

第一章绪论

考点一：数据管理技术相关概念

- 1、数据的最小存取单位是**数据项**。
- 2、**数据库** (Database, **DB**)，统一管理的相关数据的集合。DB 的方法，包括 DB 的建立、查询、更新及各种数据控制。
- 3、**数据库管理系统** (Database Management System, **DBMS**)，位于用户与操作系统之间的一层数据管理软件，为用户或应用程序提供访问。DBMS 总是基于某种数据模型，可以分为层次型、网状型、关系型、面向对象型 DBMS。
- 4、**数据库系统** (Database System, **DBS**)，实现有组织地、动态地存储大量关联数据，方便多用户访问的计算机软件、硬件和数据资源组成的系统，即采用了数据库技术的计算机系统。

考点二：数据管理经历的阶段

- 1、数据管理技术共经历了三个阶段：**人工管理阶段**、**文件系统阶段**和**数据库系统阶段**。
 - (1) 人工管理阶段:数据不保存、数据需要由应用程序自己进行管理、基本上没有文件概念、数据不共享。
 - (2) 文件系统阶段:数据可以长期保存、文件系统管理数据、文件已经多样化、数据的存取基本上以记录为单位。缺点有：① 数据共享性差，数据冗余度大；② 数据和程序缺乏独立性。
 - (3) **数据库系统阶段**：① 数据结构化；② 数据共享性高、冗余度小、易扩充；③ 数据独立性高；④ 统一的数据管理和控制。
- 2、统一的数据管理和控制：数据的安全性保护、数据的完整性控制、数据库恢复和并发控制。

考点三：信息世界涉及的基本概念

- 1、**实体**：客观存在并可相互区别的事务，可以是具体的人、事、物，也可以是抽象的概念或联系。
- 2、属性：实体所具有的某一特性，一个实体可以有若干属性来刻画。
- 3、码 (key)：唯一标识实体的属性集。
- 4、域：属性的取值范围。
- 5、实体集：同型实体的集合。
- 6、实体型：具有相同属性的实体必然具有共同的特征和性质，用实体名及其属性名集合来抽象和刻画同类实体。
- 7、联系：事务内部以及事务之间具有联系。两个实体型之间的联系有三类：一对一联系 (1:1)；一对多联系 (1: n)；多对多联系 (m: n)。

考点四：概念模型的表示方法

- 1、概念模型用于信息世界的建模，与具体的 DBMS 无关，这种信息结构也不依赖于具体的计算机系统。
- 2、用 E-R 图来描述现实世界的概念模型，E-R 图提供了表示实体型、属性和联系的方法。
 - (1) 实体型：用矩形表示，矩形框内写明实体名。

- (2) 属性：用椭圆形表示，并用无向边将其与相应的实体连接起来。
- (3) 联系：用菱形表示，菱形框内写明联系名，并用无向边分别志有关实体连接起来，同时无向边旁标上联系的类型（1:1，1: n 或 m: n）。

考点五：数据模型

1、数据模型的三个要素

- (1) 数据结构：描述系统的静态特性
 - (2) 数据操作：描述系统的动态特性
 - (3) 数据的约束条件：是一组完整性规则的集合。
- 2、不同的数据模型具有不同的数据结构形式，根据数据之间的联系可以将数据模型划分为：层次模型、网状模型、关系模型和面向对象数据模型。其中层次模型和网状模型统称为非关系模型。

考点六：层次模型

- 1、层次模型：用**树型结构**表示实体间联系的数据模型。
- 2、层次数据模型可以：① 直接表示一对多（包括一对一）的联系；② **层次模型表示多对多联系**，必须首先将其分解成一对多联系，分解方法有两种：冗余结点法和虚拟结点法。
- 3、冗余节点法的优点：结构清晰，允许结点改变存储位置；缺点：需要额外占用存储空间，有潜在的不一致性。
- 4、虚拟结点法的优点：减少对存储空间的浪费，避免产生潜在的不一致性；缺点：结点改变存储位置可能引起虚拟结点中指针的修改。

考点七：网状模型

- 1、网状模型：用**有向图结构**表示实体类型及实体间联系的数据模型。
- 2、网状数据模型的**完整性约束**：
 - (1) 插入操作允许插入尚未确定双亲结点值的子女结点值。
 - (2) 删除操作允许只删除双亲结点值
 - (3) 更新操作只需更新制定记录即可
 - (4) 查询操作可以有多种方法，可根据具体情况选用。
- 3、网状数据模型的**缺点**：① 其 DDL，DML 语言极其复杂。② 结构比较复杂，而且随着应用环境的扩大，数据库的结构就变得越来越复杂，不利于最终用户掌握。

