

第一章.习题和答案

1.1 试述数据、数据库、数据库系统的概念。

答：数据(Data)：描述事物的符号记录称为数据。数据的种类有数字、文字、图形、图像、声音、正文等。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。现代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。例如：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(Database)：数据库是长期储存在计算机内的、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，并可被各种用户共享。

数据库系统(Database System)：数据库系统是指在计算机系统中引入数据库后的系统构成，一般由数据库、数据库管理系统(及其开发工具)、应用系统、数据库管理员构成。解析数据库系统和数据库是两个概念。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。

1.2 试述数据库管理系统的概念及其主要功能。

答：数据库管理系统(Database Management System, DBMS)：数据库管理系统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 的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。

1.3 试述使用数据库系统有什么好处。

答：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。使用数据库系统的好处很多，例如，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的负担，等等。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应用逻辑的设计，而不必为数据管理的许许多多复杂的细节操心。还有，当应用逻辑改变，数据的逻辑结构也需要改变时，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。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，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，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。

1.4 试述文件系统与数据库系统的区别和联系。

答：文件系统与数据库系统的区别是：文件系统面向某一应用程序，共享性差，冗余度大，数据独立性差，记录内有结构，整体无结构，由应用程序自己控制。数据库系统面向现实世界，共享性高，

冗余度小，具有较高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的安全性、完整性、并发控制和恢复能力。

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。但是 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的文件系统来实现的。

1.5 举出适合用文件系统而不适合用数据库系统的应用例子；再举出适合用数据库系统的应用例子。

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系统比较合适。早期功能比较简单、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。例如，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。

1.6 试述数据库系统的特点。

答：（1）数据结构化。（2）数据的共享性高、冗余度低，易扩充。（3）数据独立性高。（4）数据由 DBMS 统一管理和控制。

1.7 试述数据模型的概念、数据模型的作用和数据模型的三个要素。

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。这些概念精确描述了系统的静态特性、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：是一组完整性规则的集合。

1.8 试述概念模型的作用。

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。

1.9 试述数据库系统三级模式结构，以及这种结构的优点。

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。外模式，亦称子模式或用户模式，是数据库用户(包括应用程序员和最终用户)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征的描述，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。

1.10 解释以下术语：模式、外模式、内模式。

答：模式也称逻辑模式，是数据库中全体数据的逻辑结构和特征的描述，是所有用户的公共数据视图。它是数据库系统模式结构的中间层，既不涉及数据的物理存储细节和硬件环境，也与具体的应用程序、所使用的应用开发工具及高级程序设计语言无关。一个数据库只有一个模式。

外模式也称子模式或用户模式，它是数据库用户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，是与某一应用程序有关的数据的逻辑表示。外模式通常是模式的子集，一个数据库可以有多个外模式。

内模式也称存储模式，一个数据库只有一个内模式。它使用一个物理数据模型，全面描述了数据库中数据存储的全部细节和存取路径，是数据在数据库内部的表示。

1.11 解释数据与程序的物理独立性、数据与程序的逻辑独立性。

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 / 模式的映像做相应改变，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 / 内模式映像做相应改变，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。

1.12 为什么数据库系统具有数据与程序的独立性。

答：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。

1.13 试述数据库系统的组成。

答：数据库系统一般由硬件系统、数据库集合、数据库管理系统及相关软件、数据库管理员和用户组成。

1.14 DBA 的职责是什么。

答：数据库管理员负责数据库系统建立、维护和管理。数据库管理员的职责包括：定义并存储数据库的内容、监督和控制数据库的使用、负责数据库的日常维护，必要时重组或改进数据库。

第二章.习题和答案

2.1 试述关系模型的三个组成部分。

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。

2.2 试述关系数据语言的特点和分类。

答：关系数据语言可以分为三类：1、关系代数语言。2、关系演算语言：元组关系演算语言和域关系演算语言。3、SQL：具有关系代数和关系演算双重特点的语言。这些关系数据语言的共同特点是，语言具有完备的表达能力，是非过程化的集合操作语言，功能强，能够嵌入高级语言中使用。

2.3 试述关系模型的完整性规则。在参照完整性中，为什么外部码属性的值也可以为空？什么情况下才可以为空？

答：实体完整性规则是指若属性 A 是基本关系 R 的主属性，则属性 A 不能取空值。

若属性(或属性组)F 是基本关系 R 的外码，它与基本关系 S 的主码 Ks 相对应(基本关系 R 和 S 不一定是不同的关系)，则对于 R 中每个元组在 F 上的值必须为：或者取空值(F 的每个属性值均为空值)；或者等于 S 中某个元组的主码值。即属性 F 本身不是主属性，则可以取空值，否则不能取空值。

2.4 关系与普通的表格、文字有什么区别。

答：关系是一种规范化的二维表格，在关系模型中，对关系作了下列规范性限制：（1）每一个列的分量必须来自同一个域，必须是同一类型的数据。（2）不同的列可来自同一个域。（3）列的顺序可以任意交换，元组的顺序可任意。（4）关系中不允许出现相同的元组。（5）关系中每一分量必须是不可分割的数据项。

2.5 试述等值连接与自然连接的区别和联系。

答：连接运算符是“=”的连接运算称为等值连接。它是从关系 R 与 S 的广义笛卡尔积中选取 A, B 属性值相等的那些元组。自然连接是一种特殊的等值连接，它要求两个关系中进行比较的分量必须是相同的属性组，并且在结果中把重复的属性列去掉。

2.6 阐述笛卡尔积、等值连接、自然连接三者之间的区别。

答：笛卡尔积对两个关系 R 和 S 进行乘操作，产生的关系中元组个数为两个关系中元组个数之积。等值联接则是在笛卡尔积的结果上再进行选择操作，从关系 R 和 S 的笛卡儿积中选择对应属性值相等的元组。自然连接则是在等值联接(以所有公共属性值相等为条件)的基础上再行投影操作，去掉重复的公共属性列。当两个关系没有公共属性时，自然连接就转化为笛卡尔积。

2.7 关系代数的基本运算有哪些，如何用这些基本运算来表示其他运算

答：并、差、笛卡尔积、投影和选择 5 种运算为基本的运算。其他 3 种运算，即交、连接和除，均可以用这 5 种基本运算来表达。

2.8 设有如下关系： 学生(学号，姓名，性别，专业，出生日期) 教师(教师编号，姓名，所在部门，职称) 授课(教师编号，学号，课程编号，课程名称，教材，学分，成绩)

1) 查找学习“数据库原理”课程且成绩不及格的学生学号和任课教师编号；

2) 查找学习“英语”课程的“计算机应用”专业学生的学号、姓名和成绩。

答: (1) $\Pi_{\text{教师编号}}(\sigma_{\text{课程名称}='数据库原理' \text{ and } \text{成绩} \leq 60}(\text{授课}))$

(2) $\Pi_{\text{学号, 姓名, 成绩}}((\sigma_{\text{专业}='计算机应用'}(\text{学生})) \bowtie (\sigma_{\text{课程名称}='英语'}(\text{授课})))$

2.9 设有如下关系:

S(S#, SNAME, AGE, SEX)/*学生(学号, 姓名, 年龄, 性别)*/

C(C#, CNAME, TEACHER)/*课程(课程号, 课程名, 任课教师)*/

SC(S#, C#, GRADE)/*成绩(学号, 课程号, 成绩)*/

查询: (1) 教师“程军”所授课程的课程号和课程名; (2) “李强”同学不学课程的课程号; (3) 至少选修了课程号为 k1 和 k5 的学生学号; (4) 选修课程包含学号为 2 的学生所修课程的学生学号。

答: (1) $\Pi_{C\#, CNAME}(\sigma_{TEACHER='程军'}(C))$

(2) $\Pi_{C\#}(C) - \Pi_{C\#}((\sigma_{NAME='李强'}(S)) \bowtie SC)$

(3) $\Pi_{S\#, C\#}(SC) \div \Pi_{C\#}(\sigma_{C\#='k1' \cup C\#='k5'}(C))$

(4) $\Pi_{S\#, C\#}(SC) \div \Pi_{C\#}(\sigma_{C\#=2}(SC))$

2.10 设有如下关系:

图书关系 B(图书编号 B#, 图书名 T, 作者 A, 出版社 P);

读者关系 R(借书证号 C#, 读者名 N, 读者地址 D);

借阅关系 L(C#, B#, 借书日期 E, 还书标志 BZ);

BZ='1' 表示已还; BZ='0' 表示未还;

查询: (1) “工业出版社”出版的图书名; (2) 查询 99 年 12 月 31 日以前借书未还的读者名与书名。

答: (1) $\Pi_T(\sigma_{P='工业出版社'}(B))$

(2) $\Pi_{N, T}(\Pi_{C\#, B\#}(\sigma_{E < '99/12/31' \cap BZ='0'}(L)) \bowtie R \bowtie B)$

2.11 简述关系模型的优点。

答: (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) 无论实体还是实体之间的联系都用关系来表示。对数据的检索结果也是关系(即表), 概念单一, 其数据结构简单、清晰; (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, 简化了程序员的工作和数据库开发建立的工作; (4) 关系模型具有丰富的完整性, 如实体完整性、参照完整性和用户定义的完整性, 大大降低了数据的冗余和数据不一致的概率。

2.12 为什么关系中不允许有重复元组。

答: 因为关系是一个元组的集合, 而集合中的元素用于代表现实中一个存在, 不允许重复出现, 因此在关系模型中对关系作了限制, 关系中的元组不能重复, 可以用键来标识唯一的元组。

2.13 描述键、候选键、主键和外键的定义。

答: 键: 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：关系中存在一个以上的键，均称为候选键。

主键：用户选作元组标识的一个候选键为主键。

外键：某个关系的主键相应的属性在另一关系中出现，此时该主键就是另一关系的外键，如有两个关系 S 和 SC,其中 S#是关系 S 的主键，相应的属性 S#在关系 SC 中也出现，此时 S#就是关系 SC 的外键。

2.14 什么是数据库的完整性约束条件，可分为哪几类。

答：完整性约束条件是指数据库中的数据应该满足的语义约束条件。一般可以分为六类：静态列级约束、静态元组约束、静态关系约束、动态列级约束、动态元组约束、动态关系约束。静态列级约束是对一个列的取值域的说明，包括以下几个方面：（1）对数据类型的约束，包括数据的类型、长度、单位、精度等；（2）对数据格式的约束；（3）对取值范围或取值集合的约束；（4）对空值的约束；（5）其他约束。静态元组约束就是规定组成一个元组的各个列之间的约束关系，静态元组约束只局限在单个元组上。静态关系约束是在一个关系的各个元组之间或者若干关系之间常常存在各种联系或约束。

常见的静态关系约束有：（1）实体完整性约束；（2）参照完整性约束；（3）函数依赖约束。

动态列级约束是修改列定义或列值时应满足的约束条件，包括下面两方面：（1）修改列定义时的约束；（2）修改列值时的约束。动态元组约束是指修改某个元组的值时需要参照其旧值，并且新旧值之间需要满足某种约束条件。动态关系约束是加在关系变化前后状态上的限制条件，例如事务一致性、原子性等约束条件。

2.15 如何理解数据库中的参照完整性。

答：参照完整性，简单地说就是控制数据一致性，尤其是不同表之间关系的规则。建立永久关联的目的就是利用关联表之间的制约机制互相参照，控制表间数据的一致性和完整性。这种制约机制会在用户执行插入、修改或删除记录等编辑记录的操作时，发挥其限制作用。参照完整性有效限制了对表中数据的非法编辑。具体说参照完整性是指不允许在相关数据表中引用不存在的记录。参照完整性的规则包括：

- 1、更新规则：主表关键字段值被改时，对子表相应关键字段值的更新制约机制。
- 2、删除规则：主表记录删除时，对子表关键字相匹配的记录的制约机制。
- 3、插入规则：向子表插入记录的制约机制。

第三章.习题和答案

3.1 试述数据库设计过程。

答：数据库设计过程的六个阶段：(1) 需求分析；(2) 概念结构设计；(3) 逻辑结构设计；(4) 数据库物理设计；(5) 数据库实施；(6) 数据库运行和维护。这是一个完整的实际数据库及其应用系统的设计过程。不仅包括设计数据库本身，还包括数据库的实施、运行和维护。设计一个完善的数据库应用系统往往是上述六个阶段的不断反复。

3.2 试述数据库设计过程各个阶段中的设计描述。

答：各阶段的设计要点如下：(1) 需求分析：准确了解与分析用户需求（包括数据与处理）。(2) 概念结构设计：通过对用户需求进行综合、归纳与抽象，形成一个独立于具体 DBMS 的概念模型。(3) 逻辑结构设计：将概念结构转换为某个 DBMS 所支持的数据模型，并对其进行优化。(4) 数据库物理设计：为逻辑数据模型选取一个最适合应用环境的物理结构（包括存储结构和存取方法）。(5) 数据库实施：设计人员运用 DBMS 提供的数据库语言、工具及宿主语言，根据逻辑设计和物理设计的结果建立数据库，编制与调试应用程序，组织数据入库，并进行试运行。(6) 数据库运行和维护：在数据库系统运行过程中对其进行评价、调整与修改。

3.3 试述数据库设计过程中结构设计部分形成的数据库模式。

答：数据库结构设计的不同阶段形成数据库的各级模式，即：(1) 在概念设计阶段形成独立于机器特点，独立于各个 DBMS 产品的概念模式，在本篇中就是 E-R 图；(2) 在逻辑设计阶段将 E-R 图转换成具体的数据库产品支持的数据模型，如关系模型，形成数据库逻辑模式，然后在基本表的基础上再建立必要的视图，形成数据的外模式；(3) 在物理设计阶段，根据 DBMS 特点和处理的需要，进行物理存储安排，建立索引，形成数据库内模式。

3.4 试述数据库设计的特点。

答：数据库设计既是一项涉及多学科的综合性技术又是一项庞大的工程项目。其主要特点有：(1) 数据库建设是硬件、软件和干件（技术与管理的界面）的结合。(2) 从软件设计的技术角度看，数据库设计应该和应用系统设计相结合，也就是说，整个设计过程中要把结构（数据）设计和行为（处理）设计密切结合起来。

3.5 数据字典的内容和作用是什么。

答：数据字典是系统中各类数据描述的集合。数据字典的内容通常包括：数据项、数据结构、数据流、数据存储、处理过程五个部分。其中数据项是数据的最小组成单位，若干个数据项可以组成一个数据结构。数据字典通过对数据项和数据结构的定义来描述数据流和数据存储的逻辑内容。数据字典的作用：数据字典是关于数据库中数据的描述，在需求分析阶段建立，是下一步进行概念设计的基础，并在数据库设计过程中不断修改、充实、完善。

3.6 需求分析阶段的设计目标是什么？调查的内容是什么。

答：需求分析阶段的设计目标是通过详细调查现实世界要处理的对象（组织、部门、企业等），充分了解原系统（手工系统或计算机系统）工作概况，明确用户的各种需求，然后在此基础上确定新系统的功能。调查的内容是“数据”和“处理”，即获得用户对数据库的如下要求：(1) 信息要求，指用户

需要从数据库中获得信息的内容与性质，由信息要求可以导出数据要求，即在数据库中需要存储哪些数据；（2）处理要求，指用户要完成什么处理功能，对处理的响应时间有什么要求，处理方式是批处理还是联机处理；（3）安全性与完整性要求。

