

# UNIX 分时 系统程序员手册

(第二卷)

《计算机研究与发展》编辑部

# UNIX 分时 系统程序员手册

第二卷

《计算机研究与发展》编辑部

# 目 录

## 一般著作

1. UNIX第七版概述 ..... ( 1 )  
UNIX系统各种功能的简要介绍。
2. UNIX分时系统 ..... ( 15 )  
最原始的UNIX论文,根据CACM上的稿子重印。

## 开始

3. UNIX初学者入门——第二版 ..... ( 28 )  
对该系统最基本的用法作介绍。
4. UNIX正文编辑程序的辅导介绍 ..... ( 45 )  
用编辑程序开始工作的方便途径。
5. UNIX高级编辑 ..... ( 57 )  
编辑程序的进一步工作。
6. UNIX shell介绍 ..... ( 76 )  
有关命令解释程序shell功能的介绍。
7. LEARN—UNIX的计算机辅助教学系统第二版 ..... ( 91 )  
介绍一种计算机辅助教学程序,它帮助新用户掌握浏览文件、编辑程序和资料准备软件基本知识。

## 资料准备

8. 在UNIX系统上的文献打印 ..... (103)  
介绍格式加工工具的基本用途。也介绍了可以用来加工大多数资料(包括在本卷中的那些资料)所需要的格式加工标准程序包,“-ms”。
9. 一个数学排版系统 ..... (110)  
介绍一种完成高质量数学排版工作,并且容易学会的语言EQN。附文《数学排版用户指南》
10. tbl—表格编排程序 ..... (131)  
可以方便地说明用于排版的表格材料的一种程序,这种程序易于学习和使用。
11. UNIX系统上倒排索引的一些应用 ..... (138)  
介绍了程序REFER,这个程序根据数据库自动地填写文献目录引文。  
附文《更新出版目录》
- 12(1). NROFF/TROFF用户手册 ..... (155)  
基本的格式加工程序。

- 12(2). 指导性的例子 ..... (187)
13. TROFF指导 ..... (194)
- 为那些真正要知道这些事情的人介绍TROFF。

## 程序设计

14. C程序设计语言参考手册 ..... (209)
- 关于C语言的语法和语义的正式报告。应该补充以B.W.Kernighan和D.M.Ritchie的《C程序设计语言》一书,该书中有辅导性的介绍和许多例子。
15. C程序的检查程序Lint ..... (234)
- 检查C程序的语法错误、类型不匹配、可移植性问题和大量可能有的错误。
16. make——一个维护计算机程序的程序 ..... (242)
- 为了确保用最小的努力就能正确地编译大型程序而需要的必不可少的工具。
17. UNIX程序设计——第二版 ..... (250)
- 介绍对于操作系统和标准I/O库的程序设计接口。
18. ADB系统概述 ..... (266)
- 怎样使用ADB调试程序。

## 支持工具和语言

19. YACC: 又一个编译程序的编译程序 ..... (283)
- 把一种语言的BNF形式说明和用C写的语义动作转换成该语言的一种编译程序。
20. Lex——一个词法分析程序的生成程序 ..... (306)
- 为一组正则表达式创建一种识别程序; 每个正则表达式后面可以跟随任意的C代码, 这些代码将在找到这个正则表达式时执行。
21. 一个可移植的Fortran77编译程序 ..... (323)
- 第一个Fortran77的编译程序, 并且至今仍然是最好的编译之一。
22. RATFOR——一个合理的Fortran的预处理程序 ..... (337)
- 把带有类似C控制结构和外形的Fortran转换成真正的、粗陋的Fortran。
23. M4宏处理程序 ..... (350)
- M4是一个用于C、RATFOR、COBOL以及它有权决定的其它语言的前端程序的宏处理程序。
24. SED——一个非交互式正文编辑程序 ..... (356)
- 处理大量输入的编辑程序的变种。
25. AWK——一个模式扫描和处理语言第二版 ..... (365)
- 使得许多数据变换和选择操作容易规定些。
26. DC——交互台式计算器 ..... (373)
- 如果你不需要浮点运算的话, 那么这是一台超级HP计算器。
27. BC——任意精度的台式计算机语言 ..... (379)
- 提供中级表示、控制流和内部函数的DC的前端接口。

|                               |       |
|-------------------------------|-------|
| <b>28. UNIX汇编程序参考手册</b> ..... | (388) |
| 最终要死亡的语言。                     |       |

