

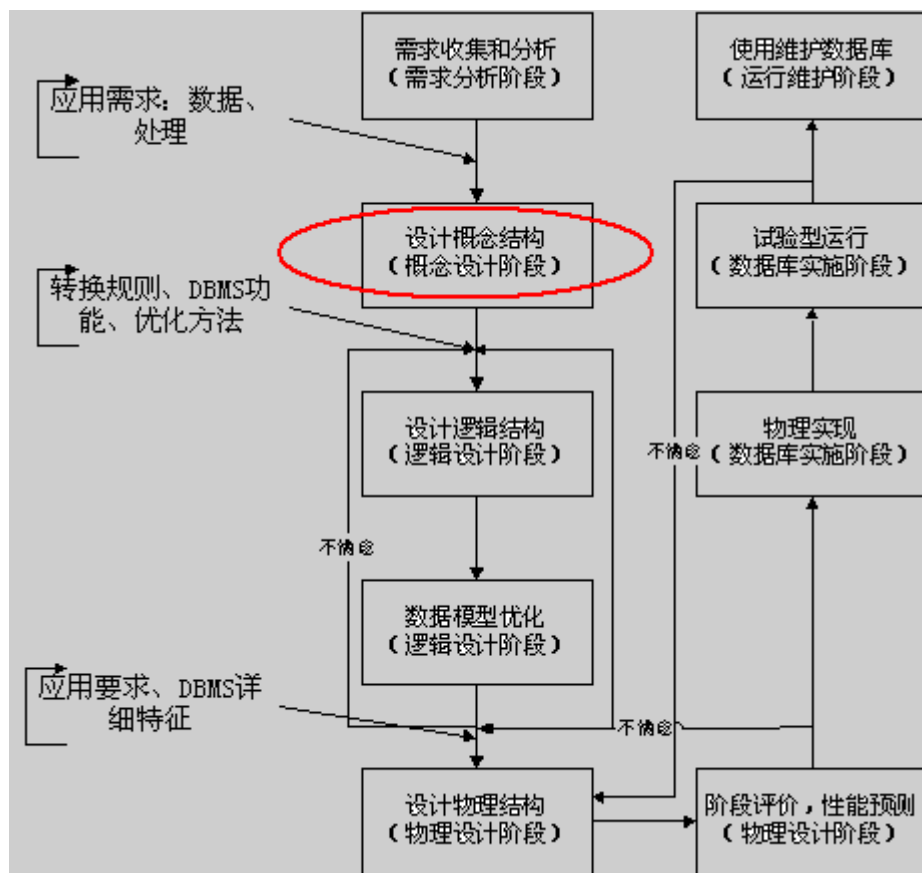
数据库设计的基本步骤

- (1) 需求分析阶段：需求收集和分析，得到数据字典和数据流图。
- (2) 概念结构设计阶段：对用户需求综合、归纳与抽象，形成概念模型，用 E-R 图表示。
- (3) 逻辑结构设计阶段：将概念结构转换为某个 DBMS 所支持的数据模型。
- (4) 数据库物理设计阶段：为逻辑数据模型选取一个最适合应用环境的物理结构。
- (5) 数据库实施阶段：建立数据库，编制与调试应用程序，组织数据入库，程序试运行。
- (6) 数据库运行和维护阶段：对数据库系统进行评价、调整与修改。

1 数据库设计概述

数据库设计是指对于一个给定的应用环境，构造最优的数据库模式，建立数据库及其应用系统，使之能够有效地存储数据。
数据库设计的基本步骤：

- 需求分析
- 概念结构设计
- 逻辑结构设计
- 物理结构设计
- 数据库的建立和测试
- 数据库运行和维护。



数据库各阶段设计描述

设计阶段	设计描述
需求分析	数据字典、全系统中数据项、数据流、数据存储的描述
概念结构设计	概念模型 (E-R图) 数据字典
逻辑结构设计	关系数据模型
物理设计	存储安排、方法选择、存取路径建立
实施阶段	编写模式、装入数据、数据库试运行
运行维护	性能监测、转储/恢复、数据库重组与重构

2 概念结构设计

在早期的数据库设计，在需求分析阶段后，就直接进行逻辑结构设计。由于此时既要考虑现实世界信息的联系与特征，又要满足特定的数据库系统的约束要求，因而对于客观世界的描述受到一定的限制。同时，由于设计时要同时考虑多方面的问题，也使设计工作变得十分复杂。1976年 P. P. S. Chen 提出在逻辑结构设计之前先设计一个概念模型，并提出了数据库设计的实体—联系方法 (Entity—Relationship Approach)。这种方法不包括深的理论，但提供了一个简便、有效的方法，目前成为数据库设计中通用的工具。

有许多商业软件支持 E-R 模型，如 Sybase 公司的 PowerDesigner DataArchitect (最新版本 v9.5.1 for Windows)、微软公司 Microsoft InfoModeler (VisioModeler) 等。