3.7 什么是数据库的概念结构？试述其特点和设计策略。

答：概念结构是信息世界的结构，即概念模型，其主要特点是：（1）能真实、充分地反映现实世界，包括事物和事物之间的联系，能满足用户对数据的处理要求，是对现实世界的一个真实模型；（2）易于理解，从而可以用它和不熟悉计算机的用户交换意见，用户的积极参与是数据库设计成功的关键；（3）易于更改，当应用环境和应用要求改变时，容易对概念模型修改和扩充；（4）易于向关系、网状、层次等各种数据模型转换。概念结构的设计策略通常有四种：（1）自顶向下，即首先定义全局概念结构的框架，然后逐步细化；（2）自底向上，即首先定义各局部应用的概念结构，然后将它们集成起来，得到全局概念结构；（3）逐步扩张，首先定义最重要的核心概念结构，然后向外扩充，以滚雪球的方式逐步生成其他概念结构，直至总体概念结构；（4）混合策略，即将自顶向下和自底向上相结合，用自顶向下策略设计一个全局概念结构的框架，以它为骨架集成由自底向上策略中设计的各局部概念结构。

3.8 试述数据库概念结构设计的重要性和自底向上方法的设计步骤。

答：重要性：数据库概念设计是整个数据库设计的关键，将在需求分析阶段所得到的应用需求首先抽象为概念结构，以此作为各种数据模型的共同基础，从而能更好地、更准确地用某一 DBMS 实现这些需求。

设计步骤：概念结构的设计方法有多种，其中最经常采用的策略是自底向上方法，该方法的设计步骤通常分为两步：第 1 步是抽象数据并设计局部视图，第 2 步是集成局部视图，得到全局的概念结构。

3.9 什么是数据库的逻辑设计？试述其设计步骤。

答：数据库的逻辑结构设计就是把概念结构设计阶段设计好的基本 E-R 图转换为与选用的 DBMS 产品所支持的数据模型相符合的逻辑结构。设计步骤为：（1）将概念结构转换为一般的关系、网状、层次模型；（2）将转换来的关系、网状、层次模型向特定 DBMS 支持下的数据模型转换；（3）对数据模型进行优化。

3.10 试述数据库物理设计的内容和步骤。

答：数据库在物理设备上的存储结构与存取方法称为数据库的物理结构，它依赖于给定的 DBMS。为一个给定的逻辑数据模型选取一个最适合应用要求的物理结构，就是数据库的物理设计的主要内容。数据库的物理设计步骤通常分为两步：（1）确定数据库的物理结构，在关系数据库中主要指存取方法和存储结构；（2）对物理结构进行评价，评价的重点是时间效率和空间效率。

3.11 试述数据库实施部署维护的内容和步骤。

答：数据库的实现主要包括了数据库实际结构的建立、数据装入、应用程序开发和调试、测试、试运行和数据库维护，具体步骤分为：（1）选择合适的数据库，编写 SQL 语句用以创建数据库、创建表和定义视图（2）导入数据，将一部分真实数据加载入表中（3）应用软件的开发测试（4）数据库的试运行，包括功能性测试、性能测试和非功能性测试（5）定期备份数据库，检测数据库的安全性和完整性，根据需求重构数据库。

第四章.习题和答案

4.1 为医院设计一个 E-R 图并构建合适的表，医院有很多病人和医生，同每个病人相关的是一系列各种检查和化验的记录。

答：见图 4-1

patients (patient_id, name, insurance, date-admitted, date-checked-out)

doctors (doctor_id, name, specialization)

test (test_id, testname, date, time, result)

doctor-patient (patient_id, doctor_id)

test-log (test_id, patient_id) performed-by (testid, doctor_id)

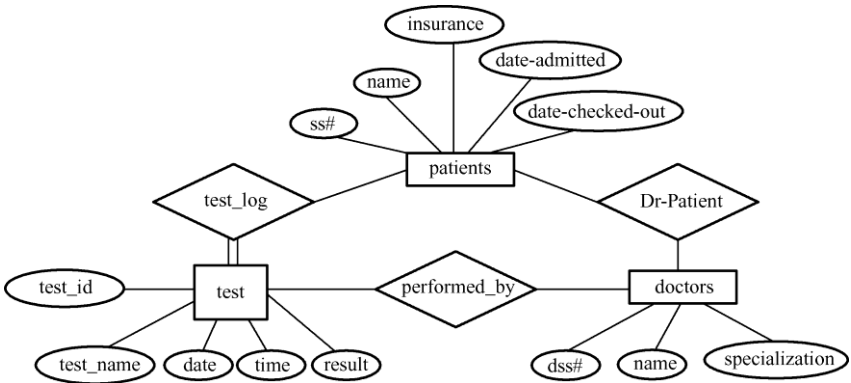


图 4-1 医院 E-R 图

4.2 解释弱实体集和强实体集之间的区别。

答：强实体集中每个实例都能被实体集的主键唯一标识，而弱实体集中每个实例不能用实体集的属性唯一标识。对于从强实体集导出的模式，强实体集的主码就是生成的模式的主码。对于弱实体集导出的模式，该模式的主码由其所依赖的强实体集的主码和弱实体集的分辨符组合而成。

4.3 既然我们可以很简单通过增加一些合适的属性将弱实体集转变为强实体集，那我们为什么还要弱实体集呢？

答：弱实体集存在的理由有：（1）避免数据重复和由强实体的重复键导致的不一致性（2）弱实体集反映了一个实体的逻辑结构，该实体依赖其他实体。（3）当强实体集被删除时，弱实体集会自动被删除。（4）弱实体集可以与它们的强实体集一起被物理存储。

4.4 给出聚集的概念，举出这个概念的两个应用实例。

答：聚集是一种抽象，通过这种抽象，联系被当作高层实体看待。举例如图 4-2 、4-3

4.5 为车辆保险公司设计一个 E-R 图。每个客户有一辆或多辆车，每辆车可以关联 0 次或任意次事故的记录。

答：见图 4.4。

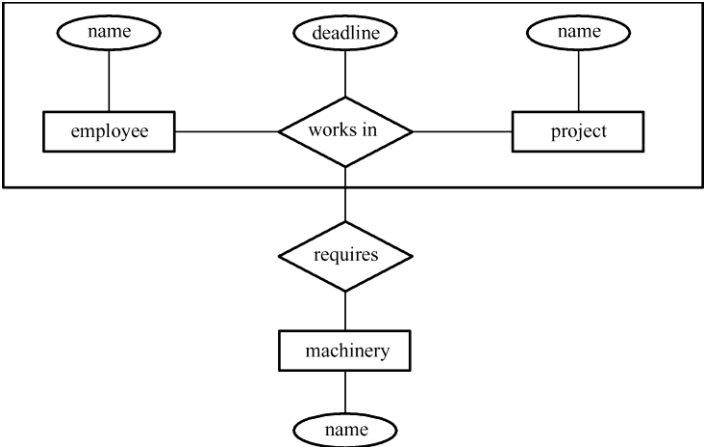


图 4-2 聚集示例一

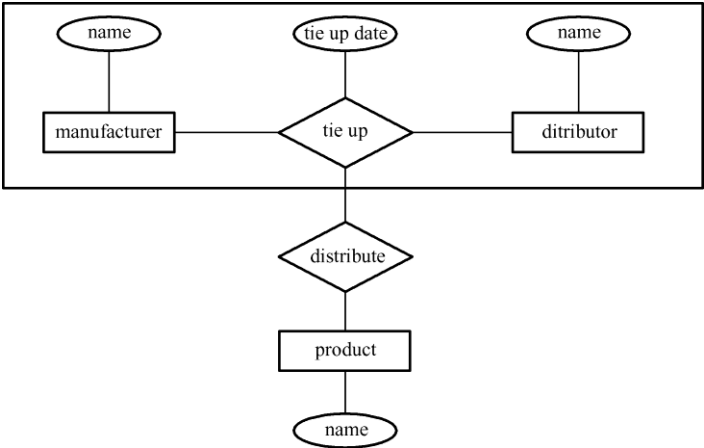


图 4-3 聚集示例二

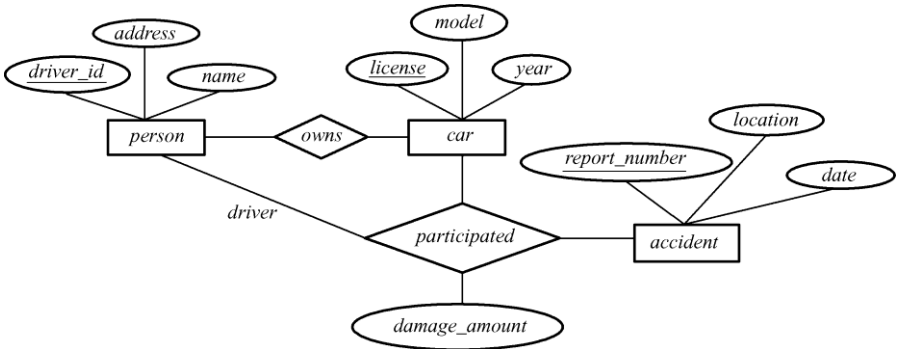


图 4-4 车辆保险公司 E-R 图

4.6 考虑一个用于记录学生各门课程考试成绩的数据库。(1) 构造一个将考试建模成实体的 E-R 图，为以上的数据库设计一个三元联系。(2) 构造另一个 E-R 图，其中只用二元联系来连接 students 和 course_offerings。要求在特定“学生和课程”对之间只有一个联系，而且可以表示出学

