数据库原理简答题总结

第一章 数据库概论

1.人工管理阶段数据管理的特点：
（1） 数据不保存在机器中
（2） 无专用的软件对数据进行管理
（3） 只有程序的概念，没有文件的概念
（4） 数据面向程序

2.文件系统阶段数据管理的特点：
（1） 数据可长期保存在外存的磁盘上
（2） 数据的逻辑结构和物理结构有了区别
（3） 文件组织已呈多样化。有索引、链接和散列文件
（4） 数据不再属于某个特定的程序，可重复使用。

3.文件系统显露出三个缺陷：
（1） 数据冗余性
（2） 数据不一致性
（3） 数据联系弱

4.数据库阶段的管理方式具有以下特点：
（1） 采用复杂的数据模型表示数据结构
（2） 有较高的数据独立性
（3） 数据库系统为用户提供方便的用户接口
（4） 系统提供四方面的数据控制功能
（5） 对数据的操作既可以以记录为单位，又可以以数据项为单位

5.数据描述三个领域之间的关系：
从事物的特性到计算机中的数据表示，经历了三个领域：现实世界、信息世界、机器世界。
（1） 现实世界：存在于人们头脑之外的客观世界，称为现实世界。
（2） 信息世界：是现实世界在人们头脑中的反映。
（3） 机器世界：信息世界的信息在机器世界中以数据形式存储。
信息世界中数据描述的术语有：实体、实体集、属性、实体标识符
机器世界中数据描述的术语有：字段、记录、文件、关键码
它们的对应关系是：

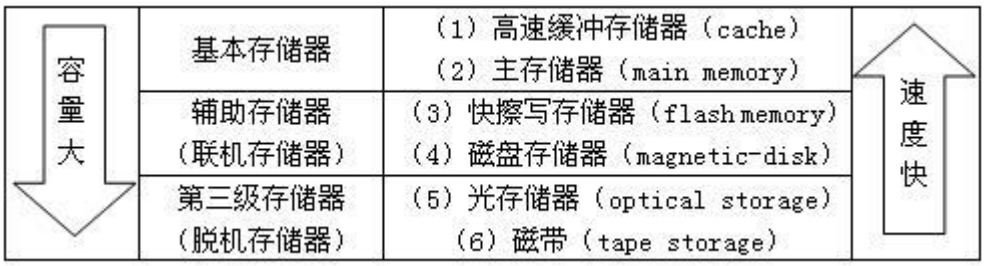
信息世界	机器世界
实体	记录
属性	字段（数据项）
实体集	文件
实体标识符（键）	关键码

在数据库中每个概念都有类型和值之区分，类型是概念的内涵，值是概念的外延

6.数据描述的两种形式：
数据描述有物理描述和逻辑描述两种形式。
物理数据描述指数据在存储设备上的存储方式，物理数据是实际存放在存储设备上的数据。
逻辑数据描述指程序员或用户用以操作的数据形式，是抽象的概念化数据。
数据管理软件的功能之一，就是要把逻辑数据转换成物理数据，以及把物理数据转换成逻辑数据。

7.物理存储介质层次：

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源



8.数据模型的种类：
目前广泛使用的数据模型可分为两种类型：概念数据模型、结构数据模型
概念数据模型：是独立于计算机系统的模型，完全不涉及信息在系统中的表示，只是用来描述某个特定组织所关心的信息结构;
它是现实世界的第一层抽象，是用户和数据库设计人员之间进行交流的工具;
这一类中著名的模型是“实体联系模型”，简称“ER”模型。
结构数据模型：是直接面向数据库的逻辑结构;
它是现实世界的第二层抽象，涉及到计算机系统和数据库管理系统;
这一类中的例子有层次、网状、关系、面向对象等模型。

9.结构数据模型的三个组成部分：
数据结构、数据操作、数据完整性约束是结构数据模型的三个组成部分。
数据结构：是指对实体类型和实体间联系的表达和实现
数据操作：是指对数据库的检索和更新（插、删、改）两类操作的实现
数据完整性约束：给出数据及其联系应具有的制约和依赖规则。

10.层次模型的特点：
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。
层次模型的特点是：记录之间的联系通过指针实现，查询效率较高。
缺点是：（1）只能表示 1:N 联系
（2）由于树型结构层次顺序的严格复杂，引起数据的查询和更新操作也很复杂，因此编写应用程序也很复杂。

