

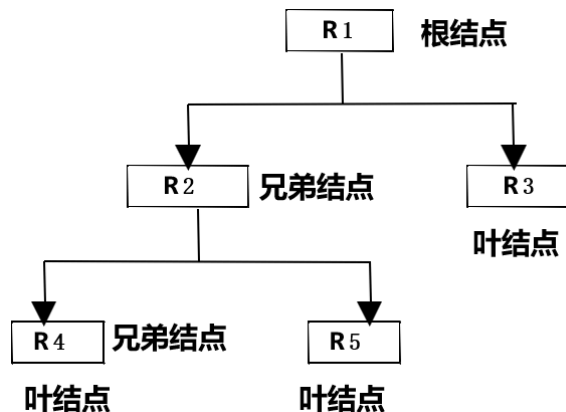
## 第二章 数据模型

### 2.1 串讲归纳

#### Important

本章较为简单但重要，主要介绍了数据库的四大种类，旨在了解数据库发展的大体脉络及数据库的各自特点，文中或有不了解的内容，无妨，之后会详细讲解。

#### 2.1.0 预备知识



#### 2.1.1 数据建模

要把现实世界的事物转化为计算机中可有效管理的数据，需要怎么做？

第一步——将客观事物抽象为概念模型

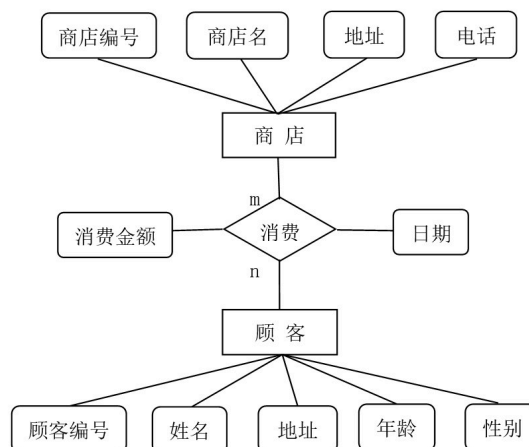
第二步——将概念模型转为数据库管理系统支持的数据模型

#### 2.1.2 数据模型三要素

(1) 数据结构 (2) 数据操纵 (3) 完整性约束

#### 2.1.3 概念模型

直接举例子：（想构建商店消费的一个数据库，首先抽象为下面的概念模型）

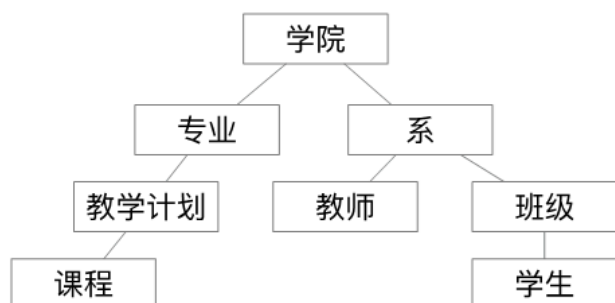


这是概念模型的一种表示方法：**实体-联系模型**

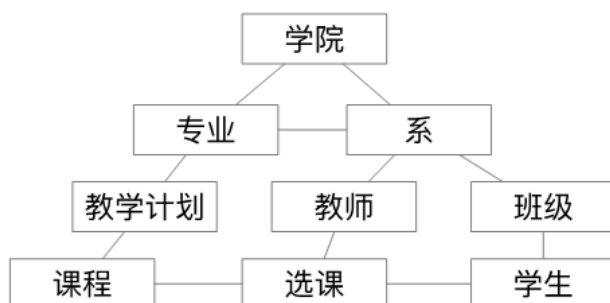
(1) 一些基本概念：实体【顾客、商店】、属性【姓名等】、码（用于唯一标识实体）、实体型（实体的类型特征）、实体集（同一类型的实体集合）、联系（实体内部的联系、实体之间的联系【例子中属于多对多】）

