

计算机资料之二

DJS-21 机

压缩纸带方案

北京工业大学

计算站

1975. 11

目 录

一 引子.....	1
二 特征孔的定义.....	1
三 实现算法.....	6
四 扩展算法.....	1 2
五 附录.....	1 5
附录A: 特征孔定义表.....	1 5
附录B: 830厂纸带压缩后输入时 内存占用情况.....	1 6
附录C: 参考程序.....	1 6

二、引子

考虑到DJS-21机的指令书写格式为

124 2444 124 2444

(1)

因此，在五单位纸带上，为了将一个指令单元穿孔，需占15排孔：一排指令特征孔，十四排指令孔。一个程序的全部指令，按此格式穿孔时，纸带相当长。例如，830厂编制的DJS-21机ALGOL60系统就有三盘纸带。

如果拥有该计算机的用户，磁带机尚未配上，或者磁鼓虽有但不稳定，或者由于欲对原ALGOL系统作某种修改试验，导致经常需要通过光电机输入编译纸带，这时，用户就会感到非常不方便。

鉴于上述考虑，就希望设计一个方案，以产生出一条与原纸带等效的压缩纸带。新纸带在长度上大大短于原纸带，这样，就使输入过程大大简化了，速度加快了，同时，纸带的复制及保管也更容易了。这些就是设计压缩纸带的出发点。

压缩纸带并不是新的概念，例如DJS-6机的ALGOL系统纸带就是压缩了的纸带；七机部十六所的DJS-6机FORTRAN系统纸带也用压缩纸带。关键的问题在于设计什么样的压缩方案。下一节我们就来讨论这一问题，并给出一个目前我们采用的设计方案。

三、特征孔的定义

在正规格式的穿孔中，曾定义了下列特征孔：

- 00001 : 数特征孔
- 00100 : 指令特征孔
- 00110 : 地址特征孔
- 00111 : 结束特征孔

本压缩方案定义了以下特征孔：

1. 开始特征孔(00010)



该特征孔用于纸带的首部，在第一次遇到该特征孔之前的所有孔形，光电输入程序均不予“理睬”。

该特征孔的第二次以后的出现则为信息孔或空特征孔(见2)。

2 空特征孔(00001, 00010)

光电输入程序视之为无信息的空排, 就好象硬件视00000为黑纸带那样, 但是它们只能出现在特征孔处才作为空, 在信息孔位置则分别表示1和2(十六进制)。

3 结束特征孔(00011)

该特征孔用于纸带的结尾, 光电输入程序(注: 此处及以下提及此名称时均指光电输入扩展程序)遇该特征孔时执行

024 1000

以停止光电输入。

4 地址特征孔(00100~00111)

这四个特征孔相当于正规方式下的00110。但它们的后两位是地址的两位高位。由于地址形式为2444, 故在地址特征孔后面紧接着三排孔是地址码的低十二位, 例如:



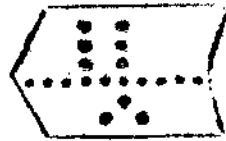
a 地址 2657



b 地址 0610



c 地址 1005



d 地址 3050

由以上例子可见, 信息“0”的表示方法采用10000码(结合于特征孔中的0例外, 如b例0610中第一个0)。

5 指令的分类

定义: 除最左八位不同时为0外, 其余各位均为0的单元称为空

0 指令

定义：除左指令最右位和右操作码七位共八位不同时为0外，其余各位均为0的单元称为右0指令。

定义：除左操作码最右二位及左地址位共16位不同时为0外，其余各位均为0的单元称为左0指令。

定义：指令的最左十位同时为0，且又不属于上面定义三类指令，则称为特殊指令。（当然也不属于全0单元）。

定义：指令单元各位均为0时，称为全0单元。指令单元各位均为1时，称为全1单元。

定义：上述各类指令之外的单元称为正常指令。

6. 正常指令特征孔 (01000~01011)