11.网状模型的特点：
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。
网状模型的特点是：记录之间联系通过指针实现，M:N 联系也容易实现，查询效率较高。
缺点是：编写应用程序比较复杂，程序员必须熟悉数据库的逻辑结构。

12.关系模型的特点：
关系模型的主要特征是用二维表格结构表达实体集，用外键表示实体间联系。
特点是：关系模型与层次、网状的最大差别是用关键码而不是用指针导航数据，表格简单，用户易懂，编程时不涉及存储结构、访问技术等细节。

13.数据库体系结构中的三级结构、两级映象：
数据库的体系结构分为三级：内部级、概念级、外部级。
外部级：最接近用户，是单个用户所能看到的数据特性。单个用户使用的数据视图的描述称为“外模式”。
概念级：涉及到所有用户的数据定义，是全局的数据视图。全局数据视图的描述称为“概念模式”。
内部级：最接于物理存储设备，涉及到实际数据存储的结构。物理存储数据视图的描述称为“内模式”。
为实现这三个抽象级别的联系和转换，DBMS 在级级结构之间提供两个层次的映象：外模式/模式映象，模式/内模式映象。

14.二级数据独立性：
数据独立性是指：应用程序和数据之间相互独立，不受影响。分为物理独立性和逻辑独立性。
（1） 物理数据独立性：如果数据库的内模式要进行修改，即数据库的存储设备和存储方法有所变化，那么模式/内模式映象也要进行相应的修改，使概念模式尽可能保持不变。也就是对内模式的修改尽量不影响概念模式。
（2） 逻辑数据独立性：如果数据库的概念模式要进行修改，如增加记录类型或增加数据项，那么外模式/模式映象也要进行相应的修改，使外模式尽可能保持不变。也就是概念模式的修改尽量不影响外模式和应用程序。

15.DBMS 的主要功能：
（1） 数据库的定义功能：DBMS 提供数据定义语言（DDL）定义数据库的三级结构及其相互之间的映象、完整性、安全控制等约束。
（2） 数据库的操纵功能：DBMS 提供数据操纵语言（DML）实现对数据库中数据的操作。
（3） 数据库的保护功能：DBMS 对数据库的保护主要通过数据库的恢复、数据库的并发控制、数据库的完整性控制、数据库的安

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

全性控制等四个方面实现。
（4） 数据库的存储管理：DBMS 的存储管理子系统提供了数据库中数据和应用程序的一个界面，其职责是把各种 DML 语句转换成低层的文件系统命令，起到数据的存储、检索和更新的作用。
（5） 数据库的维护功能：DBMS 中实现数据库维护功能的实用程序主要有数据装载程序、备份程序、文件重组织程序、性能监控程序。
（6） 数据字典（DD）：数据库系统中存放三级结构定义的数据库称为数据字典，对数据库的操作都要通过访问 DD 才能实现。

16.DBMS 的组成：
DBMS 是由两大部分组成：查询处理器和存储管理器。
（1） 查询处理器有四个主要成分：DDL 编译器、DML 编译器、嵌入型 DML 的预编译器、查询运行核心程序。
（2） 存储管理器有四个主要成分：授权和完整性管理器、事务管理器、文件管理器、缓冲区管理器。

17.DBS 的组成：
DBS 是一个实际可运行的，按照数据库方法存储、维护和向应用系统提供数据支持的系统，它是数据库、硬件、软件、数据库管理员（DBA）的集合体。
（1） 数据库（DB）:是与一个特定组织各项应用有关的全部数据的集合，由应用数据的集合（物理数据库）、关于各级数据结构的描述（描述数据库）两部分组成。
（2） 硬件：包括中央处理机、内存、输入输出设备、数据通道等硬件设备。
（3） 软件：包括 DBMS、OS、各种宿主语言和应用开发支持软件等程序。
（4） DBA：DBA 是控制数据整体结构的人，负责 DBS 的正常运行。

18.DBS 的全局结构：
（1） 数据库用户。可分为四类：DBA、专业用户、应用程序员、终端用户
（2） DBMS 的查询处理器。包括四部分：DML 编译器、嵌入型 DML 的预编译器、DLL 编译器、查询运行核心程序。
（3） DBMS 的存储管理器。包括四部分：授权和完整性管理器、事务管理器、文件管理器、缓冲区管理器。
（4） 磁盘存储器中的数据结构。包括四种形式：数据文件、数据字典、索引文件、统计数据组织。

