



中国科学院研究生教学丛书



化学信息学

邵学广 蔡文生 编著

科学出版社

G252.7
S34

中国科学院研究生教学丛书

化 学 信 息 学

邵学广 蔡文生 编著



A0951767

科 学 出 版 社

2001

内 容 简 介

本书为《中国科学院研究生教学丛书》之一。

化学信息学(chemoinformatics)是化学领域中近几年发展起来的一个新的分支,是建立在多学科基础上的交叉学科,利用计算机技术和计算机网络技术,对化学信息进行表示、管理、分析、模拟和传播,以实现化学信息的提取、转化与共享,揭示化学信息的实质与内在联系,促进化学学科的知识创新。

全书介绍了计算机与计算机网络的一些基础知识及如何利用 Internet 上的化学资源,并对化学信息学在化学、生物化学领域中的应用进行了详细论述。

本书可供化学、化工、生物化学、药物化学及相关专业的广大科技工作者和高等院校师生参阅。

图书在版编目(CIP)数据

化学信息学/邵学广,蔡文生编著. -北京:科学出版社,2001

(中国科学院研究生教学丛书)

ISBN 7-03-009380-1

I. 化… II. ①邵… ②蔡… III. 化学-信息技术 IV. O6

中国版本图书馆 CIP 数据核字(2001)第 21106 号

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

新 蕾 印 刷 厂 印 刷

科学出版社发行 各地新华书店经销

*

2001 年 6 月 第 一 版 开本: 787×1092 1/16

2001 年 6 月第一次印刷 印张: 16 1/2

印数: 1—3 000 字数: 372 000

定价: 30.00 元

(如有印装质量问题,我社负责调换〈新欣〉)

前 言

自从化学学科出现以来,信息的记载、组织与交流对化学学科的发展起了重要的作用,同时也成为化学学科的一个重要组成部分。这是因为化学实验的记录资料具有长远的实践意义。在化学学科中,化学家根据百年以前的记录资料从事科学实验的例子并不希奇。另外,化学物质结构的记录与检索需要建立独特的记录系统。随着计算机技术的发展,化学家必须建立自己的信息表示、记录与管理系统,以适应时代的要求。

化学信息可分为两大组成部分,即化学物质的化学信息和媒体形式的化学信息。前者是利用科学的原理和方法通过测量得到的化学成分的相关信息,如物质的物理、化学性质,物质中各成分的定性、定量以及结构信息,分子间的相互作用信息(包括化学反应信息)等。后者是化学信息的记录形式,如图书、期刊、专利、数据库以及音像资料等,通过化学信息的传播使化学家们共享测量的原理、方法及测量结果。

化学信息学(chemoinformatics)是近几年发展起来的一个新的化学分支,它利用计算机技术和计算机网络技术,对化学信息进行表示、管理、分析、模拟和传播,以实现化学信息的提取、转化与共享,揭示化学信息的实质与内在联系,促进化学学科的知识创新。化学信息学的研究内容主要包括:(1)利用计算机技术和计算机网络技术对化学信息进行表示和计算机管理;(2)利用计算机技术对复杂的化学信息进行解析,以快捷、方便的方式最大限度地提取和利用有用信息;(3)利用计算机对化学信息和化学体系进行模拟;(4)利用计算机和计算机网络技术收集、传播和共享化学信息。

在本书的编写过程中,参考了大量的 Internet 上的资源及有关参考书,在此谨对为本书提供帮助的网站和作者表示衷心的感谢。由于作者水平所限,书中难免存在缺点和错误,敬请读者批评指正。

作 者

2001 年 2 月于中国科学技术大学

第一章 计算机基础

§ 1.1 操作系统

操作系统(operation system, OS), 是指计算机接受指令并执行指令的软件系统, 是计算机系统必备的软件系统。只有硬件而没有操作系统的计算机是无法工作的。它的功能可简单分为两类: ①对计算机进行启动, 包括启动各设备的驱动程序、启动与硬件有关的初始化程序、分配内存、为驱动程序所服务的硬件设置中断等; ②接受指令、解释指令、并执行指令。

