

# 微生物数值分类中电子计算机程序编制方法简介\*

戴振杰

(中国科学院计算中心,北京)

赵玉峰 沈舜杰 马淑坤 施艳莽

(中国科学院微生物研究所,北京)

本文介绍使用《BCY 算法语言》编制微生物数值分类程序,以及在国产 013 计算机\*\*上实施计算的过程。目前这套程序可以根据实验所得原始数据表的操作分类单元(以下简称 O. T. U.)直接计算出符合系数,并按其大小排列,直至完成。

例如用电子计算机对 11 株菌进行数值分类,首先须列出原始特征表(表 1 是对 11 株菌作 12 项测定的原始特征表),然后按大小排列符合系数(表 2),再排出单连锁群(表 3)。

表 1 原始特征表

特征号 菌株号	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	0	1	1	1	0	1	1	0	0	1
2	1	0	0	1	1	0	1	1	0	1	1	1
3	1	0	0	1	0	0	1	1	0	0	1	1
4	1	1	0	1	1	0	0	1	0	1	0	1
5	1	1	0	1	1	1	0	0	1	1	0	1
6	1	0	0	1	0	1	1	1	0	0	0	1
7	1	0	0	1	1	0	1	1	0	1	1	1
8	1	1	0	1	1	0	0	1	1	0	0	1
9	1	1	0	1	1	0	0	0	1	1	0	1
10	1	0	0	1	1	0	1	1	0	0	1	1
11	1	0	0	1	1	1	1	1	0	0	0	1