第二章 关系模型

19. 超键、主键、候选键的定义：
超键（super key）：在关系中能唯一标识元组的属性集称为关系模式的超键。
候选键（candidate key）：不含有多余属性的超键称为候选键。（候选键可以有多个）
主键（primary key）：用户选作元组标识的一个候选键称为主键。（主键是候选键中一个）

20. 关系模式、关系子模式和存储模式：
关系模型基本上遵循数据库的三级体系结构。概念模式是关系模式的集合，外模式是关系子模式的集合，内模式是存储模式的集合。
（1） 关系模式：关系模式实际上是记录类型。它的定义包括：模式名，属性名，值域名以及模式的主键。
（2） 关系子模式：是用户所用到的那部分数据的描述。除了指出用户的数据外，还应指出模式与子模式之间的对应性。
（3） 存储模式：关系存储时的基本组织方式是文件，元组是文件中的记录。存储一个关系可以用散列方法或索引方法实现。如果关系中元组数目较少，也可以用堆文件方式实现。

21. 关系模型的三类完整性规则：
（1） 实体完整性规则：这条规则要求关系中元组在组成主键的属性上不能有空值。
（2） 参照完整性规则：这条规则要求“不引用不存在的实体”。
（3） 用户定义的完整性规则：它反映某一具体应用涉及的数据必须满足的语义要求。

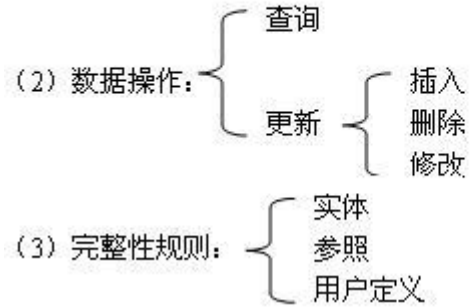
22. 参照完整性规则的形式定义：
如果属性集 K 是关系模式 R1 的主键，K 也是关系模式 R2 的外键，那么在 R2 的关系中，K 的取值只允许两种可能，或者为空，或者等于 R1 关系中某个主键值。
此规则使用时还要注意三点：
（1） 外键和相应的主键可以不同名，只要定义在相同值域上即可。
（2） R1 和 R2 也可以是同一个关系模式，表示了属性之间的联系。
（3） 外键值是否允许空，应视具体问题而定。
上述形式定义中，关系模式 R1 称为“参照关系”模式，R2 称为“依赖关系”模式。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

23. 关系模型的形式定义：

关系模型有三个组成部分：数据结构、数据操作和完整性规则。

（1） 数据结构：关系（二维表）



24. 关系查询语言根据其理论基础的不同分为哪两类：

关系代数语言：查询操作是以集合操作为基础运算的 DML 语言。（非过程性弱）

关系演算语言：查询操作是以谓词演算为基础运算的 DML 语言。（非过程性强）

25. 关系代数中的操作有哪些？

关系代数中的操作可分为两类：

传统的集合操作：并、差、交、笛卡尔积

扩充的集合操作：对关系进行垂直分割（投影）、水平分割（选择），关系的结合（联接、自然联接），笛卡尔积的逆运算（除法）等。

其中五个基本操作为：并、差、笛卡尔积、投影、选择。

四个常用组合操作为：交、连接、自然连接、除法

两种扩充的关系代数操作为：外连接和外部并

26. 关系演算有哪两种：

关系演算可分为元组关系演算和域关系演算。前者以元组为变量，后者以属性（域）为变量。

27. 什么是约束变量、自由变量：

如果元组变量未使用存在量词$\exists$或全称量词$\forall$符号定义，那么称为自由元组变量，否则称为约束元组变量。

约束变量类似于程序设计语言过程内部定义的局部变量，自由变量类似于过程外部定义的外部变量或全局变量。

28. 什么是安全运算：

在数据库技术中，不产生无限关系和无穷验证的运算称为安全运算，相应的表达式称为安全表达式，所采取的措施称为安全约束。在关系演算中约定，运算只对表达式中公式在涉及到的关系的值范围内操作。这样就不会产生无限关系和无穷验证问题，关系演算是安全的。

29. 为什么要对关系代数表达式进行优化：

查询优化是由 DBMS 对关系代数表达式进行优化组合，以提高 DBMS 的系统效率。要对关系代数进行优化的原因是：由于关系代数表达式是由关系代数操作组合而成。在关系代数操作中，执行笛卡尔积和联接运算最费时间，并且在执行过程中将产生大量的中间结果，以使系统执行效率较低。在执行前，由 DBMS 查询处理子系统先对关系代数表达式进行优化，尽可能早地执行选择和投影操作，以得到较小的中间关系，减少运算量和读外存块的次数，节省系统的执行时间，提高执行效率。