由以上定义可知，此类指令没有什么特殊性，因此4,2位代码按2444444444444444格式处理，最左两位结合于特征孔中，故而在特征孔后紧接着十排信息孔。

为了进一步压缩磁带长度，需要着重指出下面一点：

右在信息孔“0”（10000）之后尚有信息孔，则将这两排信息孔逻辑加成一排信息孔。因此，10000表示“0”“0”两排信息孔。

10111表示“0”“7”两排信息孔的叠加。

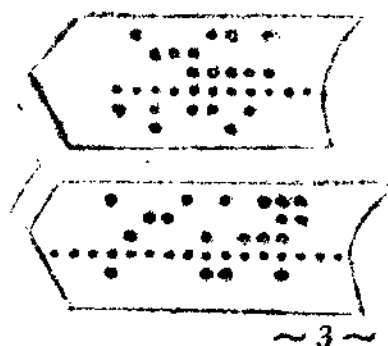
由于这一规定，在特征孔之后实际上跟着的信息孔的排数可能少于十排，（实现算法中用计数器控制就不会混乱）。

对于代码10000出现于信息孔的最后一排时，有两种情况：代表一个0（16）或两个0（16），藉助于计数器同样可避免对该代码解释的二义性。

以下为几个例子：

002 0517 034 0525

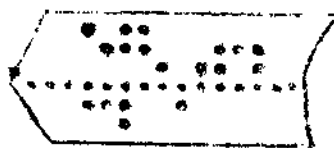
028 1104 025 0553



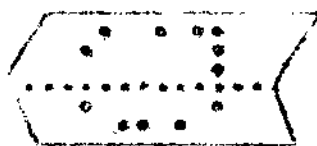
(7)

~ 3 ~

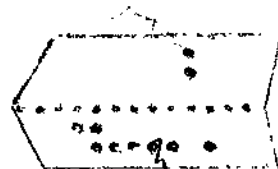
034 0590 021 0626



102 0000 004 0015



001 0000 000 0030



2 全0单元特征孔(01100)

该特征孔指出当前单元内容应为全0。该内容可由光电输入程序直接提供，故该特征孔后没有信息孔，其后必为一特征孔。

3 全1单元特征孔(01111)

该特征孔指出当前单元内容应为全1。该内容可由光电输入程序直接提供，故该特征孔后没有信息孔，其后必为一特征孔。

9 特殊指令特征孔(10000~10111)

由定义知，该类指令最左10位必为0，因此最多后跟八排信息孔。

特征孔的右三位为0~7，分别代表着后跟1~8排信息孔。

当后跟七排孔时，表示最左十四位全为0。

当后跟六排孔时，表示最左十八位全为0。

当后跟五排孔时，表示最左廿二位全为0。

当后跟四排孔时，表示最左廿六位全为0。

当后跟三排孔时，表示最左三十位全为0。

当后跟二排孔时，表示最左卅四位全为0。

当后跟一排孔时，表示最左卅八位全为0。

例子：

000 0000 000 000 $\bar{4}$

000 0000 0000 0030

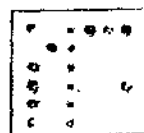
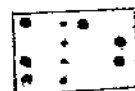
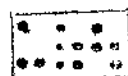
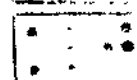
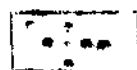
000 0000 000 070 $\bar{3}$

000 0000 004 0100

000 0020 000 25 $\bar{4}$ 0

000 073 $\bar{5}$ 13 $\bar{5}$ 35 $\bar{5}$ 5

000 0400 000 1000



10. 左θ指令特征孔(11000)

该特征孔后紧跟两排信息孔。光电输入程序将它们置于相应的位上。

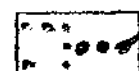
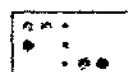
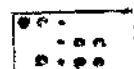
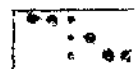
例如：

021 2000 000 0000

037 0000 000 0000

003 0000 000 0000

030 0000 000 0000



11. 左D指令特征孔(11001)

该特征孔后跟四排信息孔。

例如

000 126 $\bar{4}$ 000 0000

001 0506 000 0000

(7) 002 265 $\bar{1}$ 000 0000



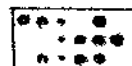
~ 5 ~

12. 右 0 指令特征孔 (11010)

