

高等院校信息技术规划教材

数据库技术应用基础

史九林 窦显玉 编著

INFORMATION TECHNOLOGY
INFORMATION TECHNOLOGY
INFORMATION TECHNOLOGY



清华大学出版社

高等院校信息技术规划教材

数据库技术应用基础

史九林 窦显玉 编著 (林道雄 陈永基 陈永基 陈永基 陈永基)

[Faint, illegible text from the reverse side of the page]

080776-91m 基盤多角形断崖

7108799-010

本公司為協助股東行使權利，特將本報告書全文刊登於本公司網站，凡欲查閱者，請逕向本公司索取，或逕向本公司索取，或逕向本公司索取。

VI. 附 註 023 X 281, 本 共

2005 年 1 月 1 日

0004 4 2 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 10

國家經濟出版社編輯部編譯，並由重慶市出版局代印。頁數，頁碼，頁類，申報，新不字文查等項件本 錢雪忠

多媒体技术与应用教程 清华大学出版社 978-7-302-17956-6 雷运俊

清华大学出版社
北京

内 容 简 介

本书以数据库技术的基本知识、基本概念和基本方法为主线,以小型数据库应用系统实例为引导,全面介绍数据库技术的主要内容,突出数据库的设计、实现和编程方法,着力培养数据库应用系统的开发能力。本书特点是,在精选课程内容的基础上,建立由总体到具体、由浅入深、由简到繁、难点分散的教材体系;实现强化实践、侧重能力、突出应用的教学目标;运用深入浅出、语言流畅、图文并茂、通俗易懂的表述方式。

本书适宜用作普通大专院校、职业技术学院等计算机应用专业及其相关专业学生数据库技术课程的教材;也可作为数据库技术人员、计算机应用系统开发人员及相关人员培训班教材,自学数据库技术人员的参考书。

本书封面贴有清华大学出版社防伪标签,无标签者不得销售。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数据库技术应用基础/史九林,窦显玉编著. —北京:清华大学出版社,2009.1

(高等院校信息技术规划教材)

ISBN 978-7-302-18199-6

I. 数… II. ①史… ②窦… III. 数据库系统—高等学校—教材 IV. TP311.13

中国版本图书馆 CIP 数据核字(2008)第 108259 号

责任编辑:袁勤勇 徐跃进

责任校对:李建庄

责任印制:杨 艳

出版发行:清华大学出版社

<http://www.tup.com.cn>

社 总 机:010-62770175

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质 量 反 馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

地 址:北京清华大学学研大厦 A 座

邮 编:100084

邮 购:010-62786544

印 装 者:北京鑫海金澳胶印有限公司

经 销:全国新华书店

开 本:185×260 印 张:17

字 数:390 千字

版 次:2009 年 1 月第 1 版

印 次:2009 年 1 月第 1 次印刷

印 数:1~4000

定 价:25.00 元

本书如存在文字不清、漏印、缺页、倒页、脱页等印装质量问题,请与清华大学出版社出版部联系调换。联系电话:010-62770177 转 3103 产品编号:026951-01

前言

foreword

数据库技术的出现虽然比计算机诞生的时间晚了近 20 年。但此后的几十年间,数据库技术与计算机技术同步发展、相互支持、共生共荣。从 20 世纪 60 年代后半期开始的几十年间,数据库技术的研究、开发和应用逐步进入高潮;成果不断涌现,不断推陈出新;经历了从层次模型到网状模型,再到关系模型的里程碑式发展过程;商品化数据库软件层出不穷,大量投放市场,犹如天女散花,遍布全球。到 20 世纪 70 年代中期,数据库之花也绽放在中国大地。现在,已不再有人怀疑数据库技术在数据存储和管理领域的重大意义和巨大作用。几乎百分之百的计算机系统都装备有数据库系统;几乎百分之百的计算机用户都应用数据库技术于自身的业务管理应用;几乎所有的计算机工作者,不管是专业的还是非专业的,都努力地去掌握数据库技术和方法,并引以为能。