30. 简述查询优化的优化策略：

（1） 在关系代数表达式中尽可能早地执行选择操作。

（2） 把笛卡尔积和随后的选择操作合并成 F 联接运算。

（3） 同时计算一连串的选择和投影操作，以免分开运算造成多次扫描文件，从而能节省操作时间。

（4） 如果在一个表达式中多次出现某个子表达式，应该将该子表达式预先计算出结果保存起来。以免重复计算。

（5） 适当的对关系文件进行预处理。

（6） 在计算表达式之前应先估计一下怎么计算合算。

31. 笛卡尔积、等值连接、自然连接三者之间有什么区别：

等值连接中有笛卡尔积运算;

自然连接是一种等值连接，它是两个关系中所有公共属性进行等值连接的结果。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

32. SQL 数据库的体系结构及术语：
SQL 数据库的体系结构也是三级，但术语与传统的关系模型不同。
关系模式称为“基本表”，存储模式称为“存储文件”，子模式称为“视图”，元组称为“行”，属性称为“列”。

33. SQL 数据库的体系结构要点是什么：
（1） 一个 SQL 数据库是表（table）的汇集，它用一个或多个 SQL 模式定义。一个 SQL 模式是表和授权的表态定义。
（2） 一个 SQL 表由行集构成，一行是列的序列，每列对应一个数据项。
（3） 一个表或者是一个基本表，或者是一个视图。（视图只保存定义，不保存数据）
（4） 一个基本表可跨一个或多个存储文件，一个存储文件也可存放一个或多个基本表。每个存储文件与外部存储器上一个物理文件对应。
（5） 用户可用 SQL 语句对视图和基本表进行查询等操作。
（6） SQL 用户可以是应用程序，也可以是终端用户。

34. SQL 的组成分成几部分：
SQL 主要分成四部分：
（1） 数据定义。（SQL DDL）用于定义 SQL 模式、基本表、视图和索引的创建和撤消操作。
（2） 数据操纵。（SQL DML）数据操纵分成数据查询和数据更新两类。数据更新又分成插入、删除、和修改三种操作。
（3） 数据控制。包括对基本表和视图的授权，完整性规则的描述，事务控制等内容。
（4） 嵌入式 SQL 的使用规定。涉及到 SQL 语句嵌入在宿主语言程序中使用的规则。

35. SQL 模式的撤消有哪两种方式：
CASCADE（连锁式）方式：执行 DROP 语句时，把 SQL 模式及其下属的基本表、视图、索引等所有元素全部撤消。
RESTRICT（约束式）方式：执行 DROP 语句时，只有当 SQL 模式中没有任何下属元素时，才能撤消 SQL 模式，否则拒绝执行 DROP 语句。

36. SQL 提供的基本数据类型有哪些？每种举两个例子：
（1） 数值型：INTEGER 长整数、SMALLINT 短整数
（2） 字符串型：CHAR（N）长度为 N 的定长字符串、VARCHAR（N）具有最大长度为 N 的变长字符串。
（3） 位串型：BIT（N）长度为 N 的二进制位串、BIT VARYING（N）最大长度为 N 的变长二进制位串
（4） 时间型：DATE 日期、TIME 时间
SQL2 允许用户使用“CREATE DOMAIN”语句定义新的域。

37. 完整性约束主要有哪三种子句：
完整性约束主要有三种子句：主键子句（PRIMARY KEY），检查子句（CHECK）和外键子句（FOREIGN KEY）

38. 什么是视图，它与表的区别是什么：
在 SQL 中，外模式一级数据结构的基本单位是视图，视图是从若干基本表和（或）其他视图构造出来的表。我们创建一个视图时，只是把其视图的定义存放在数据字典中，而不存储视图对应的数据，因此，视图被称为“虚表”，这是它与表的主要区别。

39. 对于视图元组的更新操作（INSERT、DELETE、UPDATE）有哪三条规则：
（1） 如果一个视图是从多个基本表使用联接操作导出的，那么不允许对这个视图执行更新操作。
（2） 如果在导出视图的过程中，使用了分组和聚合操作，也不允许对这个视图执行更新操作。
（3） 如果视图是从单个基本表使用选择、投影操作导出的，并且包含了基本表的主键或某个候选键，那么这样的视图称为“行列子集视图”，并且可以被执行更新操作。
SQL2 中，允许更新的视图在定义时，必须加上“WITH CHECK OPTION”短语。