操作系统有许多类型, 其中最常用的是 DOS, UNIX, Windows, X-Window。本节将简单介绍 DOS 和 UNIX 的基本命令, 有关 Windows 和 X-Window 系统的使用, 读者可参考系统提供的在线帮助进行学习。

§ 1.1.1 DOS

DOS(disk operation system)是最常用的操作系统之一。它由三个基本程序组成: IO.SYS, MSDOS.SYS 和 COMMAND.COM, 在系统加电时依次读入内存。IO.SYS 由两个独立的模块组成。第一个模块是 BIOS, 其内容是互相连接的一组用于控制台、辅助端口、打印机、数据分块设备的驻留设备驱动程序以及一些在系统启动时运行的与硬件有关的初始化程序。第二个模块称为 SYSINIT, 由 BIOS 调用, 它首先确定计算的内存大小, 然后为 DOS 的核心 MSDOS.SYS 分配内存并使之初始化, 找出设备驱动程序清单并使每个驱动程序初始化, 同时为各驱动程序所服务的硬件设置中断向量。DOS 的核心被初始化以后, 所有驻留驱动程序均可使用。最后, SYSINIT 打开 CONFIG.SYS 文件(如果存在), 此文件包含各种命令, 使用户配置 MS-DOS 的运行环境。COMMAND.COM 为命令解释程序。

DOS 有不同的版本, 目前较流行的版本为 MS-DOS 6.22。MS-DOS 6.22 的全部命令集见表 1-1, 每个命令的含义与使用方法可用 HELP 命令得到帮助, 使用方法为“HELP 命令名”。在 MS-DOS 的命令集中最常用到的命令有以下几个:

1. DIR

显示指定目录的文件和子目录。

命令格式:

DIR [drive:][path][filename][/P][/W][/A[:attributes]][/O[:sortorder]][/S][/B][/L][/C]

其中[]中的内容为任选项, 如果不指定, 计算机将采用省缺值。

2. MKDIR (MD)

创建目录。

表 1-1 DOS 6.22 命令集

ANSI.SYS	EMM386.EXE	Multi-config
Append	Erase	Nlsfunc
Attrib	Exit	Numlock
Batch commands	Expand	Path
Break	Fasthelp	Pause
Buffers	Fastopen	Power
Call	Fc	POWER.EXE
Cd	Fcbs	Print
Chcp	Fdisk	Prompt
Chdir	Files	Qbasic
Chkdsk	Find	RAMDRIVE.SYS
CHKSTATE.SYS	For	Rd
Choice	Format	Rem
Cls	Goto	Ren
Command	Graphics	Rename
CONFIG.SYS commands	Help	Replace
Copy	HIMEM.SYS	Restore
Country	If	Rmdir
Ctty	Include	ScanDisk
Date	Install	Set
Dblspace	Interlnk	Setver
Debug	INTERLNK.EXE	SETVER.EXE
Defrag	International commands	Share
Del	Intersvr	Shell
Deltree	Keyb	Shift
Device	Label	SIZER.EXE
Device drivers	Lastdrive	Smartdrv
Devicehigh	Lh	SMARTDRV.EXE
Dir	Loadfix	Sort
Diskcomp	Loadhigh	Stacks
Diskcopy	Md	Submenu
DISPLAY.SYS	Mem	Subst
Dos	Memmaker	Switches
Doskey	MenuColor	Sys
Dosshell	MenuDefault	Time
DRIVER.SYS	Menuitem	Tree
Drivparm	Mkdir	Type
Drvspace	Mode Commands	Undelete
DrvspaceTips	More	Unformat
DRVSPACE.SYS	Move	Ver
Echo	Msav	Verify
Edit	Msbackup	Vol
EGA.SYS	Mscdex	VSafe
Emm386	Msd	Xcopy

命令格式:

MKDIR [drive:]path

MD [drive:]path

3. CHDIR (CD)

改变当前目录或显示当前目录。

命令格式:

CHDIR [drive:][path]

CD [drive:][path]

CHDIR [..]

