

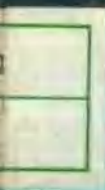
计算机软件工程丛书

系统分析与设计

王勇领

25-3

清华大学出版社



计算机工程丛书

系统分析与设计

王勇领

清华大学出版社

内 容 简 介

本书是原《计算机数据处理系统分析与设计》一书的修订更新版，也是计算机软件专业课程“系统分析与设计”的教材。作者近年在帮助不少企业、事业单位开发计算机信息系统中积累了丰富经验的基础上，对原著进行了大量的修改和补充，写成了这本更加实用更加先进的专著。本书专门论述开发一个计算机应用系统，诸如管理信息系统、决策支持系统、办公自动化系统、计算机通讯网络系统等需要的方法和工具。书中介绍了开发计算机应用系统的整个过程，着重介绍结构化系统分析、结构化系统设计，结构化程序设计的基本思想、原则和方法，以及如何应用这些方法去建立一个计算机应用系统，并且书写出一份规范化的软件工程文档。其目的是使所建立的系统不但能满足用户的要求，而且具有较高的运行效率、较强的可靠性和较高的可修改性。

本书内容丰富、新颖，深入浅出，简明易懂，不仅有理论，还有开发计算机应用系统的实例和软件文档的式样。作者还把其中一些有关管理信息系统的较为通用的源程序复制在一张软盘上，全书各章均有习题。它既可作为高等院校计算机软件专业的教材，也可供从事计算机应用的科研和工程技术人员参考。

(京)新登字 158 号

系统分析与设计

王勇领

责任编辑 贾仲良

☆

清华大学出版社出版

北京 清华园

清华大学印刷厂印刷

新华书店总店科技发行所发行

☆

开本：787×1092 1/16 印张：15.5 字数：368 千字

1991 年 10 月第 1 版 1991 年 10 月第 1 次印刷

印数：0001—8000

ISBN 7-302-00886-8/TP·321

定价：7.50 元

前 言

1986年3月,清华大学出版社出版了笔者写的一本名为《计算机数据处理系统分析与设计》的书。四年多来,有不少高等院校用它作为计算机专业和非计算机专业本科生、研究生的教材,此外,一些由中央部委或其它单位举办的系统分析师学习班也采用作教材。各校的课程名称也不尽相同,多数称为“系统分析与设计”,也有称为“信息系统分析”或“管理信息系统分析”等等。笔者认为“系统分析与设计”比较合适。当初把那本书定名为《计算机数据处理系统分析与设计》的主要原因是恐读者不清楚是什么系统的分析与设计,现在看来这种担心是多余的。一方面,“系统分析”、“系统设计”早已经成为计算机领域中的专业术语;另一方面按国际标准组织ISO的定义,“数据处理”就是普遍的计算机应用。鉴于上述情况,将《计算机数据处理系统分析与设计》的修订本改名为《系统分析与设计》。

建立计算机应用系统,无论是管理信息系统、决策支持系统、办公自动化系统或通讯网络系统,所采取的步骤基本上是相同的,只是方法可以不同,除了本书所介绍的结构化系统分析与设计技术(SADT)之外,还有软件需求工程方法学(Software Requirements Engineering Methodology),用户软件工程(User Software Engineering)等等。经常有人问哪种方法最好?笔者认为各有千秋,还是不急于定论为好。至于还有一些人认为结构化方法不容易开发大的计算机信息系统,这种看法恐怕过于片面。英国国家信息系统就是用此方法开发的,国内几个大型企业系统也是如此。只要系统开发人员真正掌握某一种方法论的基本原理和技巧,加上自己的创造能力,总是能够成功地建立起用户所需要的计算机应用系统。

自《计算机数据处理系统分析与设计》一书出版四年多来,许多读者和听过笔者讲课的学员反映,希望增加关于程序设计方法学和软件文档方面的内容,以及一个或几个较为具体的实例,笔者也有同感。从1986年那本书出版至今,经过进一步实践,笔者悟出了不少道理,有了一些新的想法。因此,这次修订,除对第四章“数据字典”进行了较大的修改外,还对系统设计部分作了大幅度的修改,这有助于读者更易掌握系统设计技术。另外,还增加了程序设计方法和软件开发文档两个部分,以使读者对计算机应用系统的开发方法有一个全面的了解。

笔者在整理开发过的几个信息系统的资料时,发现都是资料太多,尤其是图、表、数据字典、源程序等等,真可谓浩瀚如海。重新把它们整理后写进此书颇费功夫,难以下笔。因此,只好将其中一个信息系统加以简化,给读者提供一个示范性的例子,也许不能满足读者要求,笔者在此深表歉意。作为弥补,笔者将把其中一些有关管理信息系统的较为通用的源程序复制在软盘上,供读者在开发管理信息系统时参考。这些源程序是夏天同志用汉字FoxBASE编写的,并附有每个数据库文件的数据结构。对