40. SQL 语言有哪两种使用方式：
一种是在终端交互方式下使用，称为交互式 SQL；
另一种是嵌入在高级语言的程序中使用，称为嵌入式 SQL，而这些高级语言可以是 C、PASCAL、COBOL 等，称为宿主语言。

41. 嵌入式 SQL 的实现有哪两种处理方式：
一种是扩充宿主语言的编译程序，使之能处理 SQL 语句；
另一种是采用预处理方式。目前多数系统采用后一种。

42. 在宿主语言的程序中使用 SQL 语句有哪些规定：
（1） 在程序中要区分 SQL 语言与宿主语句。所有 SQL 语句前必须加上前缀标识“EXEC SQL”，并以“END_EXEC”作为语句的结束标志。（结束标志在不同的宿主语言中不同）
（2） 允许嵌入的 SQL 语句引用宿主语言的程序变量（共享变量），但有两条规定：

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

1） 引用时，这些变量前必须加冒号“：”作为前缀，以示与数据库中变量有区别。
2） 这些变量由宿主语言的程序定义，并用 SQL 的 DECLARE 语句说明。

43. 用游标机制协调 SQL 的集合处理方式所用的 SQL 语句有哪些：
与游标有关的 SQL 语句有下列四个：
（1） 游标定义语句（DECLARE）
（2） 游标打开语句（OPEN）
（3） 游标推进语句（FETCH）
（4） 游标关闭语句（CLOSE）

44. SQL DML 的嵌入使用技术：
（1） 若是 INSERT、DELETE、UPDATE 语句，则不必涉及游标，只要加上前缀标识和结束标志就能嵌入宿主语言程序中使用。
（2） 若是已知查询结果肯定是单元组的 SELECT 语句，则不必涉及游标，也可加上前缀标识和结束标志后嵌入宿主语言程序中使用，但此时应该在 SELECT 语句中增加一个 INTO 子句，指出找到的值应送到相应的共享变量中去。
（3）若是已知查询结果为多个元组的 SELECT 语句，则必须涉及到游标，用游标机制把多个元组一次一个地传送给宿主程序处理。

第四章 关系数据库的模式设计

45. 什么是关系数据库：
关系数据库是以关系模型为基础的数据库，它利用关系来描述现实世界。一个关系既可以用来描述一个实体及其属性，也可以用来描述实体间的联系。关系实质上是一张二维表。

46. 一个关系模型有哪两个方面内容：
一个关系模型包括外延和内涵两个方面的内容。
外延就是通常所说的关系，或实例，或当前值。它与时间有关，随着时间的推移在不断变化。（由于元组的插入、删除、修改引起的）
内涵是与时间独立的，包括关系、属性、及域的一些定义和说明，还有各种数据完整性约束。

47. 数据完整性约束分为哪两类：
数据完整性约束分为静态约束和动态约束。
静态约束：包括各种数据之间的联系（数据依赖），主键的设计和关系值的各种限制等等。这一类约束是如何定义关系的有效数据问题。
动态约束：主要定义如插入、删除、和修改等各种操作的影响。

48. 关系数据库设计理论主要包括哪些内容：
关系数据库设计理论主要包括三个方面的内容：数据依赖、范式、模式设计方法。其中数据依赖起着核心的作用。

49. 数据库使用过程中存在的问题是什么：
数据冗余、更新异常、插入异常、删除异常。

50. 函数依赖（FD）的定义：
设有关系模式 R（A1，A2，……，An）（即 R（U）），X，Y 是 U 的子集，r 是 R 的任一具体关系，如果对 r 的任意两个元组 t1，t2，由 t1[X]=t2[X]导致 t1[Y]=t2[Y]，则称 X 函数决定 Y，或 Y 函数依赖于 X，记为 X→Y，X→Y 为模式 R 的一个函数依赖。
或者说，对于 X 的每一个具体值，都有 Y 惟一的具体值与之对应，即 Y 值由 X 值决定，因而
这种数据依赖称为函数依赖。

51. 函数依赖的逻辑蕴涵、FD 的闭包 F+：
设 F 是关系模式 R 的一个函数依赖集，X，Y 是 R 的属性子集，如果从 F 中的函数依赖能够推出 X—>Y，则称 F 逻辑蕴涵 X—>Y，记为 $F\models X\rightarrow Y$。
被 F 逻辑蕴涵的函数依赖的全体构成的集合，称为 F 的闭包，记为 F^+。$F^+=\{X\rightarrow Y \mid F\models X\rightarrow Y\}$