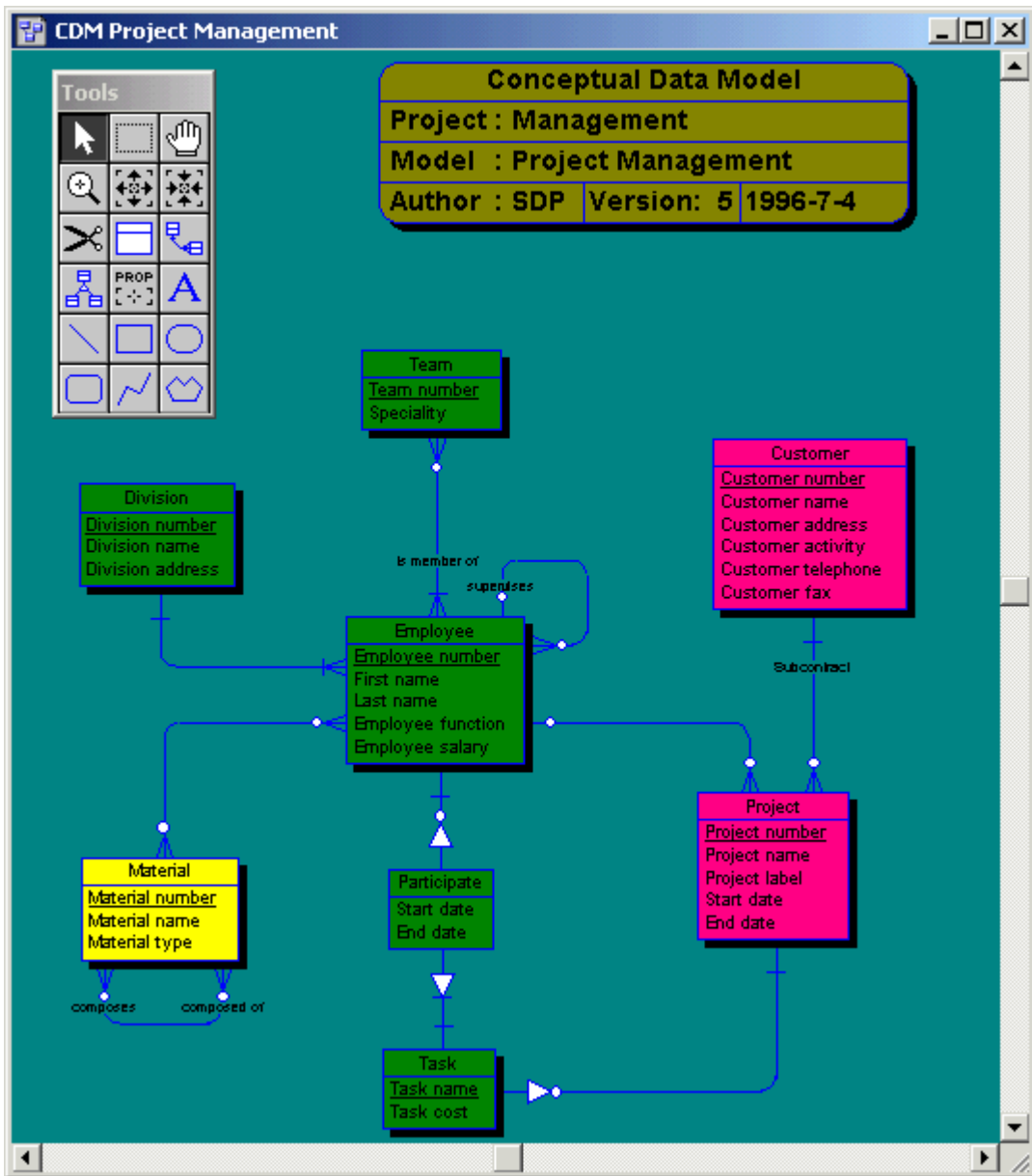
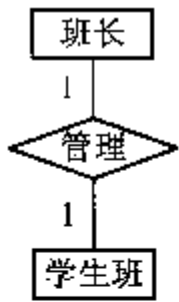


图 S-designer DataArchitect 5.1 设计的 E-R 模型

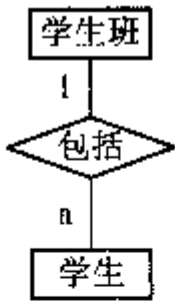
使用 E-R 模型来进行概念模型的设计通常分两步进行，首先是建立局部概念模型，然后综合局部概念模型，成为全局概念模型。

2.1 E-R 模型基本符号

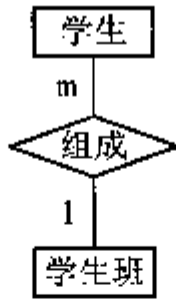
- 实体的表示：用长方形
- 联系的表示：用菱形，1:1、1:n (m:1)、(m:n)
- 属性的表示：用椭圆形



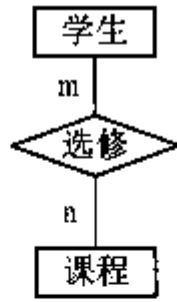
(a)



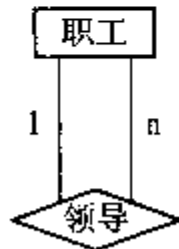
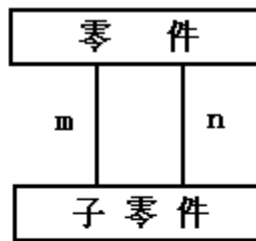
(b)



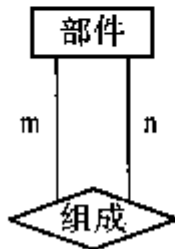
(c)



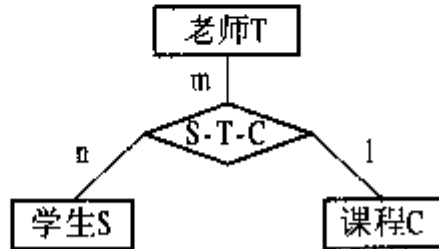
(d)



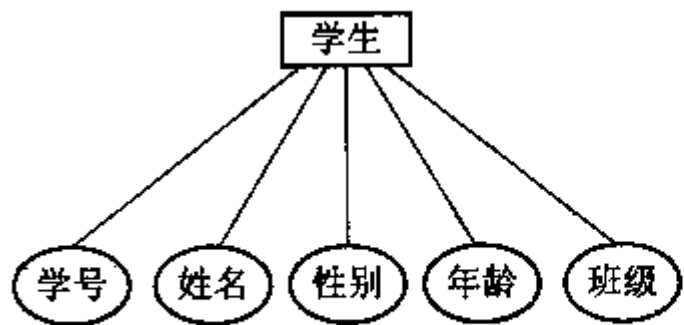
(e)



