

高科学技术丛书

CC-DOS V4.0

操作系统 高级技术分析

钱培德 著

吉林科学技术出版社

高科学技术丛书

CC-DOS V4.0

操作系统高级技术分析

钱培德 著

吉林科学技术出版社

高科学技术丛书

CC-DOS V 4.0 操作系统高级技术分析

钱培德 著

责任编辑:珂 丽

封面设计:杨玉中

出版 吉林科学技术出版社

850×1168 毫米 32 开本 21.25 印张

发行 吉林省新华书店

插页 4 501,000 字

1991 年 4 月第 1 版 1991 年 4 月第 1 次印刷

印数:1—3000 册 定价:18.70 元

印刷 吉林大学印刷厂

ISBN7-5384-0795-2/TP·8

内 容 简 介

本书是国内第一本对汉字操作系统进行高级技术分析的著作。本书以 CC-DOS 操作系统的最新版本——CC-DOS V4.0 为蓝本,对汉字操作系统作了全面的、系统的分析,并首次揭示了 CC-DOS V4.0 所采用的全部高级技术。书中介绍了汉字操作系统的总体结构与总体设计思想,提出了系统各部分的实现方法和策略。书中还分析了系统的信息管理和存贮管理机制,给出了系统采用的高级技术。本书深入地阐述了系统的汉字 I/O 管理模块,对它们作了详尽的分析和论述。书中提供了 CC-DOS V4.0 各模块的流程图和程序清单,为了便于读者学习参考,还对程序中的每条指令加了注释。

本书适用于计算机研究、设计、使用和维护人员,尤其对汉字操作系统的高水平开发工作有较大的启发和帮助。本书还可以作为大专院校计算机专业研究生和本科高年级的教材或教学参考书。

序

MS-DOS(PC-DOS)是目前国内外广泛使用的 IBM-PC 计算机的主操作系统,它可以支撑极为丰富的应用软件。CC-DOS 则是 MS-DOS 的中文版本,除了与 MS-DOS 保持完全兼容以外,它还增加了处理中文信息的能力。经过多年的开发、应用、改进和不断完善,CC-DOS 已成为 IBM-PC 计算机上处理中文信息的一个事实上的标准开发环境。

CC-DOS 已在我国广泛使用,但几年来深感缺乏全面而详尽地剖析 CC-DOS 内部结构、设计方法及实现技术的参考书籍。《CC-DOS V4.0 操作系统高级技术分析》一书正是为此目的而撰写的。作者潜心研究 CC-DOS 已有多年,积其心得写成是书,特点有三:

第一,分析深入,说明透彻。书中对 CC-DOS V4.0 每个模块都给出了程序清单、注释及算法流程,无论是设计思想还是其实现技术均有深入的分析说明。

第二,条理清楚,脉络分明。全书从系统总述到存贮管理、输入输出等无不有条不紊,有分有合,形成一个结构良好的整体。

第三,联系实际,便于使用。本书既是计算机应用开发人员的必备读物,又可作为高校计算机系统软件教学的参考书。

全书取材新颖、内容丰富、结构清晰、行文通畅,不失为一部好书,故乐而为之序。

张福炎

1990 年 5 月序于南京大学

前 言

当今的时代是信息化的时代，现代社会是充满信息的社会。信息反映了现实世界的物理状况。近十年来，信息科学已经兴起，成为一门新兴的学科。它研究信息的来源、产生、获取、识别、转换、组织、存贮、处理、检索、表达、评价和提供与信息有关的理论和方法论。近年来，社会信息量急剧增大，每年已达 10 万亿字节，处理量也随之加大，处理形式也多样化，从数据到文字、图形、图象及自然语言。由于社会对共享信息资源的需要，就要求高速大容量传递和交换信息，要求有既经济又高可靠的信息系统。这样的信息系统必须依赖于计算机才能实现，所以，用计算机实现对信息的处理与管理，已经势在必行。