该软盘感兴趣的读者可直接写信与清华大学出版社软件部联系，地址如下：邮政编码 100084，北京，清华大学出版社软件部。

本书作为计算机专业高年级教材，其先修课是操作系统和数据库原理，主讲课为 72 学时，可在一个学期讲完。若条件许可，最好给学生一个比较简单的系统作为上机实践的例子，上机的机时越充足，教学效果就越好。

作者 王勇领

1990 年 7 月 7 日

于北京计算机学院

目 录

第一部分 信息系统.....	1
第一章 计算机与信息处理.....	1
1.1 引言	1
1.2 信息系统	2
第二部分 结构化系统分析.....	18
第二章 结构化系统分析的基本概念.....	18
2.1 系统分析员在系统开发过程中面临的问题	18
2.2 结构化系统分析	22
第三章 数据流程图.....	30
3.1 数据流程图的基本符号	30
3.2 自顶向下地扩展	34
第四章 数据字典.....	42
4.1 数据字典的建立	42
4.2 数据字典的使用	55
第五章 数据存储结构规范化.....	60
5.1 问题的提出	60
5.2 规范化形式	62
5.3 第三范式设计举例	74
第六章 数据存取要求的分析.....	80
6.1 数据存取要求的特性	80
6.2 数据存取要求的基本类型	81
6.3 立即存取分析	86
第七章 处理逻辑的分析与表达.....	94
7.1 表达一个处理逻辑时存在的问题	94
7.2 结构式语言	96
7.3 判断树	101
7.4 判断表	103
7.5 三种表达工具的比较	108
7.6 结构化系统分析资料	109
第三部分 结构化系统设计.....	111
第八章 结构化系统设计的基本概念.....	111
8.1 什么是结构化系统设计	111
8.2 结构化系统设计的目标	113
第九章 结构化系统设计的基本方法.....	118

9.1	结构图	118
9.2	模块的设计原则	122
9.3	设计策略	128
9.4	一个设计例子	130
第四部分	结构化程序设计	135
第十章	结构化程序设计的基本概念	135
10.1	什么是结构化程序设计	135
10.2	程序的基本结构	137
第十一章	结构化程序设计技术	151
11.1	自顶向下的模块化设计	151
11.2	逐步求精	153
第五部分	软件文档与案例	167
第十二章	软件文档	167
12.1	软件文档和软件工具概述	167
12.2	可行性研究报告	168
12.3	系统开发计划	160
12.4	系统分析说明书	160
12.5	系统设计说明书	170
12.6	程序设计说明书	171
12.7	系统测试说明书	171
12.8	系统使用和维护说明书	172
12.9	系统开发总结报告	173
12.10	软件文档格式(图、表、一览表)	174
第十三章	CE 工程公司管理信息系统	183
13.1	可行性研究	183
13.2	系统开发计划	185
13.3	系统分析	186
13.4	系统设计	198
13.5	程序设计	208
第十四章	企业管理信息系统分析	212
14.1	计算机企业管理系统的基本结构	212
14.2	市场经营管理信息系统	213
14.3	生产管理信息系统	221
14.4	财务管理信息系统	232
14.5	人事管理信息系统	239

第一部分 信息系统

第一章 计算机与信息处理

1.1 引言

一般来说, 计算机应用系统有如下几大类: 信息处理、辅助设计与辅助制造、仿真、专家系统与人工智能等。其中信息处理的领域最为广泛, 所有的管理工作, 办公自动化和事务工作都属于信息处理。计算机信息处理系统的主要任务是对大量的数据进行处理, 为用户及时提供各种有用的信息。

在计算机术语中, 我们经常使用“数据”这个词, 实际上它具有很深的内涵, 按照国际标准化组织 (International Standard Organization, 以下简称 ISO) 的定义, 数据是对事实、概念或指令的一种特殊表达形式, 它可以用人工或自动化装置进行通信、翻译或处理。按照 ISO 的定义, 数据不仅是指数字, 还可以是指文字、图形符号及图形, 甚至是声音。现代的计算机可以接收几乎所有种类的数据。在我们把数据输入到计算机中去以后, 我们的兴趣并不仅仅是把这些数据原封不动地再取出来, 而是想要计算机对这些数据进行处理, 然后为我们提供有用的、新的信息。按 ISO 的定义, “信息”是对人有用的、能够影响人们行为的数据。例如, 飞机航班时刻表被保存在计算机里, 它是原始数据, 如果某一旅客想知道明天飞往加拿大的 CA991 航班是否还有两个空座位 (如果有的话, 就订两张票)。那么, 他想得到的就是信息。

