

CHAMPION

考 研 辅 导 丛 书
KAO YAN FU DAO CONG SHU



考 研 热 门 专 业

读 书 笔 记 及 内 部 资 料 汇 编

KAO YAN RE MEN ZHUAN YE
DU SHU BI JI JI NEI BU ZI LIAO
HUI BIAN

全国知名重点大学联合编写组 编

计 算 机 分 册

學苑出版社

考研热门专业 读书笔记及内部资料汇编

计算机分册

· 全国知名重点大学联合编写组 编

本册主编：魏龙华 柴雪晴

本册编者：柴雪晴 魏龙华 邹宝威

胡景松 李慧明 钟志诚

学苑出版社

图书在版编目(CIP)数据

考研热门专业读书笔记及内部资料汇编(全7册)/全国知名重点大学联合编写组编.
-北京:学苑出版社,2002.4
ISBN 7-5077-1239-7

I. 考… II. 全… III. 研究生-入学考试-学习参考资料 IV. G642.47

中国版本图书馆 CIP 数据核字(2002)第 071169 号

**考研热门专业读书笔记及内部资料汇编
(计算机分册)**

学苑出版社出版发行

北京市万寿路西街 11 号 100036

北京诚顺达印刷有限公司印刷

850×1168 16 开本 47.5 印张 1494 千字

2002 年 4 月北京第 1 版 2002 年 4 月北京第 1 次印刷

印数:3000 册 定价:82.00 元

编写说明

目前, 考研已经成为竞争社会的一道风景而备受人们关注。面对报考人数的逐年激增, 以及由此引发的“考研难”的问题, 我们不禁思考: 如何尽我们之所能, 为广大考生做点实事, 这就是我们精心策划此套图书的深层动因。

据统计, 近几年来考研报名人数比较集中的专业有法学、计算机、金融学、西医、企业管理学、会计学、控制理论与控制工程、工商管理、通信工程、法律硕士、工商管理硕士等。每到考试前夕, 大多数考生都陷入市场上公共课目的辅导用书“汗牛充栋”和专业辅导材料无处可寻的尴尬境地。为此, 本书编写组在长期调研的基础上, 策划推出《考研热门专业读书笔记及内部资料汇编》丛书。首批推出法学、计算机、金融学、西医、企业管理、会计学、控制理论与控制工程专业七个分册, 该丛书规模庞大, 针对性强, 信息含量高, 具有明确的指导意义和参考价值, 堪称专业课复习中不可多得的工具与助手。

本书的编写工作, 数十位专家参与其中, 历经多次修改, 精益求精, 正是他们艰辛的劳动使本书高质量地呈现在读者面前。具体说来, 本书具备以下几大特点:

一、权威性强 本丛书在编写之前, 对考生进行了大量真实的调查, 综合调查结果并结合北大、清华、人大、武大、浙大等著名高等院校的专家的意见, 初步确定了结构框架, 经过多次科学、严格的论证, 组织高水平的作者队伍组建编写组, 书稿不断完善之后, 又由相关专业权威学者多次审稿, 从而保证了书稿的权威性, 并突出了其学术价值。对于书中所涉及的考试信息, 如名师介绍、历年试题等, 我们也从多个角度进行了核实, 力争准确无误。

二、实用性强 本套丛书并没有像其他丛书一样以统一的结构和风格问世, 而是依照不同专业的特点, 在结构设计上突出专业特色, 风格各异, 但从总体上看, 又有一个共同的特点, 即实用性强。这主要表现在四个方面: 一是“点”与“线”相结合, 在理清专业课知识线索的基础上, 重点讲解常考的知识点; 二是例题与原理相结合, 复习专业课既要掌握理论, 又要学会运用原理解决问题, 本丛书尽可能用例题解释原理, 就是基于这点考虑; 三是试题与笔记相结合, 大量历年考试真题保证了考生有针对性地进行复习, 专业课程读书笔记的系统整理, 节省了考生宝贵的时间和精力; 四是内部信息与复习资料相结合, 不但帮助考生复习, 而且为考生报名考试提供了技巧性的指导。

三、资料全面新颖 正如前面所述, 本丛书不仅包括考生普遍急需的各科复习笔记, 而且还辅之以到目前为止我们所能搜集到最新最全的相关高校考研真题、内部信息和报考指导。所谓“复习笔记”, 并不仅仅局限于对教材线索的梳理, 而是做到了“源于教材, 高于教材”, 在知识线索中穿插了众多的重点、难点问题解析。对文科相关专业的前沿学术热点进行了专题总结式的梳理。本套丛书主要是为 2003 年参加考研的同学准备的, 在突出重点的同时, 我们注重将最新的信息融入其中, 特别是在考试真题的选编上, 本书编写人员克服了一系列困难, 将包括 2002 年在内的近几年考研真题提供给考生, 这一开创性的工作保证了本书的时效性。

当然, 由于条件所限, 本丛书中的信息不可能涵盖全国所有高校, 而只能选取其中的典型, 有所侧重, 在“特殊性”中找“共性”。因此, 考生在使用本书的过程中, 宜结合自己所报考院校的教学特点和出题思路, 有的放矢地开展复习工作。希望读者朋友们将您对本书的意见反馈给我们, 以便我们在以后的修订中加以完善, 服务更多的同学。

最后, 预祝考生朋友们马到成功!

编者

2002 年 4 月

目 录

第一篇 数据库系统	(1)	第五章 函数	(162)
第一章 数据库系统的引论	(2)	附录Ⅱ 考研真题重点难点精解精析	(164)
第二章 层次与网状数据库	(9)	第三部分 代数系统	(168)
第三章 关系数据库	(14)	第六章 代数系统	(168)
第四章 SQL 语言	(21)	第七章 格和布尔代数	(177)
第五章 关系数据理论	(28)	附录Ⅲ 考研真题重点难点精解精析	(181)
第六章 数据库设计	(35)	第四部分 图 论	(186)
第七章 关系系统及查询优化	(45)	第八章 图论	(186)
第八章 数据库保护	(47)	第九章 树及其应用	(195)
全真试题集	(51)	附录Ⅳ 考研真题重点难点精解精析	(199)
第二篇 计算机网络	(60)	全真试题集	(208)
第一章 计算机网络概述	(60)	第四篇 程序设计 (C 语言)	(228)
第二章 数据通信技术基础和物理层技术与标准	(65)	第一章 C 语言概论	(228)
第三章 数据链路层技术与标准	(73)	第二章 数据类型、运算、表达式	(230)
第四章 局域网的数据链路层	(78)	第三章 数组、结构体、联合体	(232)
第五章 网络层技术和标准	(84)	第四章 函数与变量	(234)
第六章 网络互连	(88)	第五章 指针	(236)
第七章 传送层	(95)	第六章 预处理文件	(237)
全真试题集	(99)	第七章 程序控制与算法设计	(238)
第三篇 离散数学	(109)	全真试题集	(244)
第一部分 数理逻辑	(110)	第五篇 数据结构	(253)
第一章 命题逻辑	(110)	第一章 数据结构基本概念	(253)
第二章 谓词逻辑	(122)	第二章 数组	(257)
附录Ⅰ 考研真题重点难点精解精析	(132)	第三章 链接表	(266)
第二部分 集合论	(139)	第四章 栈与队列	(272)
第三章 集合及其运算	(139)		
第四章 关系及其运算	(150)		

