

第一章 绪论

1. 答：

数据

描述事物的符号记录称为数据。描述事物的符号可以是数字，也可以是文字、图形、图像、声音、正文、流数据、HTML、XML等等，数据的种类十分丰富。数据与其语义是不可分的。

解析：

在现代计算机系统中数据的概念是广义的。早期的计算机系统主要用于科学计算，处理的数据是整数、实数、浮点数等传统数学中的数据等。现在计算机能存储和处理的对象十分广泛，表示这些对象的数据也越来越复杂。

数据与其语义是不可分的。500这个数字可以表示一件物品的价格是500元，也可以表示一个学术会议参加的人数有500人。还可以表示一袋奶粉重500克。

数据库技术是管理数据的技术。数据是数据库管理的基本对象。因此我们首先要知道什么是数据，知道数据有多种形式。

数据库

数据库是长期储存在计算机内、有组织的、可共享的数据集合。数据库中的数据按一定的数据模型组织、描述和储存，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。

解析：

1) 简单地讲，数据库的数据具有永久储存、有组织和可共享三个特点。

2) 数据模型：是数据库的核心概念。每个数据库中数据的都是按照某一种数据模型来组织的。

数据库系统

数据库系统：数据库系统（DBS）是指在计算机系统中引入数据库后的系统构成。数据库系统由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员构成。

解析：

数据库系统和数据库是两个概念。

数据库系统是一个人-机系统，数据库是数据库系统的一个组成部分。

但是在日常工作中人们常常把数据库系统简称为数据库。希望读者能够从人们讲话或文章的上下文中区分“数据库系统”和“数据库”，不要引起混淆。

数据库管理系统

数据库管理系统(DBMS)是位于用户与操作系统之间的一层数据管理软件。用于科学地组织和存储数据、高效地获取和维护数据。DBMS 主要功能包括数据定义功能、数据操纵功能、数据库的运行管理功能、数据库的建立和维护功能。

解析：

DBMS 是一个大型复杂的软件系统。是计算机中的基础软件。目前，专门研制 DBMS 的厂商及其研制的 DBMS 产品很多。著名的有美国 IBM 公司的 DB2 关系数据库管理系统、

IMS 层次数据库管理系统；美国 ORACLE 公司的 ORACLE 关系数据库管理系统；美国微软公司的 SQL SERVER 关系数据库管理系统等。

主要的国产数据库管理系统有：北京人大金仓公司的金仓数据库系统 KingbaseES；武汉达梦公司的达梦数据库系统；神通公司的 Oscar 等。

2. 答：

使用数据库系统的好处很多，包括：可以大大提高应用开发的效率，方便用户的使用，减轻数据库系统管理人员维护数据库系统的负担等。

使用数据库系统的好处是由数据库管理系统的特点或优点决定的。

解析：

1) 使用数据库系统可以大大提高应用开发的效率。因为在数据库系统中应用程序不必考虑数据的定义、存储和数据存取的具体路径，这些工作都由 DBMS 来完成。用一个通俗的比喻，使用了 DBMS 就如有了一个好参谋好助手，许多具体的技术工作都由这个助手来完成。开发人员就可以专注于用户需求的理解，应用逻辑的设计，而不必为管理数据的许许多多复杂的细节操心。

2) 当应用逻辑改变，数据的逻辑结构需要改变时，由于数据库系统提供了数据与程序之间的独立性。数据逻辑结构的改变是 DBA 的责任，开发人员不必修改应用程序，或者只需要修改很少的应用程序。从而既简化了应用程序的编制，又大大减少了应用程序的维护和修改。

3) 使用数据库系统可以减轻数据库系统管理人员维护系统的负担。因为 DBMS 在数据库建立、运用和维护时对数据库进行统一的管理和控制，包括数据的完整性、安全性，多用户并发控制，故障恢复等等都由 DBMS 执行。

总之，使用数据库系统的优点是很多的，既便于数据的集中管理，控制数据冗余，可以提高数据的利用率和一致性，又有利于应用程序的开发和维护。读者可以在自己今后的工作中结合具体应用，认真加以体会和总结。