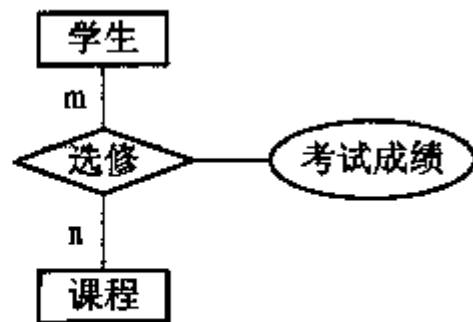
(f)



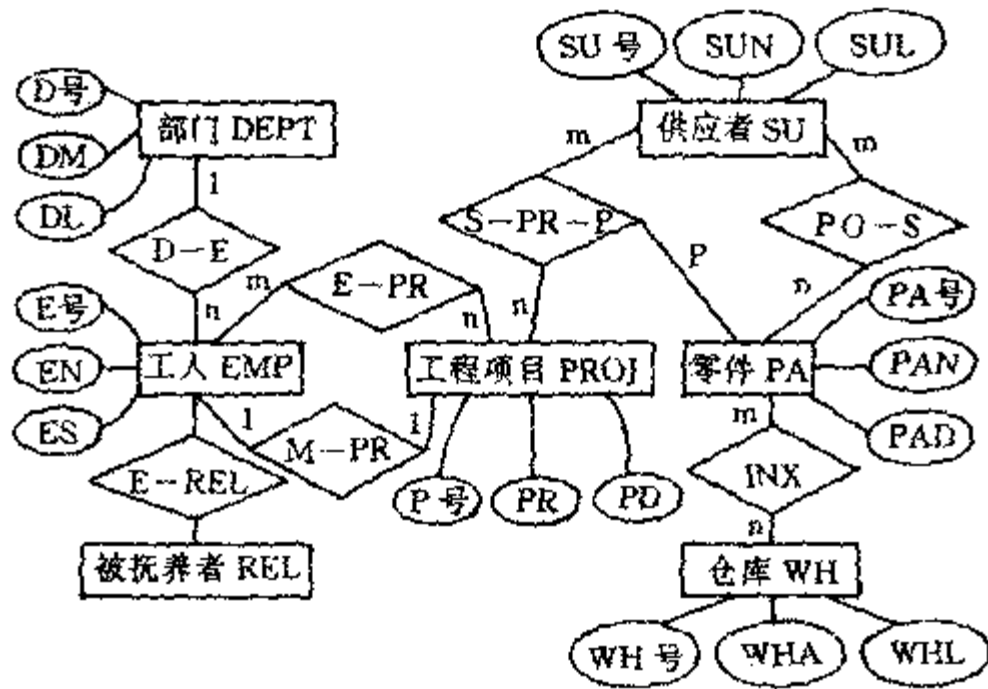
(g)



(a)



(b)



E-R 图具有以下几个特性:

- 一个联系集合可以定义在两个或两个以上的实体集合上, 例如老师--学生--课程的联系集合 S-T-C, 就是定义在三个实体上。
- 一个联系集合也可以定义在一个实体集合上, 例如零件下又分有子零件, 每个零件又可由 m 个子零件组成, 每个子零件又可组合成 n 个零件。

- 对于给定的实体集合，可以定义一个以上的联系集合，例如工程项目--工人可以定义两个联系集合，其中一个表示工程项目和工人的联系，另一个表示工程项目和工人中的工程项目负责人的联系。前者是 $n:m$ 的联系，后者是 $1:1$ 的联系。
- 实体联系图可以表示一个实体类型对另一个实体类型的存在的依赖性，例如工人这一实体下反映其被抚养者的关系，就是依赖关系，这种联系用箭头表示，说明抚养者这个实体的存在取决于工人中的相应的那个实体。

2.2 初步 E-R 图设计

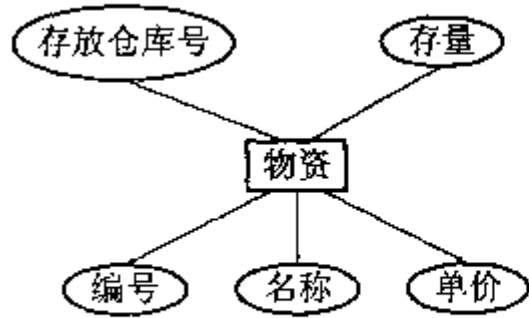
在数据分析的基础上，就可以着手设计概念结构。设计初步 E-R 图的步骤：

- 先设计局部 E-R 图，也称用户视图
- 综合各局部 E-R 图，形成总的 E-R 图，即用户视图的集成。

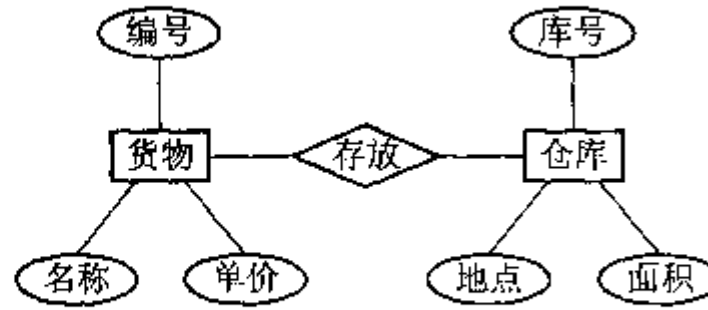
在设计初步 E-R 图时，要尽量能充分地把组织中各部门对信息的要求集中起来，而不需要考虑数据的冗余问题。

局部概念模型设计是从用户的观点出发，设计符合用户需求的概念结构。局部概念模型设计的就是组织、分类收集到的数据项，确定哪些数据项作为实体，哪些数据项作为属性，哪些数据项是同一实体的属性等。确定实体与属性的原则：

- 能作为属性的尽量作为属性而不要划为实体；
- 作为属性的数据元素与所描述的实体之间的联系只能是 $1:n$ 的联系；
- 作为属性的数据项不能再用其他属性加以描述，也不能与其他实体或属性发生联系。