## 一、框图及程序

### (一) 框图 (见下页)

## (二) 程序

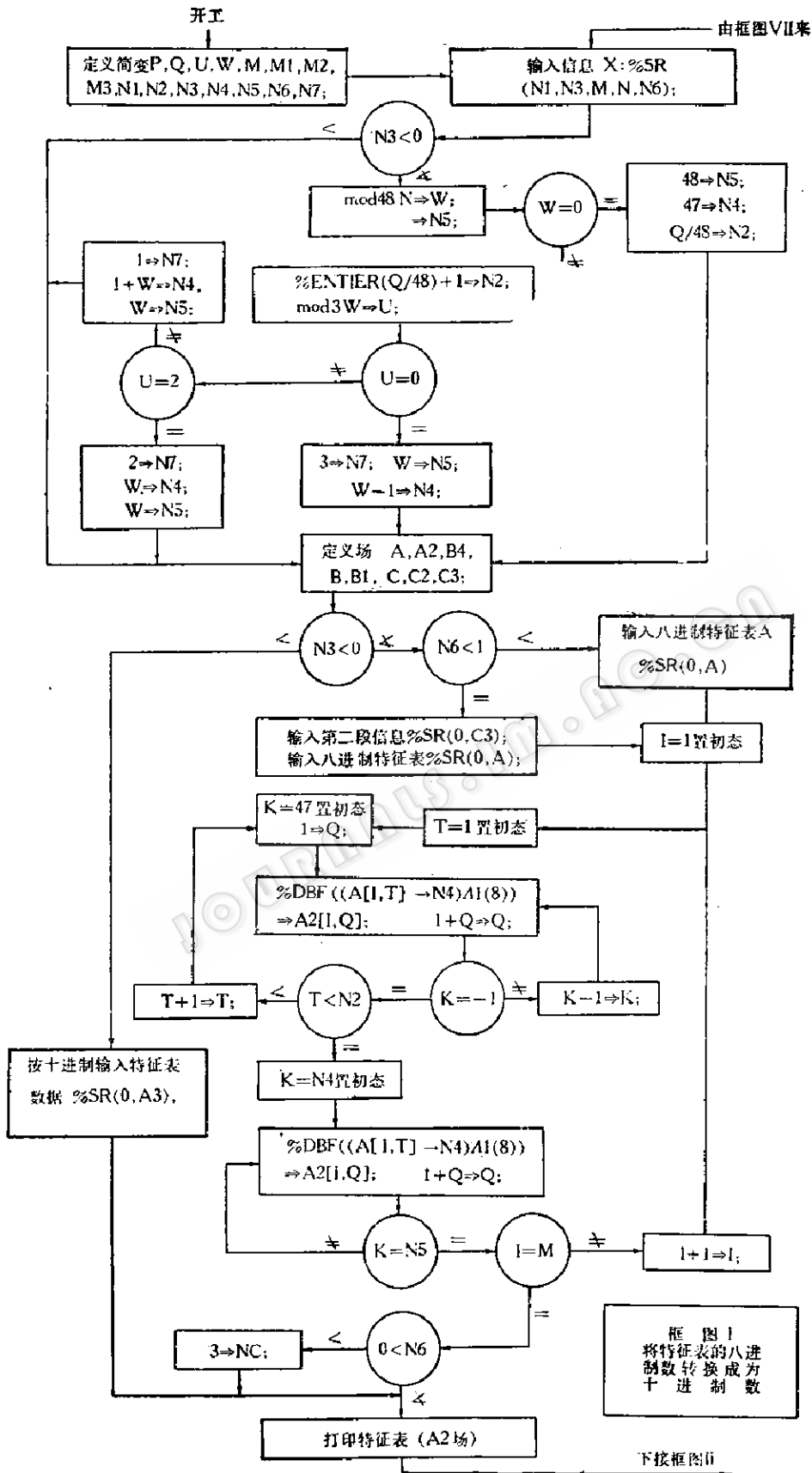
### 程 序

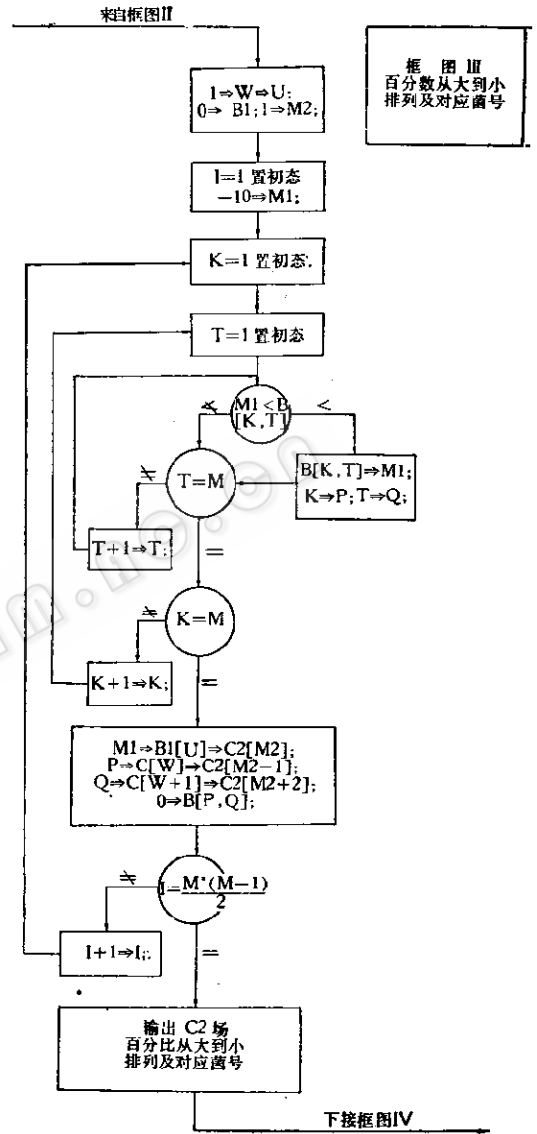
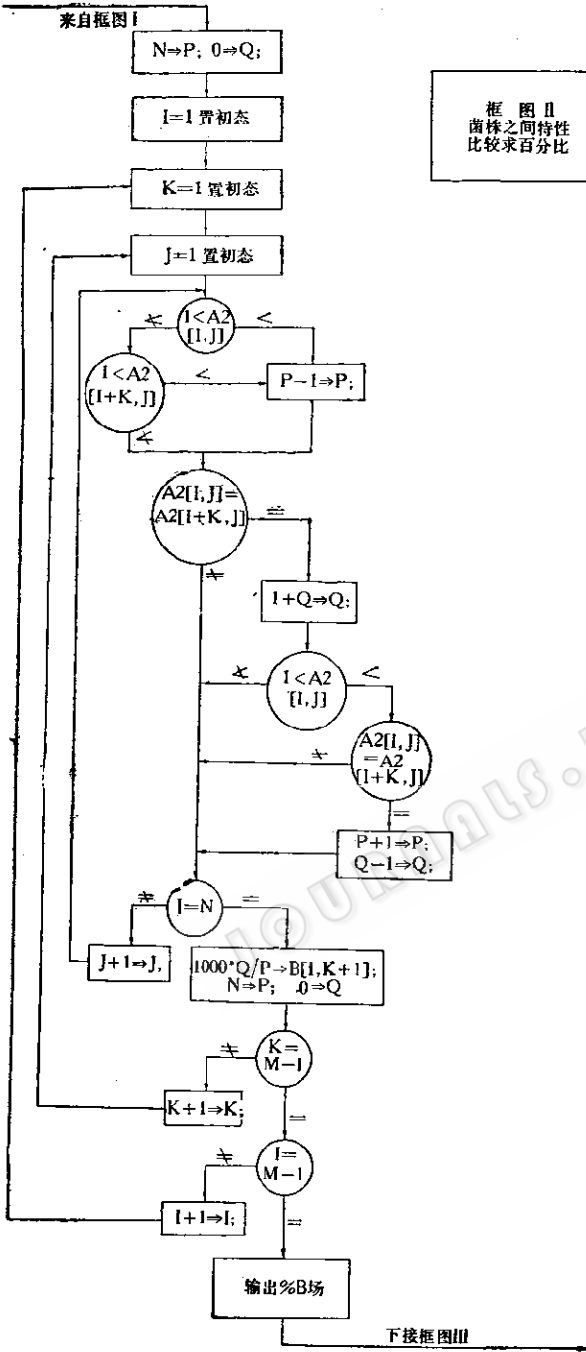
```
#YE 001
#HA 01 #KS #JB P, Q, U, W, M, MI, M2, M3, N,
NI, N2, N3, N4, N5,
#HA 02 N6, N7;
```