为什么有这些好处，可以结合第 5 题 数据库管理系统的特点来回答。

3. 答：

文件系统与数据库系统的区别：

文件系统面向某一应用程序，共享性差、冗余度大，独立性差，纪录内有结构、整体无结构，应用程序自己控制。

数据库系统面向现实世界，共享性高、冗余度小，具有高度的物理独立性和一定的逻辑独立性，整体结构化，用数据模型描述，由数据库管理系统提供数据安全性、完整性、并发控制和恢复能力。

读者可以参考《概论》书中表 1.1 中的有关内容。

文件系统与数据库系统的联系是：

文件系统与数据库系统都是计算机系统中管理数据的软件。

解析：

文件系统是操作系统的重要组成部分，而 DBMS 是独立于操作系统的软件。我们不能独立购买一个文件系统，但需要独立购买 DBMS 软件产品。而 DBMS 是在操作系统的基础上实现的，数据库中数据的组织和存储是通过操作系统中文件系统来实现的。因此，

DBMS 的实现与操作系统中的文件系统是紧密相关的。例如，数据库实现的基础是文件，对数据库的任何操作最终要转化为对文件的操作。所以在 DBMS 实现中数据库物理组织的基本问题是如何利用或如何选择操作系统提供的基本的文件组织方法。这里我们就不具体展开了。读者可以参考《概论》书第十二章《数据库管理系统》中有关内容，第十二章是 DBMS 实现技术的概述。有兴趣的读者还可以进一步学习数据库管理系统实现的有关课程。

4. 答：

·适用于文件系统而不是数据库系统的应用例子

数据的备份，软件或应用程序使用过程中的临时数据存储一般使用文件比较合适。

功能比较简单、比较固定的应用系统也适合用文件系统。

·适用于数据库系统而非文件系统的应用例子

目前，几乎所有企业或部门的信息系统都以数据库系统为基础，都使用数据库。如一个工厂的管理信息系统，其中会包括许多子系统，如库存管理系统、物资采购系统、作业调度系统、设备管理系统、人事管理系统等等，还比如学校的学生管理系统，人事管理系统，图书馆的图书管理系统等等都适合用数据库系统。

因此，数据库系统已经成为信息系统的基础和核心。

希望同学们能举出自己了解的应用例子。

5. 答：

数据库系统的主要特点有：

(1) 数据结构化

数据库系统实现整体数据的结构化，这是数据库的主要特征之一，也是数据库系统与文件系统的本质区别。

解析：

注意这里“整体”两个字。在数据库系统中，数据不再仅仅针对某一个应用，而是面向全组织，可以支持许多应用；不仅数据内部是结构化的，而且整体是结构化的，数据之间是具有联系的。数据的存取单位即一次可以存取数据的大小也很灵活。可以小到某一个数据项（如一个学生的姓名），大到一组记录（成千上万个学生记录）。而在文件系统中，数据的存取单位只有一个：记录，如一个学生的完整记录。

(2) 数据的共享性高，冗余度低，易扩充

数据库的数据不再面向某个应用而是面向整个系统，因此可以被多个用户、多个应用、用多种不同的语言共享使用。由于数据面向整个系统，是有结构的数据，不仅可以被多个应用共享使用，而且容易增加新的应用，这就使得数据库系统易于扩充，我们称之为弹性大。

解析：

数据共享可以大大减少数据冗余，节约存储空间，同时还能够避免数据之间的不相容性与不一致性。

所谓“数据面向某个应用”是指数据结构是针对某个应用设计的，只被这个应用程序或应用系统使用。可以说数据是某个应用的“私有资源”。

所谓“弹性大”是指系统容易扩充也容易收缩，即应用增加或减少时不必修改整个数据库的结构，或者只要做很少的修改。

我们可以取整体数据的各种子集用于不同的应用系统，当应用需求改变或增加时，只要重新选取不同的子集或加上一部分数据便可以满足新的需求。

(3) 数据独立性高

数据独立性包括数据的物理独立性和数据的逻辑独立性。

数据库管理系统的模式结构和二级映像功能保证了数据库中的数据具有很高的物理独立性和逻辑独立性。

解析：

所谓“独立性”即相互不依赖。数据独立性是指数据和程序相互不依赖。即数据的逻辑结构或物理结构改变了，程序不会跟着改变。数据与程序的独立，把数据的定义从程序中分离出去，加上数据的存取又由 DBMS 负责，简化了应用程序的编制，大大减少了应用程序的维护和修改。