## 实现、维护及其它

|                                           |       |
|-------------------------------------------|-------|
| <b>29. 启动UNIX系统——第七版本</b> .....           | (396) |
| 怎样配置并且让你的系统运行。                            |       |
| <b>30. 重新生成系统软件</b> .....                 | (404) |
| 当你必须改变系统时要做的事情。                           |       |
| <b>31. UNIX系统的实现</b> .....                | (410) |
| 系统在内部是怎样实际工作的。                            |       |
| <b>32. UNIX I/O系统</b> .....               | (421) |
| I/O系统实际是怎样工作的。                            |       |
| <b>33. UNIX 系统C语言编译程序介绍</b> .....         | (428) |
| PDP-11编译程序在内部是怎样工作的。                      |       |
| <b>34. 可移植C语言编译程序综述</b> .....             | (440) |
| 可移植的C编译程序在内部是怎样工作的。                       |       |
| <b>35. UNIX<sup>TM</sup>系统的拨号网络</b> ..... | (463) |
| 介绍UNIX系统间交换文件的程序——UUCP。                   |       |
| <b>36. UUCP实现介绍</b> .....                 | (469) |
| UUCP是怎样工作的，怎样管理它。                         |       |
| <b>37. 论UNIX操作系统的安全性</b> .....            | (480) |
| 怎样破坏UNIX的一种暗示，怎样避免这样做。                    |       |
| <b>38. 口令保护：历史回顾</b> .....                | (483) |
| 不诚实的家伙过去怎样能破坏口令算法，现在为什么不行了，至少不那么容易了。      |       |

# UNIX 第七版概述

贝尔实验室

## A. UNIX系统第七版的主要特点

针对更大的系统

设备可编址到 $2^{31}$ 字节, 文件到 $2^{30}$ 字节。某些实用程序需要128K内存(指令和地址空间分离)。

可移植性

操作系统和绝大多数实用程序的代码已彻底进行了修改, 以尽量减少对于特定机器硬件的依赖。

Fortran77

新的标准语言的F77编译程序在目标级上与C兼容。一个Fortran构造程序STRUCT把老的、难看的Fortran变换成RATFOR, 这是一个便于和F77一起使用的结构化的变种。

shell

全新的SH程序提供字符串变量、捕捉处理、结构化的程序设计、用户统计、可设置的检索路径、多级文件名生成等等功能。

资料准备

TROFF照排机实用程序是标准的。NROFF(用于终端)目前完全与TROFF兼容。MS宏程序包为通常的格式加工和打印格式布局提供一致的命令。TBL表格材料提供了易学的语言。REFER填写来自数据库的文献引述。

UNIX到UNIX文件复制

它在任意两台机器间进行脱机的文件传送。

数据处理

SED数据流编辑程序平行地完成长度不加限制的数据流多重编辑功能, AWK报告生成程序完成自由字段模式选择和算术运算。

程序开发

MAKE控制复杂软件的重建, 使得重新编译的工作尽可能小。

调试

ADB进行事后检查和断点调试, 管理分离的指令和数据空间、浮点运算等等。

C语言

目前该语言提供可明确定义的数据类型, 一般性初始化, 分程序结构, 长整数, 联合类型性, 显式类型转换。LINT检验程序进行严格的类型检验, 查出可能的错误和可移植性问题, 甚至跨越了分开编译的函数。

词法分析生成程序

LEX把正则表达式和语义动作的形式说明转换成可识别的子程序。类似于YACC。

图形学

适于各种设备的简单的绘图实用程序、制图子程序和一般性的绘图过滤程序现在都标准化了。

标准的I/O程序包

极其有效的缓冲字符流I/O与格式化的I/O结合在一起。

其它

操作系统和实用程序得到了增强, 并且在许多方面的限制有所放松。因数量太多这里就不一一列举了。

## B. 硬件

第七版UNIX操作系统运行在DEC PDP11/45或11/70上, 它至少具有如下设备:

128K到2M字的可控制存储; 不用奇偶校验。

磁盘: RP03, RP04, RP06, RK05 (一个以上) 或类似设备。

控制台打字机。

时钟: KW11-L或KW11-P型。

也很希望有下列设备:

诸如DL11或DH11那样的通信控制器。

全双工96字符ASCII码终端。

用作系统后援的9磁道磁带或附加磁盘。

系统通常在9磁道带上。所指定的最小内存存储或盘空间足够运行和维护UNIX。更多的用来联机维持所有资源,或处理大量用户、大型数据库、多样化的设备配件或大程序。常驻代码占12—20K字,这取决于配置;系统数据占据10—28K字。