52. 候选键、主属性、非主属性：
设有关系模式 R（A1，A2，……，An），F 是 R 的一个函数依赖集，X 是{A1，A2，……，An}的一个子集。如果
① X→A1A2……An∈F+，且
② 不存在 X 真子集 Y，使得 Y→A1A2……An 成立，则称 X 是 R 的候选键。
包含在任何一个候选键中的属性称为主属性，不包含在任何一个候选键中的属性称为非主属性。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

53. 函数依赖的推理规则：
设有关系模式 R（A1，A2，……，An）和属性集 U= A1，A2，……，An，X，Y，Z，W 是 U 的一个子集，F 是 R 的一个函数依赖集，推理规则如下：

（1） 自反律：如果 $Y \subseteq X \subseteq U$，则 $X \rightarrow Y$ 在 R 上成立。
（2） 增广律：如果 $X \rightarrow Y$ 为 F 所蕴涵，$Z \subseteq U$，则 $XZ \rightarrow YZ$ 在 R 上成立。
（3） 传递律：如果 $X \rightarrow Y$ 和 $Y \rightarrow Z$ 在 R 上成立，则 $X \rightarrow Z$ 在 R 上成立。

FD 的其他三个推理规则：

（4） 合并律：如果 $X \rightarrow Y$ 成立，那么 $X \rightarrow YZ$ 成立。
（5） 伪传递律：如果 $X \rightarrow Y$ 和 $WY \rightarrow Z$ 成立，那么 $WX \rightarrow Z$ 成立。
（6） 分解律：如果 $X \rightarrow Y$ 和 $Z \subseteq Y$ 成立，那么 $X \rightarrow Z$ 成立。

54. 什么是平凡的 FD？平凡的 FD 可根据哪一条推理规则推出？

如果 $X \rightarrow Y$，并且 $Y \subseteq X$，则称 $X \rightarrow Y$ 是平凡的 FD。根据推理规则的自反律可推出。

55. 关系模式的分解

有几个不同的衡量标准：
分解具有无损联接;
分解要保持函数依赖;
分解既要保持依赖，又要具有无损联接。

56. 什么是无损连接：

设有关系模式 R，分解成关系模式 $\rho = \{R_1, R_2, \cdots\cdots R_k\}$，F 是 R 的一个函数依赖

对 R 中满足 F 的每一个关系 r 都有：$r = \pi_{R1}(r) \bowtie \pi_{R2}(r) \bowtie \cdots\cdots \pi_{RK}(r)$，则称

ρ 是无损联结分解。

57. 试叙保持函数依赖的定义：

设 F 是属性集 U 上的一个函数依赖集，Z 是 U 上的一个子集，F 在 Z 上的一个投影定义为：

$\pi_Z(F) = \{X \rightarrow Y \mid X \rightarrow Y \in F^+ \text{且} XY \subseteq Z\}$

设关系模式 R 的一个分解为 $\rho = \{R_1, R_2, \cdots\cdots R_k\}$，F 是 R 的一个函数依赖集，如果

$$F^+ = \left(\bigcup_{i=1}^{k} \pi_{Ri}(F)\right)$$

则称为分解 ρ 保持函数依赖。

58. 第一范式（1NF）：
如果关系模式 R 的所有属性的值域中每一个值都是不可再分解的值，则称 R 是属于第一范式模式。

59. 第二范式（2NF）：
如果关系模式 R 为第一范式，并且 R 中每一个非主属性完全函数依赖于 R 的候选键，则称 R 是第二范式模式。

60. 第三范式（3NF）：
如果关系模式 R 是第一范式，且每个非主属性都不传递依赖于 R 的候选键，则称 R 是第三范式的模式。

61. BCNF：
如果关系模式 R 是第一范式，且每个属性都不传递依赖于 R 的候选键，那么称 R 是 BCNF 的模式。从 BCNF 的定义可明显地得出如下结论：
（1） 所有非主属性对键是完全函数依赖。
（2） 所有主属性对不包含它的键是完全函数依赖。
（3） 没有属性完全函数依赖于非键的任何属性组。
如果模式 R 是 BCNF，则它必定是第三范式，反之，则不一定。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

62. 模式设计方法的原则：
关系模式 R 相对于函数依赖集 F 分解成数据库模式 ρ＝｛R1，R2，……Rk｝，一般应具有下面三个特性：
（1） ρ 中每个关系模式 Ri 是 3NF 或 BCNF
（2） 保持无损联结
（3） 保持函数依赖集
（4） ρ 中模式个数最少和属性总数最少。