CD [..]

4. RMDIR (RD)

删除目录。

命令格式:

RMDIR [drive:]path

RD [drive:]path

备注: 使用此命令删除目录前应首先删除目录中的所有文件和子目录。

5. COPY

文件拷贝。

命令格式:

COPY [/Y|/-Y] [/A|/B] source [/A|/B] [+ source [/A|/B] [+...]][destination [/A|/B]] [/V]

6. XCOPY

目录拷贝(不包括隐含文件和系统文件)。

命令格式:

XCOPY source [destination] [/Y|/-Y] [/A|/M] [/D:date] [/P] [/S] [/E][/V] [/W]

7. MOVE

文件移动。

命令格式:

MOVE [/Y|/-Y] [drive:][path]filename [[drive:][path]filename[...]]destination

备注: MOVE 命令可用于目录名称的更改。

8. RENAME (REN)

更改文件名。

命令格式:

RENAME [drive:][path]filename1 filename2

REN [drive:][path]filename1 filename2

9. TYPE

显示文本文件的内容。

命令格式：

TYPE [drive:][path]filename

10. MORE

逐屏显示文本文件的内容。

命令格式：

MORE < [drive:][path]filename

command-name | MORE

11. DEL (Erase)

删除文件。

命令格式：

DEL [drive:][path]filename [/P]

ERASE [drive:][path]filename [/P]

12. UNDELETE

恢复用 DEL 删除的文件。

命令格式：

UNDELETE [[drive:][path]filename] [/DT/DS/DOS]

UNDELETE [/LIST/ALL/PURGE[drive]]/STATUS/LOAD/UNLOAD/S[drive]
/Tdrive[-entries]]

13. DELTREE

删除目录以及目录内的文件和子目录。

命令格式：

DELTREE [/Y] [drive:]path [[drive:]path[...]]

14. FORMAT

格式化磁盘。

命令格式：

FORMAT drive: [/V[:label]] [/Q] [/U] [/F:size]/[B/S] [/C]

FORMAT drive: [/V[:label]] [/Q] [/U] [/T:tracks /N:sectors] [/B/S]/[C]

FORMAT drive: [/V[:label]] [/Q] [/U] [/1] [/4] [/B/S] [/C]

FORMAT drive: [/Q] [/U] [/1] [/4] [/8] [/B/S] [/C]

15. FDISK

为硬盘分区。

命令格式:

FDISK [/STATUS]

16. UNFORMAT

恢复用 FORMAT 命令格式化的磁盘。

命令格式:

UNFORMAT drive: [/L] [/TEST] [/P]

17. DATE, TIME

显示日期或时间并提示修改(如果需要)。

命令格式:

DATE

DATE [mm-dd-yy]

TIME

TIME [hours:[minutes[:seconds[.hundredths]]][A|P]]

18. EDIT

编辑文本文件。此命令需要 QBasic.EXE 的存在。

命令格式:

EDIT [[drive:][path]filename] [/B] [/G] [/H] [/NOHI]

19. 运行环境的设置

CONFIG.SYS 文件是位于根目录下的含有一系列特殊命令的文本文件, 这些命令设置计算机的硬件以供 MS-DOS 使用。当 MS-DOS 启动时, 首先执行此文件。

CONFIG.SYS 的常用命令有: Buffers, Files, Country, Install, Device, Lastdrive, Devicehigh, Numlock, Dos, Shell, Drivparm, Stacks, Fcbs, Switches(只能用于 config.sys), Break, Rem, Set(可用于 config.sys, 此外也可在 DOS 提示符下使用), Include, Menucolor, Menudefault, Menuitem, Submenu (用于多重设置, multi-configuration)。多重设置的详细内容请用 “help Multi-Config Commands” 获得。