还没有承诺在PDP-11/34,11/40和11/60硬件上提供第七版UNIX。

## 一、基本软件

基本软件包括分时操作系统连同实用程序,一个机器语言汇编程序以及程序设计语言C的编译程序,这些软件足够编写和运行新的应用软件以及维护或修改UNIX系统本身。

### 1.1 操作系统 UNIX

这是基本的驻留代码,所有别的一切都依赖于它。它提供系统调用并维护文件系统。UNIX系统设计思想和系统设施的概括说明见1974年7月的“Communications of ACM”。更广泛的介绍在1978年7—8月的“Bell System Technical Journal”中。包括如下一些性能:

- 可重入的用户进程代码
- 指令和数据空间分离
- 由于是同一组,因而有用于合作项目的“group”存取权限
- 闹钟延时
- 用于调试和测量的时钟中断采样和进程间的监督
- 用于机器对机器通信的多路I/O。

## C、软件

大多数可用作UNIX命令的程序都罗列出来了。除了游戏以外,所有列出的软件的源代码和打印的手册都已发行。几乎所有的代码都是用C写的。命令是一个闭体,不需要额外收集信息,除非特别标明如“交互的”。通过重新定向输入就能从一个简单预备的正文中构造一个交互程序且运行之。大多数倾向交互使用(例如,编辑程序)的程序允许转义成命令级(shell)。大多数文件处理命令也能完成从标准输入到标准输出(“过滤程序”)的工作。shell的管道线设施可以用来直接把这样的过滤程序连接到别的程序的输入或输出上。

### DEVICES (设备)

所有I/O逻辑上都是同步的。I/O设备就是文件系统中的文件。通常,不可见的缓冲处理使所有物理记录结构和设备特点变得透明,并且利用硬件的能力重叠I/O。独特的应用可用无缓冲的物理记录I/O。可以得到这些设备的驱动程序;其它的也很容易编写。这些设备是

- 异步接口: DH11, DL11。支持极普遍的ASCII终端。
- 同步接口: DP11。
- 自动的呼叫设备接口: DN11。
- 行式打印机: LP11。
- 磁带: TU10和TU16。
- DEC带: TC11。
- 固定头磁盘: RS11, RS03和RS04。
- 密封式磁盘: RP03, RP04, RP06; 进行最小寻找等待时间调度。
- 盒式磁盘: RK05, 每个逻辑设备有一个或多个物理设备。
- 空设备。
- PDP-11的物理存储器,或在驻留系统

中的映象存储器。

- 照相制版机：经由DR11C的Graphic Systems System/1。

BOOT

使UNIX系统开始工作的过程。

MKCONF

裁剪与设备相关的系统代码以适合硬件配置。如新发行的那样，UNIX系统能直接在具有任何可接受的盘、足够容量的内存、两种时钟中的一种设备的任何可接受的CPU上生成。然后，其它改变，诸如最佳地分配目录给设备。包含浮点模拟器或在文件系统中安置设备名等，可在空闲时做。

## 1.2 用户存取控制

LOGIN

作为一个新用户签到。

- 核实口令并建立用户个人和(工程计划)组的标识。

- 适应终端特性。
- 建立工作目录。
- 通知邮件（来自于MAIL）的存在。
- 公布日期信息。
- 执行用户指定的直方图（统计）。
- 启动命令解释程序或其它初启程序。