该特征孔后跟二排信息孔

例为:

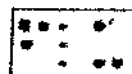
000 0000 134 0000



001 0001 020 0000



000 0000 003 0000



13. 废特征孔 (11011~11111)

在定义本特征孔系统时, 出发点是针对诸如编译程序纸带之类的常用专用纸带的。因此, 废特征孔未予定义, 故将所余之孔 (11011~11111) 全定义成废码特征孔。

该类特征孔在产生压缩纸带时没有用处, 当欲对已压缩纸带作某些修改时, 或检查复制的纸带时, 尚具用处。

14. 备注

原先在定义本系统时, 考虑将 01110 作为全 0 特征孔, 而将 01100 及 01101 作为左指令 (21 位) 的特征孔 (后跟五排信息孔)。在编制压缩程序时, 误将 01100 作为全 0 特征孔, 并将左指令特征遗忘了, 未予编入。考虑到这一疏忽对压缩量影响不大, 且当时压缩带已产生出来, 故未予改正。

三 实现算法

以下具体的谈及 DJS-21 ALGOL60 系统 (830 厂纸带) 是如何给以压缩的。

显然, 按照二中定义的格式, 是无法由人工来直接实现穿孔的。

也很显然, 这一实现算法是一个很简单的非数学问题, 且电传机又具有穿孔输出的能力, 完全可以由计算机本身来自动地产生。

但在具体实现时, 尚有若干问题。

1. 对原纸带的改造

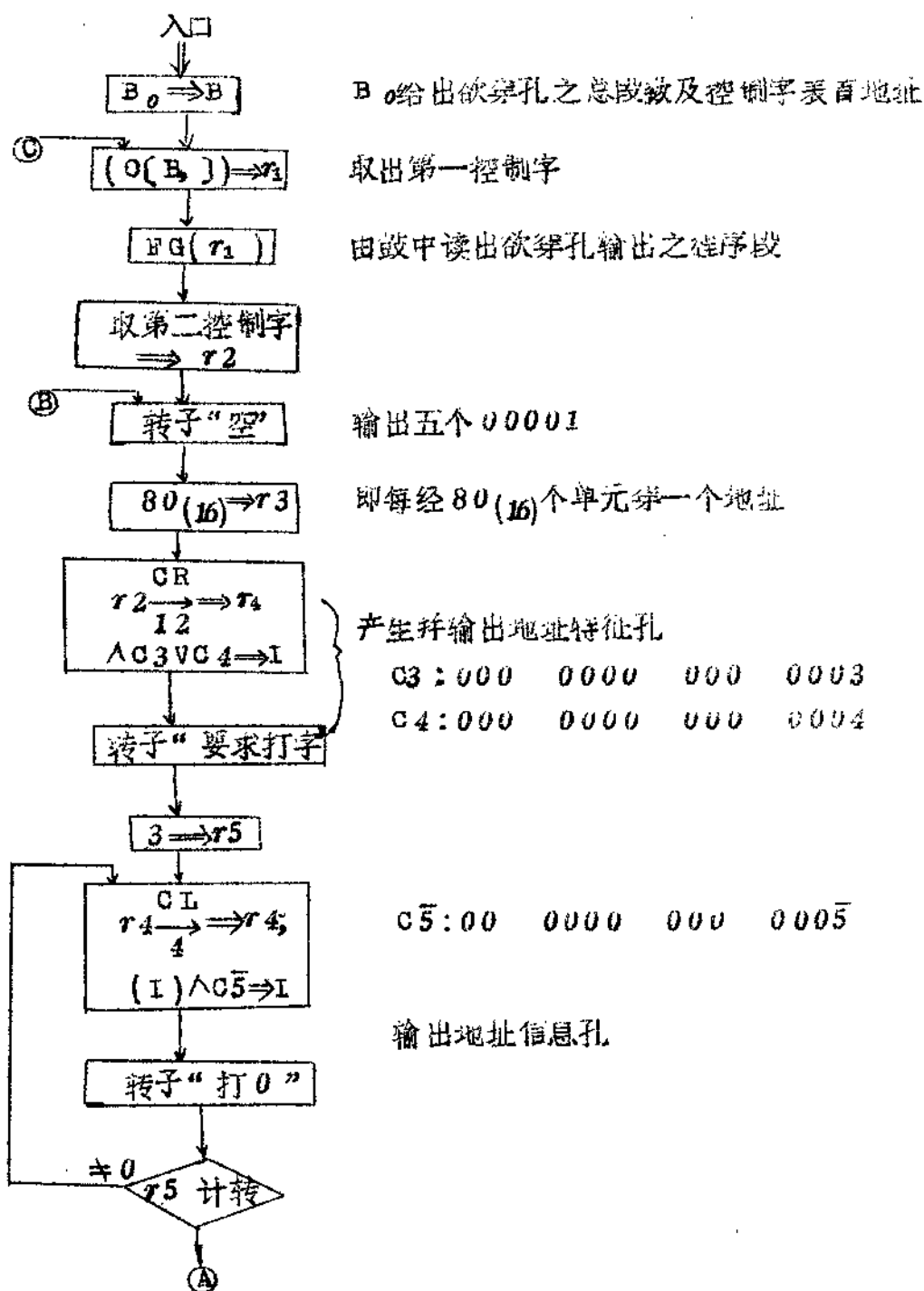
为了更有效地压缩纸带，我们将原纸带中的“联调程序”全部删除。

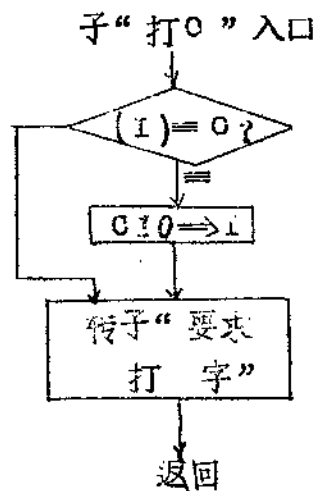
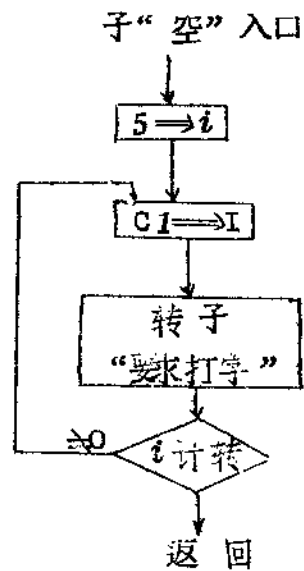
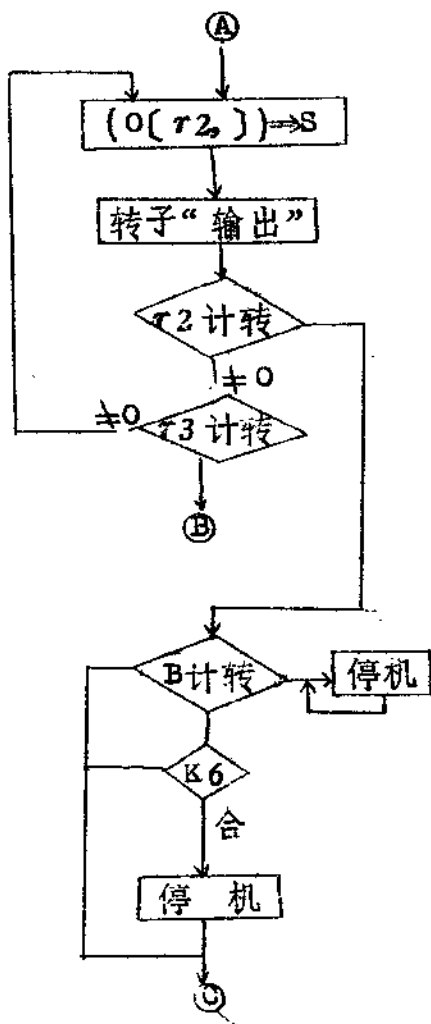
2. 存区复盖问题