生在一门课程的不同考试中的成绩。

答：图 4-5 和 4-6 均可以表示数据库的设计。

4.7 设计一个 E-R 图来跟踪记录读者最喜欢的运动队的成绩。可以保存打过的比赛，每场比赛的分数，上场的队员和每个队员在每场比赛中的统计数据。总的统计数据应该建模成派生属性。

答：见图 4-7

4.8 考虑有相同实体集重复多次出现的一个 E-R 图。简述为什么这样的冗余是应尽量避免的不良设计。

答：使用一个实体集多次会造成在模型中丢失关系。例如在图 4-8 中，参加课程的学生也是参加运动的学生，但这个模型不会表示出来。

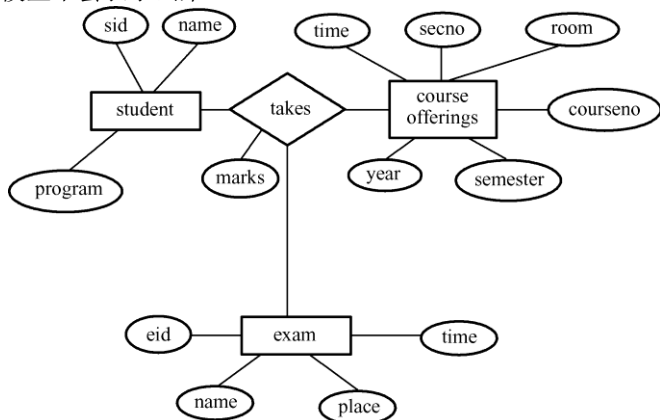


图 4-5 成绩数据库 E-R 图一

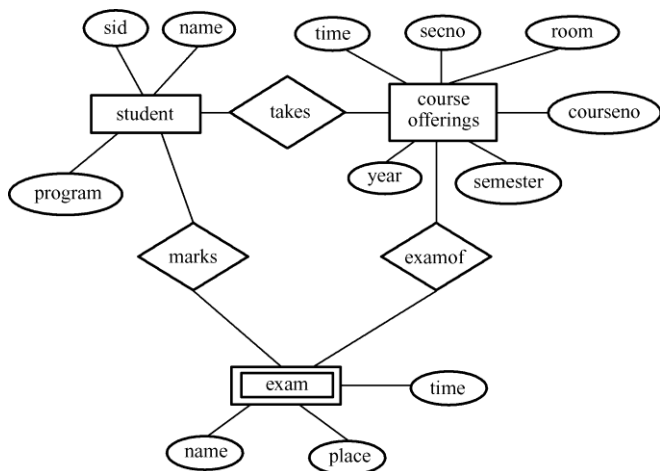


图 4-6 成绩数据库 E-R 图二

4.9 一个弱实体集可以通过在自己的属性中加入其标识实体集中的主码属性来变成一个强实体集。简述如果我们这么做会产生什么样的冗余。

答：一个弱实体集的主码可以从和它与强实体集的联系中推导出。如果我们向弱实体集中加入主

码，那么弱实体集和联系中都出现相同的主码，这是冗余的。

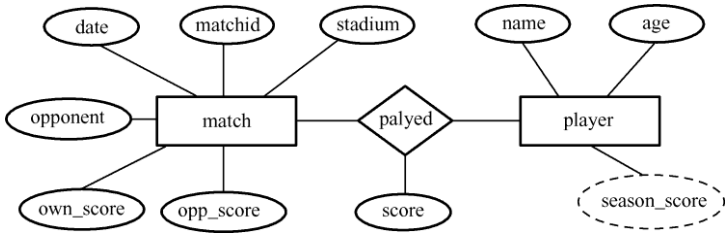


图 4-7 球队统计 E-R 图

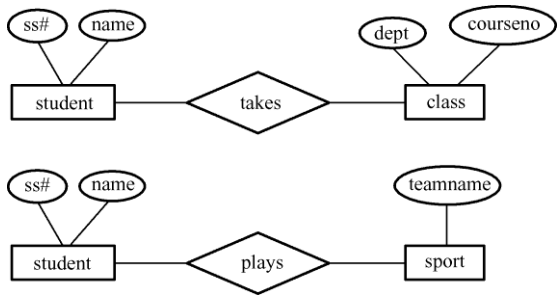


图 4-8 习题 4.8 示例 E-R 图

4.10 阐述映射基数的概念和其四种情况。

答：映射基数称为基数比率，指的是一个实体通过一个联系能同时与多少个实体相关联。

现有实体集 A 和 B，映射的基数分为四种情况：（1）一对一：A 中的一个实体至多同 B 中的一个实体相关联，B 中的一个实体也至多同 A 中的一个实体相关联。（2）一对多：A 中的一个实体可以同 B 中的任意数目实体相关联，而 B 中的一个实体至多同 A 中的一个实体相关联。（3）多对一：A 中的一个实体至多同 B 中的一个实体相关联，而 B 中的一个实体可以同 A 中任意数目的实体相关联。（4）多对多：A 中的一个实体可以同 B 中任意数目的实体相关联，B 中的一个实体也可以同 A 中任意数目的实体相关联。

4.11 E-R 图局部模式之间的冲突主要有 3 种类型，请分别阐述。

答：（1）属性冲突

这类冲突最容易产生于在不同局部模式中使用同一属性时采用了不一致的标记。例如对某个病患年龄的记录追踪，其中一个部门只需要进行记录存档，因此使用了出身年月日的标记；而另一部门则因为要对患者的年龄分布进行分析，而采用了实际岁数的存储方式。

（2）命名冲突

局部模式中使用的数据对象名字与其他模块产生冲突。同名异义和异名同义是常见的命名冲突情况。同名异义指同一名字使用在不同的局部模式中表示了不同的含义，在集成时该名字就会发生语义冲突，不知应该如何赋予该事物含义。异名同义带来的问题则一般是不一致或冗余，可以考虑协商统一命名方式解决。

（3）结构冲突

结构冲突常见的一种情况是同一实体在不同的 E-R 局部视图所包含的属性个数不完全相同，这种情况的发生是因为不同的局部模式所关注数据对象的侧重点是不完全一样的。实际情况中可考虑取属性的合集来解决这一问题。

4.12 解释局部设计、全局设计之间的关系。

答：设计局部 E-R 模式的主要工作是要确定出实体和联系的定义、属性的分配，以及根据系统的实际情况，恰当地划分出各个分系统的局部结构范围。全局设计则集成这些局部模式，形成全局的 E-R 模型。