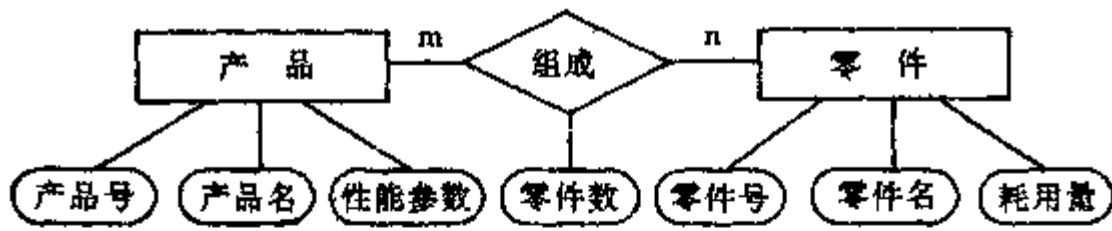


(a) 仓库作为属性

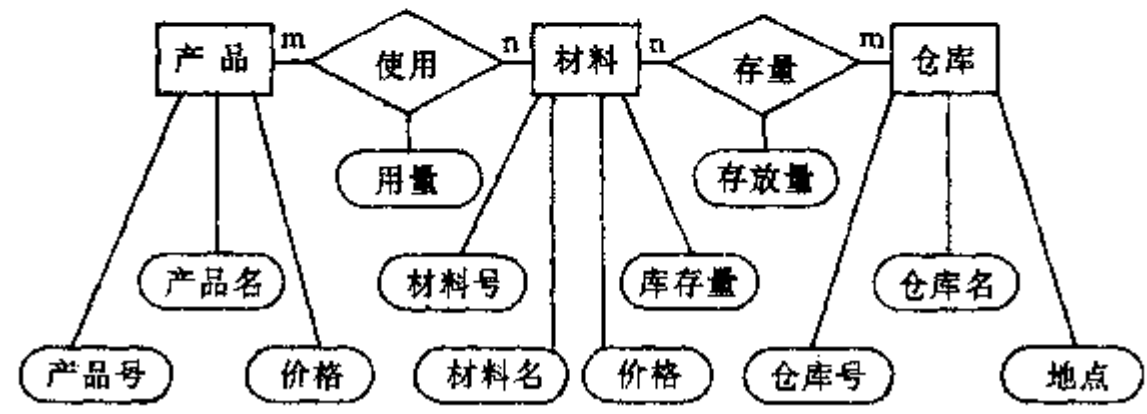


(b) 仓库作为一个实体

例 1：一个机械制造厂的简单管理系统。首先按工厂技术部门和工厂供应部门设计两个局部 E-R 图。工厂技术部门关心的是产品的性能参数，及由哪些零件组成，零件的材料和耗用量等；工厂供应部门关心的是产品的价格，使用材料的价格及库存量等。