信息是通过对数据进行处理而产生的。按 ISO 的定义, “数据处理”是指对数据进行一系列加工、储存、合并、分类、计算等操作的过程。我们用计算机对数据进行处理, 这种应用系统就是计算机信息系统。

在信息处理工作中, “系统”, 按 ISO 的定义, 是指由人、机器和各种方法构成的、用于完成一组特定功能的集合体。系统具有递归的含义, 一个系统可以从属于一个比它更大的系统。一个计算机信息系统, 包括了人和计算机系统。

信息系统从外界接收数据, 经过处理以后向人提供信息。因此, 一个信息系统的核心是数据, 数据就是资源。而有效地把数据组织到计算机中, 为用户及时地提供有用的信息, 则是计算机信息系统开发人员的一项最重要、最困难的工作。

一个计算机信息系统, 不仅要求有很高的运行效率和可靠性, 而且还要求有很强的可修改性, 即很强的可维护性。一个信息系统的可维护性越高, 它的生命周期就越长, 因为它能适应环境的变化, 易于修改。相对来说, 提高信息系统的运行效率和可靠性比较容易, 而提高它的可维护性却比较困难。因为前者可通过计算机硬件和系统软件来实

现（当然也与应用软件的开发技术有关），而后者主要和应用软件的开发技术有密切关系。本书将着重讨论开发一个计算机信息系统所需要的各种技术、方法以及软件的规范化等问题，特别是系统分析阶段和系统设计阶段中所需要的既简单、而又行之有效的方法。利用这些方法会使我们开发出来的计算机信息系统不仅具有较高的运行效率和可靠性，而且具有较强的可维护性。

1.2 信息系统

一个计算机信息系统，从硬件来说至少包括一个计算机系统，也就是有一台计算机和相应的外部设备；从系统软件来说至少有操作系统、数据库管理系统，若干种语言编译程序；从应用软件来说至少具有对数据的管理功能，这包括执行输入、处理、输出等功能。

通常有三种形式的计算机信息系统。一是单机形式，只有一台计算机以及少量的外部设备，如一台打印机，这种情况以微型计算机居多；二是集中式，一台计算机以及若干台终端设备，这种情况以非微型计算机居多，往往是中、小型机，采用联机处理方式；三是分布式，若干台计算机形成通信网络，例如若干台微型计算机构成一个局部网，或者是大中型计算机之间构成大型的数据通信网。

1.2.1 数据通信系统

当前在工业发达国家中的国家范围内，成国际，或州际范围内的计算机通信网络系统、数据通信系统的基本功能是在适当的地点和时间提供必要的数据库。什么是适当的地点和时间？这个问题只能由用户来回答，因为只有用户才知道这个系统应当在什么时候产生什么样的数据，传递到什么地方以及应当交给谁。

一个区域性的数据通信系统一般来说由三种系统组成，即用户系统、通信系统、主计算机系统，如图 1.1 所示。

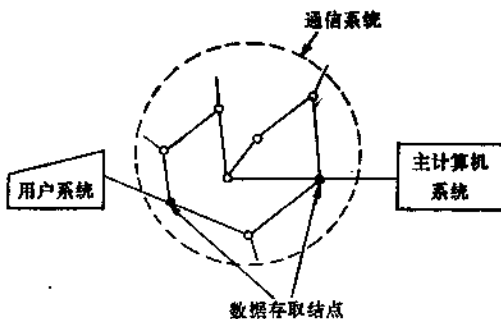


图 1.1 数据通信系统

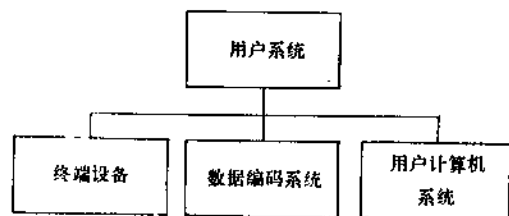


图 1.2 用户系统

一个复杂的数据通信系统，可以由若干个用户系统和若干个主计算机系统构成。

1.2.1.1 用户系统

用户系统可以由各种终端设备、数据编码系统和用户计算机系统组成。如图1.2所示。

终端设备是把数据输入到通信系统和计算机系统去的装置，它把输入的文字、

数字、图像甚至是声音转换成二进制代码的形式，然后传输给通信系统或计算机系统。终端设备还可以是数据的输出装置，把通信系统或计算机系统传递过来的二进制代码转换成文字、数字、图形等形式并打印或显示出来。

除了有基本的输入/输出功能以外，有些终端设备还有储存少量数据、编辑数据或执行简单计算的功能。具有数据存储功能的终端能够改进通信系统的交互性，数据可以用脱机的方式在终端上录入，暂时保存在某些磁性介质上。例如保存在软磁盘上，在适当的时间输入到通信系统中去。由通信系统传递过来的数据也可暂时保存起来，然后在适当的时间输出。具有编辑或计算功能的终端一般包括一个微型计算机，利用某些程序对数据进行编辑和计算。