4.13 为机动车辆销售公司设计一个概化-特殊化层次结构。该公司出售摩托车、小客车、货车和公共汽车。

答：见图 4-9。

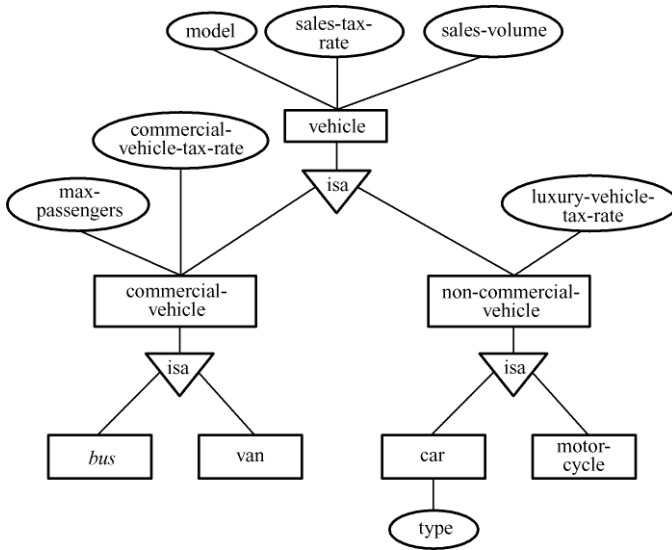


图 4-9 机动车辆销售公司 E-R 图

4.14 阐述数据建模的方法。

答：ER 模型、IDEF1X 模型、EER 模型(Enhanced Entity-Relationship Model)、扩展 ER 模型、IE 模型(Information Engineering Model)、Richard Barker's Notation、EXPRESS-G 表示法、ORM - Object-Role Modeling。

4.15 将下图(原书为图 4-10)转换为 UML。

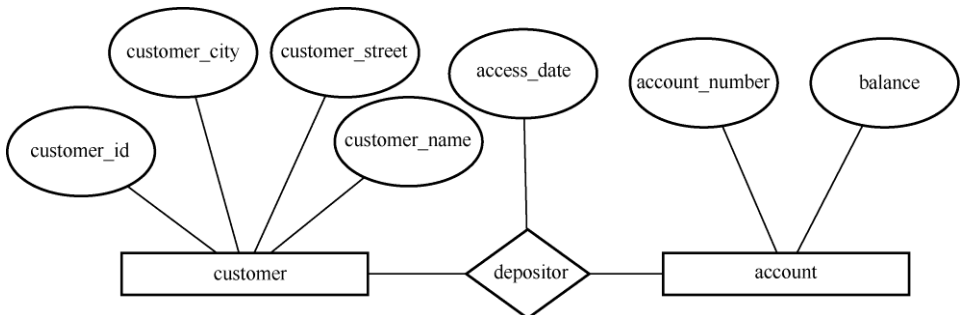
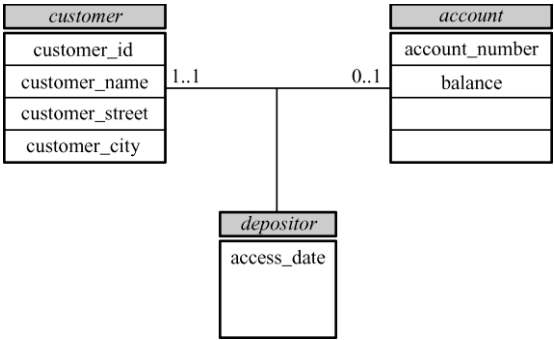


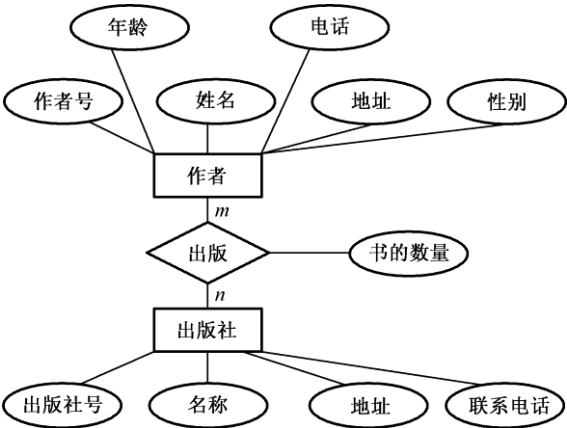
图. 客户-帐户 E-R 模型

答：见下图



4.16 现有一局部应用，包括两个实体：“出版社”和“作者”，这两个实体是多对多的联系，请读者自己设计适当的属性，画出 E - R 图，再将其转换为关系模型（包括关系名、属性名、码和完整性约束条件）。

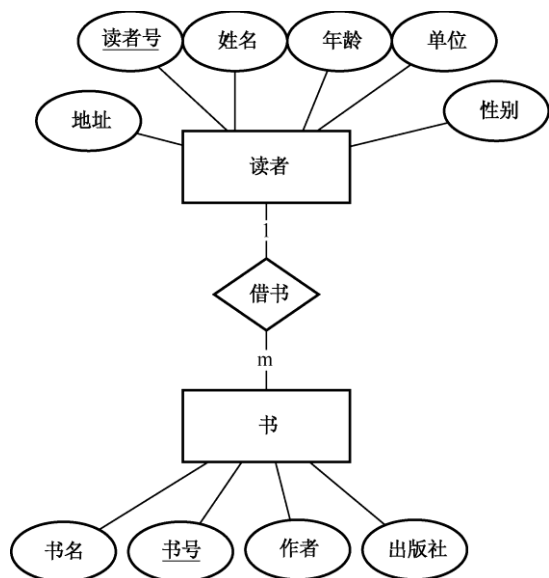
答：



关系模型为：作者（作者号，姓名，年龄，性别，电话，地址）出版社（出版社号，名称，地址，联系电话）出版（作者号，出版社号，书的数量）出版关系的主码作者号，出版社号分别参照作者关系的主码作者号和出版社关系的主码出版社号。

4.17 设计一个图书馆数据库，此数据库中对每个借阅者保存读者记录，包括：读者号，姓名，地址，性别，年龄，单位。对每本书存储：书号，书名，作者，出版社。对每本被借出的书存储：读者号、借出日期和应还日期。要求：给出 E - R 图，再将其转换为关系模型。

答：E - R 图为：



关系模型为:

读者(读者号, 姓名, 地址, 性别)

书(书号, 书名, 作者, 出版社)

借书(读者号, 书号, 借出日期, 年龄, 单位, 应还日期)

4.18 设某汽车运输公司数据库中有三个实体集。一是“车队”实体集, 属性有车队号、车队名等; 二是“车辆”实体集, 属性有牌照号、厂家、出厂日期等; 三是“司机”实体集, 属性有司机编号、姓名、电话等。

车队与司机之间存在“聘用”联系, 每个车队可聘用若干司机, 但每个司机只能应聘于一个车队, 车队聘用司机有“聘用开始时间”和“聘期”两个属性;

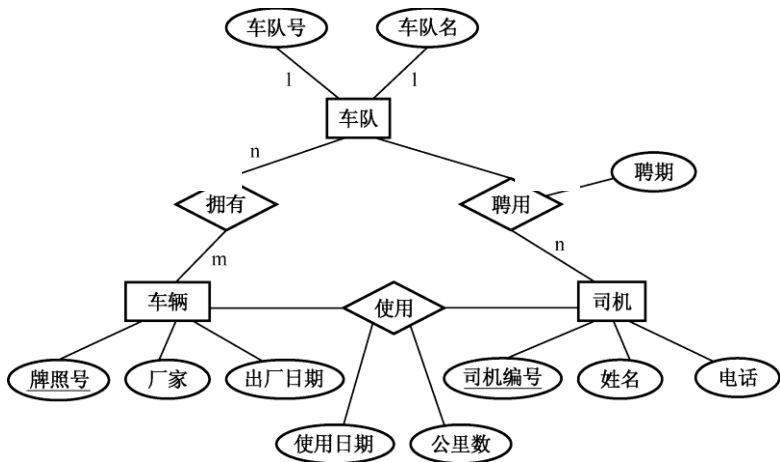
车队与车辆之间存在“拥有”联系, 每个车队可拥有若干车辆, 但每辆车只能属于一个车队;

司机与车辆之间存在着“使用”联系, 司机使用车辆有“使用日期”和“公里数”两个属性, 每个司机可使用多辆汽车, 每辆汽车可被多个司机使用。

(1) 请根据以上描述, 绘制相应的 E-R 图, 并直接在 E-R 图上注明实体名、属性、联系类型;