我国是使用汉语的国家，绝大部分信息要用汉语表示，称之为汉字信息。如果不解决计算机的汉字信息处理问题，就不可能在我国建立有效的信息系统，信息的处理和管理也将成为泡影。于是，在我国产生了一门新兴的学科——汉字信息处理。这门学科所涉及的范围很广，其中包括语言学、汉字编码、计算机体系结构和计算机软件等。事实证明，我国的现代化离不开计算机汉字信息处理技术。

我国的各种行业和各个部门中，需要各种各样的汉字信息处理系统来实现对信息的处理和管理。高性能微型计算机的问

世，已使得这些汉字信息处理系统所需的支撑硬件成为现实。然而，汉字信息处理系统还必须具有基本的系统软件——汉字操作系统。于是，汉字操作系统就作为计算机操作系统的一个新的分支，应运而生。

CC-DOS 操作系统是目前最著名的汉字操作系统，它自 1983 年问世以来，获得了很大的成功和迅速的发展。它已成为目前国内外使用最广泛的汉字操作系统。并且，它的结构已成为微机汉字操作系统的一种事实上的标准。CC-DOS 已经历了多个版本，这些版本充分体现了 CC-DOS 的发展过程，展示了一个为人试用的系统发展成为一个国内外广泛使用的系统的历程。CC-DOS 的道路并没有走完，它仍在向前发展。CC-DOS V4.0 是它的最新版本。

CC-DOS V4.0 集 CC-DOS 历届版本之所长，并吸收了许多 CC-DOS 变种的优点，再加以充分发挥。所以，这个版本与 CC-DOS 的其它版本有明显的区别。有人称 CC-DOS V4.0 为“CC-DOS 更新换代的产品”。本人认为这种说法并不过份，它能体现出我国汉字信息处理技术的水平。正因为如此，人们对 CC-DOS V4.0 格外感兴趣，企图学习它的先进技术和实现方法。但是，它的设计者为了维护该软件的经济利益，在系统内采用了一系列的加密和保密措施。这就给分析此系统设置了障碍，使系统所用的某些技术至今仍笼罩着神秘的色彩。

本人自 1983 年开始从事汉字操作系统的研究与开发，先后研制了 UNIFLEX 多用户操作系统的汉化版本、CH-DOS 汉字操作系统、PG 型汉字信息输入系统，以及通用联想式汉字操作系统。本人于 1985 年在国内首先分析了 CC-DOS V2.1，撰写了“CC-BIOS 分析”和“微型计算机汉字操作系统 CC-DOS”两本著作，首次把 CC-DOS 的内情公布于众，在计算机界产生较大的反响。这两本著作在全国发行数万册，书中的内容被许多论文和著作所引用。

本人对 CC-DOS V4.0 亦具有极大的兴趣,在经过一段时间的努力后,终于完成了对这个 CC-DOS 最新版本的分析。在分析了 CC-DOS V4.0 以后,本人觉得该系统设计新颖、独特、严谨和先进,很值得大家一学。为了适应当前国内外汉字信息处理技术飞速发展的形势,为了满足计算机科研、工程技术人员和计算机用户学习 CC-DOS V4.0 的需要,本人特撰写本书,对 CC-DOS V4.0 作出全面的、彻底的技术分析。

本书是国内第一本完整地介绍汉字操作系统的著作,它以 CC-DOS V4.0 为例,完整地 and 全面地分析汉字操作系统的结构和各模块的流程,并描述了系统的实现方法。本书力求突出普遍性、技术性和先进性,使读者能对汉字操作系统有较全面的了解和掌握 CC-DOS V4.0 所采用的先进技术。