\* 本工作承徐浩,王大超同志帮助。在编写程序过程中还得到中国科学院计算中心黄玉霞和楼金虎同志帮助。

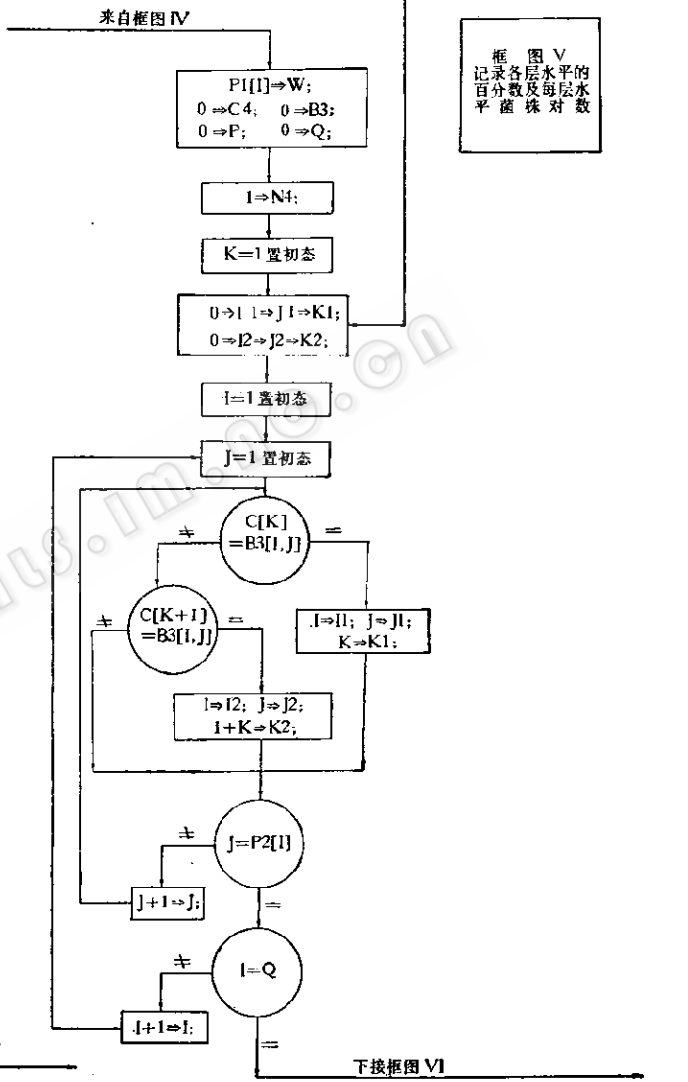
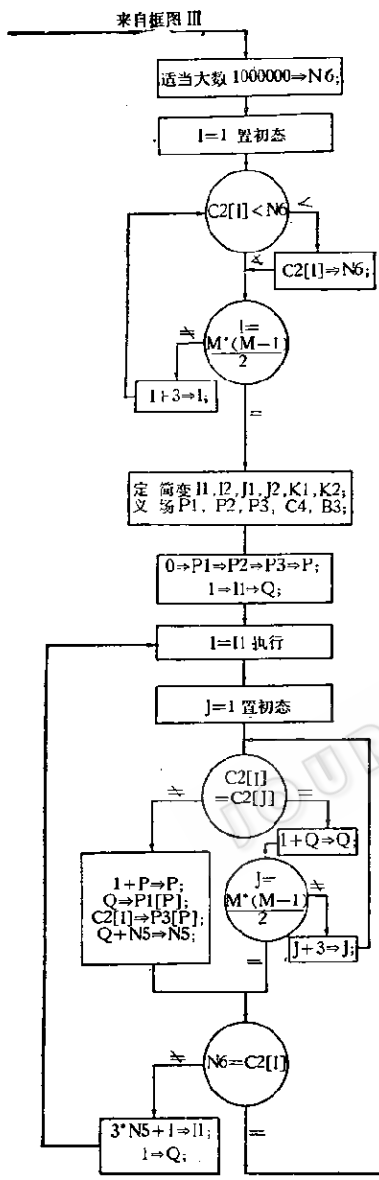
\*\* 由中国科学院计算技术研究所研制。中国科学院计算中心编印有各种有关此计算机使用的手册及资料。

本文中用到的符号所代表的意义是: #BC—步长, #CS—次数, #DU—对于, #FZ—否则, #KS—开始, #RG—如果, #JB—简变, #JS—结束, #ZE—则, #ZX—执行, #WJ—完结, 扩展 FS—数据以十进制格式按浮点输入, 扩展 F3—数据按八进制格式输入, 送毕—光电纸带输入机的中断信号, mdo<sub>x</sub>X—X-%ENTIER(X/N)\*N, %SR—输入数据标准过程, %CGSC—输出标准过程, %DBF(E)—E 的定变浮(E 的值为定整数, 函数值的浮点数), %ENTIER(E)—E 的整数函数(即不大于 E 之值的最大整数), \*BDS—场的下标变量; “+”—加号, “-”—减号, “×”—乘号, “/”—除号, “^”—逻辑乘, “= >”—赋值号, “#→”—移位(右移), “<”—小于, “=”—等号; A—保存 A1 场(八进制特征表初值), A<sub>1</sub>—八进制特征表初值场, A<sub>2</sub>—菌株分类工作场, A<sub>3</sub>—十进制特征表初值场, B—菌株间相似性百分比场, B<sub>1</sub>—百分比从大到小排列场, B<sub>3</sub>—菌株分群工作场, B<sub>4</sub>—分群后三角阵场, C—存放各 Ssm 值的相应菌号场, C<sub>2</sub>—存放 Ssm 值及相应菌株号, C<sub>3</sub>—处理缺项用场, C<sub>4</sub>—归群总场, N—菌株的特征数, N<sub>1</sub>—菌株特征表的编号, N<sub>3</sub>—特征表的进制信息(“-1”表示十进制数, “0”表示八进制数), N<sub>6</sub>—有无 nc 项信息(“0”表示无 nc 项, “1”表示有缺项信息); N<sub>2</sub>, N<sub>4</sub>, W, P, Q, K<sub>1</sub>, K<sub>2</sub>, I<sub>1</sub>, I<sub>2</sub>, J<sub>1</sub>, J<sub>2</sub> 为一般工作单元; X—作下张特征表返回地址(0040006), XX—一进入程序第二层进行地址(0040163), X<sub>4</sub>: 找 Ssm 中的最小值(0042203), X<sub>5</sub>: 一分群结束转出地址(0042331), X<sub>6</sub>: 一某水平归群未返回地址(0043640), X<sub>7</sub>: 一归群完转分群后谱三角阵地址(0043647), S<sub>0</sub>: 一十进制特征表无缺项输入地址(0040741), S<sub>1</sub>: 一中途复算地址(0040756), S<sub>2</sub>: 一打印特征表地址(0040756), YY: 一某水平分群完后进入下一水平分群地址(0043241), YYY: 一并群地址(0043012)。