(2) 将 E-R 图转换成关系模型, 画出相应的数据库模型图, 并说明主键和外键。

答: (1) E-R 图设计如下:



- (2) 转换成的关系模型应具有 4 个关系模式：
- 车队（车队号，车队名）
 - 车辆（车牌照号，厂家，生产日期，车队号）
 - 司机（司机编号，姓名，电话，车队号，聘用开始时间，聘期）
 - 使用（司机编号，车辆号，使用日期，公里数）

4.19 设某商业集团数据库中有三个实体集。一是“仓库”实体集，属性有仓库号、仓库名和地址等；二是“商店”实体集，属性有商店号、商店名、地址等；三是“商品”实体集，属性有商品号、商品名、单价。

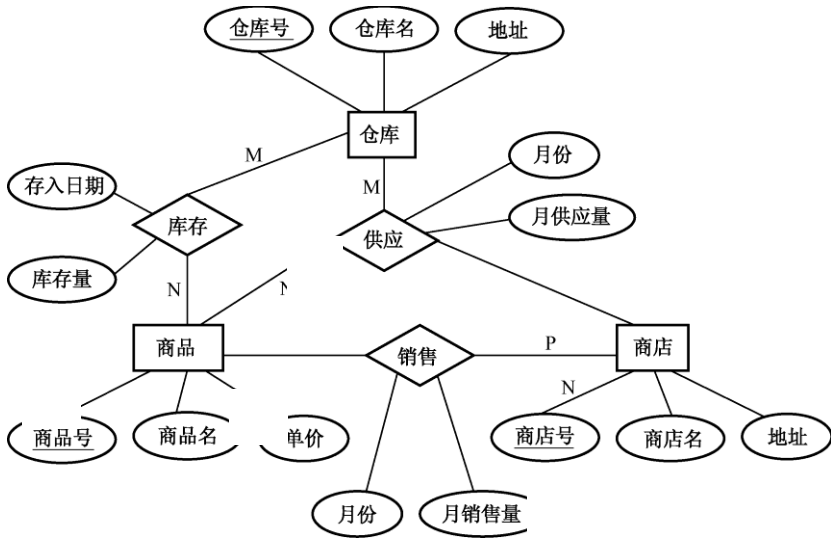
仓库与商品之间存在“库存”联系，每个仓库可存储若干种商品，每种商品存储在若干仓库中，库存有个“库存量”、“存入日期”属性；

商店与商品之间存在着“销售”联系，每个商店可销售若干种商品，每种商品可在若干商店里销售，每个商店销售一种商品有月份和月销售量两个属性；

仓库、商店、商品之间存在一个三元联系“供应”，它反映了把某个仓库中存储的商品供应到某个商店，此联系有月份和月供应量两个属性。

- (1) 请根据以上描述，绘制相应的 E-R 图，并直接在 E-R 图上注明实体名、属性、联系类型；
- (2) 将 E-R 图转换成关系模型，画出相应的数据库模型图，并说明主键和外键。

答：(1) E-R 图设计如下：



(2) 据转换规则, E-R 图可转换成 6 个关系模式:

仓库 (仓库号, 仓库名, 地址)

商品 (商品号, 商品名, 单价)

商店 (商店号, 商店名, 地址)

库存 (仓库号, 商品号, 日期, 库存量)

销售 (商店号, 商品号, 月份, 月销售量)

供应 (仓库号, 商店号, 商品号, 月份, 月供应量)

第五章.习题和答案

5.1 理解并给出下列术语的定义: 函数依赖、部分函数依赖、完全函数依赖、传递函数依赖。

答: 设 $R(U)$ 是属性集 U 上的关系模式。 X, Y 是属性集 U 的子集。若对于 $R(U)$ 的任意一个可能的关系 r , r 中不可能存在两个元组在 X 上的属性值相等, 而在 Y 上的属性值不等, 则称 X 函数确定 Y 或 Y 函数依赖于 X , 记作 $X \rightarrow Y$ 。(即只要 X 上的属性值相等, Y 上的值一定相等。)

$X \rightarrow Y$, 但 Y 不是 X 的子集, 则称 $X \rightarrow Y$ 是非平凡的函数依赖。若不特别声明, 总是讨论非平凡的函数依赖。

$X \rightarrow Y$, 但 Y 是 X 的子集, 则称 $X \rightarrow Y$ 是平凡的函数依赖。若 $X \rightarrow Y$, 则 X 叫做决定因子 (Determinant)。若 $X \rightarrow Y, Y \rightarrow X$, 则记作 $X \leftrightarrow Y$ 。若 Y 不函数依赖于 X , 则记作 $X \not\rightarrow Y$ 。

在 $R(U)$ 中, 如果 $X \rightarrow Y$, 并且对于 X 的任何一个真子集 X' , 都有 $X' \not\rightarrow Y$, 则称 Y 对 X 完全函数依赖, 记作: $X^f \rightarrow Y$ 。

若 $X \rightarrow Y$, 但 Y 不完全函数依赖于 X , 则称 Y 对 X 部分函数依赖, 记作: $X^p \rightarrow Y$ 。

如果 $X \rightarrow Y$ (非平凡函数依赖, 并且 $Y \not\subseteq X$)、 $Y \rightarrow Z$, 则称 Z 传递函数依赖于 X 。

5.2 解释 1NF、2NF、3NF、BCNF。

答: 若关系模式 R 的每一个分量是不可再分的数据项, 则关系模式 R 属于第一范式(1NF)。

若关系模式 $R \in 1NF$, 且每一个非主属性完全函数依赖于码, 则关系模式 $R \in 2NF$ 。(即 1NF 消除了非主属性对码的部分函数依赖则成为 2NF)。

关系模式 $R<U, F>$ 中若不存在这样的码 X 、属性组 Y 及非主属性 Z (Z 不是 Y 的子集)使得 $X \rightarrow Y$, $Y \rightarrow X$, $Y \rightarrow Z$ 成立, 则称 $R<U, F> \in 3NF$ 。(若 $R \in 3NF$, 则每一个非主属性既不部分依赖于码也不传递依赖于码。)

关系模式 $R<U, F> \in 1NF$ 。若 $X \rightarrow Y$ 且 Y 不是 X 的子集时, X 必含有码, 则 $R<U, F> \in BCNF$ 。

5.3 指出下列关系模式是第几范式?并说明理由。

(1) $R(X, Y, Z) \quad F = \{XY \rightarrow Z\}$

(2) $R(X, Y, Z) \quad F = \{Y \rightarrow Z, XZ \rightarrow Y\}$

(3) $R(X, Y, Z) \quad F = \{Y \rightarrow Z, Y \rightarrow X, X \rightarrow YZ\}$

(4) $R(X, Y, Z) \quad F = \{X \rightarrow Y, X \rightarrow Z\}$

(5) $R(X, Y, Z) \quad F = \{XY \rightarrow Z\}$

(6) $R(W, X, Y, Z) \quad F = \{X \rightarrow Z, WX \rightarrow Y\}$

答: (1) R 是 BCNF。

R 候选关键字为 XY , F 中只有一个函数依赖, 而该函数依赖的左部包含了 R 的候选关键字 XY 。

(2) R 是 3NF。

R 候选关键字为 XY 和 XZ , R 中所有属性都是主属性, 不存在非主属性对的候选关键字的传递依赖。

(3) R 是 BCNF。

R 候选关键字为 X 和 Y , $\because X \rightarrow YZ, \therefore X \rightarrow Y, X \rightarrow Z$, 由于 F 中有 $Y \rightarrow Z, Y \rightarrow X$, 因此 Z 是直接函数依赖于 X , 而不是传递依赖于 X 。又 $\because F$ 的每一函数依赖的左部都包含了任一候选关键字, $\therefore R$ 是 BCNF。

(4) R 是 BCNF。

R 的候选关键字为 X , 而且 F 中每一个函数依赖的左部都包含了候选关键字 X 。

(5) R 是 BCNF。

R 的候选关键字为 XY , 而且 F 中函数依赖的左部包含了候选关键字 XY 。

(6) R 是 1NF。

R 的候选关键字为 WX , 则 Y, Z 为非主属性, 又由于 $X \rightarrow Z$, 因此 F 中存在非主属性对候选关键字的部分函数依赖。

5.4 设有关系模式 $R(U, F)$, 其中:

$U = \{A, B, C, D, E, P\}, F = \{A \rightarrow B, C \rightarrow P, E \rightarrow A, CE \rightarrow D\}$

求出 R 的所有候选关键字。

答: 根据候选关键字的定义: 如果函数依赖 $X \rightarrow U$ 在 R 上成立, 且不存在任何 $X' \subset X$, 使得 $X' \rightarrow U$ 也成立, 则称 X 是 R 的一个候选关键字。由此可知, 候选关键字只可能由 A, C, E 组成, 但有 $E \rightarrow A$, 所以组成候选关键字的属性可能是 CE 。

计算可知: $(CE)^+=ABCDEP$, 即 $CE \rightarrow U$
而: $C^+=CP$, $E^+=ABE$ $\therefore R$ 只有一个候选关键字 CE 。

5.5 设有关系模式 $R(C, T, S, N, G)$, 其上的函数依赖集:

$$F=\{C \rightarrow T, CS \rightarrow G, S \rightarrow N\}$$

求出 R 的所有候选关键字。

答: 根据候选关键字的定义, R 的候选关键字只可能由 F 中各个函数依赖的左边属性组成, 即 C, S , 所以组成候选关键字的属性可能是 CS 。

计算可知: $(CS)^+=CGNST$, 即 $CS \rightarrow U$

而: $C^+=CT$, $S^+=NS$

$\therefore R$ 只有一个候选关键字 CS 。

5.6 设有关系模式 $R(U, F)$, 其中:

$$U=\{A, B, C, D, E\}, F=\{A \rightarrow D, E \rightarrow D, D \rightarrow B, BC \rightarrow D, DC \rightarrow A\}$$

(1) 求出 R 的候选关键字。

(2) 判断 $\rho=\{AB, AE, CE, BCD, AC\}$ 是否为无损连接分解?

答: (1) $(CE)^+=ABCDE$, 则 $CE \rightarrow U$, 而 $C^+=C$, $E^+=DE=BDE$, 根据候选关键字定义, CE 是 R 的候选关键字。

(2) ρ 的无损连接性判断表如下表所示, 由此判断不具有无损连接性。

R_i	A	B	C	D	E
AB	a1	a2			
AE	a1				a5
CE			a3		a5
BCD		a2	a3	a4	
AC	a1		a3		

5.7 设有关系模式 $R(U, F)$, 其中:

$$U=\{E, F, G, H\}, F=\{E \rightarrow G, G \rightarrow E, F \rightarrow EG, H \rightarrow EG, FH \rightarrow E\}$$

求 F 的最小依赖集。

答: (1) 将 F 中依赖右部属性单一化:

$$F1=\{E \rightarrow G, G \rightarrow E, F \rightarrow E, F \rightarrow G, H \rightarrow E, H \rightarrow G, FH \rightarrow E\}$$

(2) 对于 $FH \rightarrow E$, 由于有 $F \rightarrow E$, 则为多余的, 则:

$$F2=\{E \rightarrow G, G \rightarrow E, F \rightarrow E, F \rightarrow G, H \rightarrow E, H \rightarrow G\}$$

(3) 由于 $E \rightarrow G$, 所以在 $F2$ 中的 $F \rightarrow E$ 和 $F \rightarrow G$ 以及 $H \rightarrow E$ 和 $H \rightarrow G$ 之一是多余的, 则:

$$F3=\{E \rightarrow G, G \rightarrow E, F \rightarrow G, H \rightarrow G\}$$

$$\text{或 } F3=\{E \rightarrow G, G \rightarrow E, F \rightarrow G, H \rightarrow E\}$$

$$\text{或 } F3=\{E \rightarrow G, G \rightarrow E, F \rightarrow E, H \rightarrow E\}$$

$$\text{或 } F3=\{E \rightarrow G, G \rightarrow E, F \rightarrow E, H \rightarrow G\}$$

5.8 设有关系模式 $R(U, F)$ ，其中：

$U=\{A, B, C, D\}$ ， $F=\{A\rightarrow B, B\rightarrow C, D\rightarrow B\}$ ，把 R 分解成 BCNF 模式集：

- (1) 如果首先把 R 分解成 $\{ACD, BD\}$ ，试求 F 在这两个模式上的投影。
- (2) ACD 和 BD 是 BCNF 吗？如果不是，请进一步分解。

答：(1) $\Pi_{ACD}(F)=\{A\rightarrow C, D\rightarrow C\}$ $\Pi_{BD}(F)=\{D\rightarrow B\}$

(2) BD 已是 BCNF。 ACD 不是 BCNF。模式 ACD 的候选关键字是 AD 。考虑 $A\rightarrow C$ ， A 不是模式 ACD 的候选关键字，所以这个函数依赖不满足 BCNF 条件。将 ACD 分解为 AC 和 AD ，此时 AC 和 AD 均为 BCNF。

5.9 已知关系模式 $R(CITY, ST, ZIP)$ 和函数依赖集：

$F=\{ (CITY, ST) \rightarrow ZIP, ZIP \rightarrow CITY \}$

试找出 R 的两个候选关键字。

答：设 $U=(CITY, ST, ZIP)$ ， F 中函数依赖的左边是 $CITY, ST, ZIP$ ：
由于 $ZIP\rightarrow CITY$ ，去掉 $CITY$ ，故 (ST, ZIP) 可能是候选关键字。

$(ST, ZIP)^+=\{ST, ZIP, CITY\}$ ， $\therefore (ST, ZIP)\rightarrow U$ 。
又 $ST^+=ST, ZIP^+=\{ZIP, CITY\}$ ，故 (ST, ZIP) 是一个候选关键字。
由于 $(CITY, ST) \rightarrow ZIP$ ，去掉 ZIP ，故 $(CITY, ST)$ 可能是候选关键字。
 $(CITY, ST)^+=\{CITY, ST, ZIP\}$ ， $\therefore (CITY, ST)\rightarrow U$ 。

又 $CITY^+=CITY, ST^+=ST$ ，故 $(CITY, ST)$ 是一个候选关键字。
因此， R 的两个候选关键字是 (ST, ZIP) 和 $(CITY, ST)$ 。

5.10 设有关系模式 $R(A, B, C, D, E)$ ， R 的函数依赖集：

$F=\{A\rightarrow D, E\rightarrow D, D\rightarrow B, BC\rightarrow D, CD\rightarrow A\}$

- (1) 求 R 的候选关键字。
- (2) 将 R 分解为 3NF。

答：(1) 设 $U=(A, B, C, D, E)$ ，由于 $(CE)^+=ABCDE, C^+=C, E^+=BDE$
 $\therefore R$ 的候选关键字是 CE 。

(2) 求出最小依赖集 $F'=\{A\rightarrow D, E\rightarrow D, D\rightarrow B, BC\rightarrow D, CD\rightarrow A\}$
将 R 分解的 3NF: $\rho=\{AD, DE, BD, BCD, ACD\}$ 。

5.11 设有关系模式 $R(U, V, W, X, Y, Z)$ ，其函数依赖集：

$F=\{U\rightarrow V, W\rightarrow Z, Y\rightarrow U, WY\rightarrow X\}$ ，现有下列分解：

- (1) $\rho_1=\{WZ, VY, WXY, UV\}$
- (2) $\rho_2=\{UVY, WXYZ\}$

判断上述分解是否具有无损连接性。

答：(1) ρ_1 的无损连接性判断表如下所示，由此判断 ρ_1 不具有无损连接性。

Ri	U	V	W	X	Y	Z
WZ			a3			a6
VY		a2			a5	
WXY			a3	a4	a5	a6

UV	a1	a2				
----	----	----	--	--	--	--

(2) ρ_2 的无损连接性判断表如下所示，由此判断 ρ_2 具有无损连接性。

Ri	U	V	W	X	Y	Z
UVY	a1	a2			a5	
WXYZ	a1	a2	a3	a4	a5	a6

5.12 设有关系模式 $R(F, G, H, I, J)$ ， R 的函数依赖集：

$F \rightarrow I, J \rightarrow I, I \rightarrow G, GH \rightarrow I, IH \rightarrow F$