(4) 数据由 DBMS 统一管理和控制

数据库的共享是并发的共享，即多个用户可以同时存取数据库中的数据甚至可以同时存取数据库中同一个数据。为此，DBMS 必须提供统一的数据控制功能，包括数据的安全性保护，数据的完整性检查，并发控制和数据库恢复。

解析：

DBMS 数据控制功能包括四个方面：

数据的安全性保护：保护数据以防止不合法的使用造成的数据的泄密和破坏；

数据的完整性检查：将数据控制在有效的范围内或保证数据之间满足一定的关系，一定的约束条件；

并发控制：对多用户的并发操作加以控制和协调，保证并发操作的正确性；

数据库恢复：当计算机系统发生硬件故障、软件故障，或者由于操作员的失误以及故意的破坏影响数据库中数据的正确性，甚至造成数据库部分或全部数据的丢失时，能将数据库从错误状态恢复到某一已知的正确状态，亦称为完整状态或一致状态。

下面我们可以得到“什么是数据库”的一个定义：

数据库是长期存储在计算机内有组织的大量的共享的数据集合。它可以供各种用户共享，具有最小冗余度和较高的数据独立性。DBMS 在数据库建立、运用和维护时对数据库进行统一控制，以保证数据的完整性、安全性，并在多用户同时使用数据库时进行并发控制，在发生故障后对系统进行恢复。

数据库系统的出现使信息系统从以加工数据的程序为中心转向围绕共享的数据库为中心的新阶段。

6. 答：

- (1) 数据库定义功能；
- (2) 数据组织、存储和管理功能；
- (3) 数据操纵功能；
- (4) 数据库的事务管理和运行管理；
- (5) 数据库的建立和维护功能；
- (6) 其他功能，例如不同数据库之间的互访和互操作功能等等。

7. 答：

数据模型是数据库中用来对现实世界进行抽象的工具，是数据库中用于提供信息表示和操作手段的形式构架。

一般地讲，数据模型是严格定义的概念的集合。这些概念精确地描述系统的静态特性、动态特性和完整性约束条件。因此数据模型通常由数据结构、数据操作和完整性约束三部分组成。

- (1) 数据结构：是所研究的对象类型的集合，是对系统的静态特性的描述。
- (2) 数据操作：是指对数据库中各种对象（型）的实例（值）允许进行的操作的集合，包括操作及有关的操作规则，是对系统动态特性的描述。
- (3) 数据的约束条件：是完整性规则的集合，完整性规则是给定的数据模型中数据及其联系所具有的制约和依存规则，用以限定符合数据模型的数据库状态以及状态的变化，以保证数据的正确、有效、相容。

解析：

数据模型是数据库系统中最重要的概念之一。同学们必须通过《概论》的学习真正掌握数据模型的概念和作用。

数据模型是数据库系统的基础。任何一个 DBMS 都以某一个数据模型为基础，或者说支持某一个数据模型。

数据库系统中模型有不同的层次。根据模型应用的不同目的，可以将模型分成两类或说两个层次：一是概念模型，是按用户的观点来对数据和信息建模，用于信息世界的建模，强调语义表达能力，概念简单清晰；另一是数据模型，是按计算机系统的观点对数据建模，用于机器世界，人们可以用它定义、操纵数据库中的数据。一般需要有严格的形式化定义和一组严格定义了语法和语义的语言，并有一些规定和限制，便于在机器上实现。

8. 答：

概念模型实际上是现实世界到机器世界的一个中间层次。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，是数据库设计人员进行数据库设计的有力工具，也是数据库设计人员和用户之间进行交流的语言。