框图 IV  
从 C2 场中取出最小的百分数放在 N6 中



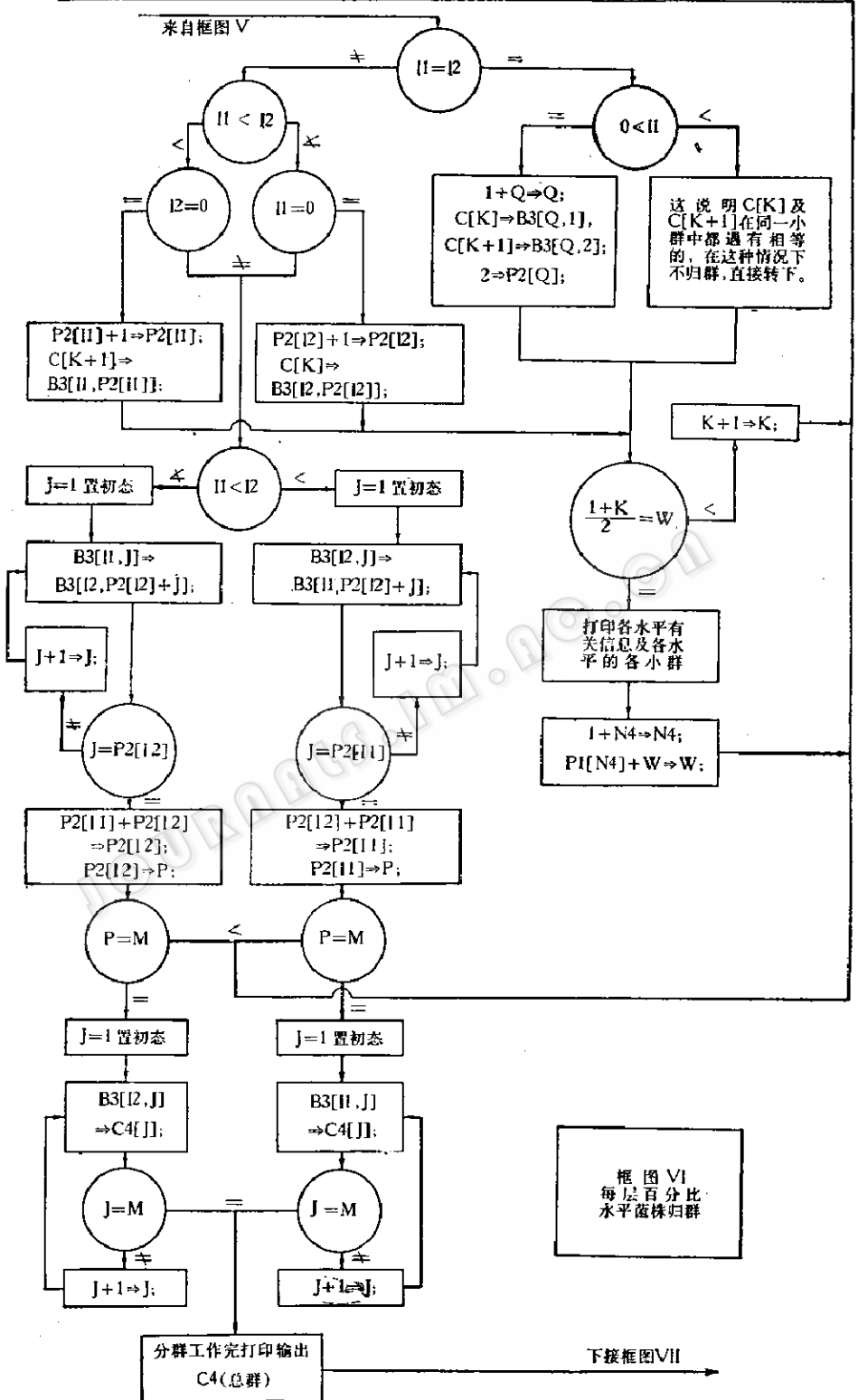
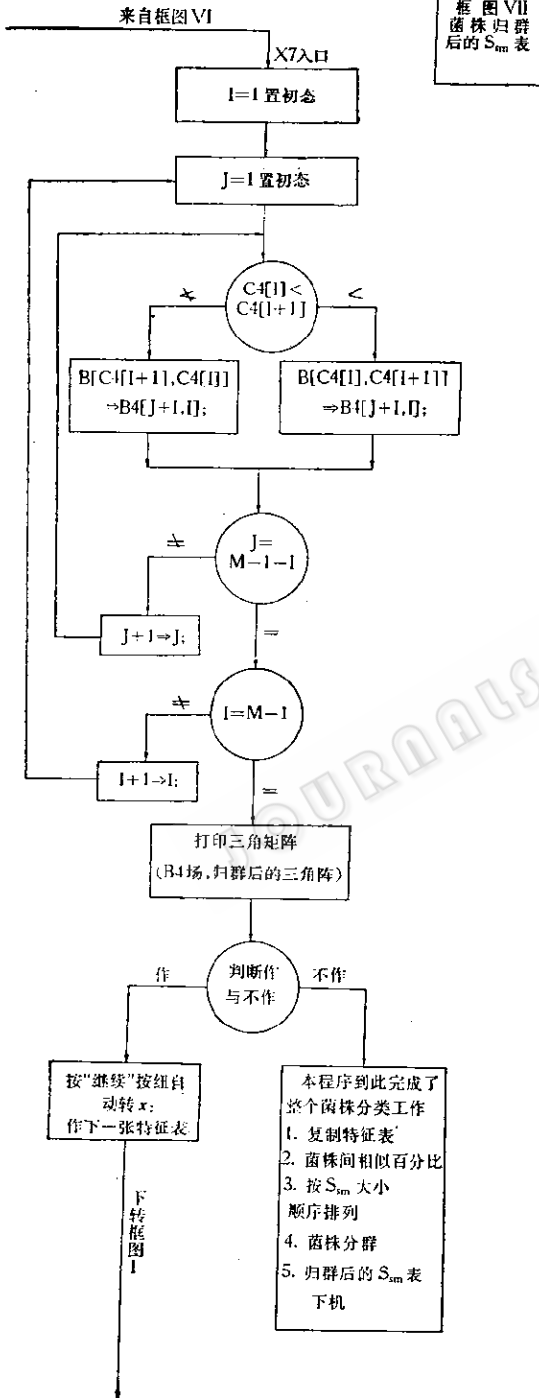