本书共分七章:第一章介绍 CC-DOS 的概况和它的用户界面,为分析该系统提供基础;第二章介绍汉字操作系统的总体结构和总体设计,并提出主要模块的实现方法;第三章分析了 CC-DOS V4.0 的信息管理机制和存贮管理模块,为对系统的进一步分析奠定基础;第四章分析系统文件的结构、DOS 的自举过程和 CC-BIOS 的自举过程,还给出了汉字系统引导程序的注释;第五章分析了显示输出管理模块中各个功能块的情况,重点介绍汉字显示输出的实现,还给出了显示输出管理模块的程序注释;第六章分析了键盘输入管理模块的各个功能块,重点介绍汉字输入的实现,还给出了键盘输入管理模块的程序注释;第七章介绍打印输出管理模块的各个功能块,重点介绍汉字打印输出的实现,还给出了打印输出管理模块的程序注释。

本书对 CC-DOS V4.0 的各个部分均给出了流程,并指出了入口和出口参数以及入口地址。对于 ROM-BIOS 中的程序模块,书中未指出入口地址,这是因为各种机型的 ROM-BIOS 程序的地址有别。

本书提供了较完整的 CC-DOS V4.0 程序清单,并把这些程

序按结构进行了划分。每个子程序前列出了功能介绍和入口参数(或出口参数)的说明。为了方便读者阅读和分析程序,每条指令后均附有简要的注释。本书中的 CC-DOS V4.0 程序清单不是源程序,而是先对 CC-DOS V4.0 的目标文件代码进行解密复原,再用 DEBUG 进行反汇编得出的程序清单。为了清楚起见,书中把数据区和工作区内容均用伪指令 DB 和 DW 等表示。另外,为了节省篇幅,对那些较大的数据区(如汉字库和输入码对照表),只给出其地址,而省略其内容。

由于 CC-DOS V4.0 的复杂性和本人水平的限制,书中定有谬误之处,敬请读者不吝指正。

钱培德

于苏州大学计算机工程系

目 录

第一章 引论	(1)
第一节 CC-DOS 的发展	(1)
一、发展过程	(1)
二、CC-DOS 的地位	(3)
三、CC-DOS V4.0 的新技术	(4)
第二节 系统的用户界面	(7)
一、引言	(7)
二、系统组成和启动	(7)
三、汉字输入的辅助操作	(10)
四、汉字打印输出的辅助命令	(14)
第二章 系统总述	(16)
第一节 系统总体设计与结构	(16)
一、引言	(16)
二、设计思想	(16)
三、系统基础与总体设计	(20)
四、总体结构	(25)
第二节 汉字代码体系	(26)
一、概述	(26)
二、汉字输入码	(27)
三、汉字机内码	(29)
四、其它汉字代码	(32)
五、CC-DOS 的汉字代码	(35)
第三节 系统实现	(38)
一、总体实现	(38)
二、汉字键盘输入的实现	(40)
三、汉字显示输出的实现	(47)

四、汉字打印输出的实现	(51)
第四节 汉字库结构的设计	(56)
一、汉字库概述	(56)
二、静态汉字库结构	(58)
三、动态汉字库结构	(60)
第三章 信息存贮管理	(68)
第一节 信息管理机制	(68)
一、文件目录表(FDT)	(69)
二、文件分配表(FAT)	(74)
三、文件分配表的使用	(78)
第二节 磁盘信息的存取	(85)
一、磁盘的逻辑结构	(86)
二、磁盘 I/O 参数表	(89)
三、逻辑扇区号的计算	(92)
四、用 DOS 中断访问磁盘	(94)
五、用 BIOS 中断访问磁盘	(105)
第三节 可执行文件的结构	(109)
一、可执行文件概述	(109)
二、COM 文件的结构	(109)
三、EXE 文件的结构	(111)
四、变体 EXE 文件的结构	(120)
第四节 存贮管理	(123)
一、存贮管理总述	(123)
二、系统的内存布局	(125)
三、存贮管理中的数据结构	(127)
四、存贮区的分配	(129)
五、存贮区的回收	(131)
六、存贮区的修改	(134)
第四章 系统自举与初始化	(137)
第一节 系统文件及其结构	(137)
一、DOS 文件	(137)
二、CC-BIOS 文件	(140)

