Chapter 2 : Relational Model

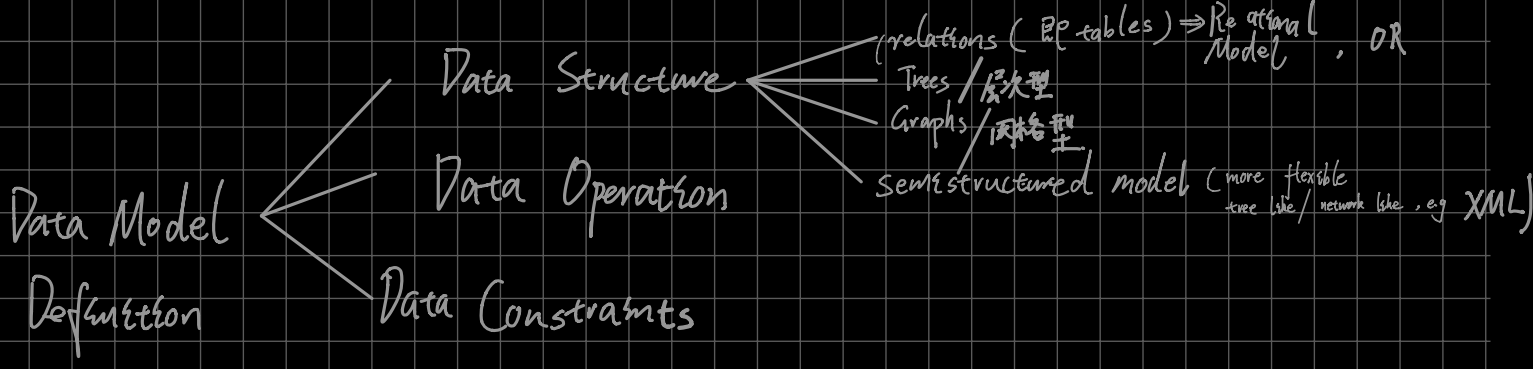
Trivial:

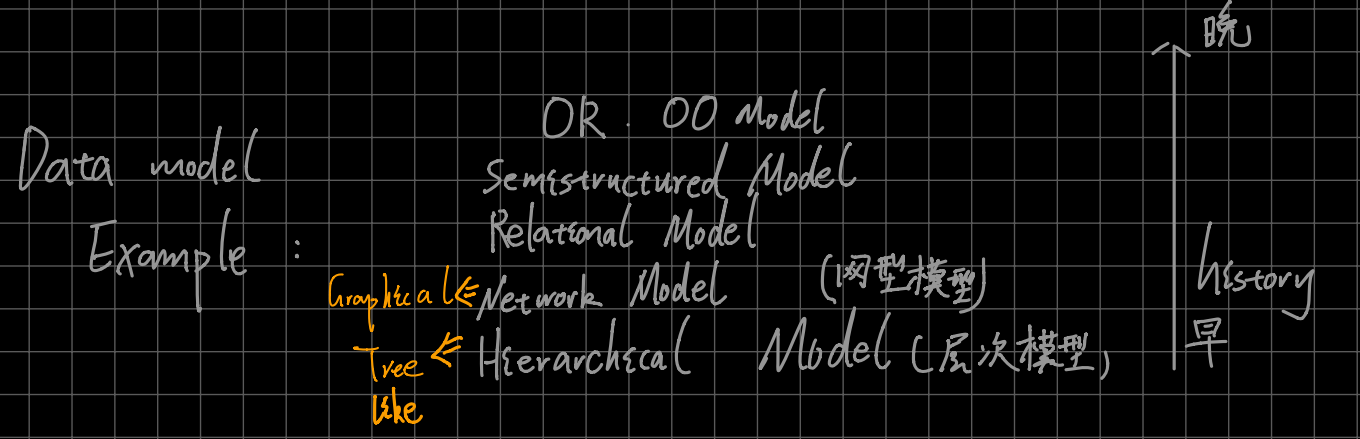
- Data Model Definition
- Relational Model Definition
- 'Key' Definition
- SQL

Key Point:

- Relational Algebra
- Constraints in Relational Algebra

Lecture 2: Relational Data Model





Why Relational Model: Relations即tables, 易理解; operation set is small; effectively used in different occasions