表 2 符合系数表

符合系数 (%)	菌株号	符合系数 (%)	菌株号	符合系数 (%)	菌株号
100.0	2 7	91.7	1 8	91.7	2 10
91.7	3 10	91.7	5 9	91.7	6 11
91.7	7 10	83.3	1 5	83.3	2 3
83.3	3 6	83.3	3 7	83.3	4 8
83.3	4 9	83.3	8 9	83.3	10 11
75.0	1 4	75.0	1 9	75.0	1 11
75.0	2 4	75.0	2 11	75.0	3 11
75.0	4 5	75.0	4 7	75.0	5 8
75.0	6 10	75.0	7 11	66.7	1 6
66.7	2 6	66.7	4 10	66.7	4 11
66.7	6 7	66.7	8 10	66.7	8 11
58.3	1 10	58.3	2 8	58.3	2 9
58.3	3 4	58.3	3 8	58.3	4 6
58.3	5 11	58.3	6 8	58.3	7 8
58.3	7 9	50.0	1 2	50.0	1 3
50.0	1 7	50.0	2 5	50.0	5 6
50.0	5 7	50.0	9 10	50.0	9 11
41.7	3 9	41.7	5 10	41.7	6 9
33.3	3 5				



- #HA 03 X:%SR[0, N1, N3, M, N, N6];
- #HA 04 %CGSC[\*5.0, N1, N3, M, N];
- #HA 05 N=>Q=>P;
- #HA 06 #RG N3<0
- #HA 07 #ZE #ZY XX
- #HA 08 #FZ;
- #HA 09 N-%ENTIER[Q/48]\*48=>W=>N5;
- #HA 10 #RG W=0
- #HA 11 #ZE #KS 48=>W=>N5;
- #HA 12 W-1=>N4;
- #HA 13 Q/48=>N2; #ZY XX
- #HA 14 #JS
- #HA 15 #FZ %ENTIER[Q/48]+1=>N2;
- #HA 16 W-%ENTIER[W/3]\*3=>U;
- #HA 17 #RG U=0
- #HA 18 #ZE #KS 3=>N7; W-1=>N4 #JS
- #HA 19 #FZ #RG U=2
- #HA 20 #ZE #KS 2=>N7; W=>N4 #JS
- #HA 21 #FZ #KS 1=>N7; W+1=>N4 #JS;
- #YE 002
- #HA 01 XX: %CGSC[\*5.0, W, N2, N3, N4, N5, N6, N7];
- #HA 02 #KS #CHA[1: M, 1: N2],
- #HA 03 A2[1: M, 1: N],
- #HA 04 B4, B[1: M, 1: M],
- #HA 05 B1[1: M\*[M-1]/2],
- #HA 06 C[1: M\*[M-1]],
- #HA 07 C2[1: 3\*M\*[M-1]/2],

表 3 单连锁群表

符合系数 (%) \ 菌号	2	7	10	3	6	11	1	8	5	9	4
2	0	0	0	0	0	0	0	0	0	0	0
7	100.0	0	0	0	0	0	0	0	0	0	0
10	91.7	91.7	0	0	0	0	0	0	0	0	0
3	83.3	83.3	91.7	0	0	0	0	0	0	0	0
6	66.7	66.7	75.0	83.3	0	0	0	0	0	0	0
11	75.0	75.0	83.3	75.0	91.7	0	0	0	0	0	0
1	50.0	50.0	58.3	50.0	66.7	75.0	0	0	0	0	0
8	58.3	58.3	66.7	58.3	58.3	66.7	91.7	0	0	0	0
5	50.0	50.0	41.7	33.3	50.0	58.3	83.3	75.0	0	0	0
9	58.3	58.3	50.0	41.7	41.7	50.0	75.0	83.3	91.7	0	0
4	75.0	75.0	66.7	58.3	58.3	66.7	75.0	83.3	75.0	83.3	0