CONFIG.SYS 文件中还可含有 “;” 和 “?” 两个特殊字符, 其中 “;” 指定该行为注释行, 不执行该行命令, 此外也可用 REM 命令指定。 “?” 指定该行的命令执行时要进行确认。它应加在要确认的命令之后, 不加空格。如要在执行 DOS=HIGH 时确认, 它应该加在等号前, 即将 DOS=HIGH 改为 DOS?=HIGH。要想确认 DEVICEHIGH /L:1,7280=MYFILE.SYS, 应改为 DEVICEHIGH? /L:1,7280=MYFILE.SYS。

当计算机系统由于某些原因不需要执行 CONFIG.SYS(或 AUTOEXEC.BAT)时, 可在开机时按 F5 键, 你将直接看到 “Starting MS-DOS ...”。如果只想跳过 CONFIG.SYS(或 AUTOEXEC.BAT)的某些命令, 可在开机时按 F8 键, MS-DOS 将让你对每一个命令确

认。当剩余的命令不需要确认时，可按 ESC 键。相反，如果想跳过所有剩余命令，可按 F5 键。

例如，某 CONFIG.SYS 文件的内容为：

```
[Menu]
menuitem=Steve
menuitem=Lisa

[Common]
dos=high
buffers=15
device=c:\dos\himem.sys

[Steve]
files=20
device=c:\dos\emm386 2048

[Lisa]
files=40
device=c:\net\network.sys
```

20. 批处理文件

批处理文件或批处理程序由一条或多条 DOS 命令和批处理命令(batch commands)构成，扩展名为 BAT。其中根目录下的 AUTOEXEC.BAT 在进入 MS-DOS 系统时自动被执行。

批处理文件中的特殊命令有：Call, If, Choice, Pause, Echo, Rem, For, Shift, Goto 等，它们的功能和用法可用 HELP 命令获得。以下为某 AUTOEXEC.BAT 文件的内容(相应于上述 CONFIG.SYS 文件的内容)：

```
c:\dos\smartdrv.exe
set temp=c:\temp
c:\dos\msav

; Go to the section that matches the current value of the CONFIG variable
goto %config%

:Steve
path=c:\dos;c:\deskspub;c:\typeset
c:\mouse\mouse.com
deskspub
; now skip Lisa's section & go to end
```

```

goto end

:Lisa
path=c:\dos;c:\network;c:\utility
doskey
net logon lisa
goto end

:end

```

§ 1.1.2 UNIX

UNIX 为美国电话电报公司(AT&T)贝尔实验室(Bell Lab.)的肯·汤普森(Ken Thompson)于 1969 年在数字设备公司(Digital Equipment Corporation)的 PDP-7 机器上开发出来的。后来发展出数个流派,主要的流派有三个:AT&T 版本、柏克莱(BSD)版本以及 SCO XENIX(为 Microsoft 公司所开发)版本。此外,各计算机公司也有不同版本,如 SGI 公司的 IRIX 版本、SUN 工作站配备的 SUN OS 版本等。近几年流行的 LINUX 系统也可以说是 UNIX 的另一版本。近来人们正在努力将不同的版本合并构成一个标准的 UNIX 系统。

UNIX 系统也像 DOS 那样由核心(kernel)部分、内部工具程序(build-in utilities)和命令解释器(shell)三部分构成。其核心部分为一连串的软件指令,用于系统和外部设备的监控。UNIX 核心部分的功能可总结为:

- (a)保证所有中断均衡地使用系统。
- (b)正确地安排数据的传送。在多用户同时使用时,保证各用户的数据传送不混乱,多用户同时使用同一数据时不发生错误等。
- (c)以最佳的方式使用内存。

UNIX 具有丰富的内部工具程序,包括 vi, pico, edit, cedit 等编辑器、数据查询排序、备忘服务、绘图、电子邮件以及有关系统、外设、用户等数据的管理软件。

Shell 的功能是查看用户所输入的信息并做相应的处理,如调用核心或内部工具。Shell 有许多种,如 SCO UNIX 提供了 Bourne Shell(sh)和 C Shell(csh)等,IRIX 6.3 中提供了 sh, rsh, csh, tcsh, wsh, xwsh 等,Red Hat LINUX 6.1 也提供了 sh, bash, csh, tcsh, ksh 等。