第五章 递归与广义表	(278)	全真试题集	(427)
第六章 树	(285)	第七篇 操作系统	(433)
第七章 搜索	(294)	第一章 操作系统概述	(434)
第八章 图	(301)	第二章 进程管理	(441)
第九章 排序	(305)	第三章 调度与死锁	(462)
第十章 散列	(313)	第四章 作业管理和用户接口	(474)
全真试题集	(317)	第五章 存储管理	(483)
第六篇 编译原理	(335)	第六章 设备管理	(506)
第一章 高级语言与编译程序概述	(336)	第七章 文件管理	(517)
第二章 文法和语言	(340)	第八章 UNIX 操作系统	(530)
第三章 词法分析	(347)	全真试题集	(568)
第四章 语法分析	(361)	第八篇 组成原理	(576)
第五章 语法分析程序的自动构造	(375)	第一章 概论	(577)
第六章 语法制导翻译和中间代码生成	(389)	第二章 计算机中数据信息的表示	(582)
第七章 符号表与错误处理	(396)	第三章 指令格式	(594)
第八章 程序运行时存储空间组织	(404)	第四章 运算方法与运算器	(619)
第九章 代码优化	(418)	第五章 主存储器与存储体系	(641)
第十章 代码生成	(425)	第六章 CPU 组织	(668)
		第七章 输入输出系统	(688)
		全真试题集	(715)
		附录 走进名校	(722)

第一篇 数据库系统

本篇主要就北京大学, 浙江大学, 同济大学, 华中理工大学 (现改名为华中科技大学) 等几所重点大学研究生招生的数据库试题进行分析, 归纳出重点、难点, 并进行阐述, 而且给出以前的考题进行分析, 使各位同学能了解考试重点, 节省复习时间。

本篇中的每一章的第一部分主要涉及一些基本概念或基础知识; 第二部分是对一些重点难点的剖析和对往年相关试题的解答; 第三部分是练习题; 第四部分则是第三部分练习题的参考答案。同时在本篇的最后部分附有上述几所学校往年的考题。

本篇的知识结构如下:

一、数据库系统

1. 数据管理发展的三个阶段
 - 人工管理阶段
 - 文件系统阶段
 - 数据库系统阶段
2. 数据模型的概念、要素、类型

二、层次数据库与网状数据库

1. 数据操纵语言
 - DBTG 的数据操纵语言 DML
 - 程序运行环境
 - DML 语句
 - IMS 的数据操纵语言 DL/1
2. DBTG 的模式、子模式及各自的基本结构
 - DBTG 的模式及其基本结构
 - 子模式 DDL
3. IMS 的数据库组织和存取方法
 - 实现层次模型的存储结构的两种基本方法
 - IMS 的数据库存取方法

三、关系数据库

1. 关系代数
 - 关系运算符
 - 传统的集合运算
 - 专门的关系运算
2. 关系演算
 - 元组关系运算
 - ALPHA 语言
 - 域关系演算

四、SQL 语言

- SQL 查询语言
 - 简单的查询
 - 表的连接查询
 - 嵌套查询

五、关系数据理论

1. 范式之间的关系
 - 函数依赖
 - 码
 - 规范化理论的提出
2. 模式的分解
 - 闭包及其运算
 - 最小函数依赖集及其算法
 - 模式的分解
 - 无损连接性
 - 函数依赖性
 - 分解算法

六、数据库设计

1. 数据库设计的特点和设计步骤
2. 需求分析
3. 概念结构设计
 - E-R 图方法
 - E-R 模型的建立
 - 局部 E-R 图
 - 总 E-R 图
4. 逻辑结构设计 E-R 图向数据模型的转化
5. 数据库物理设计
6. 数据库的实施和维护

七、关系系统及查询优化

1. 关系系统的定义及其分类
2. 全关系系统的十二条基本准则
3. 代数优化及代数表达式的变换

八、数据库保护

1. 数据库的安全性
2. 数据库的完整性
3. 数据库的并发控制
 - 并发控制的重要性
 - 基本的封锁类型
 - 活锁、死锁及其解决办法
 - 可串行性

4. 恢复 { 恢复的重要性
故障的种类
恢复的基本原理和方法

第一章 数据库系统的引论

本章要求了解数据管理的发展及各阶段特点, 涉及的一些相关的名词和相关概念; 了解数据模型的概念, 要素及其分类; 理解三级模式结构。

一、需要掌握的一些基本概念

数据、信息、数据库、数据库系统、数据库管理系统、数据模型、数据模式、文件系统。

1. 信息是以数据的形式表示的, 即数据是信息的载体; 但不是所有的数据都能表示信息, 信息是人们消化理解了的数据。另一方面, 信息是抽象的, 不随数据设备所决定的数据形式而改变; 而数据的表示方式却具有可选择性。

2. 数据指的是用符号记录下来的可以区别的信息, 这里的“符号”不仅仅指数字、字母、文字和其他特殊符号, 而且还包括图形、图像、声音等多媒体数据。

3. 数据库 DB: DB 是存储在磁盘等外部直接存取设备上的数据集合, 是按一定的数据结构组织起来的, 是一个结构化的数据集合, DB 主要是通过综合各个用户的文件, 除去不必要的冗余, 使之相互联系所形成的数据结构。

4. 数据库系统: 是指具有管理和控制数据库功能的计算机系统。

5. 数据库管理系统 DBMS: DBMS 是数据库系统的关键组成部分。任何数据操作, 包括数据库定义、数据查询、数据维护、数据库运行控制等都是在 DBMS 管理下进行的。DBMS 是用户与数据库的接口, 应用程序只有通过 DBMS 才能和数据库打交道。

6. 文件是操作系统管理数据的基本单位, 文件可以命名, 通过文件名以记录为单位存取数据, 不必关心数据的存储位置。

7. 文件系统: 用操作系统中的存取方法管理数据的软件。

8. 数据模型: 是用来描述数据的一组概念和定义。

9. 数据模式: 是对某一类数据的结构、联系和约束的描述。也称为数据库的模式。

10. 数据的物理独立性: 指当数据的物理结构或存储结构改变时, 数据的逻辑结构可以不变, 从而应用程序也不必改变。

11. 数据的逻辑独立性: 指当总体逻辑结构改变时, 通过对映像的改变而保持局部逻辑结构不变, 程序员是根据局部逻辑结构编写应用程序的, 从而应用程序也不必改变。

二、重点、难点及例题解析

(一) 数据管理发展的三个阶段

根据计算软件和硬件的发展, 数据管理技术的发展大体上分为三个阶段: 人工管理阶段、文件系统阶段和数据

库阶段。

1. 人工管理阶段 (20 世纪 50 年代中期以前)

本阶段的主要特点是:

数据不保存; 程序与数据合在一起, 因而数据没有独立性, 要修改数据必须修改程序; 程序员必须自己编程实现数据的存储结构、存取方法和输入输出, 这样就迫使程序员直接与物理设备打交道, 加大了程序设计难度, 编程效率低; 数据面向应用, 这意味着即使多个不同程序用到相同数据, 也得各自定义, 数据不仅高度冗余, 而且不能共享。

2. 文件系统阶段 (20 世纪 50 年代后期至 60 年代中后期)

本阶段的主要特点是:

程序和数据有了一定的独立性, 程序和数据分开存储; 数据文件可以长期保存在外存储器上并可以多次存取; 数据的存取以记录为基本单位, 并出现了多种文件组织, 如顺序文件、索引文件和随机文件等; 数据冗余度大; 缺乏数据独立性; 数据不能集中管理。

3. 数据库系统阶段 (20 世纪 60 年代后期开始)

本阶段的主要特点是:

(1) 数据结构化 数据库是存储在磁盘等外部直接存取设备中的数据集合, 是按一定的数据结构组织起来的。

(2) 数据共享 数据库中的数据是考虑所有用户的数据需求、面向整个系统组织的, 不同用户所使用的数据可以重叠, 同一部分数据也可多用户共享。

(3) 减少了数据冗余 在数据库方式下, 用户所使用的是数据库管理系统从数据库中映射出来的逻辑文件, 它取自于数据库中的某个子集, 并非独立存在, 从而减少了数据冗余。

(4) 有较高的数据独立性 数据独立有两个含义, 即物理数据独立性和逻辑数据独立性。所谓物理数据独立性是指数据库物理结构发生改变时, 不会影响到逻辑结构, 而用户使用的是逻辑数据, 所以不必改动程序; 所谓逻辑数据独立性是指数据库全局逻辑发生改变时, 用户也不需改动程序, 就像数据库并没发生变化一样。数据独立的好处就是数据存储方式的改变不会影响到应用程序。

(5) 提供了用户接口 在数据库系统中, 数据库管理系统作为用户与数据库的接口, 提供了数据库定义、数据库运行、数据库维护和数据安全性等控制功能, 此外还支持某种程序设计语言, 并设有专门的数据操作语言, 为用户编程提供了方便。

【例 1】填空题 (2000 年华中理工大学)

计算机管理技术的三个发展阶段是人工管理阶段、文件系统阶段和数据库系统阶段。

(二) 一些相似概念的区别和联系

1. 数据和信息

区别:数据是用符号记录下来可以区别的信息,这里的符号不仅仅指数字、字母、文件和其他特殊符号,还包括图形、图像等多媒体数据。信息是现实世界各种事物的存在特征、运动形态以及不同事物间的相互联系等诸要素在人脑中的抽象反映,是一个抽象的概念。信息源于客观事物,而后通过人们加工处理再控制客观事物,从而达到认识世界、改造世界的目的。

联系:数据是信息的载体,而信息是数据的内涵。同一信息可以有不同的数据表示形式,而同一数据也可能有不同的解释。

二者的关系可以表示为:信息 = 数据 + 数据处理

2. 数据库系统和文件系统

文件系统与数据库系统之间的主要区别:

文件系统中的文件之间不存在联系,因而从总体上看数据是没有结构的,而数据库中的文件是相互联系着的,并在总体上遵从一定的结构形式。数据库正是通过文件之间的联系反映现实世界事物间的自然联系。

联系:数据库系统从文件系统发展而来;它们都是数据组织的管理技术;它们都由数据管理软件管理数据;程序与数据之间用存取方法进行转换。

3. 数据模型和数据模式

模型和模式都是描述数据的逻辑结构,属于‘型’的范畴,但它们之间有区别。数据模型(Model)是描述数据的手段或方法,而数据模式(Schema)是用给定的数据模型对具体的数据进行描述。模型和模式的差别就如同程序设计语言和用这种语言编写的程序的差别一样。

4. 数据库系统和数据库管理系统

数据库系统:是指具有管理和控制计算机功能的计算机系统。

数据库管理系统:是一个通用的软件系统,由一组计算机程序组成。

它能够对数据库进行有效的管理。

数据库管理系统是数据库的一个组成部分。

5. 数据的逻辑独立性和数据的物理独立性

数据独立有两个方面的含义,即物理数据独立性和逻辑数据独立性。所谓物理数据独立性是指数据库物理结构(包括数据的组织和存储、存取方法、外部存储设备等)发生改变时,不会影响到逻辑结构,而用户使用的是逻辑数据,所以不必改动程序。

6. 主语言和数据操作语言(DML)

数据操纵语言 DML 实现对数据库的操作,包括检索、插入、删除和修改四种基本操作命令。

主语言用于操纵内存中的数据,如对取入内存的数据库数据进行各种运算和处理等。

用户使用主语言和嵌入主语言的 DML 语句编写应用程序使用数据库。

7. 模式数据描述语言(schema DDL)和数据库模式模式数据描述语言(Schema DDL)描述数据库的整体

数据结构和完整性约束条件等。

数据库模式指数据库管理的数据实体的类型、特征、实体间的联系的表示。

数据库模式分为三个级别的模式:物理模式、逻辑模式和视图级模式。

模式 DDL 用来定义数据库模式中的逻辑模式。

【例 2】名词解释(1999 年中国人民大学)

数据库(5分):是按一定的数据结构组织起来的,是一个结构化的数据集合。

【例 3】单项选择题(1999 年浙江大学)

建立在计算机外部设备上的结构化、有联系的数据集合是(A)

- A. 数据库(DB) B. 数据库管理系统
- C. 数据结构 D. 数据文件

【例 4】(1999 年浙江大学)计算机用于数据管理的软件系统是(B)

- A. 数据库系统 B. 数据库管理系统(DBMS)
- C. 应用程序 D. 计算机语言

【例 5】(1992 年浙江大学)数据库系统和数据库管理系统是不同的概念,分别说明之,两者有什么联系?

参考答案:数据库系统是指具有管理和控制数据库功能的计算机系统。

数据库管理系统是指数据库系统中专门用于数据管理的软件。

数据库管理系统是指数据库系统的关键组成部分。

【例 6】(1996 年浙江大学)

数据库系统中,什么叫数据与程序的独立性,什么叫数据与程序的逻辑独立性?

数据与程序的独立性指数据存储方式的改变,不会引起应用程序的改变。

所谓逻辑数据独立性指数据库全局逻辑发生改变时,用户也不需改动程序,就像数据库并没有发生变化一样。

(三) 数据模型的概念及它的三个要求和三种数据模型

根据模型应用目的的不同,可将模型分为两个层次即概念模型和数据模型。

数据从现实世界到计算机数据库里的具体表示要经历三个领域,即现实世界、信息世界、计算机世界。

(1) 现实世界

现实世界是指客观存在的世界中的事实及其联系。在目前数据库方法中把客观事实抽象成信息世界的实体,然后再将实体描述成计算机世界的记录。

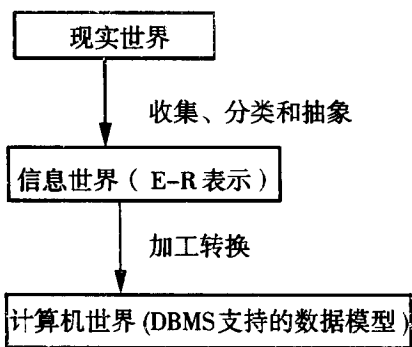
(2) 信息世界

信息世界(或概念世界)是现实世界在人们头脑中的反映,是对客观事物及其联系的一种抽象描述,一般采用实体联系方法(E-R方法)表示。

(3) 计算机世界

计算机世界可称作数据世界,是在信息世界基础上的进一步抽象,使用的方法为数据模型的方法。

为了把现实世界中的具体事物抽象和组织为 DBMS 支持的数据模型, 人们常常首先将现实世界抽象为信息世界, 然后将信息世界转换为机器世界, 也就是说, 首先把现实世界中的客观对象抽象



为某一种信息结构, 这种信息结构并不依赖于具体的计算机系统, 不是某一个 DBMS 支持的数据模型, 而是概念级的模型, 然后再把概念模型转换为计算机上某一 DBMS 支持的数据模型 (如图)

1. 数据模型的三个要素和 E-R 图的要点

(1) 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。

· 数据结构

数据结构是所研究的对象类型 (Object type) 的集合。这些对象是数据库的组成成分。一般可分为两类: 一类是与数据类型、内容、性质有关的对象; 一类是与数据之间联系有关的对象。