- #HA 08 C3[1: M\*N];
- #HA 09 0=>C3;
- #HA 10 #RG N3<0 #ZE #ZY SO
- #HA 11 #FZ; #RG N6<1 #ZE %SR[O, A]
- #HA 12 #FZ #KS %SR[O,C3]; %SR[O, A] #JS;
- #HA 13 #KS
- #HA 14 #DU I=1 #BC 1 #CS M #ZX
- #HA 15 #KS I=>Q;
- #HA 16 #DU T=1 #BC 1 #CS N2 #ZX
- #HA 17 #KS #RG T<N2
- #HA 18 #ZE #DU K=47 #BC -1 #CS 48 #ZX
- #HA 19 #KS %DBF[[A[I, T] #->K] ^1[8]] = >A2[I, Q],
- #HA 20 I+Q=>Q
- #HA 21 #JS
- #YE 003
- #HA 01 #FZ #DU K=N4 #BC -1 #CS N5 #ZX
- #HA 02 #KS %DBF[[A[I, T] #->K] ^1[8]] = >A2[I, Q];
- #HA 03 I+Q=>Q #JS
- #HA 04 #JS
- #HA 05 #JS
- #HA 06 #JS;
- #HA 07 #RG N6<1
- #HA 08 #ZE #ZY S2
- #HA 09 #FZ;
- #HA 10 0=>W;
- #HA 11 #DU I=1 #BC 2 #CS M\*N #ZX
- #HA 12 #RG O<C3[I]
- #HA 13 #ZE #KS W+2=>W;
- #HA 14 3=>A2[C3[I], C3[I+1]]
- #HA 15 #JS
- #HA 16 #FZ #KS %CGSC[\*5.0, N1, W];
- #HA 17 #ZY S2
- #HA 18 #JS
- #HA 19 SO:%SR[O, A2];
- #HA 20 S1;
- #HA 21 S2:%CGSC[\*5.0, N1, N 3, M, N, N6, N7, N2 N4];
- #YE 004
- #HA 01 %CGSC[\*1.0, A2];
- #HA 02 N=>P; O=>Q;
- #HA 03 #DU I=1 #BC 1 #CS M-1 #ZX
- #HA 04 #DU K=1 #BC 1 #CS M-1 #ZX
- #HA 05 #KS #DU J=1 #BC 1 #CS N #ZX
- #HA 06 #KS #RG I<A2[I, J] #ZE P-1=>P #FZ;
- #HA 07 #RG I<A2[I+K, J]
- #HA 08 #ZE P-1=>P
- #HA 09 #FZ;
- #HA 10 #RG A2[I, J]=A2[K+I, J]
- #HA 11 #ZE Q+1=>Q
- #HA 12 #FZ;
- #HA 13 #RG I<A2[I, J]
- #HA 14 #ZE #RG 42[I, J]=A2[I+K, J]
- #HA 15 #ZE #KS P+1=>P; Q-1=>Q
- #HA 16 #JS
- #HA 17 #FZ
- #HA 18 #FZ #JS;
- #HA 19 1000\*Q/P=>B[I, K+I]=>B4[I, K+I];
- #HA 20 #RG 2[8] ^%KZT[I]=2[8] #ZE
- #HA 21 %CGSC [\*3.0, I, I+K, Q, P, B[I, I+K]] #FZ;
- #YE 005
- #HA 01 N=>P; O=>Q
- #HA 02 #JS;

#HA 03 #RG 4[8] ^%KZT[1]=4[8] #ZE %CGSC  
 [3.0, B] #FZ  
 #HA 04;  
 #HA 05 1=>W=>U; 0=>B1; 1=>M2;  
 #HA 06 #DU I=1 #BC 1 #CS M\* [M-1]/2  
 #ZX  
 #HA 07 #KS-10=>M1;  
 #HA 08 #DU K=1 #BC 1 #CS M #ZX  
 #HA 09 #DU T=1 #BC 1 #CS M #ZX  
 #HA 10 #KS #RG M1<B[K, T]  
 #HA 11 #ZE #KS' B[K, T]=>M1;  
 #HA 12 K=>P;  
 #HA 13 T=#Q #JS  
 #HA 14 #FZ  
 #HA 16 #JS;  
 #HA 17 M1=>BI[U]=>C2[M2];  
 #HA 18 P=>C[W]=>C2[M2+1];  
 #HA 19 Q=>C[W+1]=>C2[M2+2];  
 #HA 20 #RG 1[8] ^%KZT[1]=1[8]  
 #HA 21 #ZE %CGSC[3.0, BI[U], P, Q] #FZ;  
 #YE 006  
 #HA 01 2+W=>W;  
 #HA 02 1+U=>U;  
 #HA 04 3+M2=>M2;  
 #HA 05 0=>B[P, Q];  
 #HA 06 #JS;  
 #HA 07 #RG 1[8] ^%KZT[1]=1[8]  
 #HA 08 #ZE %CGSC[\*4.0, C2] #FZ;  
 #HA 09 %SC[[H10, C[9[FS4.0], HO]], C2];  
 #YE 007  
 #HA 01 1000000=>N6; B4=>B;  
 #HA 02 #DU I=1 #BC 3 #CS M\* [M-1]/2  
 #ZX  
 #HA 03 #RG C2[I]<N6  
 #HA 04 #ZE C2[I]=>N6  
 #HA 05 #FZ;  
 #HA 06 #KS #JB K1, K2, I1, I2, J1, J2;  
 #HA 07 #CH P1, P2, P3, C4[1; M], B3[1; M, 1; M];  
 #HA 08 1=>Q=>I1;  
 #HA 09 0=>P1; 0=>P2; 0=>P3; 0=>C4; 0=>B4;  
 #HA 10 0=>N5=>P;  
 #HA 11 X4: #DU I=11 #ZX  
 #HA 12 #DU J=1+3 #BC 3 #CS [M-1]\*M/2  
 #ZX  
 #HA 13 #RG C2[I]=C2[J]  
 #HA 14 #ZE 1+Q=>Q  
 #HA 15 #FZ #KS 1+P=>P; Q=>P1[P]  
 #HA 16 C2[I]=>P3[P];  
 #HA 17 Q+N5=>N5;  
 #HA 18 #RG C2[I]=N6  
 #HA 19 #ZE #ZY X5  
 #HA 20 #FZ #KS