用户系统的终端设备基本上有六种，即电传打字机，显示终端，专用终端（这是为用户设计的专门处理某些业务的终端），声音输入/输出系统，按钮式电话，绘图和图像传送系统。

用户系统的数据编码系统也称为数据录入系统，一般采用批式输入/输出的处理方式，其中包括三种子系统，键盘-存储系统，自动数据编码系统和远程作业录入系统。

键盘-存储系统的工作方式是数据录入员把数据用键盘终端保存在存储介质上，例如磁带、软磁盘上，然后再成批地发送给计算机处理。

自动数据编码系统，不需要靠人把数据通过键盘打入到存储介质上，而是用阅读器或其它专用数据采集系统来实现。阅读器通常是指磁墨水字符识别装置或光字符识别装置。

远程作业录入系统，一般由一台小型或微型计算机组成，其特点是有比较大的缓存容量。该系统把磁带或软磁盘的数据，利用数据通信系统传输给主计算机，也可以接收由主计算机发来的数据处理结果，通常是由程序员向主计算机发出作业请求，主机响应后即可发送数据。远程作业录入系统的优点是有较高的数据传输率。

用户计算机系统是指由用户自己管理和操作的计算机系统，一般来说，它由以下几个部分组成：

- (1) 一个或一个以上的终端设备；
- (2) 一台或一台以上的小型或微型计算机；
- (3) 一台或一台以上的大容量存储装置
- (4) 一台行式打印机；
- (5) 一个通讯接口装置；
- (6) 系统软件；
- (7) 若干个应用程序。

以上这些组成部分通过系统软件和数据总线联结成一个用户计算机系统，这个系统应当便于用户操作和使用。

1.2.1.2 通信系统

通信系统把数据从一个地方传输到另一个地方，它由传输电磁波的通道把若干个结点联结起来。例如可以借用电话线进行数据传输，在一条通信线的两端分别设置调制解调器。在发送数据的一端，调制器把计算机代码转换成音频电信号，在接收数据的一端，解调器把音频电信号再还原成计算机代码。数据的传输方式有三种，简单方式、半

双工方式和全双工方式。简单方式只能在一条通信线上按规定的方向进行数据传输，也就是一端只能做信号发送点，另一端只能做信号接收点。半双工方式允许数据在同一条通信线上进行双向传输，也就是同一个端点既是信号发送点，也是信号接收点，但在同一时间里，在这条通信线上只允许进行单方向的数据传输。全双工方式是在同一条通信线上允许在同一时间里进行双方向数据传输。由于通道、结点、调制解调器的质量不同，因此数据传输的速率也不同。低速一般在每秒钟传输 300 个二进制代码以下，中速在每秒钟传输 300~9600 个二进制代码之内，高速在每秒钟传输 9600 个二进制代码以上。

通信系统中的处理机实际上就是数字电子计算机，它由专门的程序执行通信控制和处理功能。这些计算机可以设置在通信系统的结点处或是设置在数据存取结点处，也就是通信系统与用户系统的接口处和通信系统与主计算机的接口处，如图 1.1 所示。作为通信系统与主计算机接口用的通信处理机称为“前端处理机”（Front-End Processor），用于管理前端处理机工作的软件包，称为“通信监控程序”（Communication Monitor Program）。

用户在使用通信系统时，必须充分考虑到“通信协议”（Protocol），它是专门管理通信设备运行的一组协议，用于管理通信系统与用户、与主计算机之间的接口。一般有四级通信协议。第一级是物理线路，第二级是通信链控制，第三级是路由控制，第四级是存取控制。物理线路通信协议用来确定和某一个通信网络相联结一定要建立的电气线路、功能装置和技术方法。通信链控制协议确定了在网络中数据由一个结点传输到另一个结点的过程。路由控制协议确定了在网络中数据从一个发送点到另一个终点所要经过的路由选择过程。存取控制协议确定了应用程序与用户系统联结的过程。

1.2.1.3 主计算机系统

主计算机系统包括硬件和软件。硬件一般是指数据存储装置和处理机，输入/输出数据总线等。系统软件一般是指操作系统、语言编译程序、库管理系统和各种实用程序。

决定数据存储装置的重要技术指标是：访问时间、传输率和存储密度。访问时间是指寻找和读出一个数据项所需要的时间。传输率是指在一秒钟之内能够传输给数据存储系统或从数据存储系统传输出来的数据量，按字符个数计算，也就是发生数据存取动作时，一秒钟能够传输的字符个数。存储密度是指在某一存储介质中，对给定的空间能够存储的字符个数。数据存储装置有直接存取的存储设备和脱机存储设备两大类。前者可以是高速缓存或磁盘存储器，后者可以是高速打印机或计算机输出缩微胶片。