· 数据操作

数据操作是指对数据库中各种对象 (型) 的实例 (值) 允许执行的操作的集合, 包括操作及有关的操作规则。

· 数据的约束条件

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。

【例 7】(2000 年华中理工大学) 数据模型用来描述数据的一组概念和定义, 主要的三种数据模型是层次模型、网状模型、关系模型。

(2) 实体—联系 (E-R) 图是数据库设计阶段中建立数据模型的基础。提出 E-R 模型主要是作为从现实世界到其他数据模型之间转换的中间模型, 这种模型不考虑数据的具体物理存储以及存取效率这些问题, 它允许以比较有效、自然的方式描述现实世界的数据库。

涉及的主要概念有:

- 实体: 客观存在并且可相互区分的事务。
- 属性: 实体所具有的某一特性。
- 码: 惟一标识实体的属性集。
- 域: 某个 (些) 属性的取值范围。

· 实体型: 用实体各及其属性各集合来刻画同类, 实体称为实体型。

为了建立现实世界的完整模型, 常常需要对联系分类, 对于二元联系, 根据一个实体集合的实体可以和多少个另一类实体集合的实体相联系分为如下几种:

- 一对一联系 (1: 1)

每个实体集合的一个实体最多只能与另一实体集合中的一个实体有联系。比如我们确定了部门实体集合与经理实体集合之间存在一对一联系, 意味着一个部门只能由一个经理管理, 一个经理仅管理一个部门。在任何时刻现实世界应遵从这个规则, 否则此模式不是对这个现实世界的正确描述。

· 一对多联系 (1: n)

实体集合 A 中的一个实体与实体集合 B 中的多个实体有联系, 而实体集合 B 中的一个实体最多只能与实体集合 A 中的一个实体有联系, 比如部门实体集合与职员实体集合之间存在一对多联系, 即一个部门包括若干职员, 而一个职员只能属于一个部门。

· 多对多联系 (m: n)

实体集合 A 中的一个实体可与任意多个实体集合 B 中的实体相联系, 实体集合 B 中的一个实体也可与任意多个实体集合 A 中的实体相联系, 比如学生实体集合与课程实体集合之间存在多对多联系, 即一个学生可以选修多门课程, 而一门课程有多个学生选修。

作实体—联系图的具体步骤如下:

①确定所有实体集合, 用矩形方框表示实体集合, 方框内标明实体集合名称。

②选择实体集合应包含的属性, 用椭圆框表示属性, 通过无向边连接到实体集合。只有一个属性的实体集合可用属性代替, 附加到它参加的联系上。

③确定实体集合之间的联系, 用菱形框表示联系, 框内标明联系名称, 通过无向边或有向边连接到参加联系的每个实体集合。联系是有序的, 可对边编号 (一般仅对同一实体集合的联系用编号指明顺序)。

④确定实体集合的关键字, 用下划线在属性上表明关键字的属性集合。

⑤确定联系的类型, 在用无向边连接联系到实体集合时, 在边上注明是 1 或 n (多) 来指明联系的类型。在用有向边连接联系到实体集合时, 让边的箭头指向 1 的实体集合一方, 多对多因为都是多方, 故无箭头。

概念模型的表示方法最常用的是实体—联系方法, 这里仅仅介绍了 E-R 图的要点。至于如何认识分析现实世界, 从中抽取实体, 实体间的联系及建立概念模型的方法将在第六节介绍。

【例 8】数据库设计中, 在概念设计阶段可用 E-R 方法, 其设计出的图称为: (D) (1999 年浙江大学)

- A. 实物示意图
- B. 实用概念图
- C. 实体表示图
- D. 实体联系图

【例 9】实体间的联系分别是 1: 1 (或一对一联系), 1: n (或一对多联系), m: n (或多对多联系)。

【例 10】实体和属性都有“型”和“值”之分。“型”是概念的内涵, 而“值”是概念的实体。“型”表达的是个体的共性, 而“值”是个体的具体内容。

【例 11】E-R 方法, 又称为实体联系表示方法。

【例 12】E-R 模型的主要构件包括: 实体、属性、实体之间的联系 (2000 年华中理工大学)

·三种主要的数据模型及它们间的关系

数据库有不同的类型。它是根据数据模型划分的。在计算机系统中,任何一个数据库管理系统都是根据不同的数据模型有针对性地设计出来的。目前比较成熟的数据模型有:层次模型、网状模型和关系模型三种。

三种数据模型之间的根本区别在于数据之间联系的表现方式不同(即记录型之间的联系方式不同),三者分别用“树形结构”“图结构”“二维表”来表示数据之间的联系。

①层次模型

层次模型是用树结构表示记录类型及其联系的数据模型。在理解层次模型时要注意以下一些问题。

【注意】

层次模型有且仅有一个结点无父结点(该结点称为根),其他结点有且仅有一个父结点。

层次模型就像一棵倒立的树,树的结点是记录型。上一层记录型和下一层记录型之间的联系是1:n的。

在层次式数据库中查找记录,必须指定存取路径。存取路径是指从根结点开始沿途所经过的路程。

层次模型中,各记录型之间的联系非常分明,这种模型最适合表示客观世界中有严格辈份关系的事物,其缺点是不能够直接支持m:n联系的事物。

②网状模型

网状模型结点间的联系不受层次限制,可以任意发生联系,所以它的结构是结点的连通图。在理解网状模型时要注意以下一些问题。

【注意】

网状模型有一个以上的结点无父结点;至少有一个结点有多于一个父结点。

网状模型能够反映各种复杂联系情况。在网状模型的具体实现上,对于记录间的m:n联系,是将其转换m个1:n联系来处理的。

③关系模型

关系的本质是一张二维表,而关系模型的本质则是用若干个二维表表示实体及其联系。关系模型与层次模型和网状模型的理论 and 风格截然不同,如果说层次和网状模型是用“图”来表示实体及其联系,那么关系模型则是用“表”来表示实体及其联系的。

a. 关系模型表示实体 关系框架中并无内容,但只需在框架中填上具体内容,就有了数据。表里有了数据就被称为关系,每一行就是一个记录,表示一个实体;每一列是记录中的一个数据项,表示实体的一个属性。于是,这样的一张有了数据的表就描述了一个实体。

b. 关系模型表示实体之间的联系 这个问题的理解是关系模型的关键。如果在两个不同实体之间确实存在着某种自然联系,那么一定可以在那些相对应的关系中找到一个属性,这个属性能够使两个不同实体之间发生联系。

【注意】关系模型的本质是关系一体化,就是不论“实体”还是“实体之间的联系”都统一用关系来表示。

④层次数据模型与网状数据模型比较

层次数据模型和网状数据模型均属于格式化数据模型,它们有许多共同的特点,比如都用存取路径来表示数据之

间的联系;用户对数据的存取都必须按照指定的存取路径导航;数据操作时必须清楚数据库的当前位置;对数据库的操作都是一次一个记录的存取方式;程序和数据尽管层次数据模型与网状数据模型有许多类似的地方,但它们之间也存在着区别。比如

·层次数据模型中从子女到双亲的联系是惟一的,而网状数据模型则可以是不惟一。

·网状数据模型中允许有复合链的存在,而层次数据模型中则不可以。

·网状数据模型中用双向链组织数据,而层次数据模型中用单向链组织数据。

⑤网状数据模型与关系数据模型比较

任何E-R模型可以转化成网状数据模型,也可转化成关系数据模型,可以证明,任何用关系模型定义的数据库模式可以转化成网状数据库模式,反之也成立。然而,这两种模型有明显的不同,区别主要体现在以下几个方面:

a. 网状数据模型有两种基本数据结构:记录型和系;关系模型只用一种基本数据结构:关系,关系模型描述数据和操作数据更简单。

b. 网状数据模型查询数据时必须在指明怎样从一个记录型找到另一个记录型的记录,必须指明通过哪一个库中定义了的系。查询数据时从一种记录找另一种记录的操作一般称为导航。网状模型要求在查询中显式地指明导航的路径,即通过的系等等,查询语言的过程性很强。而关系数据模型通过关系的连接从一种数据找到关联的另一种数据,连接运算时利用指定任何分量间值的比较关系运算灵活地建立两个关系的关联,不必要有实现定义的路径。所以,关系数据模型的查询不需显式地导航,查询语言有更高的说明性。