#HA 21  
 #YE 008  
 #HA 01 3\*N5+1=>I1;  
 #HA 02 1=>Q;  
 #HA 03 #ZY X4 #JS  
 #HA 04 #JS;  
 #HA 05 X5:%CGSC[\*4.0, P1, P3];  
 #HA 06 PI[1]=>W;  
 #HA 07 0=>C4; 0=>B3;  
 #HA 08 0=>P=>Q;  
 #HA 09 1=>N4;  
 #HA 11 #DU K=1 #BC 2 #CS M\*M #ZX  
 #HA 12 #KS 0=>K1=>I1=>J1;  
 #HA 13 0=>K2=>I2=>J2;  
 #HA 14 #DU I=1 #BC 1 #CS Q #ZX  
 #HA 15 #DU J=1 #BC 1 #CS P2[I] #ZX  
 #HA 16 #KS #RG C[K]=B3[I, J]  
 #HA 17 #ZE #KS  
 #HA 18 I=>I1;  
 #HA 19 J=>J1;  
 #HA 20 K=>K1 #JS  
 #HA 21 #FZ #RG C[K+1]=B3[I, J]  
 #YE 009  
 #HA 01 #ZE #KS  
 #HA 02 1=>I2;  
 #HA 03 J=>J2; 1+  
 #HA 04 K=>K2 #JS  
 #HA 05 #FZ  
 #HA 06 #JS;  
 #HA 07 #RG I1=I2  
 #HA 08 #ZE #RG 0<I1  
 #HA 09 #ZE #ZY YY  
 #HA 10 #FZ #KS  
 #HA 11 1+Q=>Q;  
 #HA 12 C[K]=>B3[Q, 1];  
 #HA 13 C[K+1]=>B3[Q, 2];  
 #HA 14 2=>P2[Q];  
 #HA 15 #ZY YY  
 #HA 16 #JS  
 #HA 17 #FZ  
 #HA 18 #KS #RG I1<I2  
 #HA 19 #ZE #RG I1=0  
 #HA 20 #ZE #KS  
 #HA 21 P2[I2]+1=>P2[I2];  
 #YE 010  
 #HA 01 C[K]=>B3[I2, P2[I2]];  
 #HA 02 #ZY YY #JS  
 #HA 03 #FZ #ZY YYY  
 #HA 04 #FZ  
 #HA 05 #RG I2=0  
 #HA 06 #ZE #KS  
 #HA 07 P2[I1]+1=>P2[I1];



```

#HA 08 C[K+1]=>B3[I1, P2[I1]];
#HA 09 #ZY YY #JS
#HA 10 #FZ;
#HA 11 YYY: #RG
#HA 12 I1<I2
#HA 13 #ZE #KS
#HA 14 #DU J=1 #BC 1 #CS P2[I2] #ZX
#HA 15 B3[I2, J]=>B3[I1, P2[I1] + J];
#HA 16 P2[I1] + P2[I2]=>P2[I1];
#HA 17 0=>P2[I2];
#HA 18 #ZY YY
#HA 19 #JS
#HA 20 #FZ #KS
#HA 21 #DU J=1 #BC 1 #CS P2[I1] #ZX
#YE 011
#HA 01 B3[I1, J]=>B3[I2, P2[I2] + J];
#HA 02 P2[I1] + P2[I2]=>P2[I2];
#HA 03 0=>P2[I1]
#HA 04 #JS
#HA 05 #JS;
#HA 06 YY:
#HA 07 #RG [K+1]/2=W
#HA 08 #ZE #KS %CGSC[4.0, N4, P3[N4]];
#HA 09 I=N4=>N4;
#HA 10 P1[N4] + W=>W;
#HA 11 #KS #DU I=1 #BC 1 #CS Q #ZX
#HA 12 #DU J=1 #BC 1 #CS P2[I] #ZX
#HA 13 #RG 0 P2 I
#HA 14 #ZE %<CGSC[4.0, B3[I, J], I, J]
#HA 15 #FZ
#HA 16 #JS;
#HA 17 #RG P2[I2]<P2[I1]
#HA 18 #ZE #KS P2[I1]=>P;
#HA 19 #RG P<M
#HA 20 #ZE #ZY X6
#HA 21 #FZ
#YE 012
#HA 01 #KS #DU J=1 #BC 1 #CS M
#HA 02 #ZX
#HA 03 B3[I1, J]=>C4[J];
#HA 04 %CGSC[*4.0, C4];
#HA 05 #ZY X7 #JS
#HA 06 #JS
#HA 07 #FZ #KS P2[I2]=>P;
#HA 08 #RG P<M
#HA 09 #ZE #ZY X6
#HA 10 #FZ #DU J=1 #BC 1 #CS M #ZX
#HA 11 B3[I2, J]=>C4[J]
#HA 12 #JS #JS
#HA 13 #FZ #ZY X6; #ZY X7;
#HA 14 X6: #JS;
#HA 15 X7:

```

```

#HA 16 %CGSC[*4.0, C4];
#HA 17 #DU I=1 #BC 1 #CS M-1 #ZX
#HA 18 #DU J=1 #BC 1 #CS M-1 #ZX
#HA 19 #RG C4[I]<C4[I+J]
#HA 20 #ZEB[C4[I] C4[I+J]]=>B4[I+J, I]
#HA 21 #FZ B[C4[I+J], C4[I]]=>B4[J+I, I];
#YE 013
#HA 01 %SC[[H10, C[I1[FS4.0], HO]], C4, B4, C4];
#HA 02 #WJ;
#HA 13 0=>P=>Q=>U=>W=>M=>M1 =>
M2=>M3;
#HA 14 0=>N=>N1=>N2=>N3=>N4 => N5
=>N6=>N7;
#HA 15 0=>K1=>K2=>I1=>I2=>J1=>J2;
#HA 16 0=>A; 0=>A2;
#HA 17 0=>B; 0=>B1; 0=>B3; 0=>B4;
#HA 18 0=>C; 0=>C2; 0=>C3; 0=>C4;
#HA 19 0=>P1; 0=>P2; 0=>P3;
#HA 20 #ZY X
#HA 21 #JS #JS #JS

```

### (三) 有关本程序的几点说明

1. 设计本程序所遵循的微生物数值分类法原理见参考文献[1]、[2]、[3]、[4]和[5]。

2. 求每一对菌株间诸 O. T. U. 单元总的相似性, 本程序采用 Sokal 和 Michener 符合系数法 (matching coefficient)。按下列公式计算的结果以百分比表示, 取三位有效数字, 同时略去小数点及百分比符号。