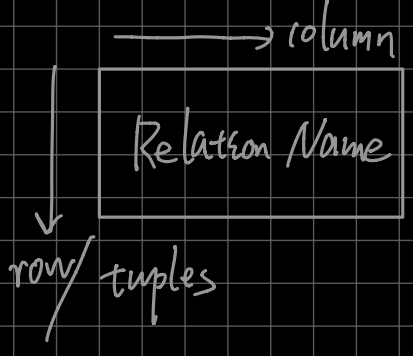
Advantages: 1. simple 2. limited but versatile

3. allow implementation of VHL (e.g. SQL)

Definition of Relation:

$R = \{(a_1, a_2 \cdots a_n) \mid a_i \in S_i\}$，其中 $S_1, S_2 \cdots S_n$ 是各个 attribute 集合.

或 $R \subseteq S_1 \times S_2 \times \cdots S_n$



Relation Schema ⇒ 关系模式 = 关系名 + 属性集合

定义: Relation_name (attribute_list)

attribute_list: set of attributes，即 attributes 之间无序且不重复。

e.g. Spy (code, name: D_2, ID: D_3)

D_2, D_3 → Domains.

△ RDB schema: set of relation schemas
(关系数据库模式 ⇒ 关系模式的集合)

△ 实际中，我们常用 attribute_list 中给出的顺序作为 standard order

Relation Instances:

定义: a set of tuples in a relation. (即关系的一个版本)

e.g current instance (当前实例) ⇒ 数据库系统仅维护当前实例

△ RDB instance: set of relation instance.

SQL:

SQL (Sequel)
- DDL : (Data Definition Language)
- DML : (Data Manipulation Language)
- DCL (Data Control Language)
- TCL (Transaction Control Language)

Relations in SQL:

Used in: → Name:
- Stored Relation ⇒ table (表)
- Calculated Relation ⇒ View (视图), 在需要时被完整/部分地构造
- Query / Updated ⇒ Temporary table (临时表), (constructed by SQL processor)

SQL Basic

Values
- Strings: 'I'm'
- DATE: '2020-09-15' (yyyy-mm-dd)
- TIME: '20:40(:20)' hh:mm(:ss)
- NULL

Data types
- Predefined: implicit type coercions.
- Constructed
  - Array
  - Multiset
  - Ref.
  - Row
- User-defined: 结构体类型 (Structured)

e.g

```
CREATE TABLE Student (
    sno  CHAR(10),
    age  INT  DEFAULT 18,
    id   INT  PRIMARY KEY)
or  id   INT  UNIQUE KEY
```

SQL Operation
- Create Table:
- Drop Table: DROP column. DROP TABLE relation
- Alter Table:

```
ALTER TABLE relation
ADD column type
```

△ Difference between PRIMARY and UNIQUE:

① 1个Relation中PRIMARY只能出现1次，但可以通过独立主键声明一次性指定多个attributes. e.g. PRIMARY(age, id)

UNIQUE可出现多次

② 任意一个tuples的指定为PRIMARY KEY的 attributes 均不能为NULL. 但 UNIQUE 可以。

Lecture 3 : Relational Algebra

Definition :

Algebra : operands + operators → expressions

Relational Algebra : relations + set operators → Relational Algebra expressions.