考点八：关系模型

- 1、关系模型：无论是实体还是实体间的联系均由单一的结构类型——关系来表示。
- 2、关系模型的数据结构：
 - (1) 关系：一张表。
 - (2) 元组：表中的一行。

- (3) 属性：表中的一列。
 - (4) **主码**：表中的某个属性组，它可以唯一确定一个元组。
 - (5) 域：属性的取值范围。
 - (6) 分量：元组中一个属性值。
 - (7) 关系模式：对关系的描述，一般表示为：关系名（属性 1，属性 2，...，属性 n）。
- 3、关系的三类完整性约束：实体完整性、参照完整性和用户定义的完整性。
- (1) 实体完整性规则：要求关系中组成主码的属性上不能有空值。
 - (2) 参照完整性规则：要求不引用不存在的实体。
 - (3) 用户定义完整性规则：由具体应用环境决定，系统提供定义和检验这类完整性的机制。
- 4、关系模型的存储结构：关系数据模型中，实体及实体间的联系都用表来表示。在数据库的物理组织中，表以文件形式存储，每一个表通常对应一种文件结构。
- 5、关系模型的优缺点：
- (1) 优点：建立在严格的数学概念的基础上；概念单一，无论实体还是实体之间的联系都用关系来表示。
 - (2) 缺点：由于存取路径对用户不透明，查询效率往往不如非关系数据模型

考点九：数据库的系统结构

1、数据系统的三级模式结构：

- (1) 外模式：也称子模型或用户模型，它是数据库用户能够看见和使用的局部数据的逻辑结构和特征的描述。
- (2) **模式**：也称逻辑模式，是数据库中全体数据逻辑结构和特征的描述。
- (3) 内模式：数据物理结构和存储方式的描述，是数据库内部的表示方法。

2、两级映象：外模式/模式映象，一般在外模式中描述。模式/内模式映象，一般在内模式中描述。两层映象保证了数据库系统中的数据能够具有较高的逻辑独立性和物理独立性。

考点十：数据库管理系统

1、数据库管理系统组成：

- (1) 数据定义语言及其翻译处理程序
- (2) 数据操纵语言及其编译（或解释）程序
- (3) 数据库运行控制程序
- (4) 实用程序

2、数据库管理系统的功能：

- (1) 数据定义
- (2) 数据操纵
- (3) 数据库运行管理
- (4) 数据组织、存储和管理
- (5) 数据库的建立和维护

(6) 数据通信接口



第二章 关系数据库简介

考点一：关系

- 1、域：是一种有相同数据类型的值的集合。
- 2、基本关系具有以下六条性质：
 - ① 列是同质的，即每一列中的分量同一类型的数据，来自同一个域。
 - ② 不同的列可出自同一个域，称其中的每列为一个属性，不同的属性要给予不同的属性名。
 - ③ 列的顺序无所谓，即列的次序可以任意交换。
 - ④ 任意两个元组不能完全相同。
 - ⑤ 行的顺序无所谓，即行的次序可以任意交换。
 - ⑥ 分量必须取原子值，即每一个分量都必须是不可分的数据项。

考点二：关系模式

- 1、关系数据库中，**关系模式是型，关系是值**。
- 2、关系模式是对关系的描述，包括几个方面：关系实质是一张二维表，表的每行为一个元组，每列为一个属性；一个关系通常是由赋予它的元组语义决定的。

考点三：关系的完整性

- 1、实体完整性规则说明如下：
 - (1) 实体完整性规则是针对基本关系而言的。一个基本表通常对应现实世界的一个实体集。例如学生关系对应于学生的集合。
 - (2) 现实世界中的实体是可区分的，即它们具有某种唯一性标识。
 - (3) 相应地，关系模型中以主码作为唯一性标识。
 - (4) 主码中的属性即主属性不能取空值。所谓空值就是"不知道"或"无意义"的值。
- 2、在实现参照完整性时需要考虑的几个方面是：
 - (1) 外码是否可以接受空值。
 - (2) 在被参照关系中删除元组时的考虑。
 - (3) 修改被参照关系中主码时的考虑。