数据库技术是一种数据管理技术,是一门计算机专业的专业基础课程,是计算机专业及其相关专业的学生都必须学习和掌握的技术。因为计算机是一种信息处理工具;信息处理的基础是数据处理,数据处理的基础是数据管理。数据库技术又是一种面向应用、实践性很强的技术;离开应用就会失去存在的价值;没有足够的实践,技术就显得苍白无力。笔者曾从事数据库技术课程教学 30 余年。近几年,又有机会参与和从事计算机应用技术型人才和职业技术人才的培养和教育工作。深感技术教学、实践教学和能力培养的深远意义和实践意义。笔者以为,职业技术教育需要处理好两个关系。一是理论与实践的关系。在强调实践教学的同时不要忽视理论教学。实践只有特殊性,理论才有一般性。从实践出发寻找理论支持是一个好的教育方式。我们反对没有实践的理论,但也反对没有理论的实践。“实践→理论→再实践”是一个合理的公式。二是案例教学与举一反三的关系。案例是“举一”、是解剖麻雀、是引导、是学会、不是目的;“反三”是能力、是可持续发展、是目的;二者必须有机结合、相辅相成。

教材是教学的脚本。有一本合适的教材是教学和学习成功的

一半。理论与实践结合、案例与举一反三相辅必须通过教材来体现和实施。笔者在本书编写过程中试图针对职业技术教育层次进行初步尝试和改革。主要考虑的是如何改革教材体系与结构、内容重点与表述、教学方法与过程等。首先,对内容进行优选。侧重数据库技术的基本问题,如基本知识、基本概念和基本原理;以实例为引导,如选择合适的DBMS实例和应用课题实例;使学生能掌握数据库技术的基本内容和基本技术,初步建立起数据库技术基础知识和思维方法;训练学生正确运用数据库方法设计和开发数据库应用系统的初步能力。第二,建立由总体到具体、由浅入深、由简到繁、难点分散的教材体系;使学生循序渐进、自然地进入数据库领域。如第1章先介绍数据库的总体概念,第2章给出一个小型应用课题并进行分析,此后各章围绕应用课题展开,第7章是一个总结。第8章是一个DBMS拓展。第三,强化实践,注重能力,特别着力于数据库设计过程的问题分析和具体DBMS实现,体现实践第一性原理。书中围绕应用实例给出每一阶段的操作过程,供学生验证、实验、模仿和练习。要求学生能借助实例达到举一反三的目的,进一步自行完成实例未给出部分的设计和实现。注意把数据库技术与数据库技术的应用融通在一起;发挥示范DBMS的实践作用。第四,为使学生抓住课程重点,每章开头列出本章要点和知识点。每章习题给出丰富的不同类型的题目,力图覆盖全章的知识要点和能力训练。通过习题回味可加深对教材内容的理解,提高实践能力。第五,在全书的表述上注意做到深入浅出、语言流畅、图文并茂、通俗易懂、重点突出。以期达到让读者“一看就懂,一学就会,一练就通,举一反三”的目的。以上可以算是本书的一些特点吧。是否合适,还是由读者说了算;还得由专家们点评认可。

本书适合作普通高等院校计算机应用专业及其相关专业学生的数据库技术应用课程教材;也可作为计算机应用系统开发人员及相关人员的学习参考书或培训教材。

全书共8章,由两人参与编写。史九林编写了第1至7章。窦显玉编写了第8章。全书由史九林策划、设计和统稿。

南京大学计算机科学与技术系徐洁磐教授是我国知名数据库专家,我的良师益友。他不仅拨冗仔细审阅了全书手稿,提出了许多改进意见和建议;而且,在本书写作过程中还给予了极大鼓励、关心支持、全力帮助和悉心指导,笔者深表感谢。在本书编写过程中还得到了三江学院王芝庆教授、金肯职业技术学院计算机与通信工程系领导的关心,以及计算机应用技术专业同仁们的支持。在此一并表示感谢。