PASSWD

改变一个口令。

- 用户能改变它私用的口令。
- 为安全起见，口令经过编码加密。

NEWGRP

改变工作用户组（计划）。防止对计划作未经授权的变化。

## 1.3 终端管理

TABS

为特定类型的终端设置合适的制表限位。

STTY

为终端的最佳控制设置任选。就它们可由输入推断而言，这些任选是由LOGIN自动设置的。

- 半双工或全双工
- 回车 + 行馈给或换行
- 制表符解释
- 奇偶性
- 大写到低写的转换
- 不作加工的和编辑好的输入
- 对于制表、换行和回车的延迟。

## 1.4 文件管理

CAT

把一个或多个文件串接到标准输出上。专门用于不作加工的打印，把数据插入管道线以及缓冲一点一点来的输出。可以工作在任何文件上而不管其内容。

CP

把一个文件复制到另一个文件中，或者把一组文件复制到一个目录中。可以工作在任何文件上而不管其内容。

PR

打印文件并且附有标题、日期以及每一页的页号。

- 多列输出。
- 几个文件的平行列汇集。

LPR

脱机打印。Spool（脱机）任意文件到行式打印机上。

CMP

比较两个文件，若有不同则要报告。

TAIL

打印输入的最后n行。

- 可以打印最后n个字符，或从n行或字符到末尾。

SPLIT

把一个大文件分裂成多个可管理的部分。偶而为编辑（ED）所需要。

DD

用于与另外的系统，尤其是IBM370，交换数据的物理文件格式翻译程序。

SUM

总计一个文件的所有字数。

## 1.5 目录和文件名的管理

RM

删除一个文件。若有其它名字与该文件联结，那末仅仅去除名字。

- 一步步地通过目录，交互地删除文件。
- 删除整个目录层次结构。

LN

“联结”另一个名字（别名）到一个已存文件。

MV

移动一个或多个文件。用于重新命名文件。

CHMOD

改变一个或多个文件的访问权限。可由文件主执行。

CHOWN

改变一个或多个文件的属主。

CHGRP

改变一个文件的同组用户号。

MKDIR

构造一个新目录。

RMDIR

删除一个目录。

CD

改变工作目录。

FIND

检索目录层次结构，寻找所有符合给定条件的文件。

• 条件包括下列一些：名字与一给定模式匹配，在一给定范围内创建日期，在一给定范围内最后使用的日期，给定权限，给定属主，给定的特别文件特征，上述条件的布尔组合。

- 任何一个目录皆可看成是根。
- 在每一个找到的文件上执行指定的命令。

## 1.6 运行程序

SH

shell或命令语言解释程序。

• 为所有可执行程序提供自变量且运行这些可执行程序。

• 重新定向标准输出、标准输入和标准错误文件。

• 管道：同时执行与另一个进程的输入相连的一个进程的输出。

• 利用下列控制手段组合复合命令：

if...then...else； case开关； while循环； 遍及一列表达式的for循环； break，continue和exit； 以括号分组。

• 初启后台进程。

• 完成shell程序，即具有可替换自变量的命令正文。

• 构造来自所有满足特定模式的文件名的自变量清单。

• 对于捕俘和中断采取特别动作。

• 为了寻找命令提供用户可设置的检索路径。

• 执行用户可设置的对于注册的直方图统计。

• 在邮件到达时有选择地告知邮件情况。

• 提供具有缺省约定的变量和参数。

TEST

为shell条件中的使用作检验。

• 字符串比较。

• 文件特性和可存取性。

• 上述条件的布尔组合。

EXPR

为核算命令自变量而作字符串计算。

• 整型运算。

• 模式匹配。

WAIT

等待异步运行进程的终结。

READ

为交互的shell过程从终端上读一行。

ECHO

打印命令行的剩余部分。对于shell程序中的诊断或提示符，或对于在一个管道线中插入数据十分有用。

## SLEEP

中止执行一给定的时间。

## NOHUP

运行一个免受挂断该终端影响的命令。

## NICE

在低（或高）优先数上运行一个命令。

## KILL

终止命名的进程。

## CRON

按指定的时间调度正常的动作。

- 动作是任意的程序。

- 时间是月，月中的天，星期内的天，小时和分的合取。可以为每一种单独指定范围。

## AT

调度一个运行任意长时间的一步动作。

## TEE

在进程间传递数据并且把一个副本转换给一个或多个文件。

### 1.7 状态询问

## LS

列出在一个或多个目录中的一个，几个或所有文件的名字。

- 按字母或时间顺序，递增或递减排序。

- 任选信息：大小、属主、伙伴、最近修改时间、最近访问时间、权限、索引节点号。

## FILE

通过查阅系统文件卷的索引及读取文件本身来决定文件中是何种信息。

## DATE

打印今日的日期和时间。具有历法和钟表特性的大量知识。

- 可以设置UNIX系统关于日期和时间的打算。