考点四：集合运算符

- 1、运算的三大要素：**对象、运算符和运算结果**
- 2、传统的集合运算是二目运算，包括并、交、差、广义笛卡尔积四种运算。
- 3、并：设关系 R 和 S 具有相同的目 n（即两个关系都有 n 个属性），且相应的属性取自同一个域，则关系 R 和关系 S 的并由属于 R 或属于 S 的元组组成。其结果关系仍为 n 目关系。
记作： $R \cup S = \{t | t \in R \vee t \in S\}$

4、差：设关系 R 和 S 具有相同的目 n，且相应的属性取自同一个域，则关系 R 和关系 S 的差由属于 R 而不属于 S 的元组组成。其结果关系仍为 n 目关系。

记作： $R - S = \{t \mid t \in R \wedge \neg t \in S\}$

5、交：设关系 R 和 S 具有相同的目 n，且相应的属性取自同一个域，则关系 R 和关系 S 的交由既属于 R 又属于 S 的元组组成。其结果关系仍为 n 目关系。

记作： $R \cap S = \{t \mid t \in R \wedge t \in S\}$

6、广义笛卡尔积：两个分别为 n 目和 m 目的关系 R 和 S 的广义笛卡尔积是一个 (n+m) 列的元组的集合。元组的前 n 列是关系 R 的一个元组，后 m 列是关系 S 的一个元组。若 R 中有 k_1 个元组，S 中有 k_2 个元组，则关系 R 和 S 的广义笛卡尔积有 $k_1 \times k_2$ 个元组。

记作： $R \times S = \{t_r t_s \mid t_r \in R \wedge t_s \in S\}$

考点五：关系运算

1、专门的额关系运算包括选择、投影、连接、除等。

2、选择：又称限制，它是在关系 R 中选择满足给定条件的诸元组。

记作： $\sigma_F(R) = \{t \mid t \in R \wedge F(t) = \text{'真'}\}$ ，其中 F 表示选择条件，取值为‘真’或‘假’。

3、投影：从 R 中选择出若干属性列成新的关系。记作： $\Pi_A(R) = \{t[R] \mid t \in R\}$

4、连接：从两个关系的笛卡尔积中选取属性间满足一定条件的元组。

记作： $R \bowtie_{A \theta B} S = \{t_r t_s \mid t_r \in R \wedge t_s \in S \wedge t_r[A] \theta t_s[B]\}$ 。

连接运算有等值连接和自然连接。

(1) 当 θ 为“=”的连接运算称为等值连接，从关系 R 和关系 S 的笛卡尔积中选取 A,B 属性值相等的那些元组；

(2) 自然连接是特殊的等值连接，要求两个关系中进行比较的分量必须是相同的属性组，并且在结果中去掉重复的属性。

5、除：给定关系 R (X,Y) 和 S (Y,Z)，其中 X,Y,Z 为属性组。R 中的 Y 与 S 中的 Y 可以有不同的属性名，但必须出自相同的域集。R 与 S 的除运算得到一新的关系 P (X)，P 与 R 中满足下列的元组在 X 属性列上的投影：元组在 X 上分量值 x 的象集 YX 包含 S 在 Y 上投影的集合。

记作： $R \div S = \{t_r[X] \mid t_r \in R \wedge Y_x \supseteq \Pi_Y(S)\}$ ，其中 Y_x 为 x 在 R 中的象集， $x = t_r[X]$ 。

考点六：关系演算

1、关系演算是以数理逻辑中的谓词演算为基础的。按谓词变元不同，关系演算还可以分为元组关系演算和域关系演算。

2、检索操作用 Get 语言实现。

3、关系演算中元组变量的用途：① 简化关系名。在处理实际问题时如果关系的名字很长，使用起来就会感到不方便，这时可以设一个较短名字的元组变量来简化关系名。② 操作条件中使用量词时必须使用元组变量。

- 4、数关系演算的另一种形式是域关系演算。域关系演算以元组变量的分量即域变量作为谓词变元的基本对象。
- 5、1975 年由 IBM 公司的 M.Zloof 提出的 QBE (Query By Example) 就是一个很有特色的域关系演算语言，该语言于 1978 年在 IBM370 上得以实现。QBE 也指此关系数据库管理系统。

获取完整资料 请下载 APP 或关注公众号



扫码下载 app



扫码关注公众号