9. 答：

实体：客观存在并可以相互区分的事物叫实体。

实体型：具有相同属性的实体具有相同的特征和性质，用实体名及其属性名集合来抽象和刻画同类实体称为实体型。

实体集：同型实体的集合称为实体集。

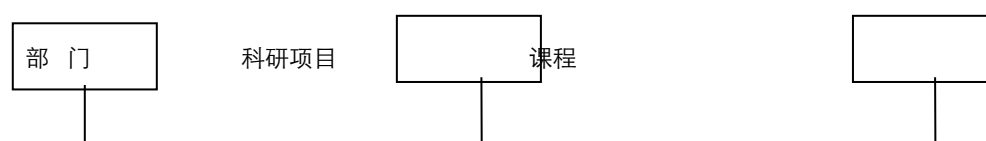
属性：实体所具有的某一特性，一个实体可由若干个属性来刻画。

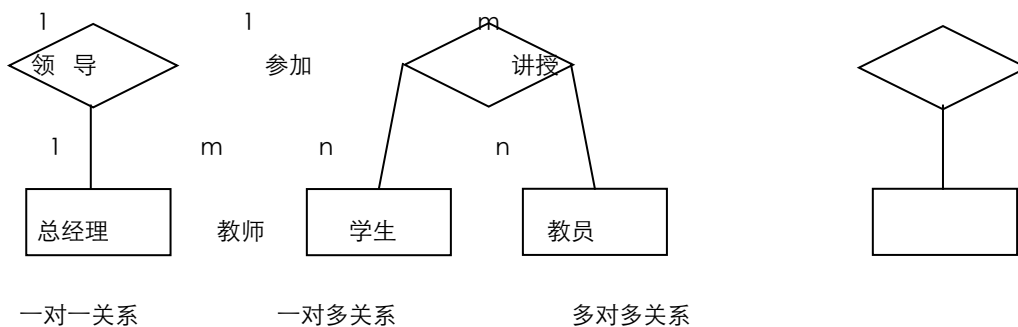
码：唯一标识实体的属性集称为码。

实体—联系图（E-R 图）：E-R 图提供了表示实体型、属性和联系的方法：

- 实体型：用矩形表示，矩形框内写明实体名。
- 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。
- 联系：用菱形表示，菱形框内写明联系名，并用无向边分别与有关实体连接起来，同时无向边旁标上联系类型（1:1, 1:n 或 m:n）。

10. 答：





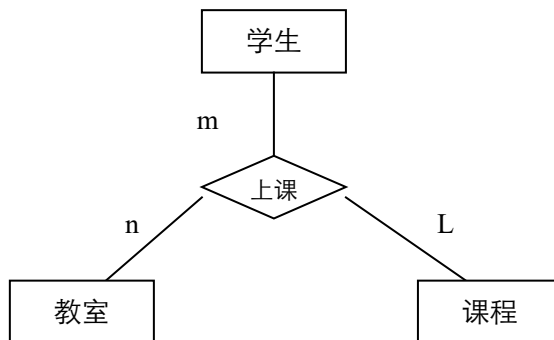
解析:

部门和总经理是两个实体。一个部门或者一个公司只有一个总经理，一个总经理只在一个公司工作，不能兼任两个以上公司的总经理，所以部门和总经理两个实体之间是一一对应的联系。

一个科研项目可以由多个教员和多个学生承担。如果规定一个教员只能参加一个项目，一个学生也只能参加一个项目。按照这样的语义，科研项目课程和教员、学生三者之间是一对多的联系。