(2) 实体-联系方法——E-R图【例子就是】

## 2.1.4 常见数据模型



(1) 层次数据模型



(2) 网状数据模型

### (1) 层次数据库——树

代表

IBM公司的IMS (Information Management System, IMS) 数据库

特征

- 结点的双亲（也叫父节点）是唯一的
- 只能直接处理一对多的实体联系
- 每个结点表示一个记录类型，记录类型之间的联系用结点之间的连线（有向边）表示
- 各个记录类型、同一记录类型中各个字段不能同名
- 每个记录类型可以定义一个排序字段，也称为码字段
- 任何记录值只有按其路径查看，没有一个子女记录值能够脱离双亲记录值而独立存在

#### 层次模型的完整性约束

- 无相应的双亲结点值就不能插入子女结点值（看父脸色）
- 如果删除双亲结点值，则相应的子女结点值也被同时删除（删父除子）
- 更新操作时，应更新所有相应记录，以保证数据的一致性（一更全更）

优点

简单清晰；查询效率高；完整性约束良好

缺点

分层次不适用；易冗杂（比如一个学生（子）属于n门课程（父），得造n棵树）；操作限制多；查询子结点必须通过父结点。

(2) 网状数据库——图

代表

美国IDS (Integrated Data Storage, 集成数据存储)

特征

作为图，基本无限制，自由得多；层次模型实际上是网状模型的一个特例

优点

能够更直接地描述现实世界（如一个结点可以有多个双亲，结点之间可以有多种联系）；具有良好的性能，存取效率较高。

缺点

结构复杂，随数据增长愈发复杂；语法复杂；记录之间联系是通过存取路径实现，用户必须了解系统结构的细节。

(3) 关系数据库——表

student							
姓名	学号	班级	年龄	性别	住址	课号	电话
张三	100	计91	20	男	上海杨浦	上海	89150
李四	200	计92	19	男	上海徐汇	上海	88888
王五	300	计93	18	女	上海浦东	上海	77777
赵六	400	计94	19	女	上海静安	上海	99999
刘七	500	计95	21	男	上海普陀	上海	88666

course			
课程名	课号	地点	教师
DB	1	5101	周老师
DB	2	5102	钱老师
DM	3	5103	金老师

grade		
学号	课号	分数
100	1	99
200	1	98
300	2	97

代表

Oracle、MySQL等

特征

关系必须是规范化的，满足一定的规范条件（最基本的规范条件：关系的每一个分量必须是一个不可分的数据项, 不允许像下图这样表中还有表）

主修专业	联系方式		
	手机号	Email	微信号
信息安全	18301200745	liyong@ qq. com	liyong@ ruc
⋮	⋮	⋮	⋮

关系模型中的一些基本概念

关系术语	一般表格的术语
关系名	表名
关系模式	表头（表格的描述）
关系	（一张）二维表
元组	记录或行
属性	列
属性名	列名
属性值	列值
分量	一条记录中的一个列值
非规范关系	表中有表（大表中嵌有小表）

完整性约束条件

实体完整性；参照完整性；用户定义的完整性

优点

建立在严格的数学概念的基础上；概念单一；关系模型的存取路径对用户是透明的（具有更高数据独立性，安全保密性，简化了工作）

缺点

查询效率不高

(4) NoSQL数据库——非关系型

*i*

Note

NoSQL = Not only SQL = Structured Query Language

包括键值、列族、文档、图、向量 数据库

优势：模式灵活、高可拓展、高可用

2.2 题库

- 1【单选题】不是关系模型中的完整性约束的是(C)
- A、实体完整性B、用户定义的完整性C、属性完整性D、参照完整性
- 2【单选题】下列不是基本关系具有的性质的是(D)
- A、分量必须取原子量B、不同的列可以出自同一个域C、列和行的顺序无所谓D、元组间的候选码可以相同
- 3【单选题】概念模型是现实世界的第一层抽象，这一类模型中，最著名的模型是(D)

A、层次模型B、网状模型C、关系模型D、实体-联系模型

4【单选题】下列哪个名字定义为“一组具有相同数据类型的值的集合”?(A)