63. 一个好的模式设计方法应符合哪三条原则：
表达性，分离性，最小冗余性。
表达性涉及到两个数据库模式的等价性问题，即数据等价和依赖等价，分别用无损联接和保持函数依赖性来衡量。
分离性是指属性间的“独立联系”应该用不同的关系模式表达。
最小冗余性要求在分解后的数据库能表达原来数据库的所有信息这个前提下实现。
关系模式设计方法基本上可以分为分解与合成两大类。

64. 多值依赖 MVD：
设 R（U）是属性集 U 上的一个关系模式，X，Y 是 U 的子集，若对 R（U）的任一关系 r，对于 X 的一个给定的值存在着 Y 的一组值与其对应，同时 Y 的这组值又不以任何方式与 U-X-Y 中的属性相关，那么称 Y 多值依赖于 X，记为 X→→Y。

65. 平凡多值依赖：
对于属性集 U 上的一个多值依赖 X→→Y，如果 Y⊆X 或者 XY=U，那么称 X→→Y 是一个平凡多值依赖。

66. 第四范式（4NF）：
设关系模式 R，D 是一个多值依赖集，如果 D 中存在一个非平凡多值依赖 X→→Y，并且 X 必是 R 的超键，那么称 R 是 4NF 模式。

第五章 数据库设计

67. 什么是软件生存期：
软件生存期是软件工程的一个重要概念。是指从软件的规划、研制、实现、投入运行后的维护，直到它被新的软件所取代而停止使用的整个期间。通常分为六个阶段：
（1） 规划阶段
（2） 需求分析阶段
（3） 设计阶段
（4） 程序编制阶段
（5） 调试阶段
（6） 运行维护阶段

68. 数据库系统的生存期：
一般分为七个阶段，即：
（1） 规划阶段
（2） 需求分析阶段 1）信息要求 2）处理要求 3）安全性和完整性要求
（3） 概念设计阶段
（4） 逻辑设计阶段 两部分：数据库逻辑设计和应用程序设计
（5） 物理设计阶段 两部分：物理数据库结构的选择和逻辑设计中程序模块说明的精确化
（6） 实现阶段
（7） 运行维护阶段

69. 数据库设计过程的输入有哪些内容：
（1） 总体信息需求
（2） 处理需求
（3） DBMS 的特征
（4） 硬件和 OS 特征

70. 数据库设计过程的输出有哪两部分：
一部分是完整的数据库结构，其中包括逻辑结构与物理结构。
另一部分是基于数据库结构和处理要求的应用程序的设计原则。

71. 常见的数据库设计方法有哪几种：
（1） 视图模式化及视图汇总设计方法

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

（2） 关系模式的设计方法
（3） 新奥尔良设计方法
（4） 基于 E-R 模型的数据库设计方法
（5） 基于 3NF 的设计方法
（6） 基于抽象语法规范的设计方法
（7） 计算机辅助数据库设计方法

72. 实用的数据库设计方法至少应包括哪些内容：
（1） 设计过程
（2） 设计技术
（3） 评价准则
（4） 信息需求
（5） 描述机制

73. 一种设计方法学需要有三种基本类型的描述机制：
（1） 实现设计过程的最终结果将用 DBMS 的 DDL 表示。
（2） 信息输入的描述。
（3） 在信息输入和 DDL 描述之间的其它中间步骤的结果的描述。

74. 数据库设计中的规划阶段的主要任务：
是进行建立数据库的必要性及可行性分析，确定数据库系统在组织中和信息系统中的地位，以及各个数据库之间的联系。

75. 需求分析阶段的任务：
需求分析阶段应该对系统的整个应用情况作全面的、详细的调查，确定企业组织的目标，收集支持系统总的设计目标的基础数据和对这些数据的要求，确定用户的需求，并把这些要求写成用户和数据库设计者都能接受的文档。

76. 需求分析的步骤：
大致可分为三步来完成，即需求信息的收集、分析整理和评审。

77. 数据字典由哪几部分组成：
（1） 数据项
（2） 数据结构
（3） 数据流
（4） 数据存储
（5） 加工过程

78. 数据抽象：
抽象是对实际的人、物、事或概念的人为处理，它抽取人们关心的共同特性，忽略非本质的细节，并把这些特性用各种概念精确地加以描述，这些概念组成了某种模型。
抽象有两种形式，系统状态抽象（抽象对象）和系统转换抽象（抽象运算）。