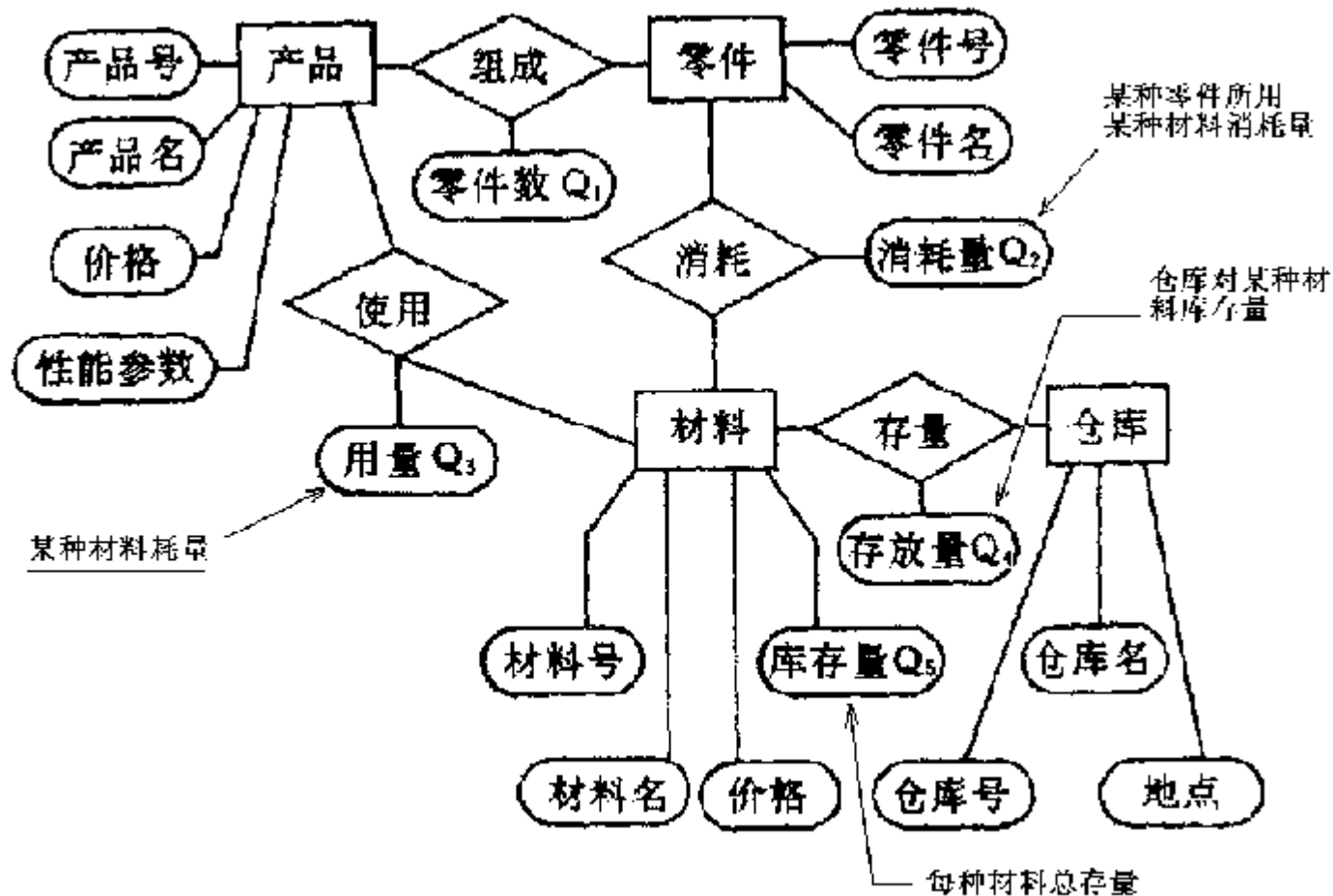
技术部门



供应部门

例 1 图 局部 E-R 图

综合这两个分 E-R 图，得到初步 E-R 图。



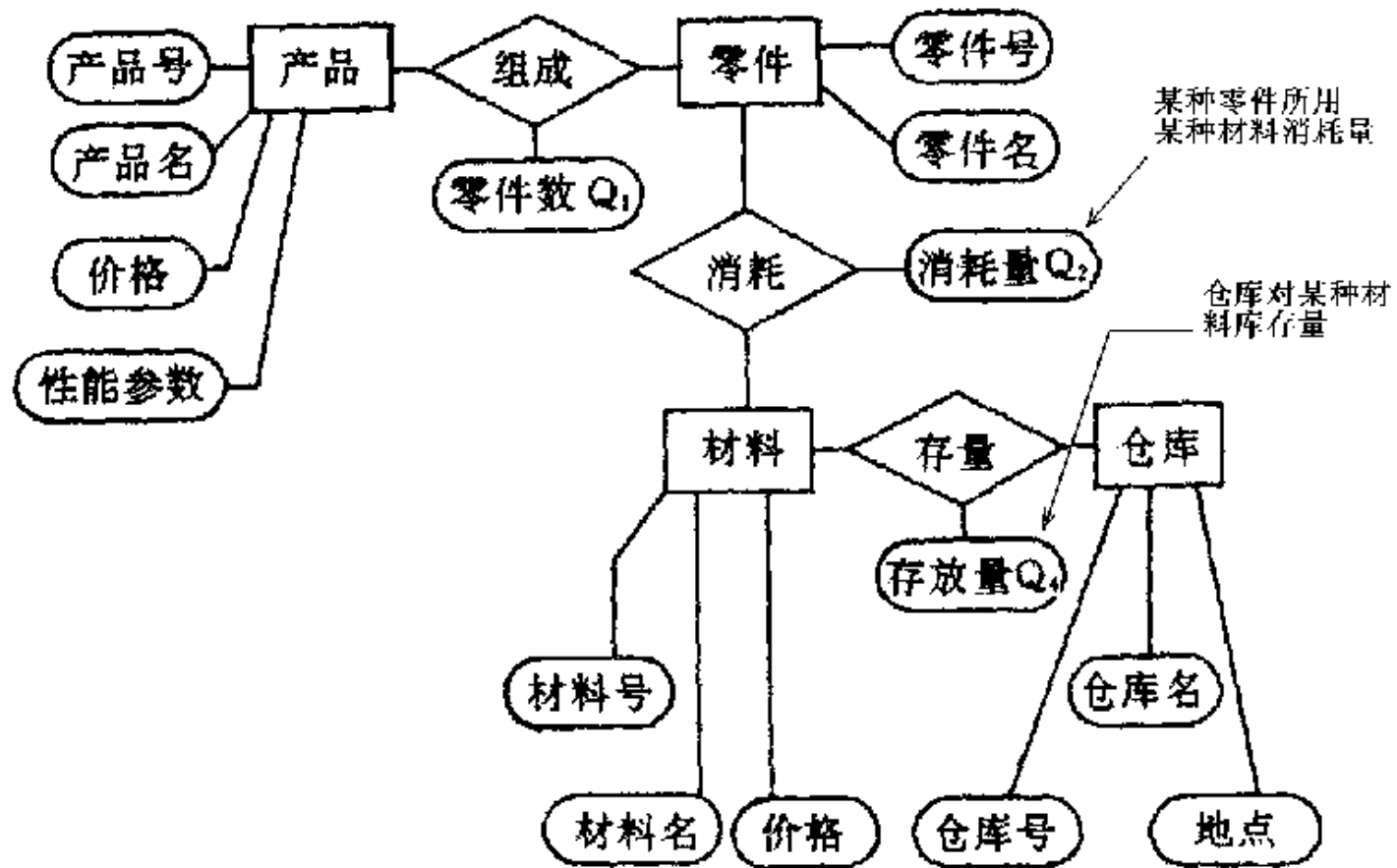
例 1 图 初步 E-R 图

初步 E-R 图是现实世界的纯粹表示,可能存在冗余的数据和实体间冗余的联系。所谓冗余的数据是指可由基本数据导出的数据,冗余的联系是指可由基本联系导出的联系。

2.3 基本 E-R 图设计

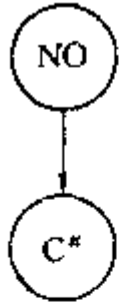
初步 E-R 图由于存在冗余的信息,会破坏数据库的完整性,给数据库的管理带来麻烦,以至引起数据不一致的错误。因此,必须消除数据上的冗余和联系上的冗余,消除冗余后的 E-R 图,称为基本 E-R 图。