由于原程序是复盖型的，压缩后的纸带希望一次输入，当然不能允许复盖型的结构。为此，对各段程序在输入时在内存中的位置作适当的调整，以避免复盖式的结构，具体调整情况见附录二。

3. 算法框图

以下给出实现算法的框图。



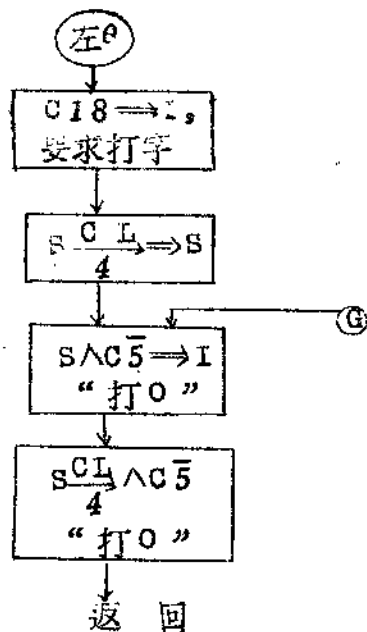
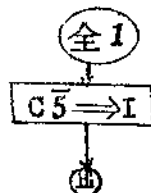
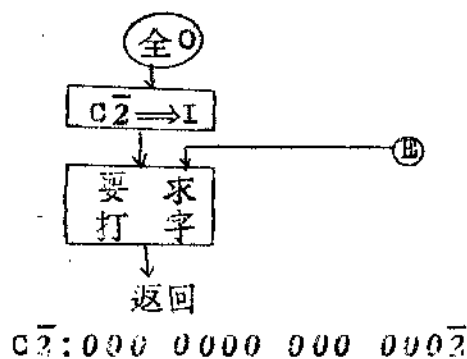
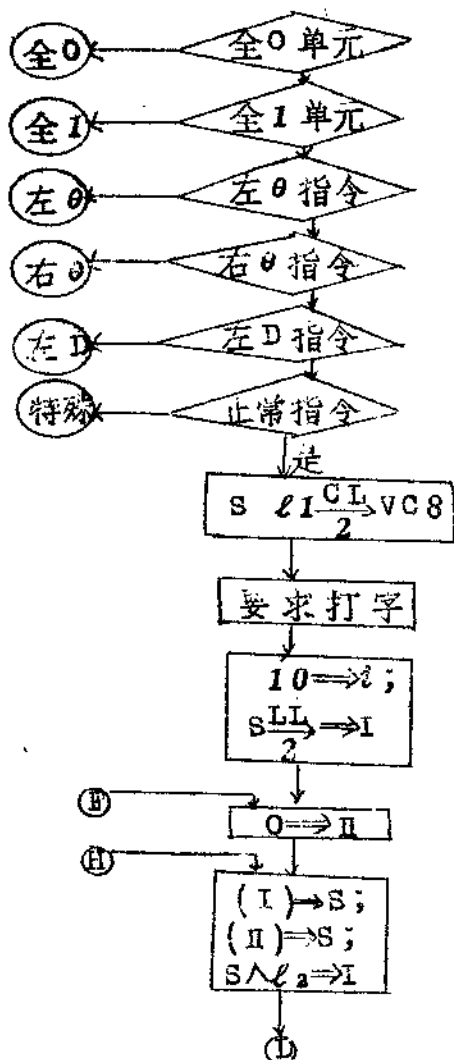


C10: 000 0000 000 0010

C1 : 000 0000 000 0001

~ 0 ~

予“输出”入口



$\ell_1 : 138 \quad 0000 \ 000 \ 0000 \ C8:000 \ 0000 \ 000 \ 0008$

$\ell_1 : 135 \quad 3800 \ 000 \ 0000 \ C18:000 \ 0000 \ 000 \ 0018$

~10~



4 框图说明

a、符号注解：

$d(B)$ ：表示有效地址 $= d + (B)$ 右址；

(B) 右址 + 变址值 $\Rightarrow B$ ；(37)