## DF

报告系统文件卷设备上的自由空间量。

## DU

打印在一个层次结构中所有文件所占据

的全部空间总和。

## QUOT

打印由用户ID使用的文件空间使用总和。

## WHO

告知谁在系统里。

- 当前注册的用户、端口和次数的清单。
- 所有注册和注销的任选的历史。

## PS

报告正在活动的进程。

- 列出你本人的或每一个人的进程。
- 告知哪些命令正在执行。
- 任选的状态信息：状态和调度信息优先数、附着的终端、正在等待的东西、大小。

## IOSTAT

打印有关系统I/O活动的统计信息。

## TTY

打印你的终端名字。

## PWD

打印你的工作目录名。

### 1.8 后援和维护

## MOUNT

把一个包含一个系统文件卷的设备安装到目录树上。防范无意义的安装。

## UMOUNT

从目录树上摘除包含在一个设备上的系统文件卷。防范摘除一个正在工作的设备。

## MKFS

在一个设备上构造一个新的系统文件卷。

## MKNOD

为一个特别文件构造一个i节点（系统文件卷登记项）。特别文件可以是物理设备、虚拟设备、物理存储器等等。

## TP, TAR

管理在磁带或DEC带上的文件档案。

TAR更新一些。

- 把文件收集到一个档案中。
- 按日更新DEC带档案。

- 替换或删除DEC带上的文件。
- 打印目录一览表。
- 从档案中检索（有关文件）。

## DUMP

按日选择或不加区分地转储储存在一个特定设备上的系统文件卷。

## RESTOR

恢复一个被转储的系统文件卷，或者有选择地从那里检索一些部分。

## SU

临时变成一个具有所有权力的特权用户，因而具有一切特权。要求一个口令。

## DCHECK, ICHECK, NCHECK

检验系统文件卷的一致性。

- 打印全部统计数字：文件数、目录数、特别文件数、已用的空间、空闲空间。
- 报告空间的重复使用。
- 恢复损耗的空间。
- 报告不可访问的文件。
- 检验目录的一致性。
- 列出所有文件名的清单。

## CLRI

从一个系统文件卷强制清除一个文件和它的空间。用于修补损坏的系统文件卷。

## SYNC

强制完成系统中未完成的I/O。用于适当地关闭（系统文件卷）。

## 1.9 记帐

报告所用的时间信息可以全部用手工清除或关闭。

## AC

发布累积的连接时间报告。

- 按用户或按天的连接时间
- 对于所有用户或选出的某些用户。

## SA

发布shell记帐报告。在每一个命令执行时给出用途信息。

- 使用的次数。
- 总的系统时间，用户时间和流失的时间。

间。

- 任选的平均值和百分比。
- 各种字段上的分类。

## 1.10 通信

### MAIL

邮寄消息给一个或多个用户。也用于读取和处理外来的邮件。邮件的出现由LOGIN和SH最恰当地公布。

- 每一种消息能一个个地单独处理。
- 消息能保留在文件中或转发。

### CALENDAR

为今日和明日的事件提供自动提醒服务。

### WRITE

建立与另一个用户的直接终端通信。

### WALL

写给所有用户。

### MESG

禁止接收来自WRITE和WALL的消息。

### CU

挂通另一个分时系统。

- 对远程机器的透明接口。
- 文件传送。
- 从本地文件中取出向远程的输入或者把远程的输出放进本地文件。
- 远程系统不一定要是UNIX系统。

### UUCP

UNIX系统到UNIX系统的复制。

- 自动排队直至传输线可用及远程机器可以传送为止。
- 在两个远程机器上复制。
- 在两个机器间的差异、邮件等。

## 1.11 基本的程序开发工具

这些实用程序中的某些用作在下一节要讲述的更高级语言的组成部分。

### AR

维护档案和程序库。考虑内务处理的效率，把几个文件组合成一个。

- 建立新档案。
- 每日更新一次档案。
- 替换或删除文件。
- 打印内容（目录）表。
- 从档案中恢复。

## AS

汇编程序。类似于PAL-11，但在细节上不同。

- 建立由以下一些内容组成的目标程序：

代码、可能只读、初始化的数据或读-写代码、不可初置的数据。

• 可重新定位的目标代码是直接可执行的而不必进一步转换。

- 目标代码通常包括一张符号表。
- 多重源文件。
- 局部标号。
- 条件汇编。

• “条件转”指令变成分支或分支加转移，这取决于距离。

## Library

这是基本的运行程序库。所有软件可以随便使用这些子程序。

- 经缓冲的一个一个字符的I/O
- 对于标准输入和输出、文件、存储器内部转换的格式化的输入和输出转换(SCANF和PRINTF)
- 存储分配程序。
- 时间转换。
- 口令加密。
- 快速分类。
- 随机数生成程序。
- 数学函数程序库，包括三角函数和反函数、幂、对数、平方根、贝塞尔函数。

