A **relational database schema** S is
  - a set of relation schemas $S = \{R_1, \ldots, R_m\}$, and
  - a set of integrity constraints IC.
- A **relational database state** of S is a set of relations such that
  - there is just one relation for each relation schema in S, and
  - all the relations satisfy the integrity constraints IC.

The Basics

- Consider a relational database schema StuEnrol that has three relation schemas:
  - Student(StudentID, Name, DoB, Email).
  - Course(No, Cname, Unit);
  - Enrol(StudentID, CourseNo, Semester, Status, EnrolDate);

Student			
StudentID	Name	DoB	Email

Course		
No	Cname	Unit

Enrol				
StudentID	CourseNo	Semester	Status	EnrolDate

- That is, StuEnrol={Student, Course, Enrol}.

The Basics

- Relational Database State – Example

STUDENT			
StudentID	Name	DoB	Email
456	Tom	25/01/1988	tom@gmail.com
458	Peter	23/05/1993	peter@gmail.com
459	Fran	11/09/1987	frankk@gmail.com

COURSE		
No	Cname	Unit
COMP1130	Introduction to Advanced Computing I	6
COMP2400	Relational Databases	6

ENROL				
StudentID	CourseNo	Semester	Status	EnrolDate
456	COMP2400	2016 S2	active	25/05/2016
458	COMP1130	2016 S1	active	20/02/2016
459	COMP2400	2016 S2	active	11/06/2016