消除冗余的方法,可以采用分析的方法。

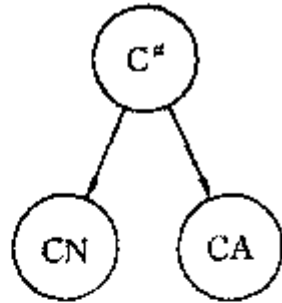


例 1 图 基本 E-R 图

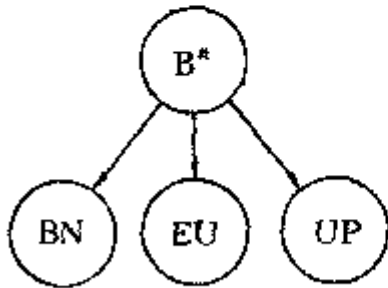
例 2: 某书店购图书汇总登记表, 经过 1NF、2NF、3NF 分析得



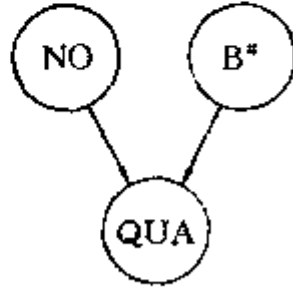
订书单汇编
(关键字 NO)



订户索引
(关键字 C*)

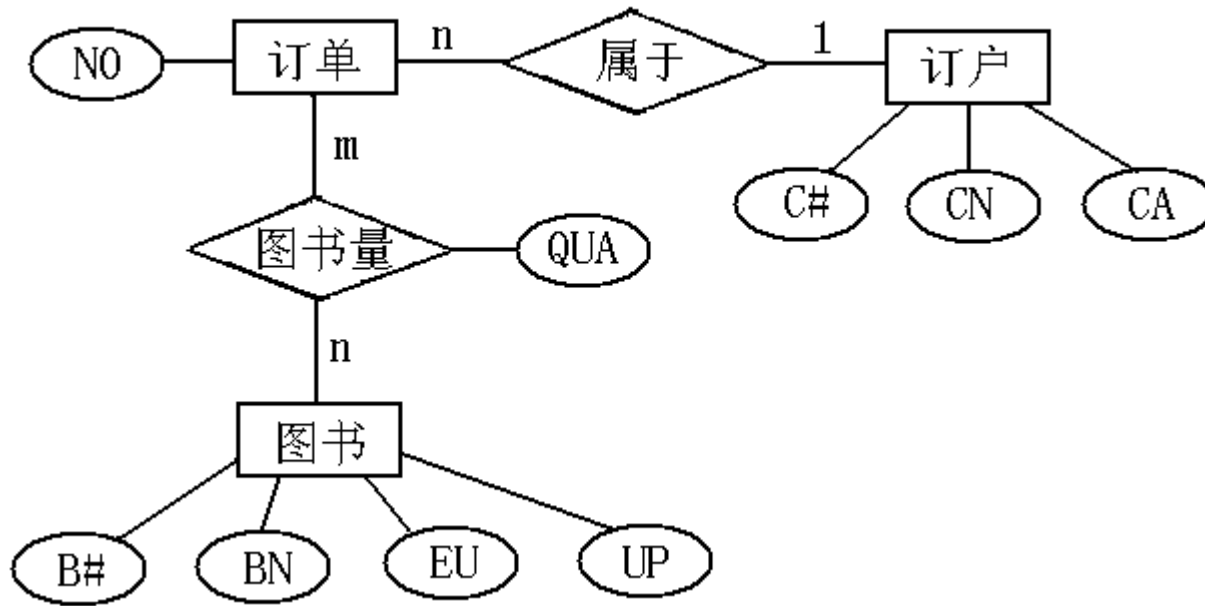


图书索引
(关键字 B*)



订书单细节
(关键字 NO, B*)

对应的基本 E-R 图



例 2 图 图书汇总登记表基本 E-R 图

基本 E-R 图满足第三范式要求。

用例 1 的分析方法，对于一个庞大的复杂系统是有困难的，可以将初步 E-R 图转换成函数依赖关系，利用规范化理论，将其分解为多个符合第三范式的函数依赖关系，再构成基本 E-R 图。

3 逻辑结构设计

任务：将基本 E-R 图转换为与选用 DBMS 产品所支持的数据模型相符合的逻辑结构。

过程：

- 将概念结构转换为现有 DBMS 支持的关系、网状或层次模型中的某一种数据模型；
- 从功能和性能要求上对转换的模型进行评价，看它是否满足用户要求；
- 对数据模型进行优化

3.1 E-R 图向关系模型的转换

原则：

- (1) 一个实体型转换为一个关系模型，实体的属性就是关系的属性，实体的键就是关系的键；
- (2) 一个联系转换为一个关系模式，与该联系相连的每个实体型的键以及联系的属性都转换为关系的属性。这个关系的键分为以下三种不同的情况：

- 若联系为 1:1，则相连的每个实体型的键均是该关系模式的侯选键。
- 若联系为 1:n，则联系对应的关系模式的键取 n 端实体型的键。
- 若联系为 m:n，则联系对应的联模式的键为参加联系的诸实体型的键的组合。