处理机是执行程序的装置，程序是由若干条指令构成的。一个处理机有两个基本部分：一个是控制器，另一个是算术-逻辑运算器。近年来出现的“微处理机”（Micro-Processor）这个词具有两种含义，一种含义指这个处理机利用微程序来解释指令；另一种含义指它是由大规模集成线路构成的处理机。一般来说，两种含义都有。

计算机的中央处理器 CPU 的运行速度大大高于数据存储装置、网络通信系统、打印机等外部设备的速度。为了解决两者之间的矛盾，可以用一组公共线路联结许多台外部设备，如磁盘机、磁带机、前端处理机等，由同一个中央处理器来控制，这就是所

谓“输入/输出数据总线”的概念。它由许多个并行的通道组成，对应一个字节中的每一个二进制代码都有一个通道，这样可以并行传输，而不需要用串行的方式来传输一个字节，提高了传输速度。输入/输出数据总线中还有各种控制线路（专用处理器），作为中央处理器和外部设备之间的接口，接收和处理来自内存存储器传输过来的数据，或是检索外部存储设备中的数据，并把它发送给内存存储器。

操作系统是系统软件中的一个重要组成部分；它是一组程序，用于执行输入/输出控制、虚拟存储管理和多道程序运行管理等功能。

语言编译程序的功能是把程序员用某种高级语言编写的源程序翻译成计算机能够运行的目标程序，所有的编译程序都有诊断和错误检测功能。

库管理系统把各种计算机程序和子程序作为文件加以管理，它把每一个程序都当做文件来维护。

实用程序也称为服务程序。例如清除某一存错区域，装配某一程序，数据排序与合并，不同存储介质上文件的复制等。这些服务程序加强了计算机的处理功能，也方便了用户。

在主计算机系统中最重要的是数据存储装置和操作系统，它们决定了系统的运行效率。

下面介绍日本国内最大的一个数据通信系统——ZENGIN 系统，它是由日本NTT数据通信公司开发。该系统将日本国内所有的银行，包括外国在日本开设的银行联结起来，形成一个巨大的银行电子转帐系统。在 ZENGIN 系统的网络中有 163 家银行、5135 家与这些银行有关的财政金融机构、以及 42303 个办公室。此工程分三期实施，第一期工程从 1973 年 4 月开始，第二期工程从 1979 年 2 月开始，第三期工程从 1987 年 11 月开始，到 1990 年初已基本建成。客户可从任何一个银行或其分行将钱立即转汇到另一个银行或其分行。在这个系统中，数据的传输量和处理量是很大的。现金转移每月要处理 877 万笔，金额达 4752 亿日元；支票、帐单等事务每月要处理 4 亿笔，金额达 4 千万亿日元以上；信用卡事务每月要处理 4.5 亿笔，金额达 1 千万亿日元；此外还有其它种类的事务等，平均每天处理 164 万笔事务。

为了确保可靠，该系统采用双中心系统，在东京有一个中心，在大阪有一个中心，其示意图如图 1.3 所示。

在东京和大阪的每一个中心分别有两台 DIPS-11 MODEL 45E 的主计算机，其处理能力很大，每小时能处理 125 万笔事务。ZENGIN 终端有两种类型，分别是 RC 和 TCP-10。终端类型的选择取决于每一个银行或金融机构的业务量的大小。

通讯网线路的传输速度有四种：分别是每秒 9600，4800，2400 和 1200 比特。联结东京和大阪中心的是高速数字线路，速度是 384000 比特/秒。

ZENGIN 系统中有两种联结系统，即单独联结系统和公共联结系统。它们的作用是将银行或财政金融机构与东京和大阪的中心联结起来。对于每一个大银行或财政金融机构，采用单独联结系统，ZENGIN 终端设置在银行的计算机中心，直接与东京和大阪中心相联；对于较小的银行或财政金融机构，则采用公共联结系统，几家银行或财政金融机构形成一组，共享同一个公共联结系统。

