

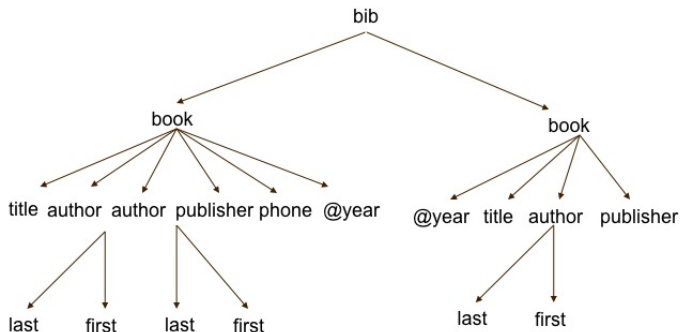
Data Model

What is a Data Model?

A **data model** is a collection of concepts for describing data. It consists of three parts:

- The **structure** of a database. For example:
 - relational model: tables
 - semistructured model: trees/graphs (XML)
- The **operations** for manipulating these structures: (retrieving, modifying...)
- The **constraints** that a database should obey (e.g. a week day should be an integer between 1-7)

Tree Structure Example



XML Representation

XML input: www.a.b/bib.xml

```
<book year="1996">
  <title> HTML </title>
  <author> <last> Lee </last> <first> T. </first></author>
  <author> <last> Smith</last> <first>C.</first></author>
  <publisher> Addison-Wesley </publisher>
  <price> 59.99 </price>
</book>
<book year="2003">
  <title> WMD </title>
  <author> <last> Bush</last> <first> G.</first></author>
  <publisher> white house </publisher>
</book>
```

Relational Data Model

- The Relational Data Model is based on the concept of a Relation.
 - The strength of the relational approach to data management comes from the formal foundation provided by the theory of relations.
- 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.
- The above paper caused a major revolution in the field of database management and earned Dr. Codd the ACM Turing Award.

Why Relational Data Model

- Used by all major commercial database systems.
- Very simple model, often matches how we think about data.
- Considered as foundation of Database Management Systems.

Relational Database – Informal Definition

- A **relational database** is a set of relations.
- A **relation** is a table consisting of a “set” of tuples (a.k.a. rows); the order of tuples in a relation does not matter.
- A **data value** in each row represents certain facts that correspond to a real-world entity or relationship.
- Each **column** has a proper name that reflects the meaning of the data items for that column.

An Example of Relation

Table: Student

Student ID	name	major	GPA	College
100	Michael	CS	3.7	SU
101	Owen	Music	3.4	SU
102	Ryan	CS	3.9	SU

attributes (a.k.a. columns): *student ID, name, major, GPA, and college.*

tuple (a.k.a. row): *101, Owen, Music, 3.4, and SU.*

Relational Database – Formal Definition

Definition (Schema)

The **schema** of a relation is denoted by $R(A_1, A_2, \dots, A_n)$, where R is the name of the relation, A_1, A_2, \dots, A_n are the **attributes** of the relation, and each attribute has a **domain** name (data type) for a set of valid values.

Question

Identify relation name, attributes, and tuples.

Student

Student ID	name	GPA	College
100	Michael	3.7	SU
101	Owen	NULL	SU
102	Ryan	3.9	SU

Relation Schema

Relation Schema = relation name and attribute list. (Optionally: data types of attributes.)

Movies (title, year, length, genre)

Movies (title:string, year:integer,
length:integer, genre:string)

Question

How to describe a database in terms of relations (tables)?

Database Schema

How do you describe a **database schema** as opposed to a **relation schema**?

Example of a Database Schema

Student(StudentName, StudentNum, Class, Major)

Course(CourseName, CourseNum, CreditHours, Department)

Prerequisite(CourseNum, PrerequisiteNumber)

Section(SectionNum, CourseNum, Semester, year, Instructor)

GradeReport(StudentNum, SectionNum, Grade)

Characteristics Of Relations

- All values are considered **atomic** (indivisible).
- A special **null** value is used to represent values that are unknown or inapplicable

Student ID	name	GPA	College
100	Michael	3.7	SU
101	Owen	NULL	SU
102	Ryan	3.9	SU

Characteristics Of Relations (Cont')

- no duplicate tuples in a relation (a set of tuples)
- no ordering of tuples in a relation

name	Student ID	GPA	College
Michael	100	3.7	SU
Owen	101	NULL	SU
Ryan	102	3.9	SU

Student ID	name	GPA	College
100	Michael	3.7	SU
101	Owen	NULL	SU
102	Ryan	3.9	SU

Summary of Terminologies for a Relational Database

- Relation
- Database
- Relation Schema
- Database Schema
- Attributes
- Tuples
- Instance of a relation