- (1) 求出 R 的所有候选关键字。
- (2) 判断 $\rho = \{FG, FJ, JH, IGH, FH\}$ 是否为无损连接分解？
- (3) 将 R 分解为 3NF，并具有无损连接性和依赖保持性。

答：(1) 从 F 中看出，候选关键字中至少包含 J 和 H (因为它们不依赖于谁)，计算：
令 $X = \{JH\}$ ， $X(0) = JH$ ， $X(1) = IJH$ ， $X(2) = GIJH$ ， $X(3) = FGIJH$
 \therefore 候选关键字只有 JH 。

(2) ρ 的无损连接性判断表如下所示，由此判断 ρ 不具有无损连接性。

Ri	F	G	H	I	J
FG	a1	a2			
FJ	a1		a3	a4	a5
JH			a3		a5
IGH		a2	a3	a4	
FH	a1		a3		

(3) 求出最小依赖集 $F' = \{F \rightarrow I, J \rightarrow I, I \rightarrow G, GH \rightarrow I, IH \rightarrow F\}$

\therefore 满足 3NF 且具有依赖保持性的分解为：

$\rho = \{FI, JI, IG, GHI, IHE\}$

ρ 的无损连接性判断结果如下所示，由此判断 ρ 不具有无损连接性。

Ri	F	G	H	I	J
FI	a1	a2		a4	
JI		a2		a4	a5
IG		a2		a4	a5
GHI	a1	a2	a3	a4	
IHE	a1	a2	a3	a4	

令 $\rho = \rho \cup \{JH\}$ ， JH 是 R 的候选关键字。
 $\therefore \rho = \{FI, JI, IG, GHI, IHE, JH\}$ 具有无损连接性和依赖保持性

5.13 设有关系模式 $R(A, B, C, D, E)$ ，其上的函数依赖集：

$F = \{A \rightarrow C, C \rightarrow D, B \rightarrow C, DE \rightarrow C, CE \rightarrow A\}$

- (1) 求 R 的所有候选关键字。
- (2) 判断 $\rho=\{AD, AB, BC, CDE, AE\}$ 是否为无损连接分解?
- (3) 将 R 分解为 BCNF, 并具有无损连接性。

答：(1) 从 F 中看，候选关键字至少包含 BE(因为它们不依赖于谁)，而 $(BE)^+=ABCDE$
 \therefore BE 是 R 的惟一候选关键字。

(2) ρ 的无损连接性判断结果如下所示，由此判定 ρ 不具有无损连接性。

Ri	A	B	C	D	E
AD	a1		a3	a4	
AB	a1	a2	a3	a4	
BC		a2	a3	a4	
CDE	a1		a3	a4	a5
AE	a1		a3	a4	a5

(3) 考虑 $A \rightarrow C$

\because AC 不是 BCNF(AC 不包含候选关键字 BE) 将 ABCDE 分解为 AC 和 ABDE, AC 已是 BCNF。
进一步分解 ABDE, 选择 $B \rightarrow D$, 把 ABDE 分解为 BD 和 ABE, 此时 BD 和 ABE 均为 BCNF。 $\therefore \rho=\{AC, BD, ABE\}$

5.14 建立一个关于系、学生、班级、学会等信息的关系数据库。

学生：学号、姓名、出生年月、系名、班号、宿舍区。

班级：班号、专业名、系名、人数、入校年份。

系：系名、系号、系办公地点、人数。

学会：学会名、成立年份、办公地点、人数。

语义如下：一个系有若干专业，每个专业每年只招一个班，每个班有若干学生。一个系的学生住在同一宿舍区。每个学生可参加若干学会，每个学会有若干学生。学生参加某学会会有一个入会年份。

请给出关系模式，写出每个关系模式的极小函数依赖集，指出是否存在传递函数依赖，对于函数依赖左部是多属性的情况讨论函数依赖是完全函数依赖，还是部分函数依赖。指出各关系模式的候选码、外部码，有没有全码存在？

答：(1) 关系模式如下：

学生：S(Sno, Sname, Sbirth, Dept, Class, Rno)

班级：C(Class, Pname, Dept, Cnum, Cyear)

系：D(Dept, Dno, Office, Dnum)

学会：M(Mname, Myear, Maddr, Mnum)

(2) 每个关系模式的最小函数依赖集如下：

A、学生 S (Sno, Sname, Sbirth, Dept, Class, Rno) 的最小函数依赖集如下: $Sno \rightarrow Sname, Sno \rightarrow Sbirth,$
 $Sno \rightarrow Class, Class \rightarrow Dept, DEPT \rightarrow Rno$

传递依赖如下：

由于 $Sno \rightarrow Dept$, 而 $Dept \rightarrow Sno, Dept \rightarrow Rno$ (宿舍区)

所以 Sno 与 Rno 之间存在着传递函数依赖。

由于 $Class \rightarrow Dept, Dept \rightarrow Class, Dept \rightarrow Rno$

所以 Class 与 Rno 之间存在着传递函数依赖。

由于 $Sno \rightarrow Class$, $Class \rightarrow Sno$, $Class \rightarrow Dept$

所以 Sno 与 Dept 之间存在着传递函数依赖。

B、班级 C(Class, Pname, Dept, Cnum, Cyear)的最小函数依赖集如下:

$Class \rightarrow Pname$, $Class \rightarrow Cnum$, $Class \rightarrow Cyear$, $Pname \rightarrow Dept$.

由于 $Class \rightarrow Pname$, $Pname \rightarrow Class$, $Pname \rightarrow Dept$

所以 Class 与 Dept 之间存在着传递函数依赖。

C、系 D(Dept, Dno, Office, Dnum)的最小函数依赖集如下:

$Dept \rightarrow Dno$, $Dno \rightarrow Dept$, $Dno \rightarrow Office$, $Dno \rightarrow Dnum$

根据上述函数依赖可知, Dept 与 Office, Dept 与 Dnum 之间不存在传递依赖。

D、学会 M(Mname, Myear, Maddr, Mnum)的最小函数依赖集如下:

$Mname \rightarrow Myear$, $Mname \rightarrow Maddr$, $Mname \rightarrow Mnum$

该模式不存在传递依赖。

(3) 各关系模式的候选码、外部码, 全码如下:

A、学生 S 候选码: Sno; 外部码: Dept、Class; 无全码

B、班级 C 候选码: Class; 外部码: Dept; 无全码

C、系 D 候选码: Dept 或 Dno; 无外部码; 无全码

D、学会 M 候选码: Mname; 无外部码; 无全码

5.15 下面的结论哪些是正确的? 哪些是错误的? 对于错误的请给一个反例说明。

(1) 任何一个二目关系是属于 3NF。

答: 正确。因为关系模式中只有两个属性, 所以无传递。

(2) 任何一个二目关系是属于 BCNF。

答: 正确。按 BCNF 的定义, 若 $X \rightarrow Y$, 且 Y 不是 X 的子集时, 每个决定因素都包含码, 对于二目关系决定因素必然包含码。详细证明如下: (任何二元关系模式必定是 BCNF)。

证明: 设 R 为一个二目关系 $R(A1, A2)$, 则属性 A1 和 A2 之间可能存在以下几种依赖关系:

A、 $A1 \rightarrow A2$, 但 $A2 \not\rightarrow A1$, 则关系 R 的码为 A1, 决定因素都包含码, 所以, R 是 BCNF。

B、 $A1 \rightarrow A2$, $A2 \rightarrow A1$, 则关系 R 的码为 A2, 所以决定因素都包含码, R 是 BCNF。

C、R 的码为 $(A1, A2)$ (即 $A1 \rightarrow A2$, $A2 \rightarrow A1$), 决定因素都包含码。R 是 BCNF。

(3) 任何一个二目关系是属于 4NF。

答: 正确。因为只有两个属性, 所以无非平凡的多值依赖。