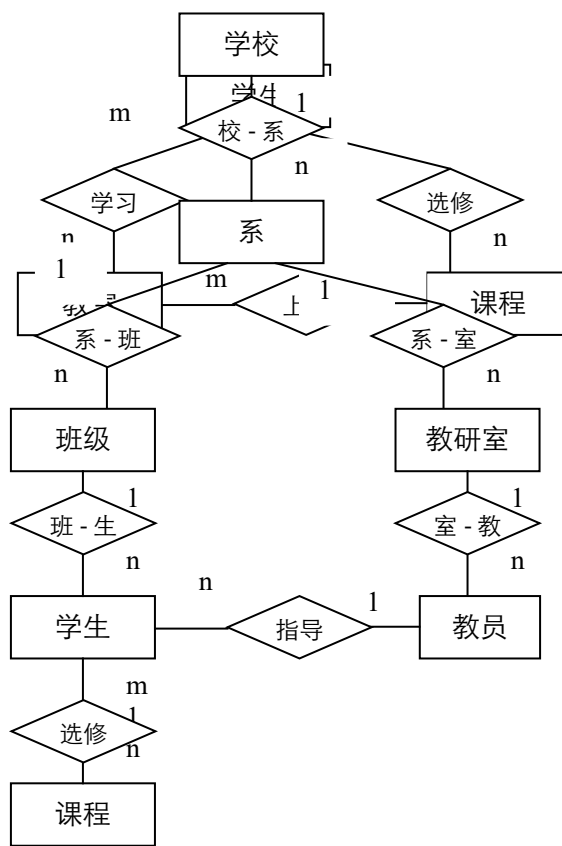
如果一门课程可以由多个教员讲授，一个教员可以讲授多门课程，则课程实体和教员实体之间是多对多的联系。

11. 答:



三个实体型之间的多对多联系和三个实体型两两之间的三个多对多联系是不等价。因为它们拥有不同的语义。

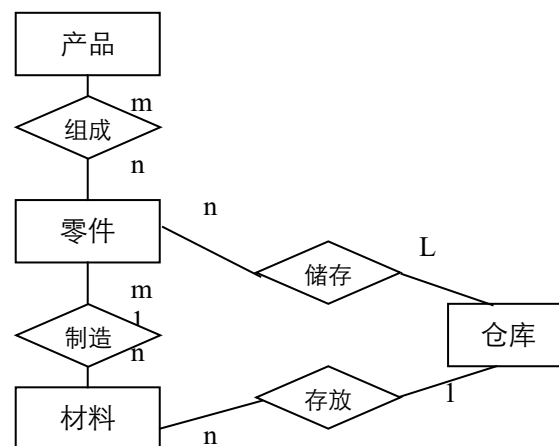
12. 答:



解析：

在画 ER 图时，读者可以按照习题中对问题的描述一步一步画出每一句话中涉及的实体，再根据给出的实际语义，画出实体之间的联系。例如每个教研室有若干教员，每个班有若干学生，可以画出教研室和教员，班级和学生之间一对多的联系，从“有的教授和副教授每人各带若干研究生”一个研究生一般指定一个导师，这是通常的规则，所以可以画出教员和学生之间一对多的联系。

13. 答：



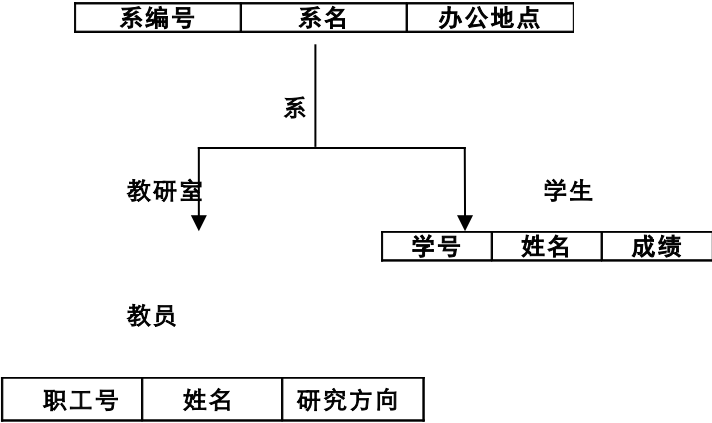
解析：

对实体之间联系的语义描述有时不是直截了当的，需要我们从对现实世界的整体描述中进行分析，导出实体之间的某种联系。就如本题中，“零件和仓库的联系”就要从以下描述中分析：“零件按所属的不同产品分别放在仓库中”。因为一个产品由多种零件组成的，所以一个仓库中放多种零件。反过来一种零件是放在一个仓库还是多个仓库中呢？因为一种零件可以用在多种产品上，这些零件按所属的不同产品分别放在仓库中，于是我们知道一种零件可以放在多个仓库中。所以零件和仓库之间是多对多的联系。