【例13】(2000年华中理工大学)网状数据库系统的一个存取命令一次存取一个记录。

【例14】(1999年浙江大学)在数据库的数据模型中有(A)

- A. 网状模型、层次模型、关系模型
- B. 网状模型、关系模型、链表模型
- C. 数值型、字符型、逻辑型
- D. 单用户型、多用户型、网络型

【例15】(1999年浙江大学)在关系数据库系统中,一个关系相当于(A)

- A. 一个数据文件
- B. 一条记录
- C. 一个关系数据库记录
- D. 一个关系代数运算

【例16】层次模型不能直接表示(C)

- A. 1:1关系
- B. 1:m关系
- C. m:n关系
- D. 1:1和1:m

【例17】网状型层次型和关系数据库划分原则是(D)

- A. 记录长度
- B. 文件的大小
- C. 联系的复杂程度
- D. 数据之间的联系

【例18】非关系模型中数据结构的基本单位是基本层次联系

【例19】试述三种基本数据模型的优缺点。

(1) 层次模型的优点是:

①结构清晰,表示各结点之间的联系简单;

②容易表示如“家族关系”等现实世界的层次结构的事物及其之间的联系。

缺点是:

①不能表示两个以上实体型之间的复杂联系和实体型之间的多对多联系;

②严格的层次顺序使数据插入和删除操作变得复杂,如父结点的删除导致子结点的删除。

(2) 网状模型的优点是:

①能够表示实体之间的多种复杂联系。

缺点是:

①网状模型比较复杂,需要程序员熟悉数据库的逻辑结构;

②在重新组织数据库时容易失去数据独立性。

(3) 关系模型的优点是:

①使用表的概念,简单直观;

②直接表示实体之间的多对多联系;

③具有更好的数据独立性;

④具有坚实的理论基础。

缺点是:

①关系模型的联结等操作开销较大,需要较高性能计算机的支持。

3. 数据库系统的三级模式结构

数据库的结构指数据库管理的数据实体的类型、特征、实体间的联系的表示,也称为数据库模式(schema)或数据库的内涵。

数据库中结构的定义可以在多个抽象级别进行,形成多个级别的数据库模式。

绝大多数数据库系统在总的体系结构上都具有以下三级模式的结构特征:

模式是数据在数据库系统的内部表示,即对数据的物理结构和存储方式的描述,是低级描述,一般由 DBMS 提供的工具或语言完成。用户可以在多大程度上直接定义内模式,在具体的 DBMS 中各不相同,当前的许多 DBMS 产品自身可以完成大部分内模式的定义工作,几乎不需用户介入。

内模式又称存储模式,具体描述了数据如何组织存储在存储介质上。内模式是系统程序员用一定的文件形式组织起来的一个个存储文件和联系手段;也是由他们编制存取程序,实现数据存取的故称内模式为系统程序员视图。

外模式(External Schema):用户模式,是逻辑模式的子集或变形,它是与某一应用有关的数据。

4. 三级组织结构的区别与联系

模式是内模式的逻辑表示;内模式是模式的物理实现;外模式则是模式的部分抽取。三个模式反映了对数据库的三种不同观点:模式表示概念级数据库,体现了对数据库的总体观;内模式表示了物理级数据库,体现了对数据库的存储观;外模式表示了用户级数据库,体现了对数据库的用户观。总体观和存储观只有一个,而用户观可能多个,有一个应用,就有一个用户观。

在三个模式间存在着两种映射:一是“外模式/模式”

间的映射,这种映射把用户数据库与概念数据库联系起来;另一种映射是“模式/内模式”间的映射,这种映射把概念数据库与物理数据库联系起来。

三级模式中,只有内模式才是真正存储数据的,而模式和外模式仅是一种逻辑表示数据的方法。

【例 20】模式是由子模式组合得到的,子模式是由模式的部分抽取得到的。

【例 21】(2000 年华中理工大学)物理数据独立性是指应用程序独立于 (B)

A. 内模式 B. 模式 C. 子模式

【例 22】(1993 年浙江大学)什么叫模式,子模式?试述模式在数据库中的重要作用。

模式是数据库中全体数据的逻辑结构和特性的描述,所有用户的公共数据视图。

子模式是个别用户看到的数据视图,是与某一应用有关的数的逻辑表示。

【例 23】数据库三级组织结构,也称 SPARC 分级结构,从内到外分三个层次描述,分别称为内模式、概念模式、外模式。

【例 24】内模式,又称存储模式,具体描述了数据是如何组织存储在存储介质上,故称内模式为系统程序员视图。

【例 25】数据库系统的三级结构中,总体观和存储观有一个,用户观有多个,一个应用对应一个用户观。

【例 26】数据库系统三级结构的关系,可以概括为 A

(A) 模式是内模式的逻辑表示

(B) 模式是内模式的物理实现

(C) 模式是外模式的部分抽取

(D) 外模式是内模式的物理实现

【例 27】内模式是系统程序员用一定的文件形式组织起来的一个个存储文件和联系手段。

在数据库系统领域,一般不必深入到二进位或字节的级别看待数据,而是从文件级(本书称为物理级)开始,因为数据库系统往往是建立在文件系统基础之上的。三级抽象模式在数据库系统中都存储于数据目录(系统数据字典)中,是数据目录的最基本的内容,DBMS 通过数据目录来管理和访问数据模式。

数据库按 SPARC 公布的标准分为三级结构:概念模式、外模式和内模式。

三、练习题

(一) 填空题

1. 经过处理和加工提炼而用于决策或其他应用活动的数据称为_____。

2. 数据本质上是对_____的一种符号化表示,即用一定的符号表示_____。数据是_____的载体,而_____是数据的内涵。

3. 数据库管理大体上可分为_____阶段_____阶段和_____阶段。

4. 数据库管理系统的主要功能有_____、_____

_____、数据库的运行管理和数据库等 4 个方面。

5. 数据库语言包括_____和_____两大部分,前者负责描述和定义数据库的各种特性,后者用于说明对数据进行的各种操作。

6. 数据库的模式有_____和_____,前者直接与操作系统或硬件联系,后者是数据的完整表示。

7. 数据库系统与文件系统的本质区别在于_____。

8. 数据库是一个_____化的数据集合,主要是通过综合各个用户的文件,除去不必要的冗余,使之相互联系所形成的_____。

9. 在数据库系统阶段,数据库技术具有以下五个特点_____,_____,_____有较高的数据独立性和提供了方便的用户接口。

10. DBMS 有两种类型的程序设计语言,如“FORTRAN、C”等属于_____型,FOXPRO 则属于_____型。

11. 数据库三个模式之间存在两种映射,一是_____间的映射,另一映射是_____间的映射。

12 概念模式,简称_____,是对数据库的_____描述,并不涉及_____故称为 DBA 视图。

13. DBMS 是指_____,它是位于_____和_____之间的一层管理软件。

14. 非关系模式中数据结构的基本单位是_____。

15. _____结构是对数据系统的静态特性的描述,_____是对数据库系统的动态特性的描述。

16. DBA 是指_____,其职责是维护和管理_____,使之始终处于最佳状态。

17. 现实世界的事物反映到人的头脑中经过思维加工成数据,这一过程要经过三个领域,依次是_____,_____,_____。

18. 按照数据结构的类型来命名,数据模型分为_____,_____和_____。

19. 数据模型是由_____,_____,_____三部分组成的。

20. 数据独立性是指_____和_____是相互独立的。

21. 数据冗余可能导致的问题有_____和_____。

22. 从外部视图到子模式的数据结构的转换是由_____实现的;模式与子模式之间的映像是由_____实现的;存储模式与数据物理组织之间的映像是由_____实现。

(二) 选择题 (每小题仅有一个正确答案)