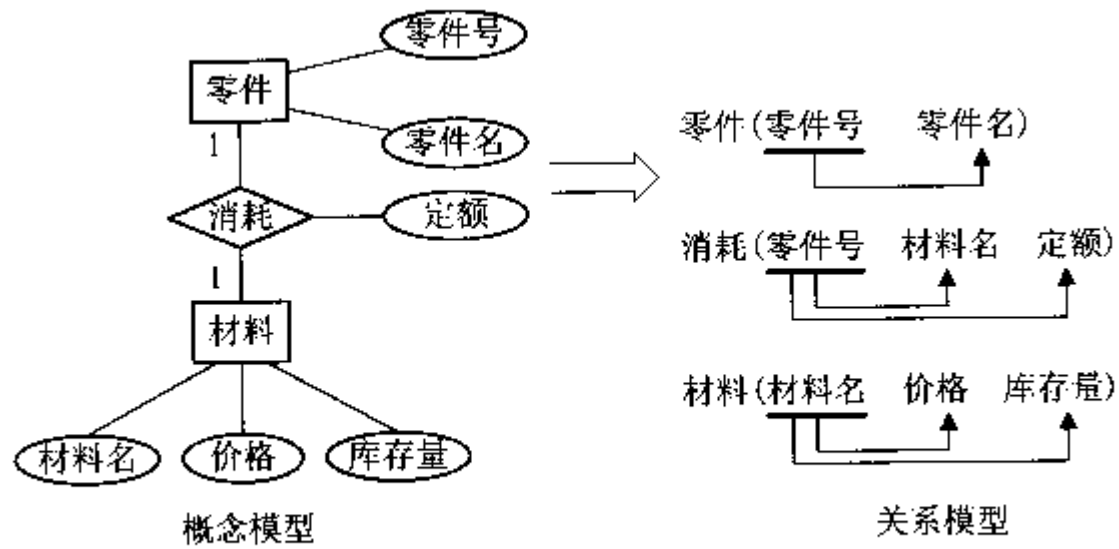


图 18.5 零件和材料的概念模型到关系模型的转换

图 联系为 1:1 时

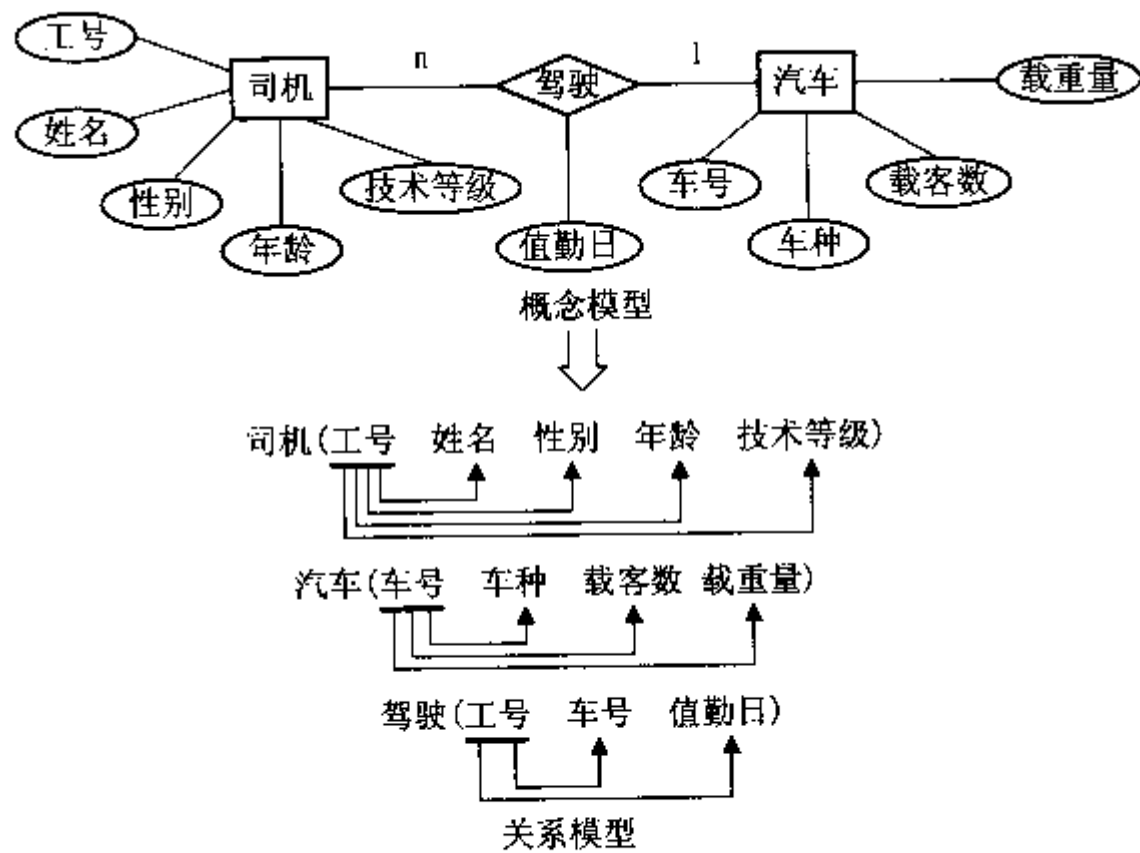


图 18.6 司机和汽车之间的概念模型到关系模型的转换

图 联系为 1:n 时

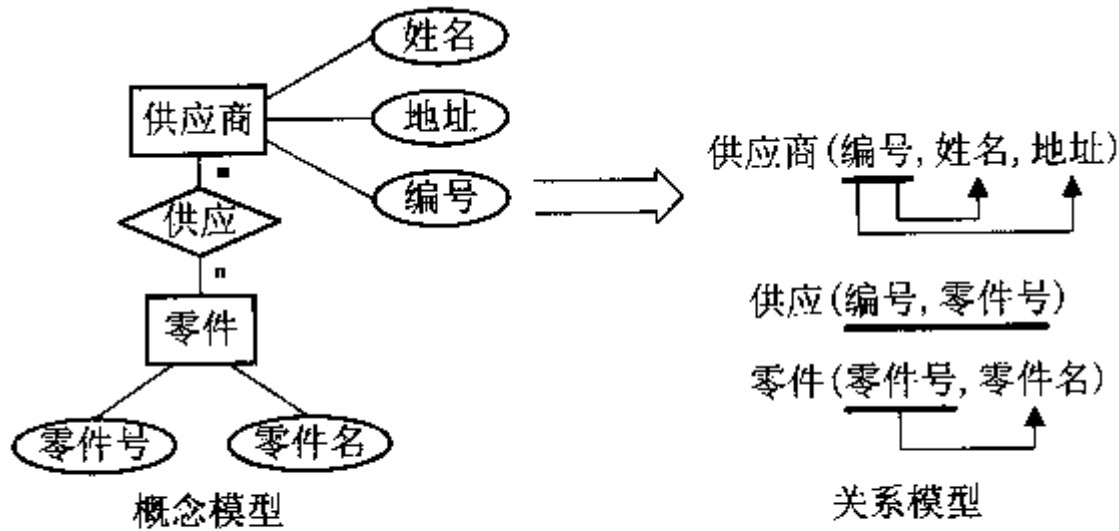


图 18.7 供应商和零件之间的概念模型到关系模型的转换

图 联系为 m:n 时

(3) 一些特殊联系的处理，可分为以下几情况：

- 当一个实体的存在是依赖于另一个实体的存在时，两个实体之间的联系便代表了两个实体间的一种所有关系。
- 当联系定义在同一个同型实体上时，联系转化为一个关系模式，与该联系相连的实体型的键以及联系的属性转换为关系模式的属性。