由于笔者水平有限,时间仓促,书中疏漏和错误在所难免,恳望读者批评指正。

编 者

2008年9月于南京

目录

Contents

第 1 章 数据库系统基本概念 1

1.1 什么是数据库 1

1.1.1 信息、数据和数据管理 1

1.1.2 信息处理与数据处理 6

1.1.3 数据库系统的初步印象 7

1.1.4 数据库技术的发展 9

1.2 数据库系统的体系结构 10

1.2.1 三级模式 11

1.2.2 二级映射 12

1.2.3 3 种记录 12

1.2.4 数据独立性 14

1.3 数据库语言 15

1.3.1 数据描述语言 15

1.3.2 数据操纵语言 16

1.3.3 程序设计语言 16

1.4 数据库管理系统 17

1.4.1 数据库管理系统的功能 17

1.4.2 数据库管理系统的组成 18

1.4.3 数据库管理系统的分类 18

1.5 数据库系统活动过程 19

1.6 数据库技术的特点 20

习题 1 22

第 2 章 数据库应用系统 25

2.1 关于对数据库的几个认识 25

2.2 数据库应用系统 26

2.2.1 数据库应用系统构成 26

2.2.2 数据库应用系统开发 27

2.3	一个数据库应用实例课题——教学管理	28
2.3.1	课题背景	28
2.3.2	课题要求	28
2.4	实例课题需求分析	29
2.4.1	需求调查	29
2.4.2	数据需求分析	29
2.4.3	功能需求分析	32
2.4.4	数据与功能关系分析	34
2.5	实例课题解决方案的选择	35
2.5.1	解决方案选择的意义	35
2.5.2	解决方案选择的内容	35
2.5.3	实例课题的解决方案	36
2.6	数据库系统 Visual FoxPro	36
2.6.1	简介 VFP	36
2.6.2	VFP 的启动和关闭	38
2.6.3	VFP 项目管理器	38
	习题 2	40
第 3 章	数据模型设计及数据库创建	42
3.1	数据库数据主要特性	42
3.2	数据抽象过程——4 个世界	43
3.2.1	现实世界	43
3.2.2	概念世界	44
3.2.3	逻辑世界	45
3.2.4	物理世界	45
3.3	数据模型的基本概念	46
3.3.1	实体与联系	47
3.3.2	约束条件	49
3.3.3	操作方式	49
3.3.4	数据模型的设计步骤	50
3.4	概念模型设计	50
3.4.1	设计要点	50
3.4.2	E-R 图方法	51
3.4.3	概念设计实例	52
3.5	关系数据模型	53
3.5.1	关系模型	53
3.5.2	关系数据库模型	55
3.5.3	关系模型的特点	55

3.6	关系模型设计	55
3.6.1	从概念模型到关系模型	55
3.6.2	关系模型设计实例	57
3.6.3	教学管理数据库的数据模型	61
3.7	关系数据库创建	61
3.7.1	创建 VFP 数据库	61
3.7.2	在数据库上创建数据表	63
3.7.3	创建联系和表间约束条件	66
3.8	数据装入	68
3.8.1	数据装入方法	68
3.8.2	实例数据库表批装入演示	69
3.8.3	多媒体数据的输入	71
3.8.4	创建完成的教学管理数据库表内容展示	73
习题 3		75
第 4 章	关系数据库操作	77
4.1	数据库操作的一般概念	77
4.1.1	数据库操作的种类	77
4.1.2	数据库操作的分解	78
4.2	关系代数	79
4.2.1	关系代数的构成元素	79
4.2.2	关系代数运算规则	80
4.2.3	关系代数的复合运算	83
4.2.4	关系代数的意义	85
4.3	VFP 数据库操作	86
4.3.1	VFP 数据库存储概念	86
4.3.2	打开和关闭操作	87
4.3.3	记录操作	89
4.3.4	浏览	93
4.3.5	文件操作	98
习题 4		102
第 5 章	查询、SQL 与视图	104
5.1	查询	104
5.1.1	查询的概念	104
5.1.2	单表查询设计	105
5.1.3	多表查询设计	110