$d(B)$ ：表示有效地址 $= d + (B)$ 右址；

$\longrightarrow$ ：表示移位，箭头上表示移位方式，下面表示移位位数（+进），移位方式中第一字母表示移位类型（L 表逻辑移，A 表算术移，C 表循环移），第二字母表示移位方向（L 表左移，R 表右移）。

b、“要求打字”子程序的注解

该子程序借用系统中动态子程序中的对应程序。故实际上在入口处，需先调出动态子程序的后半部。

c、常数注解

L 1 : 125 3800 000 0000

L 2 : 138 0000 000 0000

C 1 : 000 0000 000 0001

C 3 : 000 0000 000 0003

C 4 : 000 0000 000 0004

C 7 : 000 0000 000 0007

C 8 : 000 0000 000 0008

C 2 : 000 0000 000 0002

C 5 : 000 0000 000 0005

C 10 : 000 0000 000 0010

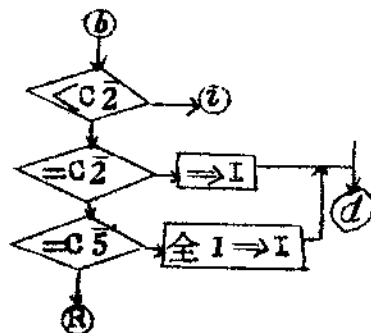
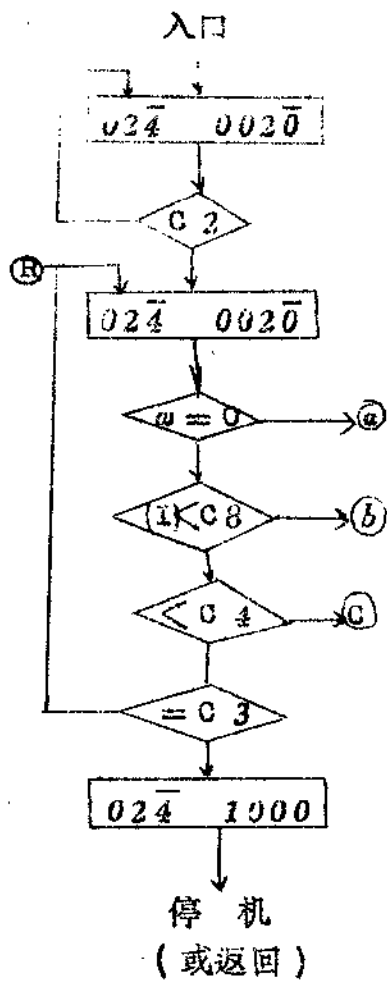
C 18 : 000 0000 000 0018

C 19 : 000 0000 000 0019

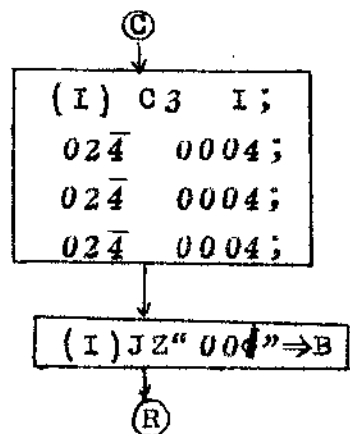
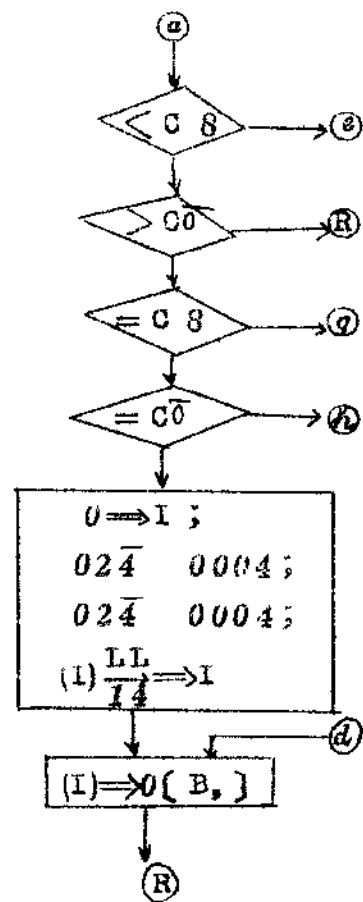
四 扩展算法