第二节 DOS 的自举与初始化	(145)
一、DOS 自举总述	(145)
二、引导记录的执行	(148)
三、DOS 输入输出模块的自举	(150)
四、DOS 内核模块的自举	(152)
五、命令处理模块的自举	(153)
第三节 CC-BIOS 的自举与初始化	(153)
一、CC-BIOS 自举总述	(153)
二、汉字系统引导程序的执行	(154)
三、键盘输入管理模块的自举	(160)
四、显示输出管理模块的自举	(167)
五、打印输出管理模块的自举	(171)
第四节 汉字系统引导程序的注释	(176)
第五章 显示输出管理模块	(187)
第一节 模块总述	(187)
一、概述	(187)
二、模块结构和调用方法	(188)
三、字库组织	(190)
四、视频存贮区	(192)
五、模块的总体流程	(194)
第二节 视频初始化	(196)
一、概述	(196)
二、视频初始化的实现	(197)
三、西文系统的视频初始化	(199)
第三节 光标功能的实现	(202)
一、光标位置的读取	(202)
二、光标的产生和定位	(202)
三、RRAM 映象区的浮动	(205)
四、光标属性的定义	(212)
第四节 屏幕的滚动	(213)
一、屏幕滚动的概念	(213)
二、屏幕滚动的实现	(214)

第五节 字符的读出与显示	(222)
一、字符的读出	(222)
二、字符的显示	(223)
三、字符的 TTY 方式显示	(240)
第六节 提示行管理与其它	(243)
一、提示行管理	(243)
二、汉字库字模的读写	(247)
三、光标的建立与删除	(249)
四、象元的读出与写入	(250)
五、视频状态的读出	(254)
六、屏幕颜色的定义	(255)
七、当前显示页的选择	(257)
第七节 显示输出管理模块的程序注释	(260)
第六章 键盘输入管理模块	(340)
第一节 模块总述	(340)
一、概述	(340)
二、键盘中断处理程序	(341)
三、键盘缓冲区及其指针	(342)
四、输入码对照表	(345)
五、主要工作区	(347)
第二节 主体流程的分析	(348)
一、模块主体流程	(348)
二、各功能块的流程	(349)
第三节 字符输入模块的分析	(353)
一、字符输入模块的主体	(353)
二、字符处理子程序	(357)
第四节 功能符的处理	(361)
一、功能符处理子程序	(361)
二、对 CTRL+Fi 符的处理	(362)
三、对 ALT+Fi 符的处理	(368)
四、输入制表符处理	(372)
第五节 输入码的处理	(375)

一、区位码处理程序	(375)
二、区位码转换程序	(379)
三、拼音-首尾码处理程序	(382)
四、拼音-首尾码转换程序	(391)
第六节 高频字统计和字典功能	(398)
一、高频字统计处理	(398)
二、字典功能的实现	(400)
第七节 外部输入码的处理	(404)
一、外码模块的连接与装入	(404)
二、外码的字典功能处理	(405)
三、外码的转换程序	(406)
第八节 词组输入的处理	(414)
一、词组输入处理概述	(414)
二、词组输入码对照表	(415)
三、主要工作区	(416)
四、词组处理程序	(417)
五、词组选择处理程序	(421)
六、词组恢复符处理程序	(424)
第九节 键盘输入管理模块的程序注释	(425)
第七章 打印输出管理模块	(524)
第一节 模块总述	(524)
一、概述	(524)
二、汉字字型的变换	(525)
三、模块的结构	(526)
第二节 主体流程和工作区	(527)
一、主体流程的分析	(527)
二、各功能块的流程	(528)
三、主要工作区	(532)
第三节 打印参数的定义	(537)
一、总述	(537)
二、打印字型的定义	(537)
三、字间距的定义	(538)