A、域B、元组C、基数D、分量

5【单选题】在(C)中,一个结点可以有多个双亲,结点之间可以有多种联系。

A、以上都有B、层次模式C、网状模型D、关系模型

6.【单选题】下列属于键值(Key-Value)存储数据库的是( C)

A、MongoDBB、HBaseC、RedisD、Neo4J

7.【单选题】下列属于文档型数据库的是(D)

A、HBaseB、Neo4JC、RedisD、MongoDB

#### 💡 Tip

主要类型	NoSQL数据库产品
键值数据库	Redis, Memcached ( 内存键值存储 )
	Dynamo ( 最终一致性键值存储 )
	LevelDB、RocksDB (持久化键值存储)
列族数据库	Cassandra, Hbase, Google BigTable
🍪 文档数据库	MongoDB 【之后会介绍】
图数据库	Neo4J
向量数据库	Milvus

8.【单选题】关系数据库的数据模型是B\_\_。

A、结构模型B、关系模型C、网状模型D、层次模型

9.【单选题】关系模型中,码是(C)。

A、以上都不是B、可以由多个任意属性组成C、由一个或多个属性组成,其值能够惟一标识关系中一个元组D、至多由一个属性组成

10.【单选题】下面不属于关系数据库管理系统的是(A)

A、IMSB、MySQLC、SQL ServerD、Oracle

11.【单选题】(A)的存取路径对用户透明,从而具有更高的数据独立性,更好的安全保密性,也简化了程序员的工作和数据库开发建立的工作。

A、网状模型B、关系模型C、以上都有D、层次模式

#### 💡 Tip

为什么透明反而更保密了?

释: 这里说的透明指的是指帮助简化了管理和权限控制, 比如你打开一个数据库应用, 只需要看到表格、数据, 而不需要知道这些数据是怎么存储在硬盘上的。不想让你看到的你照样看不到, 而管理者凭借更容易的管理模式可以更方便地管理, 所以保密性、数据独立性也就更强。

- 12.【单选题】关系模型是用(B)结构来描述数据间的联系。
- A、立体表B、二维表C、层次D、三维表
- 13.【单选题】区分不同实体的依据是(A)
- A、属性B、概念C、对象D、名称
- 14.【单选题】下列哪个不是关系模型的组成[B]
- A、关系完整性的约束B、关系运算集合C、关系操作D、关系数据结构
- 15.【单选题】关系模型中常用的关系操作包括查询操作和(A)操作两大部分
- A、更新B、删除C、修改D、插入
- 16.【填空题】用有向图结构表示实体类型及实体间联系的数据模型称为**网状模型**，数据之间的联系通常用**图**实现
- 17.【填空题】用树型结构表示实体类型及实体间联系的数据模型称为**层次模型**，上一层的父结点和下一层的子结点之间的联系是\_\_的联系。
- 18.【填空题】**关系模型**是目前最常用也是最重要的一种数据模型。采用该模型作为数据的组织方式的数据库系统称为**关系型数据库**。
- 19.【填空题】关系的每一个分量必须是一个 **不可分**的数据项。
- 20.【填空题】关系模式是静态的，**稳定的**\_\_
- 21.【填空题】关系模型中有三类完整性约束:实体完整性、用户定义的完整性和参照完整性。其中**实体完整性**和参照完整性被称作是关系的两个不变性。
- 22.【填空题】关系的完整性约束条件包括三大类:分别是实体完整性、**参照完整性**\_\_和用户定义的完整性。
- 23.【填空题】关系模式的所有属性是这个关系模式的候选码称为**全码**。
- 24.【填空题】关系模型中常用的关系操作包括查询操作和**更新**（更新、插入、删除、修改）操作两大部分。
- 25.【填空题】关系操作的特点是**集合**操作。