要想进入一个 UNIX 系统必须有一个合法的账号和相应的密码,它是由系统管理员建立的。一旦有了账号和相应的密码,就可以通过终端或远程登录方式进入系统。打开终端的电源或远程登录连通以后,系统将显示一定的信息,一般包括机器名称、操作系统的信息和系统管理员设置的信息。最后显示 login:_, 在 login:后输入账号名并按回车(Return 或 Enter)键,系统将显示 Password:_, 输入密码后再按回车键即可进入系统。

UNIX 的命令要比 DOS 命令更丰富,要获得有关帮助,可采用 MAN(Manual)命令。如果已知命令名获取使用方法和功能,可使用“man 命令名”。如果不知道命令名,想查找某一类命令,可使用“man -k Keyword”。以下为 UNIX 系统中几个经常用到的

命令:

1. ls

显示目录的详细内容, 同 DOS 的 DIR。

命令格式:

ls [-RadLCxmlnogrtucpFbqisflAM] [names]

[]内为命令的可选项, 由于 UNIX 命令的可选项很多, 在此不一一介绍, 具体功能请使用帮助工具查看。

2. mkdir

创建子目录, 同 DOS 的 MKDIR 和 MD。

命令格式:

mkdir [-m mode] [-p] dirname ...

3. cd (pwd)

改变或显示当前目录。cd 用于改变工作目录, pwd 用于显示当前目录。同 DOS 的 CD。

命令格式:

cd [directory]

4. rm

删除文件, 同 DOS 的 DEL 和 ERASE。

命令格式:

rm [-f] [-i] file ...

rm -r [-f] [-i] dirname ... [file ...]

5. rmdir

删除目录, 同 DOS 的 DELTREE。

命令格式:

rmdir [-p] [-s] dirname ...

备注: 此命令与 RMDIR 和 RD 一样, 要求被删除的目录必须是空目录。删除有文件或子目录的目录时, 使用 “rm -rf” 更方便。

6. cp

文件拷贝, 同 DOS 的 COPY。

命令格式:

cp [-firRp] file1 [file2 ...] target

备注: target 必须指定, 当前目录 “.” 不能省略。

7. mv

文件或目录的移动或更改名称，同 DOS 的 MOVE 和 REN。

命令格式：

```
mv [-if] file1 [file2 ...] target
```

8. ln

文件的关联。

命令格式：

```
ln [-sif] file1 [file2 ...] target
```

9. who

显示当前在线的用户。

命令格式：

```
who [-mu] -s [-bHlprt] [file]
```

```
who [-mTu] [-abdHlprt] [file]
```

```
who -qn x [file]
```

```
who am i
```

```
who am I
```

10. w

显示当前在线的用户或指定的用户以及他们当前的工作内容。

命令格式：

```
w [-fhlsuW] [user]
```

11. du

查看指定目录的磁盘占用情况，显示文件或目录的大小。

命令格式：

```
du [-arklmsxL] [name ...]
```

12. df

查看磁盘占用情况。

命令格式：

```
df [-befiklnPqrtV] [-w fieldwidth] [-F FStype] [filesystem ...]
```

13. passwd

更改密码和密码属性。

命令格式：

```
passwd [name]
```

```
passwd [-l|-d] [-n min] [-f] [-x max] [-w warn] name
```

```
passwd -s [-a]
passwd -s [name]
```

14. date

显示或更改日期和时间，同 DOS 的 DATE 和 TIME。

命令格式：

```
date [-u] [+format]
date [-a [-] sss.fff]
date [-u | -n] [[mmdd]HHMM | mmddHHMM[cc]yy] [.ss]
```