1. 关于信息和数据,下面_____的论述是正确的

- A. 信息与数据,只有区别,没有联系
- B. 信息是数据的载体
- C. 同一信息用同一数据表示形式
- D. 数据处理本质上就是信息处理

2. 文件系统与数据库系统的最大区别是_____

- A. 数据共享
- B. 数据独立
- C. 数据结构化
- D. 数据冗余

3. 数据库的概念模型独立于_____

- A. 具体的机器和 DBMS
- B. E-R 图
- C. 信息世界
- D. 现实世界

4. 计算机系统中任何软件必须在_____的支持下才能

工作

- A. 应用程序
- B. 操作系统
- C. 数据库
- D. 硬件

5. 数据库系统的核心是_____

- A. 编译系统
- B. 数据库
- C. 操作系统
- D. 数据库管理系统

6. 数据库系统的最大特点是_____

- A. 数据的三级抽象和二级独立性
- B. 数据共享性
- C. 数据的结构化
- D. 数据独立性

7. 数据库管理系统的工作不包括_____

- A. 定义数据库
- B. 对已定义的数据库进行管理
- C. 为定义的数据库提供操作系统
- D. 数据通信

8. 下面关于数据库管理系统的论述,_____是正确的

- A. 数据库管理系统是用户与应用程序的接口
- B. 应用程序只有通过数据库管理系统才能和数据库联系
- C. 数据库管理系统用 DML 来定义三级模式
- D. 数据库管理系统用 DDL 来实现对数据库的各种操作

作

9. 数据库系统中的软件是指_____

- A. 数据库管理系统
- B. 应用程序
- C. 数据库
- D. 数据库管理员

10. 应用数据库的主要目的是为了_____

- A. 解决保密问题
- B. 解决数据完整性问题
- C. 共享数据问题
- D. 解决数据量大的问题

11. 数据库的特点之一是数据的共享,严格地讲,这里的数据共享是指_____

- A. 同一个应用中的多个程序共享一个数据集合
- B. 多个用户、同一种语言共享数据
- C. 多个用户共享一个数据文件
- D. 多种应用、多种语言、多个用户相互覆盖地使用数据集合

12. 在数据库中存储的是_____

- A. 数据
- B. 数据模型
- C. 数据以及数据之间的联系
- D. 信息

13. 数据库系统的核心是_____

- A. 数据库
- B. 数据库管理系统
- C. 数据模型
- D. 软件工具

14. 数据库管理系统是_____

- A. 操作系统的一部分
B. 在操作系统支持下的系统软件
C. 一种编译程序
D. 一种操作系统
15. 下述关于数据库系统的正确叙述是____
A. 数据库中只存在数据项之间的联系
B. 数据库的数据项之间和记录之间都存在联系
C. 数据库的数据项之间无联系, 记录之间存在联系
D. 数据库的数据项之间和记录之间都不存在联系
16. 以下关于数据模型的描述, ____不正确
A. 数据模型表示的是数据库本身
B. 数据模型表示的是数据库的框架
C. 数据模型是客观事物及其联系的描述
D. 数据模型能够以一定的结构形式表示出各种不同数据之间的联系
17. 实体是信息世界中的术语, 与之对应的数据库术语为____
A. 文件 B. 数据库
C. 字段 D. 记录
18. E-R 图所表示的实体及其联系, 实际上是____
A. 信息模型 B. 关系模型
C. 数据模型 D. 实体模型
19. 实体集中, 各实体值相应的属性有着____的域
A. 不同 B. 相同
C. 各自 D. 固定
20. 网状模型在具体实现上, 对于记录间的 $m:n$ 关系, 可以将其转换为____联系
A. $1:1$ B. $n:m$
C. $1:n$ D. 不能转换
21. 下列关于实体的说法, ____不正确
A. 实体是指现实世界中存在的一切事物
B. 实体靠联系来描述
C. 实体所具有的性质统称为属性
D. 实体和属性是信息世界表达概念的两个不同单位
22. DBTG 属于____
A. 层次模型 B. 数据模型
C. 关系模型 D. 网状模型
23. 数据库三级模式体系结构的划分, 有利于保持数据库的____
A. 数据独立性 B. 数据安全性
C. 结构规范化 D. 操作可行性
24. 三个模式反映了对数据库的三种不同观点, 以下说法中正确的是____
A. 内模式表示了概念级数据库, 体现了对数据库的总体观
B. 外模式表示了物理级数据库, 体现了对数据库的存储观
C. 外模式表示了用户级数据库, 体现了对数据库的用户观
D. 外模式表示了用户级数据库, 体现了对数据库的存储观
25. 数据库管理系统能实现对数据中数据的查询、插入、修改和删除等操作, 这种功能称为____
A. 数据定义功能 B. 数据管理功能
C. 数据操纵功能 D. 数据控制功能
26. 下面命令中____不是 DML 的基本操作
A. 排序 B. 插入
C. 修改 D. 检索
27. 通过指针链接表示和实现体之间联系的模型是____
A. 关系模型 B. 层次模型
C. 网状模型 D. 层次和网状模型
28. 有关“采用映射技术的好处”, ____是不对的
A. 保证了数据的独立性
B. 保证了数据共享
C. 方便了用户使用数据库
D. 保证了数据库的开放性
29. 数据库系统提供了多种语言, 下面____不是其中的一种
A. DDL B. SCL (自主型程序设计语)
C. DML D. HL (宿主型程序设计语言)
30. 在数据库中, 产生数据不一致的根本原因是____
A. 数据存储量太大
B. 没有严格保护数据
C. 未对数据进行完整性控制
D. 数据冗余
31. 为使程序员编程时即可使用数据库语言又可使用常规的程序设计语言, 数据库系统把数据库语言嵌入到____中
A. 编译程序 B. 操作系统
C. 中间语言 D. 宿主语言
32. 数据的管理方法主要有____。
A. 批处理和文件系统
B. 文件系统和分布式系统
C. 分布式系统和批处理
D. 数据库系统和文件系统
33. 在数据库的三级模式结构中, 描述数据库中全体数据的全局逻辑结构和特征的是()
A. 外模式 B. 内模式
C. 存储模式 D. 模式

参考答案

- (一) 1. 信息 2. 信息 信息 信息 信息 3. 人工管理 文件系统 数据库 4. 数据定义功能, 数据操纵功能 5. 数据描述语言、数据操纵语言 6. 存储模式、概念模式 7. 数据库系统实现了整体数据的结构化 8. 结构、数据结构 9. 数据结构化 数据共享 减少了数据冗余 10. 宿主 自主 11. 外模式/模式 模式/内模式 12. 模式、整体逻辑、物理存储 13. 数据库管理系统 用户 操作系统 14. 基本层次联系 15. 数据 操作 16. 数据库管理员 数据库 17. 现实世界 信息世界 计算