代表关系的变量

代表有限关系的常量

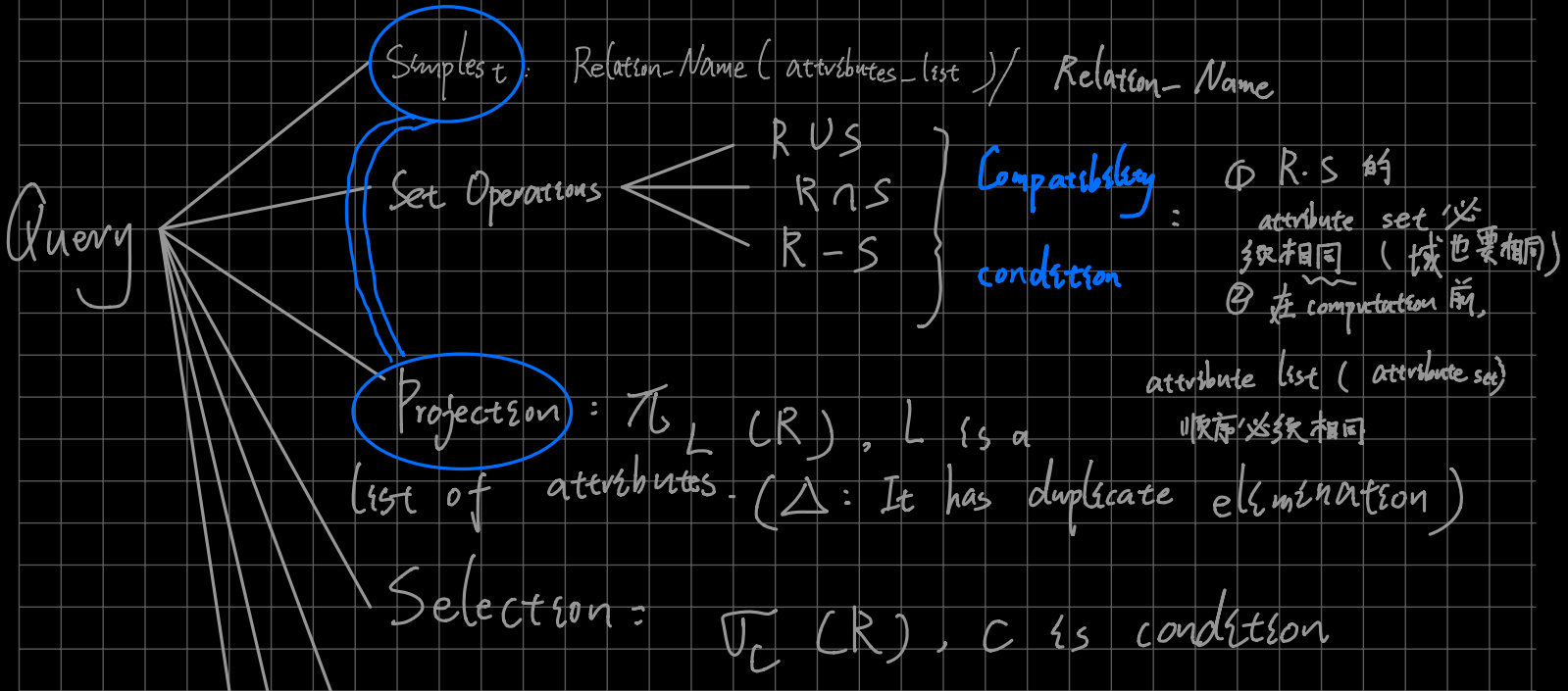
special relational

⇔ Query ⇔ data Operation

△ SQL ⇔

syntactically sugared Relational Algebra (RA) + Relational calculus (RC)

语法糖.



组合操作：

笛卡尔积及笛卡尔积的特殊形式(Join)：在信息孤岛中建立连接的桥梁.

Cartesian Product: $R \times S$, pairing tuples in all possible combinations. (Unrelated tuples is meaningless)

Theta-Join: $R \bowtie_\theta S \Leftrightarrow \sigma_\theta(R \times S)$, θ is an equation (条件)

△ Theta-Join 并不进行合并属性列去重.

Natural Join: $R \bowtie S$, θ = (shared attributes of R and S) (此时有 Duplicate Elimination, 保留一个为 shared attributes 数据)

($\bowtie$ 操作将 R, S 模式中相同属性且属性有相同值的元组配对. ⇒ 相等的共有属性)

Quotient: $T = R \div S$ ($R(X, Y)$, $S(Y)$)

Constraint: S 的属性集 $\subseteq$ R 属性集.

$$T = \{t \mid t \in \pi_X(R) \text{ AND } \forall s \in S\ (ts \in R)\}$$

- t: 结果表中元素
- $\pi_X(R)$: R 属性子集 X, 即除 S 中属性外.
- $\forall s \in S\ (ts \in R)$: 对任意 S 中元素, ts 均 $\in R$. (意思 $S \subseteq R$)

方法：先将 $\pi_Y(R)$ 中所有 $S(Y)$ 中出现的行选出，

再看这些行：$t \in \pi_X(R)$ 是否 $\forall s$, $Xs \in R$.

若是，将该 t 加入 $T(X)$ 中.