四、行间距的定义	(540)
五、行宽和工作方式的定义	(543)
第四节 打印驱动的实现	(545)
一、主体流程的分析	(545)
二、西文方式下的打印输出处理	(547)
三、汉字方式下的打印输出处理	(551)
第五节 行信息的打印输出	(559)
一、缓冲区内容的输出	(559)
二、字模的读取与变换	(563)
三、输出行的纵扩处理	(577)
第六节 屏幕硬拷贝	(581)
一、总体流程	(581)
二、字符方式下的屏幕硬拷贝	(582)
三、图形方式下的屏幕硬拷贝	(584)
四、一列信息硬拷贝的实现	(588)
第七节 打印输出管理模块的程序注释	(595)

第一章 引 论

第一节 CC-DOS 的发展

一、发展过程

随着 IBM-PC 微型计算机在我国的引进和应用,广大用户迫切需要 IBM-PC 微型计算机的汉字操作系统,以满足汉字信息处理的要求。与 IBM-PC 一起引进的还有仓颉汉字系统,这种汉字系统需要在 IBM-PC 系统内加插一块仓颉汉卡,这无疑会增加用户的经济负担;另外,仓颉汉字系统采用仓颉码作为输入汉字的手段,这种编码相当繁琐,一般用户不易掌握和记忆;再则,此系统输出的是繁体汉字,不符合国内使用简化字的情况。所以,这种汉字系统未能获得广泛应用。

为了满足广大 IBM-PC 用户的要求,为了推动我国的计算机应用发展,电子工业部第六研究所于 1983 年推出了具有我国特色的汉字操作系统 CC-DOS。由于 CC-DOS 具有功能齐全、使用方便和不需要增加硬件等优点,因此它一经问世,就受到广泛的欢迎。

最初推出的是 CC-DOS V1.0 和 CC-DOS V1.1,它们在 IBM-PC 及其兼容机上运行。它们具有“区位”、“国标”、“音韵”、“电报”、“首尾”和“拼音”六种汉字输入方式。由于受到当时 IBM-PC 的存贮资源限制,这六种汉字输入方式中,只有“区位”和“国标”输入方式能支持 GB2312 国标字库中的一、二级

汉字(619个图形符号和6763个汉字)。其它四种输入方式只支持GB2312国标字库中的一级汉字(3755个汉字)。另外,系统中尚有不少地方不够完善,例如,系统没有词组输入功能等。

随着IBM-PC/XT微型计算机和PC-DOS V2.0的推出,电子工业部第六研究所于1984年开发出了CC-DOS V2.0和CC-DOS V2.1。这两种版本具有“区位”、“首尾”、“拼音”和“快速”四种汉字输入方式,这些输入方式均可支持GB2312国标字库中的一、二级汉字。这两个版本克服了第一版本的缺点,它们具有词组输入功能,能够造入新的汉字和定义词组,可以打印出24×24点阵的高质量汉字。这意味着CC-DOS进入了成熟阶段,它赢得了众多的用户,使CC-DOS成为拥有最多用户的汉字操作系统。

随着应用程序对内存需求量的不断增加,以及汉卡技术的发展,在PC-DOS V3.0的基础上,CC-DOS V3.0于1985年问世了。这个版本充分利用PC-DOS V3.0灵活的系统重构能力,大大增强了汉字输入输出的处理能力。它支持汉卡,把汉字库从内存移到了汉卡上,从而为应用程序提供了更多的内存资源。它可以支持包括彩色打印机在内的多种打印机。它对输入方式作为调整,加入了查询方式及笔形码,增加了字典查询功能,对拼音码作了改进和提高,允许一字多音。例如“车”,其拼音由“che”和“ju”输入这两者之任意一个,均可得到“车”,这就解决了拼音输入方式中的一字多码的问题。另外还增加了拼音的声调,以减少重码个数。

CC-DOS问世后,我国的计算机汉字信息处理技术发展较快,电子工业部第六研究所采用一系列新技术,于1987年重新设计和开发出了CC-DOS V4.0。CC-DOS V4.0集中了CC-DOS以往各版本的优点,汇集了当时国内的汉字信息处理新技术,它是CC-DOS更新换代的产品。这个版本的汉字输入方式扩充到十余种,其辅助输入手段丰富,配有词组输入、模糊输入、字