#### 💡 Tip

在做判断时，你大概率会被几个词弄懵：主关键字、主码、候选关键字、候选码、主属性、非主属性。好，相信听到这你已经懵了，现在我提前给出通俗的解释：

**一个关系模式中，必须有关键字来表示元组，而关键字分为候选关键字（=候选码）和主关键字（=主码），候选关键字共同唯一标识元组，主关键字同样唯一标识元组（由候选关键字选出）且只能有一个。主属性：属于任何候选关键字的属性。非主属性：不属于任何候选关键字的属性。**

🧠 候选关键字和主关键字都能唯一标识元组，但主关键字的作用是明确“默认”选择一个主导标识符。

相信接下来做相关判断，你已经心如明镜了。

- 26.【判断题】关系操作的特点是集合操作方式。对
- 27.【判断题】参照完整性是指基本关系R的主属性不能取空值。错
- 28.【判断题】关系中的两行可以完全相同。错
- 29.【判断题】关系模式就是一个二维表。错
- 30.【判断题】关系模型是基于图的数据模型。错

- 31.【判断题】关系模式和关系是相同的。 错
- 32.【判断题】关系模式的主关键字可以有多个。 错
- 33.【判断题】关系和普通表的区别有两个方面，一是属性列不可再分，二是元组不会有重复。 对
- 34.【判断题】实体完整性规则是针对基本关系而言的。 对
- 35.【判断题】某一属性组能唯一的标识一个实体,则称该属性组为候选码。 对
- 36.【判断题】在一个关系模式中,有可能没有非主属性。 对
- 37.【判断题】一个关系中可以有多个候选码和主码。 错
- 38.【判断题】在参照完整性中,外码不可以取空值。 错

#### 💡 Tip

- 候选关键字 (Candidate Key)

**唯一性：**能够唯一标识元组。

**最小性：**不能再去掉任何一个属性，否则不再具有唯一性。

- 主关键字 (Primary Key)

**唯一性：**主关键字的值必须唯一。

**非空性：**主关键字的值不能为 Null。

- 39.【判断题】任何一个关系模式一定有关键字。 对
- 40.【判断题】关系模型用二维表表示实体及联系。 对
- 41.【判断题】关系操作的对象和结果都是集合。 对
- 42.【判断题】关系是关系模式在任何时刻的状态或内容。 错

#### 💡 Tip

**关系是关系模式在某一时刻的状态或内容。**

- 43.【判断题】任何关系数据库系统都应该支持实体完整性和参照完整性。 对
- 44.【判断题】**关系模式是静态的、稳定的,而关系是动态的、随时间不断变化的。** 对
- 45.【简答题】什么是概念模型?试述概念模型的作用。

概念模型用于信息世界的建模，是现实世界到信息世界的第一层抽象，一个中间层次。是数据库中设计人员进行数据设计的有力工具，也是数据库设计人员和用户之间交流的语言。

- 46.【简答题】定义并解释概念模型中以下术语:实体，实体型，实体集，实体之间的联系。

**实体：**客观存在并可以相互区分的事物。

**实体型：**用实体名及其属性名集合来抽象和刻画同类实体。

**实体集：**同型实体的集合。

**实体之间的联系：**包括实体内部的联系和实体之间的联系。前者指组成实体的各属性之间的联系；后者指不同实体集之间的联系，有一对一、一对多等等。

47.【简答题】关系数据库的优缺点

优点：建立在严格的数学概念的基础上；概念单一；存取路径对用户透明，具有更高数据独立性，安全保密性，简化了管理人员的工作。

缺点：由于存取路径对用户隐蔽，查询效率不如层次数据库和网状数据库高；为提高性能，必须对用户的查询请求进行优化，增加了开发数据库管理系统的难度。

① Note

为什么一会隐蔽一会透明？

在计算机中，存取路径对用户隐蔽=存取路径对用户透明。他们二者想表达的意思都是存取路径对于用户或者程序员来说是不可见的，碰见这两种说法要知道是一个意思。

48.【简答题】什么是数据模型？

现实世界数据特征的抽象，信息世界中表示实体类型及实体之间联系的模型。

**“学友”工作室出品**

**更多学科复习资料、竞赛学习等内容**

欢迎关注华理“以学会友ing”公众号