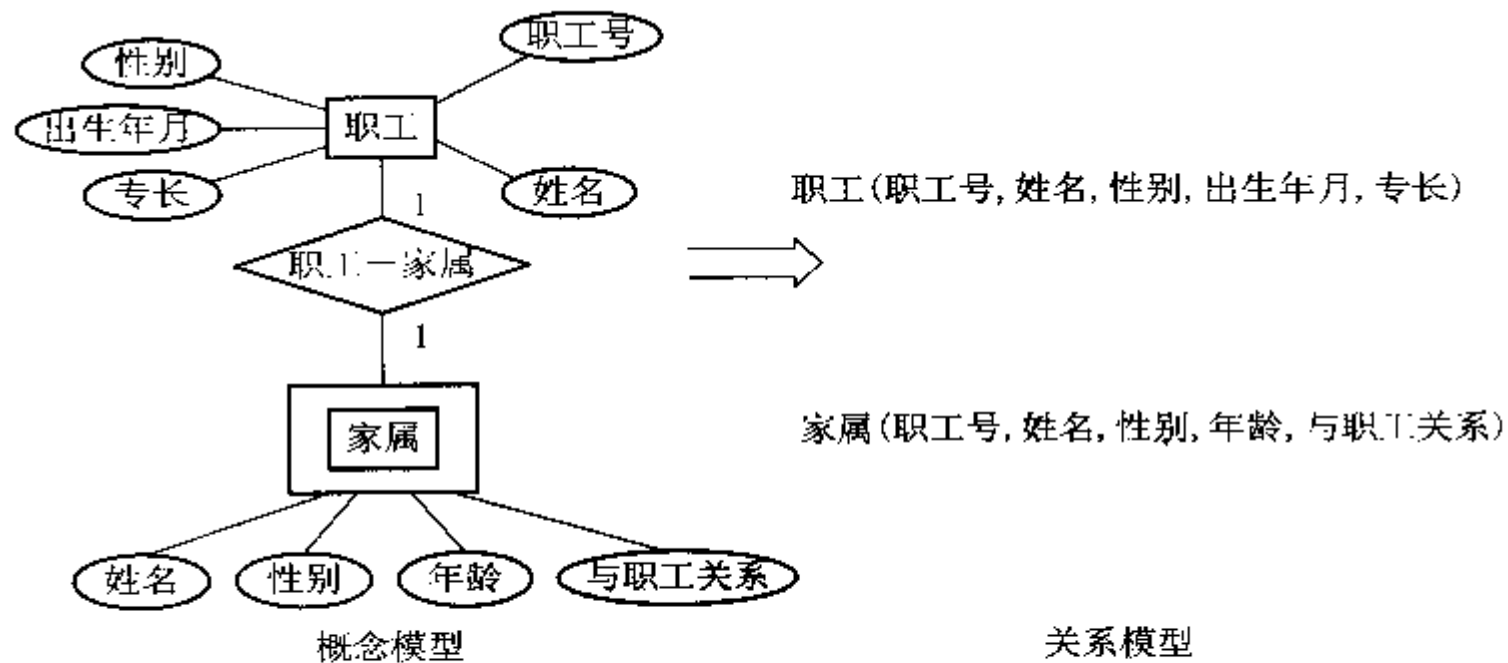


图 18.8 职工与家属之间的概念模型到关系模型的转换

图 实体存在依赖关系

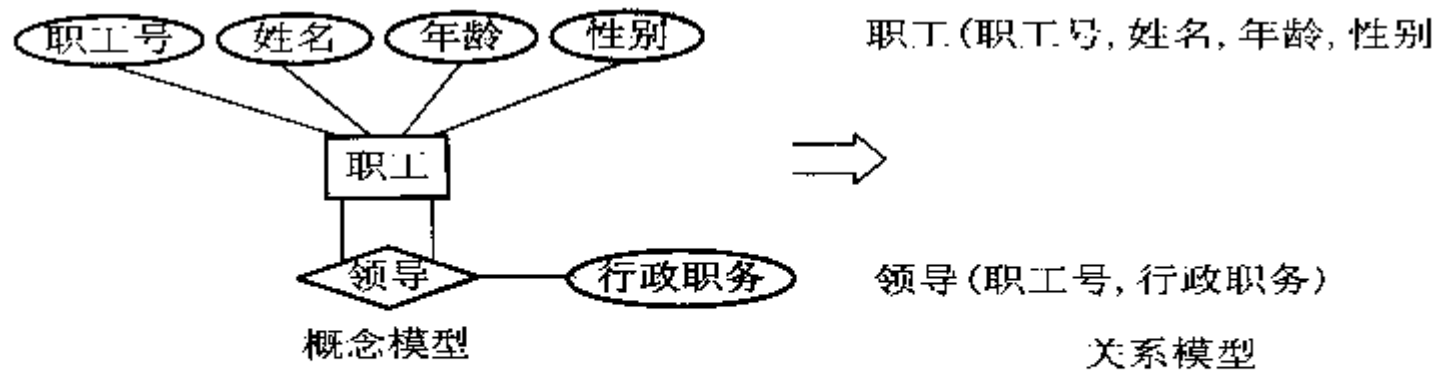


图 18.9 领导和职工构成的概念模型到关系模型的转换

图 联系定义在同一同型实体上

3.2 模型评价：功能评价，性能评价

3.3 模型优化：减少连接运算，垂直和水平分割，使用快照，减少数据占用空间

4 物理结构设计

5 数据库的建立和测试

6 数据库的运行和维护

7 数据库保护