CMPUT 291

What is a Database ?

- database is a very large, integrated collection of data
- Models real-world enterprise
  - Entities (e.g students, courses)
  - Relationship (e.g students **taking** 291)

University	student data, course
Hospital	Patient data, facility
Bank	Account,data

What is a DBMS?

A *Database Management System (DBMS)* is a software package designed to store and manage databases.

数据库管理系统被称为DBMS，是一种操作和管理数据库的大型软件，用于建立 使用和维护数据库，简称DBMS。它对数据库进行统一的管理和控制，以保证数据库的安全和完整性

Non-database Approach

- One data set per program
- programmer defines (and implements) storage structures, access methods, etc.

Why we use DBMS?

- Reduced redundancy
- Less risk of inconsistency (减少数据不一致的风险）
- Uniform data administration （统一数据管理权限）
- concurrent access, recovery from crashes （并行访问，不容易因为数据库崩溃而无法恢复）

- Data independence （数据独立性）

schema：a description of the data contents, structures and other aspects.

（对数据内容、结构和其他方面的描述）

Three schema level

- external schema (view)
  - An external schema describes the part of the database which specific user is interested in. It hides the unrelated details of the database from the user.

    外部层级描述了特定用户感兴趣的数据库的部分。它向用户隐藏了与数据库无关的细节

- conceptual schema (tables)
  - The conceptual schema describes the Database structure of the whole database for the community of users.

    概念层级为用户群体描述了整个数据库的数据库结构。

- Physical schema
  - The internal schema defines the physical storage structure of the database. The Physical schema is a very low-level representation of the entire database.

    内部模式定义了数据库的物理存储结构。物理模式是整个数据库的一个非常低级的表示。

Data independence 分为 Logical and physical data independence 两种

Logical data independence 是防止数据发生逻辑结构上的变化，而 physical data independence是防止数据发生物理结构发生变化。总的来说Data independence是为了数据库不会受储存结构和访问策略发生变化而崩溃 data independence也是作为 DBMS的最大的一个优点。

DBMS的作用：

- **Data definition facilities （数据的定义机制）**
  - **provides a data definition language (DDL) （有对应的数据库语言（DDL））**
  - **stores the definitions in a user-accessible catalog (data dictionary) （在一个用户可访问的目录（数据字典）中存储定义）**
- **Data manipulation facilities （数据的操作机制）**
  - **provides a query language for storing, retrieving and updating data. （提供了一种用于存储、检索和更新数据的查询语言并更新数据。）**
- **Facilities for integrity constraints （完整性约束的机制）**
  - **does the validation check for integrity constraints before updates （在更新前对完整性约束进行验证检查）**
  - **different kinds of constraints (不同种类的有对应的限制)**
    - **primary key constraints (entity integrity) entity的完整性**
    - **foreign key constraints (no dangling references)**
    - **check constraints**

Data model:

- conceptual level
  - hides detail

Relational database: (a set of relations (tables))

- basic idea
  - organize data as a set of tables
  - view each as a set of row

Relation: concise of

- instance (table contents, with row and columns)

- schema (table structure, with name and type of columns)

Domain: the set of values from which the values of an attribute are drawn.

Integrity constrains (ICs)

Primary key Constrains

- A set of field is a key of a relation if it is both
  - unique: no two distinct types can have same values in all key fields
  - Minimal: no subset of a key is a key
- A relation can have more than one key
  - candidate key: all keys of the relation
  - primary key: one defined by DBA
- super key: 1st condition holds but 2nd may not

Foreign key:

- set of fields in one relation that “refers” to a tuple in another relation; a FK must correspond to a key(primary or candidate) of the other relation

View:

A view is just a relation, but we store a definition, rather than a set of tuple.

(Views can be dripped using the **DROP VIEW** command.)

External level:

- Applications can access data through some view
  - Different views of data for different categories of users
  - A view is computed (from data in the conceptual level)
- Mapping from external to conceptual schema is done by DBMS
- Conceptual schema can be changed without changing application

- Referred to as **logical data independence**

ER modelling:

THE world is described in three terms

- entity: a distinguishable object
- entity set: a set if entities of the same type
- Relationship: represents the fact that certain entities are related to each other
- Attribute: describes a property of an entity or a relationship (underlined is unique)