5.2	SQL 语言	112
5.2.1	SQL 语言简介	112
5.2.2	SQL 数据查询语句	113
5.2.3	SQL 数据修改语句	120
5.2.4	SQL 数据定义语句	122
5.3	视图	125
5.3.1	视图的概念	125
5.3.2	视图的定义	125
5.3.3	视图的操作	129
5.3.4	视图合成	131
5.3.5	视图的性质	131
5.3.6	视图与数据独立性	132
5.4	关系数据库系统体系结构	132
习题 5		133
第 6 章	数据库编程	135
6.1	程序和程序设计	135
6.1.1	什么是程序	135
6.1.2	程序设计语言	136
6.1.3	程序设计技术	136
6.1.4	开始编程	136
6.2	VFP 程序设计语言	138
6.2.1	VFP 数据	138
6.2.2	VFP 运算	138
6.2.3	常用 VFP 命令语句	140
6.2.4	VFP 流程控制语句	141
6.3	VFP 结构程序设计	145
6.4	VFP 面向对象程序设计	149
6.4.1	什么是面向对象程序设计方法	149
6.4.2	对象和类	153
6.4.3	类的运用	157
6.4.4	类的设计	158
6.5	VFP 编程工具	162
6.5.1	表单向导	162
6.5.2	生成器	164
6.5.3	再举一个表单设计例子	168
6.5.4	报表	172
6.6	应用系统集成和菜单设计	175

6.6.1	应用系统集成	175
6.6.2	VFP 菜单设计	176
习题 6		178
第 7 章 数据库设计与数据库管理		180
7.1	数据库设计概述	180
7.1.1	什么是数据库设计	180
7.1.2	数据库设计模型与路线图	181
7.1.3	数据库设计原则	183
7.1.4	数据库设计工具	183
7.2	数据库设计	184
7.2.1	需求分析	184
7.2.2	概念设计	192
7.2.3	逻辑设计	196
7.2.4	物理设计	203
7.3	数据库管理	206
7.3.1	数据服务管理	206
7.3.2	数据库性能管理	207
7.3.3	数据库安全管理	208
7.3.4	故障恢复管理	209
习题 7		210
第 8 章 SQL Server 2000 简介		213
8.1	SQL Server 2000 的运行环境	213
8.1.1	SQL Server 简介	213
8.1.2	客户端/服务器体系结构	214
8.1.3	SQL Server 集成	215
8.1.4	安装 SQL Server 2000	217
8.1.5	SQL Server 主要的管理工具	224
8.2	在 SQL Server 上建立示例数据库	225
8.2.1	数据库与表	225
8.2.2	SQL Server 数据库对象	229
8.2.3	创建和删除数据库	230
8.2.4	表的创建和删除	240
8.2.5	创建连接	244
8.2.6	数据装入	248

251	8.3 T-SQL 和示例数据库查询	251
251	8.3.1 T-SQL 语言	251
252	8.3.2 T-SQL 查询设计	252
252	8.3.3 示例数据库的 T-SQL 查询	252
256	习题 8	256
257	第 8 章 SQL Server 2000 简介	257
257	8.1 SQL Server 2000 的运行环境	257
257	8.1.1 SQL Server 简介	257
257	8.1.2 客户端\服务器体系结构	257
257	8.1.3 SQL Server 集成	257
257	8.1.4 安装 SQL Server 2000	257
257	8.1.5 SQL Server 主要的管理工具	257
257	8.2 在 SQL Server 上建立示例数据库	257
257	8.2.1 数据库引擎	257
257	8.2.2 SQL Server 数据库对象	257
257	8.2.3 创建和删除数据库	257
257	8.2.4 数据库的备份和还原	257
257	8.2.5 数据库的连接	257
257	8.2.6 数据库的安全	257