机世界 18. 层次模型 网状模型 关系模型 19. 数据结构 数据操作 完整性约束 20. 用户的应用程序 存储在外存上的数据库中的数据 21. 浪费存储空间及修改麻烦 潜在的数据不一致性 22. 应用程序 DBMS 操作系统的存取方法

(二) 选择题 1.D 2.C 3.A 4.B 5.D 6.C
7.C 8.B 9.A 11.D 12.C 13.B 14.B 15.B 16.A
17.D 18.A 19.B 21.B 22.D 23.A 24.A 25.C
26.A 27.D 28.A 29.C 30.D 31.D 32.D 33.D

第二章 层次与网状数据库

本章要求了解层次与网状数据的基本概念, 并能加以区分。重点掌握型与值的概念间的区别联系; 层次数据库的重点是掌握它的物理数据库和逻辑数据库的区别以及四种存储结构; 网状层次模型中, 它的数据操纵语言中 DML 语句中的 FIND 语句的使用, 它的模式、子模式及基本结构是重点。

一、基本概念

1. 数据项: 是指 DBTG 模式中最小的数据单位。
2. 记录: 是数据项的有序集合。
3. 系: 表示记录间一对多的联系。
4. 系型: 是惟一的首记录型和若干个相关联的属记录型的集合, 表示首记录和属记录型之间一对多的联系。
5. 系值: 是系型的一个实例。一个系型有若干系值, 每一系值由一个首记录值带上系型中诸属记录型的相关联的各个记录值组成。在一个系值中, 首记录的值总是一个属记录的值可以是 K 个 ($K \geq 0$)。因此在一个系型中系值的数目就是首记录值的数目。
6. 奇异系: 是以系统本身作首记录的系。
7. 码: 为区别整个数据库中的记录, 而给记录指定的一个惟一标识符。
8. 字段: IMS 中数据不可分割的最小单位。
9. 片段: 若干字段组成片段。
10. 物理数据库记录型: IMS 中片段型的层次序列结构, 记为 PDBR 型。
11. 一个数据库记录: 一个根片段的值及其所有后代片段值构成比 PDBR 型的一个值。
12. 一个物理数据库 (PDB): 指一个 PDBR 型的全部值的有序集。
13. 一个数据库: 是若干个 PDB 的集合。
14. 层次序列: 根片段经过各级后代片段到达某一片所经过的路径称为此片段的层次路径。
15. 层次序列码: 一个片段值的层次路径上各片段的序号与相应的各片段值的排序码连接而成的字符串。
16. PDB 的层次序列: 某个 PDB 所有片段值的层次序列码按升序排列的结果。
17. LDB (逻辑数据库) 是用逻辑关系联合一个或多个物理数据库中的部分片段型所构成的。
18. PSB (程序说明块): 用户应用程序所使用的数据的逻辑结构描述。
19. 运行单位: 是 DBTG 中一个程序连同它的数据在处理机中动态的执行过程。

20. 用户工作区: 是 DBMS 按照此运行单位所启用的子模式建立的, 为子模式中说明的每一个记录分配相应的存储空间。

21. 记录的当前值: 运行单位能识别的每一种记录型最后被存取的记录值。

22. 系的当前值: 运行单位能识别的每一种系型中最后被存取的记录值。

23. 域的当前值: 运行单位能识别的每个域的最后被存取的记录值。

24. 运行单位的当前值: 运行单位能识别的所有记录型中最后被存取的记录值。

二、重点、难点分析及例题解析

(一) 重点概念的理解

1. 型和值的概念

(1) 实体有型和值之分, 记录有型和值之分, 系同样也有型和值之分

系代码	名称	电话
-----	----	----

为一个记录型, 而

800	计算机	45101
-----	-----	-------

为一个记录值。

在系的组织中应注意:

一个记录型可以作为几个系的属记录,

一个记录型可以作为一个系的首记录, 也可以同时作为另一个系的属记录。

一个记录型可以是几个系的首记录。

两个记录型之间允许有多个系存在, 即多重关系。

一个属记录的值不能属于同一个系的不同系值, 比如一个学生不可能同时就读多个专业。由于系是表示记录型之间一对多的联系。因此:

①一个首记录值属于 (决定) 这个系的一系值。

②一个属记录值最多属于这个系的一个系值。

在 DBTG 中, 一对一、一对多的联系可以直接表示; 对于实体型之间多对多的联系就必须进行转换, 引进一个联结记录, 把多对多的联系转换成两个 (或多个) 一对多的联系。

(2) 物理数据库记录型和一个数据库记录

物理数据库记录型是 IMS 中片段型的层次序列结构。

一个物理数据库记录是一个根片段的值及其所有后代片段值。

IMS 限制一个 PDBR 型的层数不超过 15 层, 片段型的

数目不超过 255 个。一个片段型可以有任意个片段值，但规定任何一个子女片段值是不能在没有双亲片段值的条件下存在。

2. 物理数据库 (PDB) 和逻辑数据库 (LDB)

一个物理数据库 PDB 就是一个 PDBR 型的全部值的有序集。

在 IMS 中的物理数据库中的“物理”二字含义和通常的概念不同。通常，物理和逻辑相对应，例如，数据库的物理结构常常指存储结构而与逻辑结构相对应。而这里的“物理”不是指存储结果，仍具有通常的逻辑的概念。并带有“实际存在”的含义。

在现实世界中，许多数据模型是网状结构的，IMS 为了解决应用中网状结构的要求，同时消去数据冗余，引入了逻辑数据库的概念。

逻辑数据库 LDB 是用逻辑关系联合一个或多个物理数据库的部分片段型构成的一个新的层次结构。

二者既有联系，又有区别：

(1) 一个逻辑数据库是用户看到的数据视图，用户可以像对物理数据库那样对逻辑数据库进行处理。所以从用户角度看，逻辑数据库和物理数据库处于同等地位，都是数据模型的一部分。从概念上讲它属于模式，而不属于子模式。LDB 也是用 PBD 来描述的，称为逻辑 PBD。

装入逻辑数据库的过程由装入作为此逻辑数据库基础的物理数据库和建立所需的逻辑指引元两个步骤组成。

(2) 但是逻辑数据库和物理数据库毕竟是不同的。逻辑 PBD 中仅定义数据结构而无物理存取方法的说明。IMS 并不按照该 PBD 所述的结构再去重复存储数据。LDB 中的各个片段实际存在于原来的各个 PDB 中。也就是说，它不是凭自己资格而存在的。它是用一个或几个已存在的物理数据库来定义的。这样既给用户提供了另一种数据视图，又避免了数据冗余。

使用 LDB 的好处有：

- 提供了减少冗余度的一种手段。
- 可以处理多对多的实体联系。
- 将一个复杂的网状结构表示成树结构。
- 使用户可以不同的方式看待数据。

【例 1】 (1998 年北京大学) DBTG 系统中的系值：DBTG 系统中的每一个系值由一个首记录值带上系型中诸属记录型的相关联的各个记录值组成。在一个系值中，首记录的值总是一个。

【例 2】 (2000 年华中理工大学) 某一时刻，DBTG 中一个记录型的当前值同其所在种类的当前值 (B)

- A. 不可能相同 B. 可能相同 C. 必定相同

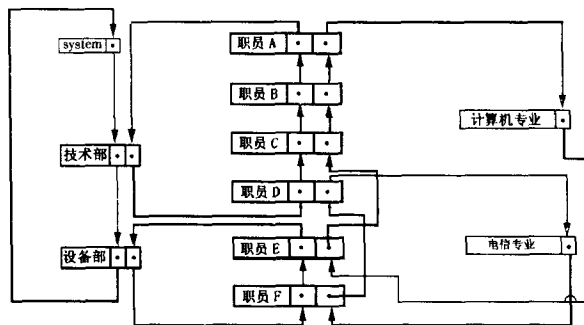
【例 3】 (1999 年同济大学) 什么是物理数据库 (PDB)，什么是逻辑数据库 (LDB)？IMS 引入逻辑数据库的目的何在？