79. 对象的两种形式：
（1） 聚集：的数学意义就是笛卡尔积的概念。通过聚集，形成对象之间的一个联系对象。
（2） 概括：是从一类其它对象形成一个对象。对于一类对象{O1，O2，……，On}可以概括成对象 O，那么 Oi 称为 O 的其中一个。

80. 依赖联系：
在现实世界中，常常有某些实体对于另一些实体具有很强的依赖关系，即一个实体的存在必须以另一个实体的存在为前提。我们通常把前者称为弱实体。在 ER 图中，用双线框表示弱实体，用指向弱实体的箭头表明依赖联系。

81. 子类、超类：
某个实体类型中所有实体同时也是另一实体类型中的实体。此时，我们称前一实体类型是后一实体类型的子类，后一实体类型称为超类。在 ER 图中，带有子类的实体类型（超类）以两端双线的矩形框表示，并用加圈的弧线与其子类相连，子类本身仍用普通矩形框表示。
子类具有一个很重要的性质：继承性。它可继承超类上定义的全部属性，其本身还可包含其它另外的属性。

82. ER 模型的操作：
（1） 实体类型的分裂：垂直分割、水平分割
（2） 实体类型合并：分裂的逆过程。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

（3） 联系类型的分裂
（4） 联系类型的合并

83. 采用 ER 方法的数据库概念设计分成哪三步：
（1） 设计局部 ER 模式：1）确定局部结构范围 2）实体定义 3）联系定义 4）属性分配
（2） 设计全局 ER 模式：1）确定公共实体类型 2）局部 ER 模式的合并 3）消除冲突。
（3） 全局 ER 模式的优化：1）实体类型的合并 2）冗余属性的消除 3）冗余联系的消除

84. 冲突分为哪三种：
属性冲突，包括属性域的冲突、属性取值单位冲突。
结构冲突，包括：
（1） 同一对象在不同应用中的不同抽象。
（2） 同一实体在不同局部 ER 图中属性组成不同。
（3） 实体之间的联系在不同的局部 ER 图中呈现不同的类型。
命名冲突，包括属性名，实体名，联系名之间的冲突：同名异义、异名同义

85. ER 模型向关系模型的转换：
ER 模型中的主要成分是实体类型和联系类型。
对实体类型，将每个实体类型转换成一个关系模式，实体的属性即为关系模式的属性，实体标识符即为关系模式的键。
对联系类型，就视 1：1、1：N、M：N 三种不同的情况做不同处理。
（1） 对 1：1 可在两个实体类型转换成的两个关系模式中任意一个关系模式的属性中加入另一个关系模式的键和联系类型的属性。
（2） 对 1：N，则在 N 端实体类型转换成的关系模式中加入 1 端实体类型转换成的关系模式的键和联系类型的属性。
（3） 对 M：N，则将联系类型也转换成关系模式，其属性为两端实体类型的键盘加上联系类型的属性，而键为两端实体键的组合。

86. 什么是物理设计：
对一个给定的逻辑数据模型选取一个最适合应用环境的物理结构的过程，称为数据库的物理设计。物理结构，主要指数据库在物理设备上的存储结构和存取方法。

87. 物理设计的步骤：
物理设计可分五步完成，前三步涉及到物理数据库结构的设计，后两步涉及约束和具体的程序设计。
（1） 存储记录结构设计
（2） 确定数据存储安排
（3） 访问方法的设计
（4） 完整性和安全性
（5） 程序设计

88. 在数据库系统生存期中，生存期的总开销可分为几项：
规划开销、设计开销、实现与测试开销、操作开销、维护开销。

89. 用户使用和计算机资源的操作开销是：
（1） 查询响应时间
（2） 更新事务的开销
（3） 报告生成的开销
（4） 改组频率和开销
（5） 主存储空间
（6） 辅助存储空间

90. 数据库实现阶段的主要工作：
（1） 建立实际数据库结构
（2） 试运行
（3） 装入数据

91. 数据库的重新组织设计：
对数据库的概念模式、逻辑结构或物理结构的改变称为重新组织，其中改变概念模式或逻辑结构又称为重新构造，改变物理结构则称为重新格式化。

92. 运行维护阶段的主要工作：
（1） 维护数据库的安全性和完整性控制及系统的转储和恢复。
（2） 性能的监督、分析与改进。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源

（3） 增加新功能。
（4） 发现错误，修改错误。

公众号【大学百科资料】整理，有超百科复习资料+海量网课资源