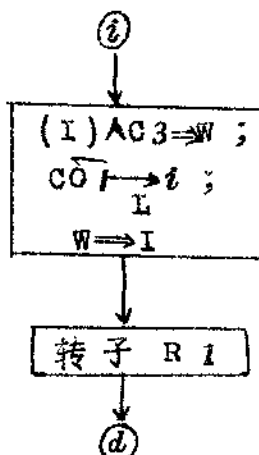
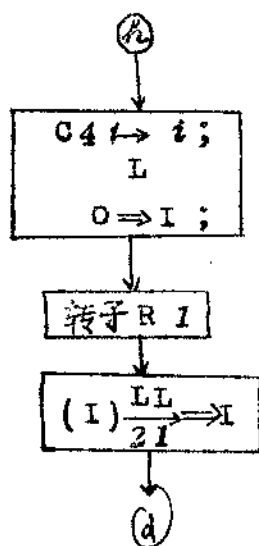
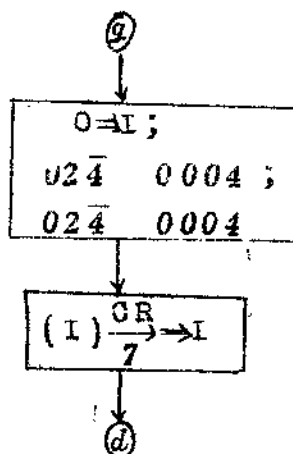
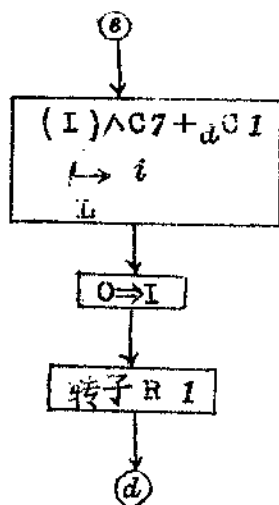
所谓扩展算法，就是实现将压缩磁带输入内存的光电输入程序。

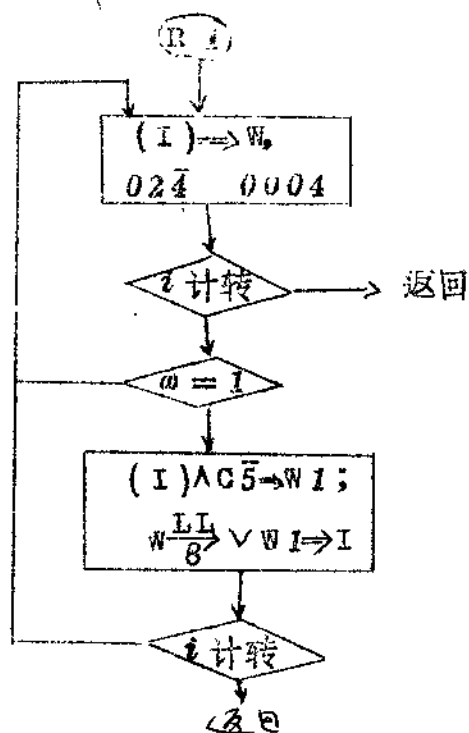
由压缩带的产生算法不难得到扩展算法的框图。



(见 = . 14 注)







该图本身很直观，不必多加解释，仅作以下注解：

符号 $i \rightarrow$ 表示填地址， $a \xrightarrow{L} i$ 表示将 a 的右址填入 i 的左址 (30)；类似地， $\xrightarrow{R}$ 为填右址 (31)。

常数 C 的给出类似于上一节。

五 附录

附录 A：特征孔定义表

00000	硬件空码	10000	特殊指令
00001	空	10001	
00010	开始、空	10010	
00011	结束	10011	
00100	地址	10100	
00101		10101	
00110		10110	
00111		10111	