第1章

chapter 1

数据库系统基本概念

计算机的一个十分重要的应用领域是信息管理,信息管理的核心技术是数据管理。数据管理的近代技术是数据库。数据库技术的出现和迅速发展使之成为计算机学科的一个重要分支。其应用的广泛性、深入性和重要性已是尽人皆知,不言而喻。

本章的主要任务是让读者对数据库技术有一个比较全面的了解和认识,建立起关于数据库的基本概念,包括基本知识和术语、基本原理和方法、基本结构和技术等几个主要方面。本章主要回答以下几个问题。

- (1) 什么是信息、数据、数据管理和数据处理?
- (2) 有哪些常见的计算机数据管理技术?
- (3) 什么是数据库?数据库系统由哪些元素组成?
- (4) 数据库系统的体系结构是如何架构的?
- (5) 什么是数据库管理系统?在数据库系统中有什么作用?
- (6) 数据库语言包括哪些语言?它们各有什么用途?
- (7) 数据库系统的活动过程是什么?
- (8) 数据库系统技术为什么能得到广泛应用?

1.1 什么是数据库

1.1.1 信息、数据和数据管理

1. 信息

信息(information)一词已被人们熟悉和广泛使用。至于什么是信息还没有公认、确切的定义。控制论创始人维纳(N. Wiener)说:“信息就是信息,它既不是物质也不是能量。”实际上,信息是现实世界中客观事物的属性表征及其运动状态的真实反映所形成的关于事物的概念。人一旦接收、认识和掌握了某事物的信息,他就具备了关于该事物的知识。现实生活中,人们常常把获得的消息、新闻、情报、广告和知识等看成信息。信息总是伴随着事物的存在而存在,事物又总是以信息顽强地表现自己,不管你是否已经认识或获取到。人总是千方百计地获取信息,以决定自己此后的社会活动,以企进一步认识

世界和改造世界。因此,信息已成为一个组织或部门的重要财富和资源。人类的社会活动形成了物质流、资金流和相伴随的信息流。信息流控制和制约着物质流和资金流的运动。

2. 数据

数据(data)是信息符号化的结果,是用以表示、存储、传输信息的一种结构化符号串。所谓“结构化”是指数据中的符号必须遵循预先规定的规则构成“串”。所谓“符号”是指文字、标点、声音元素、几何元素、时间元素等公众已经接收并广泛流行的基本符号的全体。使用不同基本符号构成的数据有不同的展现形式,如数值、文本、语音、音频、图像、图形、视频等。可以用不同形式的数据表示同一信息。通常人们把信息和数据等同视之。虽然这并不妨碍对数据和信息的理解,但实际上,信息和数据是有严格区别的。数据是表示信息的一种媒体,也称载体;信息载荷在数据之上。或者说,数据所表达的内涵和意义才是信息。近代计算机技术允许把不同形式的数据叠加、复合在一起,表示、处理和展现复杂的信息,这就是多媒体和多媒体技术。

相对于计算机信息处理,数据又被定义为计算机输入、处理和输出的对象。所以有人把数据说成是计算机领域的专门术语,是构成计算机软件的一部分。任何媒体形式的数据若要进入计算机,都必须经过数字化的过程。就是把数值、文本、语音、音频、图像、图形、视频等形式的数据转换成比特(bit,又称位)流,即二进制表示,并进行编码、存储在计算机的存储设备上,然后计算机才能通过软件对其进行处理或输出。西文编码表,汉字编码表,图形、图像、视频和声音的采样、量化、编码/压缩编码等都是信息媒体数字化的工具。因此,任何形式的数据在计算机内都统一成了二进制位串,即“0”/“1”的串。这也是计算机多媒体技术的基础。