## ADB

交互式调试程序。

- 事后转储。
- 检验任意文件，大小不限。
- 调试程序作为一个独立进程，进行交互断点调试。

- 局部和全局变量的符号参照。

- C程序的栈跟踪。

• 输出格式：八进制、十进制或十六进制1,2或4个字节的整数；单精度和双精度浮点数；字符或字符串；反汇编的机器指令

- 修补

- 检索整数、字符或浮点数。

- 处理分离的指令和数据空间。

## OD

转储任意文件。输出选择包括八进制或十进制按字，八进制按字节，ASCII，操作码，十六进制的任意组合。

- 转储范围可控制。

## LD

连接编辑。组合可重定位的目标文件。插入来自指定程序库的所要求的子程序。

- 形成的代码可以是可分享的。

• 形成的代码可以有分离的指令和数据空间。

## LOADER

装入时，按合适的次序放置目标文件名，以便依赖于其它文件的文件在它们之后进入。

## NM

打印一个目标程序的名字表(符号表)。对于被打印的格式和名字次序加以控制。

## SIZE

报告一个或多个目标文件的内存需求量。

## STRIP

从一个目标文件中除去再定位和符号表信息以节省空间。

## TIME

运行一个命令且报告在上面花费的时间信息。

## PROF

根据由时间采样执行程序所收集的统计数字，构造每个程序花费时间的直方图。使用浮点。

- 子程序调用C程序的频度和平均次数。

## MAKE

控制大程序的建立。使用一个说明源文件相关性的控制文件来构造新的版本；利用最近改变的时间以尽量减少所需的工作量。

- 了解CC, YACC, LEX, 等等。

### 1.12 UNIX系统程序员手册

#### Manual

UNIX分时系统程序员手册的机器可读的文本。

- 系统概貌。
- 所有命令。
- 所有系统调用。
- C中所有的子程序以及汇编程序库。

## 二、语言

### 2.1 C语言

#### CC

C语言的编译与/或连接编辑程序。UNIX操作系统，绝大多数子系统和C本身都是用C编写的。要知C的全面描述，请读The C Programming Language, (由B. W. Kernighan 和 D. M. Ritchie 编著, Prentice-Hall, 1978年出版)。

- 为结构化程序设计而设计的通用语言。
- 数据类型包括字符、整数、浮点数，双精度浮点数、指向所有类型的指针，由上述类型返回的函数、所有类型的数组、所有类型的结构和联合。
- 打算对全部机器设施进行独立于机器的控制的操作，包括针对存储的操作和指针运算。
- 作为参数化的代码和标准文件包含的宏预处理程序。
- 所有过程可递归，带有值参。
- 独立于机器的指针管理。
- 目标代码使用PDP-11的全部寻址功能。
- 运行时程序库访问所有系统设施。
- 可定义的数据类型。

- 所有的设备和其它特别文件。
- 系统文件卷的格式和系统软件所知的文件种类。
- 引导和维护的过程。

#### MAN

在你的终端上打印指定的手册的章节。

### 1.13 计算机辅助教学 (CAI)

#### LEARN

解释CAI正文的程序，再加上学习UNIX系统的正文。

- 这是基本文件和命令，编辑程序，高级的命令和文件，EQN, MS宏, C程序设计语言的正文。

- 分程序结构。

#### LINT

C程序的检验程序。报告诸如下述那样的有疑问的或不可移植的用法：不匹配的数据说明和过程接口；不可移植的类型转换；非法的变量，执行不到的代码，无效操作；类型不配的指针；废弃的文法。

- 分别编译的程序的充分交互式检验。

#### CB

C程序更漂亮的形式。完成合适的缩进和括号设置。

### 2.2 FORTRAN

#### F77

ANSI标准FORTRAN77的完整编译程序。

- 在目标级上与C及支援工具兼容。
- 与FORTRAN66在源代码上有选择的兼容。
- 自由格式源代码。
- 任选的下标界检验，检测未初始化的变量。
- 所有宽度的运算：2和4字节整数，4和8字节实数；8和16字节复数。

#### RATFOR

Ratfor把类似于C的那样的合理控制结构加到FORTRAN上。

- 复合语句
- if-else, do, for, while, repeat-until, break, next语句。
- 符号常量。
- 文件插入。
- 自由格式源代码。
- 象 $>$ ,  $>=$ 那样的关系的翻译。
- 产生纯正的FORTRAN以便运行。
- 可以与F77一起使用。

## STRUCT

利用成组语句, if-else, while, for, repeat-until把普通而难看的FORTRAN转换成结构FORTRAN (即, Ratfor)。

## 2.3 其它算法语言

### BAS

在风格上类似于BASIC的交互型解释程序。未编号的语句立即解释, 编号的语句在“运行”中解释。

- 语句包括注释、转储、for...next、goto、if...else...fi、列表、打印、提示、返回、运行、保留。
- 所有计算都是双精度
- 递归函数定义和调用。
- 内部函数包括log, exp, sin, cos, atn, int, sqr, abs, rnd。
- 对于复杂的程序编辑转义成ED。