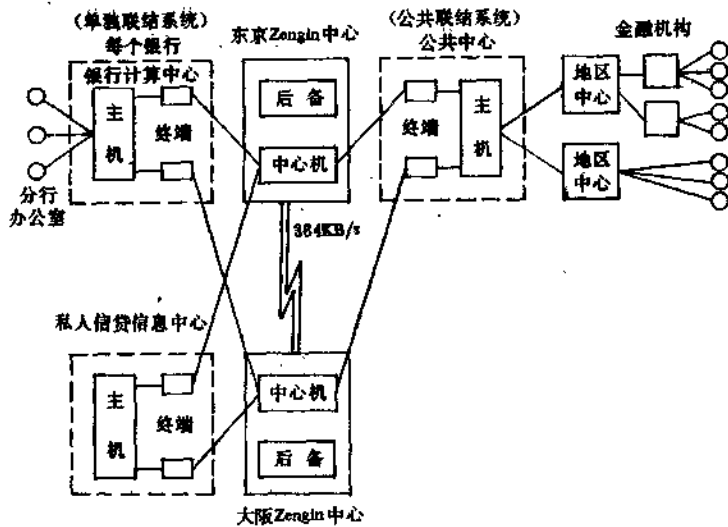


图 1.3 ZENGIN 系统

ZENGIN 系统的建成，进一步促进了日本银行和财政金融业务的发展，对发展日本经济产生了积极的影响。

1.2.2 数据管理系统

数据管理系统是计算机信息系统的一个重要组成部分，它的基本功能是按照用户的要求，从大量的数据资源中提取有信息价值的信息。例如可以通过检索、排序、合并、转换、汇总等方法获得这些数据。数据管理系统要解决两个主要的问题，一是定义各种数据的要求形式，二是如何由系统来处理这些要求。

对数据的要求可分常规的和特殊的两种。所谓常规要求是用户的日常业务活动经常发生的数据请求，可以事先预料到。例如企业领导要求知道当前企业的职工人数、某种产品的每月生产量等。所谓特殊要求是非常规性的，无法事先预料到的数据请求。例如企业领导想要知道由于某项原材料供应不足，而影响生产，以致哪些订货单位不能按期收到所订产品？应当分别赔偿它们多少钱？

对常规请求的处理，传统的方法是采用应用程序包 (Application Program Package, 简称 APP)，按照具体的应用程序要求建立文件系统，每一个文件基本上是专用的。例如职工档案、工资文件等。应用程序包无法解决对数据的特殊要求，除非临时编制程序，即使这样也不能使用户得到满意的结果。因为首先要向程序员解释这项特殊要求，然后程序员再确定在哪些文件中会有这些数据，最后再编写程序，调试和运行。花了很多时间，却往往事倍功半。

针对这种情况，从 70 年代开始，产生了一种新的数据管理方法，即数据库管理系统 (Data Base Management System, 简称 DBMS)，它强调数据是一种资源，因此要对数据资源加以组织，在统一的控制下为尽可能多的应用服务，实现数据共享，而不是为了某了特殊用户的需要设立专用的文件。

1.2.2.1 数据库方法

数据库方法用一种标准方式，即用一个软件包来执行数据的存储和检索功能。例如企业职工的劳动工时、每日生产量等数据都可以用同一个软件包存储或检索，这个软件包称为数据库管理系统（DBMS）。

数据库方法与应用程序包方法相比，有以下三项优点：

- （1）减少了数据的冗余度；
- （2）增强了数据的独立性；
- （3）简化了应用程序。

由于早期的文件系统是根据应用程序的需要而专门建立的，因此即使不同的应用程序所需要使用的数据有许多相同之处，也必须建立各自的文件，造成了大量的数据重复。例如职工档案文件中有职工姓名，工资文件中也有职工姓名，这样既浪费了存储空间，也使得数据的修改变得十分困难，很容易产生数据的不一致性，从而降低了数据的正确性。例如某个职工的姓名更改了，就要更改所有与该职工有关的文件。数据库方法把所有用户所需要的同一个数据项只存储在一个记录中，要求所有的应用程序只有向数据库管理系统提出请求，才能对这个数据项进行存储或检索。

应用程序包方法强调了数据和应用程序之间的依赖关系，数据的逻辑结构对该应用程序是优化的，但它的物理存储结构和逻辑结构是一致的，所以要想对现有的数据进行新的应用比较困难，一旦对现有的数据结构加以修改，则应用程序也要相应地修改。反过来，应用程序的改变也将影响到数据结构的改变，使得数据缺乏独立性。数据库方法使得数据的存储结构与应用程序相互独立，即使改变了数据的存储装置也不会影响应用程序。

文件系统使得应用程序的编制相当繁琐，而且缺少对数据的一致性、安全性、保密性等有效的、统一的控制手段。但是数据库方法可以减少应用程序对文件的定义、数据的编辑、有效性检验以及打印报表等工作，完全由数据库管理系统担负，这样既简化了应用程序，又增强了数据的一致性、安全性和保密性。

另外，数据库方法还减少了处理某些特殊数据要求的应用程序设计工作，它可以用数据库管理系统中的查询和报表模块来解决，不再需要专业程序员编制处理这些特殊数据要求的程序。

按照对数据的要求形式，数据管理系统可以由四个子系统组成，能够基本解决所有的应用问题。它们是：

- （1）存储和检索，由数据库管理系统完成。
- （2）处理，由应用程序完成。
- （3）输出，由查询、报表、数据管理模块完成。
- （4）输入，由数据库管理系统和应用程序完成。