3. 数据管理

计算机数据处理面临数量巨大、关系复杂的数据集群。如何管理好数据,以提高数据处理的有效性和提供优良的信息服务,成为计算机学科的一个重要研究课题。随着计算机软硬件技术的不断发展,特别是外部存储器产品的不断更新,容量不断增大,计算机数据管理(data management)技术经历了自由管理、文件管理和数据库管理3个重要发展阶段。

1) 自由管理阶段

自由管理(adaptive management)实际上是无管理。在计算机仅有内存可以存储数据的年代,程序员把数据和处理它的程序捆绑,构成一个整体。程序流程与数据两相适应,共存共亡。或者说,在这种方式下,没有统一的数据管理规则;有的只是程序员智慧和匠心之体现;任何一个程序都必须特别地考虑它所处理的数据的逻辑特性和存储特性;按地址引用数据。数据为处理它的程序私有。在数据量不大的情况下,这种方法尚可使用。

2) 文件管理阶段

把数据组织成文件进行管理是计算机数据管理技术的重大进步,至今久盛不衰。文件管理(file management)是以文件为单位的数据管理方式。文件是一种按规定规则或标准组织起来的,同类数据的集合;独立地存储在外部存储设备上。为其命名一个名字

作为标识,称文件名。此后,就可以按文件名直接引用文件;不再需要考虑那些与物理存储位置、设备特性有关的细节问题。为适应不同的应用和存储设备,文件有不同的组织方式,如流式文件和记录式文件,顺序文件、索引文件和直接文件等;有不同的存取方法,如顺序存取和直接存取。操作系统的文件管理功能(即文件管理系统)对文件进行统一管理和服务,把文件的所有物理结构和物理特性掩盖起来(又称“透明”性)。用户程序只需设计文件的逻辑结构,通过操作系统实现对文件的基本操作,如打开/建立、读写、关闭、复制、删除等。因此,几乎所有程序设计语言都把文件的逻辑结构定义和处理功能纳入其中。

文件管理方式有许多优点。主要是:

(1) 数据可以长久地保存在外部存储器上,避免了同一数据的多次输入。

(2) 数据的逻辑结构和物理结构分开管理。应用程序负责对数据逻辑结构的解读和处理,把数据的物理结构以及应用程序与文件间的数据交换交由操作系统完成,一定程度上提高了程序设计的效率。

(3) 提供了有条件的数据共享能力。也就是说,在程序设计时,只要清楚了解了它所面对文件的逻辑结构就可以使用这些文件。

但是,文件管理方式也存在许多不足。主要是:

(1) 文件与程序的相关性。文件只能由程序创建与读写。任何一个文件都必定依附于一个程序而存在。程序负责设计文件的逻辑结构,又基于文件逻辑结构而处理。

(2) 数据共享能力有限。一个文件的设计很难满足多个应用程序的不同要求。

(3) 文件只是孤立地存在,不能反映出不同数据之间的关联性。结果是,同一数据可能存储在不同的文件中。后果是,可能造成大量数据冗余,发生数据不一致性错误。

假定某学校有3个不同的应用:学生学籍管理、学生电话查询和学生邮政通信。分别为它们设计3个文件:学生信息表文件(数据结构见表1-1)、学生电话号码表文件(数据结构见表1-2)和学生地址表文件(数据结构见表1-3)。显然,在这3个文件中学号和姓名重复存储了3次,是一种数据冗余现象。数据冗余不仅占用了大量的存储空间,也为维护数据的一致性带来难度。

表 1-1 学生信息表文件

学 号	姓 名	性 别	出生年月	班 级
210806101	陈敏敏	女	1990.2	应用 061
210806102	李学好	女	1989.10	应用 061
220806101	顾家新	男	1987.12	经济 061
210806201	黄玲玲	女	1991.1	应用 062
220806201	柯向民	男	1989.3	经济 062
220806101	王怀国	男	1986.7	经济 061
220806202	徐晶晶	女	1988.8	经济 062
210806103	余美美	女	1990.11	应用 061
210806202	张全理	男	1991.9	应用 062