### DC

交互式可编程台式计算装置。存放整数或程序的既有命名的存储单元, 又有传统的栈。

- 精度不作限制的十进制运算。
- 适当对待十进制小数。
- 任意的输入和输出基数, 尤其是二进制、八进制、十进制和十六进制。
- 逆波兰运算符有+, -, \*, /, 余数,

幂, 平方根, 取, 存, 复写, 清除, 打印, 进入程序正文, 执行。

### BC

一种类似于C的、与台式计算装置DC的交互式接口。

- 具有高级文法的DC的所有功能。
- 数组和递归函数。
- 立即计算表达式。而函数在调用时计算。
- 任意精度的基本函数: exp, sin, cos, atan。
- 没有go-to的程序设计。

## 2.4 宏处理程序

### M4

通用宏处理程序。

- 面向字符流; 在正文的任何地方识别宏。
- 文法适应大多数更高级语言的功能文法。
- 能计算整数运算表达式。

## 2.5 编译程序的编译程序

### YACC

一种基于LR(1)的编译程序编写系统。在执行生成的语法分析程序期间, 可以调用任意C函数以完成代码生成或语义动作。

- BNF语法形式说明。
- 优先关系。
- 正式接收具有非BNF分解规则的二义性文法。

### LEX

词法分析程序的生成程序。在隔离每一个词法记号时可以调用任意的C函数。

- 全正则表达式, 外加左和右上下文相关性。
- 生成的词法分析程序与YACC语法分析程序接口清晰。

## 三、文 本 处 理

### 3.1 资料准备

#### ED

交互型上下文编辑程序。随机存取一个文件的所有行。

- 经由行号或模式寻找所需行。模式包括：指定的字符、不计较字符、在字符中选择、上述结构的重复、行首、行尾。

- 增加、删除、改变、复制、移动或加入行。

- 置换或分割一行的内容。

- 替换一行中一个模式的一个或所有呈现。

- 组合或分割文件。

- 在编辑过程中转义成shell（命令语言）。

- 在一个给定范围内每一种受选模式的行上执行上述任意一种操作。

- 为了绝对安全起见有任选的加密措施。

#### PTX

完成置换（上下文关键字）索引。

#### SPELL

通过把一份资料中的每一个字与一张字清单进行比较，发现拼写错误。

- 25000个单词清单包括固有名字。

- 处理通常的前缀和后缀。

- 收集单词以帮助编选局部拼写清单。

#### LOOK

在目录中寻找以指定前缀开始的词。

#### TYPO

通过统计技术发现拼写错误，不限于英文。

#### CRYPT

为安全起见加密和解密文件。

### 3.2 资料格式加工

#### ROFF

终端排字程序。非技术人员很容易学会，对于简单资料是很好的。输入由诸如

- sp2 插入两空行

- ce 下面一行居中

那样的和控制行混杂的数据行组成。ROFF被认为已过时了；只打算作非正规使用。

- 一边对齐或两边都对齐。

- 自动用连字符连接

- 一般化的、每页连续出现的头和尾，带有奇偶编页功能，计数等等。

- 对于频繁使用的控制序列（不可替换自变量）可定义宏。

- 可动态调整所有四边及页的大小。

- 斜坡式缩进及一直线式缩进。

- 绝对和相对的参数设置。

- 输出行任选的正当格式计数。

- 多重文件能力。

- 不用作为过滤程序。

#### TROFF, NROFF

高级排版程序。TROFF驱动一种Graphic Systems照排机；NROFF驱动所有类型的ASCII终端。TROFF和NROFF的风格类似于ROFF，但在适当编程后，TROFF和NROFF有能力进行精致得多的格式加工。TROFF和NROFF接受同样的输入语言。

- 所有ROFF的功能可得到或可定义。

- 彻底地可定义页格式，关键是在指定行中动态地设置“中断”。

- 维护几种独立的可定义排字环境（例如，一种用于正文，一种用于脚注，而一种用于异常精致的标题）。

- 可以随意组合任意数量的输入池。

- 具有可替换自变量的宏，以及能在行中引用的宏。

- 计算且打印数字量。

- 有条件地执行宏。

- 表格式加工功能。

- 可以英寸、厘米、全身、点、机器单

位、或上述的算术组合来表达位置。

- 对于异常困难的格式加工问题访问字符宽度计算。

- 重复打击、组合括号、水平和垂直画线。

- 动态地相对或绝对定位和大小选择，整体或字符级。

- 对于相似的特别字符、反向运动、均衡的留空等等，能利用所用终端的特性。

Graphic Systems排字机有几种102个字符、15种大小的铅字（同时有四副）的字母表。TROFF对于产品的粗略试样提供终端输出。

NROFF在具有反向行馈给的终端上产生多列输出，或传送给后处理程序COL。

虽然不熟练的人员可以很容易地训练成根据下述MS提供的那些加工过的格式开始从事资料工作，但是还是需要高度的程序设计技巧以利用TROFF和NROFF的格式加工能力。TROFF和EQN本质上与NROFF和NEQN相同，因而通常有可能定义可互换的格式，以便在实际排字之前在终端上产生近似的样张。预处理程序MS，TBL和REFER与TROFF和NROFF完全兼容。