其中，存储和检索系统是关键，它决定了处理、输出、输入三个系统的工作方式。所有这四个子系统加上操作系统和通信系统又决定了整个信息系统的运行效率。

1.2.2.2 存储和检索

存储和检索系统把数据项组织成记录，数据库管理系统规定了对记录的标准组织方

法，也规定了对记录以及记录索引的标准存储和访问方法，还规定了对记录之间关系的标准定义方法。目前在国际上流行的有三种标准的数据库方法：层次型数据库系统（IMS）、网状型数据库系统（DBTG）、关系型数据库系统。

层次型数据库系统，它的整体数据逻辑结构用若干个“数据库描述”（Data Base Description，简称 DBD）来定义，一个 DBD 可以定义一个物理数据库记录型或定义一个逻辑数据库记录型，其中包括这个数据库记录型的命名，以及组成这个数据库记录型的各个片段（Segment）的命名和描述。若干个特定 DBD 的集合，组成一个特定数据库系统的整体数据逻辑结构。每一个用户只能获得整体数据逻辑结构中的一部分，即局部逻辑结构。在 IMS 系统中，片段是存取数据库的最小命名单位，片段有型 and 值。所谓片段型相当于传统文件中的记录类型，片段值相当于具体的记录值。物理数据库记录型是具有树形联系的片段型集合。每一个物理数据库记录型含有且只含有一个根片段，以及零个或多个从属片段，片段都位于一定的层次上。

网状型数据库系统，它的整体数据逻辑结构模式用数据描述语言（Data Description Language，简称 DDL）来定义和产生数据库模式（Schema）。DDL 是描述整个数据库的一种独立的语言。模式主要包括各种记录型的定义，以及表示记录型之间联系的系型定义，对每一个 DBTG 用户，只能获得整体数据逻辑结构中的一部分，即局部数据逻辑结构。在 DBTG 系统中，数据库被划分成一个或多个部分，并给予命名。每一个命名部分称为域（Area），它是基本运行单位。

关系型数据库是根据数学中的关系理论建立起来的，是以关系模型为基础的数据库，它把关系看作是一个二维表，对一个关系的描述称为关系模式，它包含关系名、域名、属性名、属性向域的映象和属性间的依赖。关系必须是规范化的，这样才能更好地描述客观世界。目前关系型数据库尚未完全实现商品化，但是在微型机上已经基本实现。由于它简单灵活，独立性高，理论严格，所以是一种很有前途的数据库。

尽管实现的标准不同，但对记录的存取方法却是相同的。一般采用顺序存取或直接存取方法，有时也采用两者的结合，即索引顺序存取方法。

在存储和检索系统中，有三种数据定义类型：一是模式或子模式的数据定义语言（Data Definition Language，简称 DDL），它用于定义数据库；二是再组织模块，用于当模式改变时重新组织数据库；三是数据字典，它为用户或数据库管理员提供有关数据库的信息。

数据定义语言和模式的概念来自 CODASYL（Conference on Data System Language，数据系统语言协会）对数据库管理系统的定义。模式是逻辑数据库的精确定义，它包括了对数据项、记录、记录之间关系（指针）的定义。

一个数据项的定义至少有下列几条内容：

- （1）数据项名称；
- （2）别名；
- （3）类型（字符、整数、十进制小数）；
- （4）长度和小数点位置；
- （5）是否是关键字。

另外，在一个模式中还包含了以下三项主要内容：

- (1) 存取和修改权代码；
- (2) 虚拟数据项的计算参数；
- (3) 过程调用。

存取和修改权代码，用于保护数据的安全性。虚拟数据项的计算参数用于一个在数据库中实际上不存在的，但根据要求可以计算出来的数据项，例如总计或平均数。过程调用定义了连续进行数据存储或检索之前，调用某一个应用程序的条件，例如在数据库中增加一个数据之前，先要调用某一特殊程序对该数据进行有效性检验。

数据定义语言 DDL 是用于把模式定义到数据处理系统和数据库管理系统中去的一组规则，它描述了数据的特性、特殊的关系等，类似于 COBOL 语言中的数据部。这种描述具有层次的特点，即若干个数据项组成一条记录，若干个记录组成一个域，若干个域组成数据库，最后还要定义记录之间的关系。每一个数据库管理系统 DBMS 都有自己的数据定义语言 DDL。

此外，有些数据库管理系统不仅定义了模式，还定义了子模式，它是数据库的局部逻辑描述，它的目的是把一个特殊的应用程序包与数据库联系起来。与一个数据库有关的子模式允许有许多个，但只允许有一个模式。