表 1-2 学生电话号码表文件

学 号	姓 名	电 话
210806101	陈敏敏	12345678
210806102	李学好	87654321
220806101	顾家新	11223344
210806201	黄玲玲	56784321
220806201	柯向民	24680721
220806101	王怀国	55728363
220806202	徐晶晶	86261188
210806103	余美美	86868686
210806202	张全理	34598712

表 1-3 学生地址表文件

学 号	姓 名	住 址
210806101	陈敏敏	南京北京西路
210806102	李学好	上海武夷路
220806101	顾家新	济南纬一路
210806201	黄玲玲	南京龙园北路
220806201	柯向民	南京虎踞路
220806101	王怀国	北京平安里
220806202	徐晶晶	天津海河路
210806103	余美美	南京江心洲
210806202	张全理	北京三里河

由上可见,这两种数据管理方式下,程序和数据紧紧联系在一起。如果说文件管理方式下程序和数据之间的关系有所松动的话,则自由管理方式下程序和数据是牢不可破的。这种程序和数据的结构关系称之为“以程序为中心”,数据只处于辅助地位,如图 1-1(a)所示。

3) 数据库管理阶段

数据库方法是在文件基础上的一个重大发展。目的是要消除文件管理方式的缺陷,力图提供数据构造更合理、管理能力更强大、管理方法更完善、数据服务更实用的数据管理技术和方法。数据库管理方法(database management)与文件管理方法有着本质的区别。其主导思想是对数据实行整体性的统一、集中、独立创建和管理,实现从信息角度完整描述和存储一个“组织”(企业、机构、机关、部门等)。以数据集成、统一管理、最小冗余和多层次数据共享为主要特征。例如,把文件管理下的学生信息表文件(见表 1-1),学生

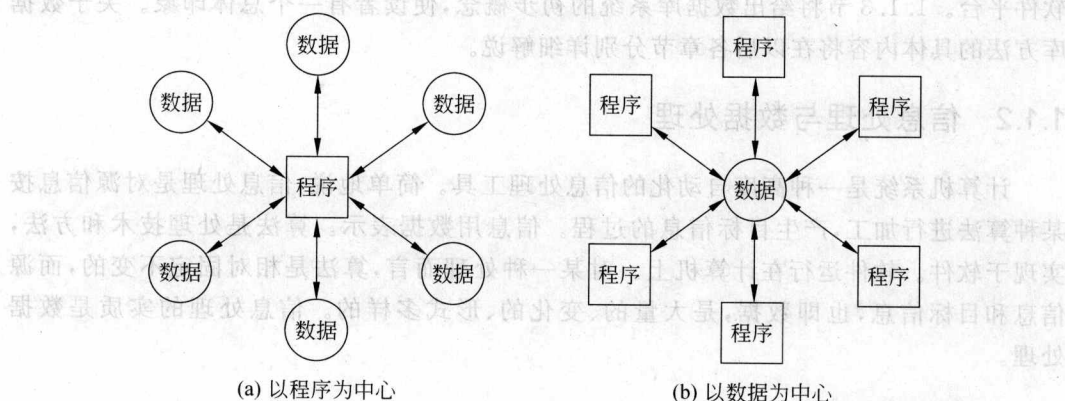


图 1-1 程序与数据结构关系示意图

电话号码表文件(见表 1-2)和学生地址表文件(见表 1-3)3 个文件集成在一个表“学生表”(数据结构见表 1-4)中来管理,显然有许多好处:消除了数据冗余,保证了数据的一致性。这只有在数据库系统方式下才能合理实现。原 3 个文件的数据的使用方法也将不同。