MS

作为与NROFF/TROFF一起使用的标准化的原稿格式加工程序包。本资料采用MS进行格式加工。

- 页数和手稿日期。
- 自动计数小标题。
- 脚注。
- 单列或双列。
- 分段、显示及缩进。
- 给方程加编号。

#### EQN

TROFF的一种数学排版预处理程序。把成簇或显示的、容易阅读的公式翻译成详尽的排版指令。公式可以按下列格式书写：

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

$i=1$  to  $N$  ( $x_{sub i} - \bar{x}$ )  $sup 2$ ,  
它产生：

$$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2$$

- 自动估计下标、小下标等等的大小变化。

- 希腊字母和特别符号的全字母表，诸如“gamma”，“GAMMA”，“积分 (integral)”等。

- 矩阵、条件选择等公式中的垂直线。

- 带有任意复杂极限的积分、求和等。

- 区分符：圆点、双圆点、圆帽、一杠等。

- 非程序员和数学公式打字员很容易学会。

#### NEQN

用于NROFF的一种EQN版本；接受与EQN同样的语言。为在NROFF了解的任何终端上显示而准备公式。

- 在终端的图象能力之内与EQN有同样的功能。

#### TBL

用于把表格和内容的简单描述翻译成详细的排版指令的NROFF/TROFF的一种预处理程序。

- 计算列的宽度。

- 处理左右列对齐，中心列和小数点对准。

- 设置专栏标题。

- 表格登记项可以是正文，它被调整得能适应这种表格。

- 能框住表格的全部或一部分。

#### REFER

在来自（未满足要求的）数据库的资料中填写文献引述。

- 当它们出现或收集在结尾处时，引述可以按任意风格打印。

- 可以按作者名顺序编号，等等。

#### TC

在Tektronix 4014显示屏上模拟Graphic Systems排版机。对于在排版前检验TROFF页格式加工很有用。

GREEK

在象DASI-300和DASI-450那样的Diablo机制终端上以及在Tektronic 4014上特选打印。

- 给出半行向前和反向动作。
- 通过重复打击近似希腊字母和其它特别

字符。

COL

对于一趟打印用反向行馈给规范化文件。

DEROFF

从输入中删除所有TROFF命令。

CHECKEQ

检验资料，提防在EQN使用中可能有的错误。

## 四、信息处理

SORT

一行行地分类或合并ASCII文件。对输入大小没有限制。

- 按递增或递减顺序分类。
- 按字典顺序或按数字键分类。
- 由定界符或字符位置确定的多码定位。
- 可以把与小写一起的大写插入字典顺序。

- 任选地压缩重复数据。

TSORT

拓扑分类——把一种编序转换成一种全序。

UNIQ

把一个文件中的相继重复行整理成一行。

• 发出的行是单一的，重复的，或两者兼有的行。

- 可以给出每一行的冗余计数。

TR

根据一种任意代码进行逐个字符翻译。

- 可以合并选出来的重复字符。
- 可以删除选出来的字符。

DIFF

报告使两个文件达成一致所必需的行改变、增加和删除。

• 可以产生一个编辑程序正文，以便把一个文件转换成另一个。

- 对照一个老版本进行两个新版本的差异

比较。

COMM

在两个已排序的文件中标识公共行。最多以三列输出，分别说明只出现在第一个文件中，出现在两个文件中，及/或只出现在第二个文件中的行。

JOIN

通过填补有标识键的记录组合两个文件。

GREP

当用在编辑程序ED中时，打印一个满足一种模式的文件中的所有行。

- 可以打印所有不匹配的行。
- 可以打印命中计数
- 可以打印每一个文件中的首次命中。

LOOK

在已排序的文件中折半查找具有特定前缀的行。

WC

计算一个文件中的行、“字”（由空格分开的字符串）、和字符数。

SED

ED的面向字符流的版本。能在一个不定长度的输入流中的每一行上执行一系列编辑操作。

- 可以由地址或一组地址界来选择行。
- 控制流程和条件测试。
- 多重输出字符流。