15. more, cat, tail, head, pr

文本文件内容查看。

命令格式：

```
more [-cdfllrsuw] [-lines] [+linenumber] [+pattern] [filename ...]
page [-cdfllrsuw] [-lines] [+linenumber] [+pattern] [filename ...]
cat [-u] [-s] [-v [-t] [-e]] file ...
tail [+number lbcrl] [file]
tail [-lbcrl] [file]
tail [+number lbcfl] [file]
tail [-lbcfl] [file]
head [-count] [file ...]
head [-n count] [file ...]
pr [[-column] [-wwidth] [-a]] [-eck] [-ick] [-drtfp] [+page] [-nck] [-ooffset]
[-llength][-sseparator] [-hheader] [file ...]
pr [[-m] [-wwidth]] [-eck] [-ick] [-drtfp] [+page] [-nck] [-ooffset] [-llength] [-sseparator]
[-hheader] file1 file2 ...
```

16. grep, egrep, fgrep

显示与指定字符串匹配的内容，常用于字符串的查找。

命令格式：

```
grep [-[AB] NUM] [-CEFGVabchiLlnqrsvwxyUu] [-e]pattern [...] filename
备注：此命令常与 more, cat, pr 等命令一起使用，如：more filename | grep pattern。
```

17. chmod, chown, chgrp

更改文件或目录的属性(permissions mode)、所有者和所属的组。

命令格式：

```
chmod [-R] mode file ...
chmod [-R] [ugoa]{+|-|=}[rwxXstl] file ...
chown [-Rfh] owner[{:|}group] file ...
```

chgrp [-Rfh] group file ...

备注：UNIX 系统中文件或目录的属性

当你用“ls -l”(l 表示 long)命令查看文件目录时，会得到如下的结果：

drwxr-xr-x	8	guest	guest	4096	May 5 20:26	guest
lrwxrwxrwx	1	root	sys	14	Aug 12 1997	cat->../sbin/cat
-rw-r--r--	1	xshao	user	6555	Aug 2 16:15	plot.c

文件或目录属性 子目录数 拥有者 小组名 文件大小 最后修改时间 文件名

属性的第一个字符最常见的有：

- d 表示本条目为目录；
- l 表示本条目为关联文件；
- 表示本条目为常规文件。

后面的 9 个字符最常见的有：

- r 表示允许读；
- w 表示允许写；
- x 表示允许执行(必须是可执行文件时才可执行)；
- 表示不允许读、写和执行。

其中前三位表示拥有者的权利，中间三位表示同一小组的权利，后三位表示其他人的权利。

chmod 的选项可用下列字母组合：u(拥有者)，g(小组)，o(其他人)，+(允许)，-(不允许)或用三位二进制数字(0 表示不允许，1 表示允许)所对应的整数表示。例如：

chmod ugo+r+w+x plot.c 或 chmod 777 plot.c 将使 plot.c 的属性变为：

-rwxrwxrwx 1 xshao user 6555 Aug 2 16:15 plot.c

chmod ugo+r+w-x plot.c 或 chmod 666 plot.c 将使 plot.c 的属性变为：

-rw-rw-rw- 1 xshao user 6555 Aug 2 16:15 plot.c

chmod o-r-w-x plot.c 或 chmod 660 plot.c 将使 plot.c 的属性变为：

-rw-rw---- 1 xshao user 6555 Aug 2 16:15 plot.c

18. tar

文件归档。

命令格式：

tar key [arguments] [name ...]

备注：key 的省缺值是磁带机，要使文件在磁盘上归档必须指定 key 为“f”。此命令常用的命令格式为：tar ctf filename.tar filenames 和 tar xtf filename.tar，前者表示将 filenames 指定的文件归档到 filename.tar，后者表示将 filename.tar 中的文件复原，“t”表示在归档和复原的过程中显示文件清单。

19. gzip, gunzip, gzcat, compress, uncompress, zcat

文件的压缩与解压。

命令格式:

```
gzip [-acdhLnNrtvV19] [-S suffix] [name ...]  
gunzip [-acdhLnNrtvV] [-S suffix] [name ...]  
gzcat [-fhLV] [name ...]  
compress [-f] [-v] [-c] [-V] [-d] [-b bits] [name ...]  
uncompress [-f] [-v] [-c] [-V] [name ...]  
zcat [name ...]
```

20. logout (exit)