表 1-4 学生表数据库

学 号	姓 名	性 别	出生年月	班 级	住 址	电 话
210806101	陈敏敏	女	1990.2	应用 061	南京北京西路	12345678
210806102	李学好	女	1989.10	应用 061	上海武夷路	87654321
220806101	顾家新	男	1987.12	经济 061	济南纬一路	11223344
210806201	黄玲玲	女	1991.1	应用 062	南京龙园北路	56784321
220806201	柯向民	男	1989.3	经济 062	南京虎踞路	24680721
220806101	王怀国	男	1986.7	经济 061	北京平安里	55728363
220806202	徐晶晶	女	1988.8	经济 062	天津海河路	86261188
210806103	余美美	女	1990.11	应用 061	南京江心洲	86868686
210806202	张全理	男	1991.9	应用 062	北京三里河	34598712

数据库管理方式从根本上改变了程序与数据的关系,把以程序为中心的程序、数据结构关系转变为“以数据为中心”,如图 1-1(b)所示。这意味着,数据不再依赖于程序的存在而存在;可以先于程序而存在。从而,把数据的逻辑结构定义、物理存储结构和维护服务管理等与应用程序设计分开处置,大大提高了应用程序的生产率和数据共享能力。数据上升为主导地位。

数据库方法的出现,一方面是因为海量数据管理的需要;另一方面是计算机技术快速发展提供了条件;特别是外存储器的容量越来越大、种类越来越多、性能越来越好、价格越来越低、可靠性越来越高。第三是因为软件技术的发展为数据库开发提供了优秀的

软件平台。1.1.3 节将给出数据库系统的初步概念,使读者有一个总体印象。关于数据库方法的具体内容将在以后各章节分别详细解说。

1.1.2 信息处理与数据处理

计算机系统是一种高度自动化的信息处理工具。简单地说,信息处理是对源信息按某种算法进行加工,产生目标信息的过程。信息用数据表示。算法是处理技术和方法,实现于软件。软件运行在计算机上。对某一种处理而言,算法是相对固定不变的,而源信息和目标信息,也即数据,是大量的、变化的、形式多样的。信息处理的实质是数据处理。

1. 信息处理

信息处理(information processing)自古有之,只是信息意识、处理方法、可用工具、应用范围和施用效果经历了由初级到高级,由简单到复杂,由落后到先进,由手工到自动的发展过程。今人谈及信息处理时,总是与计算机联系在一起,即计算机信息处理。简单地说,信息处理是“收集→输入→处理→输出→施用”信息的一个不断往复的过程。在这个过程中要使用到计算机软硬件的大多数技术,特别是数据库技术。信息、计算机硬件、软件、应用程序以及相关人员等共同构成的计算机应用系统称为计算机信息系统,简称信息系统,其主要目的是提供信息服务,如商业营销系统、学校教务管理系统、办公自动化系统、情报检索系统、图书检索系统、地理信息系统、电子商务、电子政务等都是典型信息系统的例子。

2. 数据处理

信息系统的基础是数据处理(data processing),信息处理通过数据处理得以实现。因为信息的载体是数据,对信息进行处理就是对数据进行处理。因此,一般情况下把信息处理和数据处理混为一谈,不再加以严格的区分。所谓数据处理是指对数据进行收集、记录、整理、组织、输入、存储、加工、维护、查询、传输、输出等一系列基本操作的总和。数据处理的基本环节是数据管理。

数据处理是计算机的一个庞大应用分支。它具有 4 个显著特点:

(1) 涉及的数据量十分巨大,而一个特定应用处理却只使用其中有限部分。因此,需要使用海量存储器存储数据库中的数据。

(2) 除输入输出数据外,绝大部分数据是持久的,需要长时间地存储在计算机系统中,以提供给一个组织或更大范围内的多个程序共享使用。

(3) 数据与数据之间具有比较复杂的关联关系,需要为其建立相应的结构模型以构造、存储和管理数据。

(4) 对数据的处理算法相对比较简单,侧重于诸如分类、排序、检索、统计、汇总、制表等一类的处理。