数据存储再组织模块是数据存储和检索系统中相当重要的组成部分。一般来说，用户可能会要求以下几种变动：

- (1) 改变一个数据项，或者一个记录，或者一个域的名称；
- (2) 改变一个数据项或记录的长度，或是一个域的大小；
- (3) 改变一个数据项的类型；
- (4) 改变一个数据项的小数点位置；
- (5) 从一个记录类型中增加或删除某些数据项，或者从一个域中增加或删除某些记录型；
- (6) 修改某一数据项；
- (7) 改变一个记录类型中的关键字；
- (8) 增加、修改或删除某些次关键字；
- (9) 建立或删除记录关系。

这些变动中有一些只影响了模式的变化。在多数情况下，需要对数据库本身进行重新组织，恢复数据库的使用效率，这就要靠数据存储再组织模块来实现。

模式只是为数据库管理系统运行时提供必要的数据库特性，但这还不够，往往还需要有关数据库的其它信息。因为用户经常要求有关某些数据项更多的描述性的信息，而数据管理员需要知道是谁修改了数据项或记录，又是谁正在使用某些数据项或记录，程序员和系统分析员想知道当前数据库中某些数据项的综合信息以免使新的数据和老的数据在同一数据库中产生重复现象。为了解决这些问题，使用户方便地使用这些数据库定义，可以建立数据字典，它本身就是一个文件，既可以人工建立，也可以由计算机建立，由专门的数据字典软件包来实现。

数据字典的功能有：

- (1) 解释数据的名称、格式等；
- (2) 指明数据定义所在的存储位置；
- (3) 对数据库使用情况加以统计；
- (4) 对数据库管理系统生成数据加以定义；
- (5) 对输入数据的合法性和安全性予以辅助性检查。

例如，一个数据字典中可以分别记录有关模式、用户、数据库管理人员的各种数据项。

一般来说，数据库管理系统软件由若干个模块组成，其工作流程如图 1.4 所示。

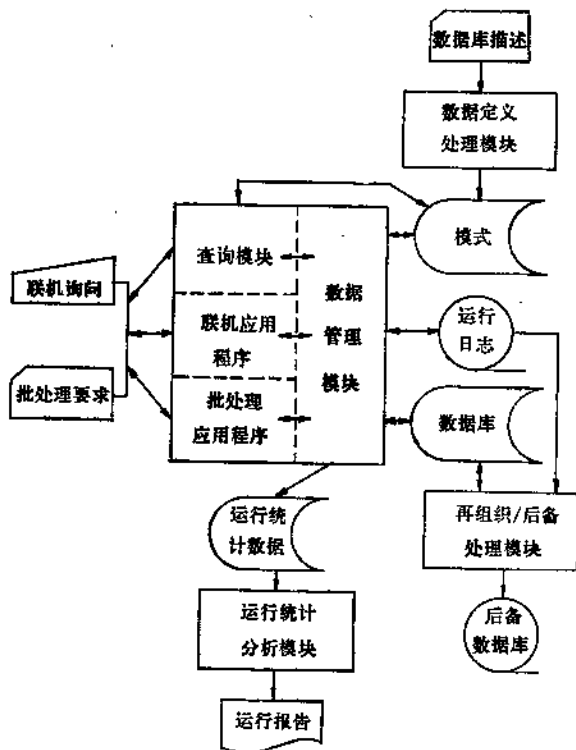


图 1.4 数据库管理系统

数据管理模块是数据库管理系统的核心模块，起着对系统的监控作用，它执行存取控制、对数据要求予以处理、记录运行日志、数据一致性检验等功能。在处理一个数据请求时，数据管理模块首先访问数据库模式，寻找该数据在数据库中的存储位置，然后通知操作系统，检索出所要的数据。

查询模块用于处理联机查询或批式请求，它把来自用户的查询请求翻译成数据库管理系统中的数据操纵语言 (Data Manipulation Language, 简称 DML)，它也把响应的结果以用户能够理解的形式显示出来。查询模块可以直接访问模式或数据字典，回答用户的查询请求。

数据定义处理模块的功能是定义数据模式，或改变数据模式。凡是定义新的数据项或改变原有的数据项，都由这个模块重新定义数据模式。

当数据的存储结构发生变化，翻要重新组织，或是要修改现有的记录片段间的存取路径，或是由于数据处理系统发生故障时，数据库管理员要求产生后备数据库，以及故障排除后重新恢复，这些工作都要由再组织/后备模块完成。

运行统计分析模块的功能是对运行统计文件进行存取并准备运行报告，该模块主要是为数据库管理员服务的。数据库管理员的职责是控制该系统，使其有效地运行。

数据库管理系统和操作系统以及用户的接口如图 1.5 所示。

图中的数据访问控制系统，用于终端用户在访问数据库时，控制数据的一致性和安全性。例如检验用户是否有权对该数据进行存取。用户接口软件包 (User Interface Package, 简称 UIP) 可以是一个文件查询系统，也可以是一个应用程序，联机用户