退出系统。

命令格式:

```
logout  
exit
```

21. UNIX 编辑器

UNIX 提供了许多编辑器,特别是安装了 X-Window 系统的 UNIX 系统,其中最常用的是 pico 和 vi 编辑器,因为它们分别是 e-mail 收发软件省缺的编辑器。关于 pico 和 vi 编辑器的使用方法和常见命令可通过“man 命令”或在线帮助获得。

§ 1.1.3 Windows

Windows 是继 DOS 之后最流行的操作系统之一。它没有复杂的指令,没有复杂的规则,使用者只需选择相应的菜单或图形化的按钮即可完成系统的管理和程序的运行。Windows 系统还具有很强的网络功能,只需简单的设置即可轻松上网。由于 Windows 系统具有丰富的在线帮助功能,关于 Windows 的使用这里不再赘述。

X-Window 是基于 UNIX 系统开发的窗口系统,具有与 Windows 相类似的功能和使用方法。

§ 1.2 计算机语言与程序设计

§ 1.2.1 计算机语言

程序设计语言分为低级语言和高级语言两大类。低级语言是指面向机器的语言,如机器语言和汇编语言;高级语言独立于机器,用高级语言编写的程序在不同类型的机器上要借助于不同的翻译(编译)程序。常用的高级语言有 Basic, Fortran, Pascal, PROLOG, C, C++, Java(J)以及在特定的运行环境下使用的语言(Matlab, Mathematica, S-plus 等)。各种语言在性质、规则、语法等方面均不相同,但它们都由如下四种成分构成:

- (1)数据成分,用以描述程序中所涉及的数据;
- (2)传输成分,用以表达程序中的数据传输;
- (3)运算成分,用以描述程序中所包含的运算;
- (4)控制成分,用以表达程序中的控制结构。

此外,某些语言系统提供了丰富的图形处理部分,用于图形的绘制与显示。这些基本成分构成了任意复杂的程序。

表 1-2 列出了一般编程需要的基本功能与相应的 Fortran, Pascal 和 C 语言语句。从表中可以看出:不同计算机语言具有很大的相似性。因此,在解决化学领域中的一般计算问题时,采用何种语言并不十分重要。

表 1-2 Fortran, Pascal 和 C 语言的常用语句比较

功 能		Fortran	Pascal	C
数 据 类 型	基本类型	character[*n], integer[*n], real[*n], double precision	char, integer, shortint, byte ... real, double, boolean	char, int, int short, unsigned ... float, double
	数组	logical[*n] integer a dimension a(m,n)	var a: array [1..m,1..n] of integer	int a[m][n]
	结构	structure, union, record, map	type, type~=set of, type~=record ... end	Class, struct, union, enum
	指针	无	有	有
运 算 符	算术运算	+, -, *, /, mod	+, -, *, /, div, mod	+, -, *, /, %, mod
	位运算	iand, ior, not, ishft,	and, or, xor, shl, shr	&, , ^, ~, >>, <<
	逻辑运算	.and., .or., .not.	and, or, not	&&, , !
	关系运算	.gt., .ge., .lt., .le., .ne., .eq.	>, >=, <, <=, <>, =, in	>, >=, <, <=, !=, ==
输入 与 输出	输入	read	read, readln	scanf, fscanf, fread
	输出	write, print	write, writeln	printf, fprintf, fwrite
	打开文件	open	assign	fopen
	关闭文件	close	close	fclose
循环 与 分支	循环	for, do, do~while	for, while~do, repeat~until	for, while, do~while,
	分支	if, if~then, if~then~else, if~then~else if~then~else, goto, if~goto, goto (1,2) a, select case ~ end select	if~then, if~then~else, if~then~else if~then~else, goto, if~goto, case~of~else	if, if~else, if~else if~else, goto, if~goto, switch~case
过程 与 函数	定义	subroutine abc(a,b,c) real function abc(a,b,c)	procedure abc(a,b,c) function abc(a,b,c):real	float abc(a,b,c)
	调用	call abc(x,y,z) f=abc(x,y,z)	abc(x,y,z); f:=abc(x,y,z);	f=abc(x,y,z);