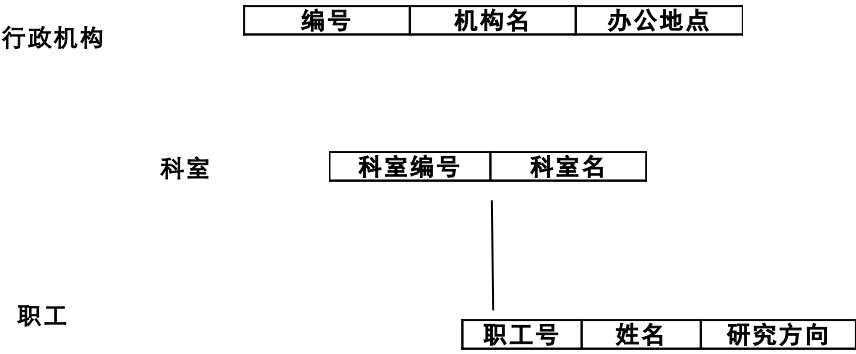
“材料和仓库的联系”则根据“原材料按照类别放在若干仓库”这句话就可以得出：一个仓库中放多种材料，而一种材料只放在一个仓库中，所以仓库和材料之间是一对多的联系。

14. 答:

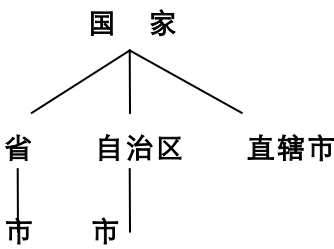
(1) 教员学生层次数据库模型



(2) 行政机构层次数据库模型

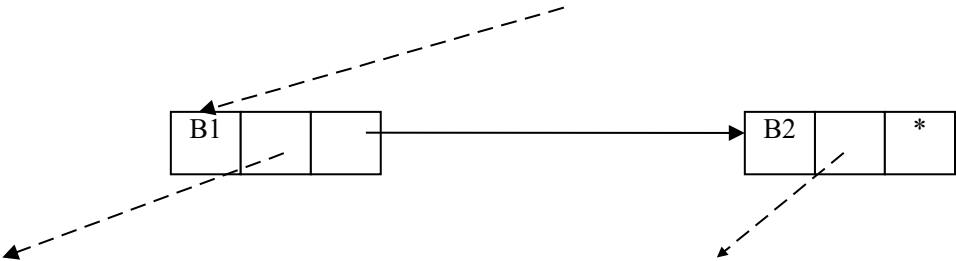


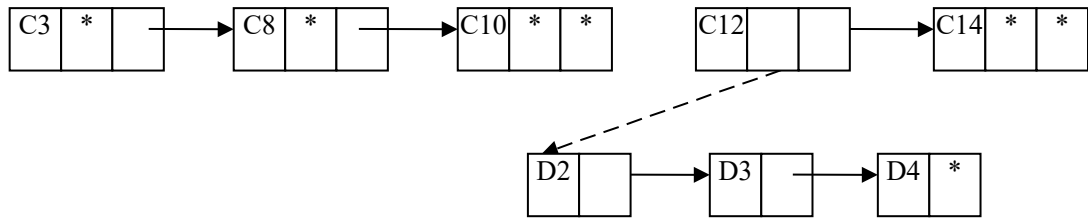
(3) 行政区域层次数据库模型



15. 答:

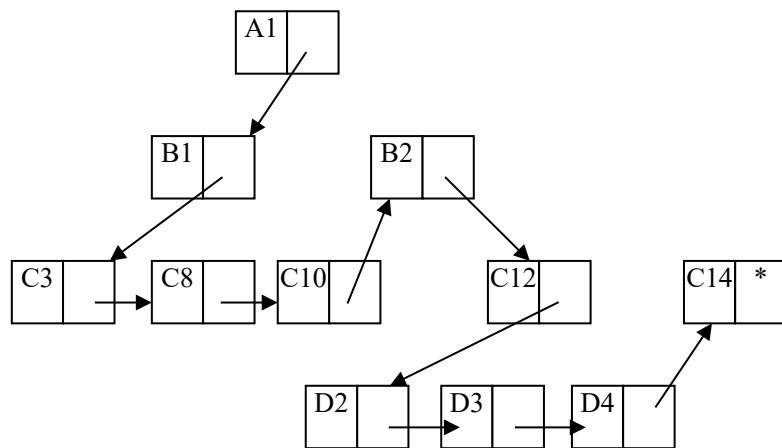
子女-兄弟链接法:





图中虚数表示子女链，记录结构的第二部分存放子女指针；实线表示兄弟链，记录结构的第三部分存放兄弟指针。星号 “*” 表示空指针。

答：层次序列链接法：

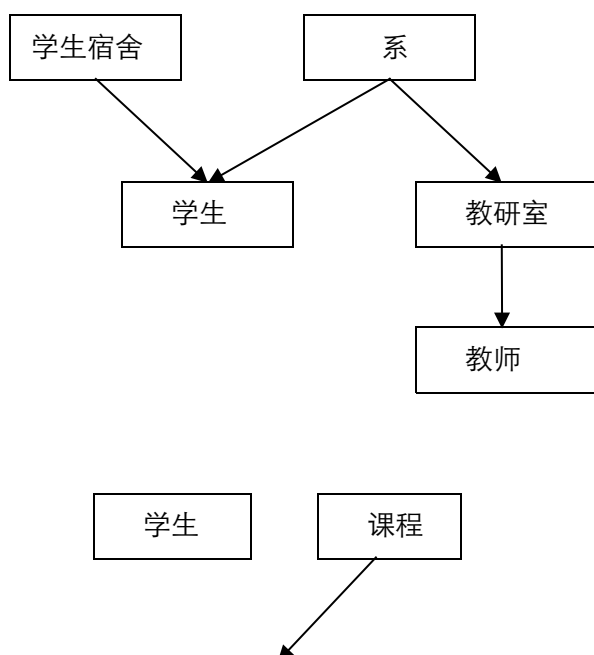


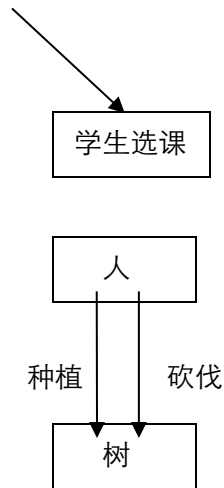
16. 答：

满足下面两个条件的基本层次联系的集合为网状模型。

- (1) 允许一个以上的结点无双亲；
- (2) 一个结点可以有多于一个的双亲。

实例：





17. 答:

层次模型的优点主要有:

- (1) 层次模型的数据结构比较简单清晰; 对具有一对多的层次关系的部门描述非常自然、直观, 容易理解, 这是层次数据库的突出优点
- (2) 层次数据库的查询效率高;
- (3) 层次数据模型提供了良好的完整性支持。

层次模型的缺点主要有:

- (1) 现实世界中很多联系是非层次性的, 层次模型不能自然地表示这类联系。
- (2) 用层次模型来表示非层次结构的方法比较笨拙, 只能通过引入冗余数据 (易产生不一致性) 或创建非自然的数据结构 (引入虚拟结点) 来解决, 对插入和删除操作的限制比较多。
- (3) 查询子女结点必须通过双亲结点。
- (4) 由于结构严密, 层次命令趋于程序化。

网状数据模型的优点主要有:

- (1) 能够更为直接地描述现实世界, 如一个结点可以有多个双亲。
- (2) 具有良好的性能, 存取效率较高。

网状数据模型的缺点主要有:

- (1) 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。
- (2) 其 DDL, DML 语言复杂, 用户不容易使用。

18. 答:

关系模型由关系数据结构、关系操作集合和关系完整性约束三部分组成。在用户观点下, 关系模型中数据的逻辑结构是一张二维表, 它由行和列组成。

- (1) **关系**: 一个关系对应通常说的一张表。
- (2) **属性**: 表中的一列即为一个属性;
- (3) **域**: 属性的取值范围;
- (4) **元组**: 表中的一行即为一个元组;
- (5) **码**: 表中的某个属性组, 它可以唯一确定一个元组;

- (6) **分量**：元组中的一个属性值；
- (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）。

19. 答：

关系数据模型具有下列优点：

- (1) 关系模型与非关系模型不同，它是建立在严格的数学概念的基础上的。
- (2) 关系模型的概念单一。无论实体还是实体之间的联系都用关系表示。操作的对象和操作的结果都是关系。所以其数据结构简单、清晰，用户易懂易用。
- (3) 关系模型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。

当然，关系数据模型也有缺点，其中最主要的缺点是，由于存取路径对用户透明，查询效率往往不如非关系数据模型。因此为了提高性能，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软件的难度。

20. 答：

数据库系统的三级模式结构由外模式、模式和内模式组成。（参见书上图 1.28）

外模式，亦称子模式或用户模式，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，是数据库用户的数据视图，是与某一应用有关的数据的逻辑表示。

模式，亦称逻辑模式，是数据库中全体数据的逻辑结构和特性的描述，是所有用户的公共数据视图。模式描述的是数据的全局逻辑结构。外模式涉及的是数据的局部的逻辑结构，通常是模式的子集。

内模式，亦称存储模式，是数据在数据库系统内部的表示，即对数据的物理结构和存储方式的描述。

数据库系统的三级模式是对数据的三个抽象级别，它把数据的具体组织留给 DBMS 管理，使用户能逻辑抽象地处理数据，而不必关心数据在计算机中的表示和存储。

为了能够在内部实现这三个抽象层次的联系和转换，数据库系统在这三级模式之间提供了两层映像：外模式／模式映像和模式／内模式映像。正是这两层映像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。

21. 答：

模式、外模式、内模式的解释参见第 20 题。

DDL：数据定义语言。用来定义数据库模式、外模式、内模式的语言。

DML：数据操纵语言。用来对数据库中的数据进行查询、插入、删除和修改的语句。

22. 答：

数据与程序的逻辑独立性：当模式改变时（例如增加新的关系、新的属性、改变属性的数据类型等），由数据库管理员对各个外模式／模式的映像作相应改变，可以使外模式保持不变。应用程序是依据数据的外模式编写的，从而应用程序不必修改，保证了数据与程序的逻辑独立性，简称数据的逻辑独立性。

数据与程序的物理独立性：当数据库的存储结构改变了，由数据库管理员对模式／内模式映像作相应改变，可以使模式保持不变，从而应用程序也不必改变。保证了数据与程序的物理独立性，简称数据的物理独立性。

数据库管理系统在三级模式之间提供的两层映像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。

23. 答:

数据库系统一般由数据库、数据库管理系统（及其开发工具）、应用系统、数据库管理员和用户构成。

24. 答:

DBA 负责全面地管理和控制数据库系统。具体职责包括：

- (1) 决定数据库的信息内容和结构；
- (2) 决定数据库的存储结构和存取策略；
- (3) 定义数据的安全性要求和完整性约束条件；
- (4) 监督数据库的使用和运行；
- (5) 数据库系统的改进和重组重构。

25. 答:

系统分析员负责应用系统的需求分析和规范说明，系统分析员要和用户及 DBA 相结合，确定系统的硬件软件配置，并参与数据库系统的概要设计。

数据库设计人员负责数据库中数据的确定、数据库各级模式的设计。数据库设计人员必须参加用户需求调查和系统分析，然后进行数据库设计。在很多情况下，数据库设计人员就由数据库管理员担任。

应用程序员负责设计和编写应用系统的程序模块，并进行调试和安装。