参考答案：物理数据库是指一个 PDBR 型的全部值的有序集。逻辑数据库是用逻辑关系联合一个或多个物理数据库中的片段所构成的一个新的层次结构。

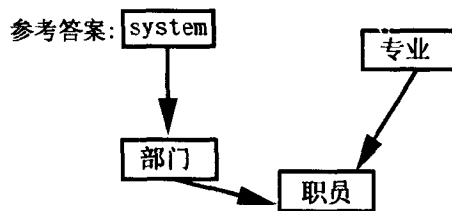
因为现实中的许多网状结构在只有 PDB 时存在大量的数据冗余、多对多的实体联系、结构复杂及用户看待数据

的方式单一等缺点。为解决这些问题引入了 LDB 的概念。

【例 4】 (2000 年华中理工) 设以网状数据模型建立的数据库中有如图所示的一批互相链接的数据记录。



a. 试给出其 DBTG 方式的数据模型结构 (5 分)



b. 指出其中有几个记录类型、几个系类型、几个记录值、几个系值 (5 分)

参考答案：有 4 个记录型：system、部门、专业和职员。其中 system 有 1 个记录值。部门有 2 个记录值。专业有 2 个记录值。职员有 6 个记录值。

有两个系型：system → 部门 → 职员 和 专业 → 职员。其第一个系型有 1 个系值，第二个系型有两个系值。

(二) DBTG 的数据操纵语言 DML

DBTG 的数据操纵语言 DML 是宿主型的，即面向某种高级语言。用户要使用主语言和嵌入主语言的 DML 语句编写应用程序使用数据库。DML 语句完成对数据库数据的检索、插入、修改和删除。主语言用于操纵内存中的数据。如对取入内存数据库数据进行各种运算和处理等。

1. 程序运行环境

DBTG 中的运行单位与操作系统的进程的概念类似。一个运行单位有程序区、变量区、UWA (用户工作区) 和系统通信区等组成。系统通信区主要包括当前状态指示表和运行状态反馈信息。

用户工作区的概念和一些当前值的概念参看前面的部分。

当前状态指示表是由 DBMS 管理和修改的，它是由 DBMS 为每一个运行单位保留的位置标识，以免每次查找都从头开始，从而提高系统的效率。在开始时，它的所有值全为空值。

数据库状态反馈信息是执行 DML 语句成功、失败或遇到例外情况的反馈信息。

对于一个运行单位，它有且仅有一个 UWA，一张当前状态指示表，一组运行状态反馈信息。

2. DML 语句

在 DML 语句中, FIND 语句是一个非常重要的语句, 也是常考的语句。它的操作对象是一个记录, 功能是查找一个记录并使它成为运行单位的当前值。FIND 语句有六种格式, 其中以前三种格式最为重要。

(1) FIND ANY 记录句

或者

FIND ANY 记录句 USING 数据项名 IN 记录名

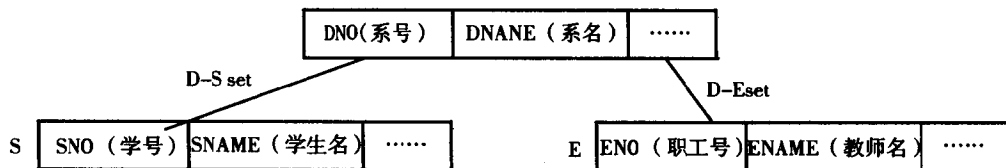
根据记录的码值或某个数据项值查询记录, 码和数据项应先赋值。

(2) FIND OWNER WITH 系名

找某个系的当前的首记录

(3) FIND $\left\{ \begin{array}{l} \text{FIRST} \\ \text{LAST} \\ \text{NEXT} \\ \text{PRIOR} \\ \text{整数} \\ \text{标识符} \end{array} \right\}$ 记录名 WITHIN 系名

在一个系值中按照指明的顺序查找属记录, 其中 ||



对如下三个运行单位分别写出各运行单位的记录的当前值系 (set) 的当前值和运行单位的当前值。(8分)

① MOVE 'S4' TO SNO IN S

FIND ANY S

② FIND OWNER WITHIN D-S (结果为 D02, 计算机系)

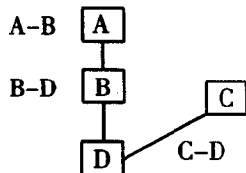
③ FIND FIRST WITHIN D-E (结果为 E110, 王刚)

参考答案: 执行上述三个语句后, 当前状态指示表的变化如下表所示

FIND 语句	D 记录	S 记录	E 记录	D-S 系	D-E 系	运行单位
1	NULL	S4	NULL	S4	NULL	S4
2	D02	S4	NULL	D02	NULL	D02
3	D02	S4	E110	D02	E110	E110

【例 6】(1992 年浙江大学)

在 DBTG 中, 有如下的数据模型和对应的数据库



号表示必须选其中一个。

若选 FIRST, 则找出系值中 (按系序的) 第一个属记录。

若选 LAST, 则找出系值中 (按系序的) 最后一个属记录。

若选 NEXT, 则找出系的当前值记录 (按系序的) 的后一个属记录。若系的当前记录已是系值中最后一个属记录, 则出现例外状态。

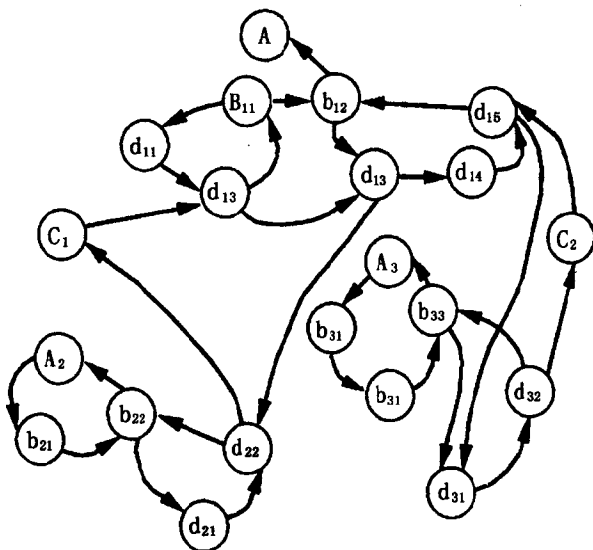
若选 PRIOR, 则找出系的当前值记录 (按系序的) 的前一个属记录。若系的当前记录是系值中第一个记录, 则为例外状态。

若用标识符, 则此标识符是程序中定义了的整型变量, 取其值按前面规定找属记录。

若给出某个整数, 如 6 (或 -6), 则是找出按 NEXT (或 PRIOR) 方向从系值首 (末) 端开始的第 n 个记录。

【例 5】(1991 年浙江大学)

用 DBTG 类 DBMS 建立了如下数据模型的数据库, 其中 DNO, SNO, ENO 分别为各个记录型的关键字。



对如下的 DML 语句, 查询的结果各是什么?

① FIND "A"

② FIND NEXT B WITHIN A-B SET

③ FIND NEXT B WITHIN A-B SET

④ FIND LAST B WITHIN B-D SET

⑤ FIND OWNER C WITHIN C-D SET

参考答案: ① A ② b₁₁ ③ b₁₂ ④ d₁₅ ⑤ C₂

(4) FIND DUPLICATE WITHIN 系名 USING 数据项