符合系数 (S<sub>sm</sub>) =

$$\frac{\text{阳性的阴性的符合数总和}[(\Sigma+) + (\Sigma-)]}{\text{总测定数(全体) - 无效测定数}(\Sigma_{nc})}$$

3. 本程序将各对菌株之间诸 O. T. U. 相似值 (S<sub>sm</sub>) 所组成的矩阵, 取三角矩阵形式排列, 其余部分 (包括矩阵对角线上诸 O. T. U. 自身之比, 即 S<sub>sm</sub> = 0) 全部置“0”, 便于检验。

4. 簇群分析 (clusters analysis) 的准备工作, 本程序是将 S<sub>sm</sub> 三角矩阵中全部数值按大小顺序重排, 重排后一百分数水平之内的诸 S<sub>sm</sub> 值, 均按两个菌株序号先小后大次序排列。如某表中 S<sub>sm</sub> 值为 88% 水平的有四对, 则应排成 2—3, 2—6, 4—9, 5—6。

5. 将已按大小排列的 S<sub>sm</sub> 值, 按各个水平顺序在同一水平内分群, 从最高水平开始, 先取

本水平内左边  $M_i$  (菌株号) 最小的第一对置首位,把与第一对有相同  $M_i$  的各对依次往后接,凡与接上的诸  $M_i$  有相同  $M_i$  数的对子,继续接上,直到和所余各对的  $M_i$  数均不相同为止,划为第一群。剩余各对中,再选取左边  $M_i$  最小的一对置第二群首位,依第一群归群的方法划第二群,余此类推,直到把本水平内的全部菌株均被划分成彼此的  $M_i$  值无联系的诸群为止。再下降一个  $S_{sm}$  水平,把上一水平各群按同样方法依先后顺序继续接纳与本群内诸  $M_i$  有相同号数的对子或已划分的群。如此继续,直到在某一水平内全部的  $M$  株菌都归成一群为止。按此水平内各菌株号  $M_i$ ,依排列顺序,成行成列重做  $S_{sm}$  值的三角矩阵。

6. 本程序已设计同时将若干份特征表进行批量处理的方案。当第一份特征表计算打印完毕,只须按动控制台上“继续”按钮,第二份特征表信息及特征表数据即自行输入,随即自动计算和打印结果。如此操作直至批量处理结束为止。

## 二、特征表数据的准备

### (一) 十进制特征表

将  $M$  株菌中的每株菌按形态生理特征测定  $N$  项列表<sup>[5]</sup>,表中横向由左向右依次排列  $N$  项测定结果,由上向下依菌株序列号共  $M$  行,排列成  $M \times N$  的矩阵表。表中用“1”代表阳性反应,用“0”代表阴性反应,用“3”代表疑似项(或缺项),用“nc”(not considered)表示。

### (二) 八进制特征表

将上述十进制特征表中全部“3”(nc 项)改为“0”,则表中  $M$  行每行全部由“0”或“1”组成的  $N$  个数,每行自左而右每三位一节化为八进制数。因为  $N$  在每个特征表中的数值不等,所以  $N/3$  的余数也不相同,化为八进制时必然会出现三种情况:第一种情况是  $N$  能被 3 整除,这时可化为  $N/3$  个八进制数;第二种情况是  $N$  被 3 除后余数为 1,此时可补两个“0”,配成一个八进制数;第三种情况是  $N$  被 3 除后余数为 2,此时最后要补一个“0”,即配成一个八进制数。

特征表中的  $N$  项,按上述三种情况处理均可化成八进制数,按顺序将每 16 个数编为一组,一组占一行,每行末尾加“;”号。不足一行者也占一行,加“;”号,以此类推。

例如十进制数:

```
111 100 010 001 010 110 001 100 011
100 100 001 010 110 111 010 001 011
11
```

化为八进制形式为:

```
7421 2614 3441 2672;
136:
```

## 三、穿孔书写格式

程序及数据均用八单位纸带按《013》计算机专用码\*,穿孔。穿孔步骤如下:

1. 留一米长空白纸带。

2. 按“\*”、《题目名字》、“;”、纸带号、“;”,送毕的顺序穿孔。《题目名字》不超过 6 个以字母打头的字母数字串;纸带号为不超过 177 的八进制数。

3. 程序的穿孔。

4. 第一段和第二段信息的穿孔。

第一段信息共 5 个数,书写格式如下:扩展 FS  $N_1; N_3; M; N; N_6$ ; 送毕。例如“扩展 FS79201; -1 (或 0); 80; 60; 0 (或 1); 送毕”这段信息,表示纸带穿孔是用 79201 号特征表,十进制(或八进制)格式穿孔输入,有 80 株菌,每株菌有 60 个特征。

第二段信息是缺项的行列位置。它的书写格式是:扩展 FS  $m_1; n_1; m_2; n_2; \dots; m_n; n_n$ ; 送毕。其中  $m_1, m_2, \dots, m_n$  表示每个  $nc$  对应的行号,  $n_1, n_2, \dots, n_n$  表示每个  $nc$  对应的列号,每一对  $m_i, n_i$  表示一个  $nc$  项的位置。特征表以十进制格式书写,或特征表没有  $nc$  项而用八进制格式书写,均不需加第二段信息,每个“送毕”后不得加任何符号(如“,”、“;”等)。

5. 特征表的穿孔。

十进制特征表的穿孔格式是:先写上“扩展 FS”,然后将特征表的全部数据自左向右按行的顺序,每个字号加“;”号,最后加“送毕”。例

如特征表为:

101.....10  
101.....30  
103.....00  
011.....00

穿孔格式即为:

扩展 FS 1; 0; 1; .....1; 0;  
1; 0; 1; .....3; 0;  
1; 0; 3; .....0; 0;  
0; 1; 1; .....0; 0; 送毕。

八进制特征表的穿孔格式是先写“扩展 F3”，然后按前述八进制格式书写，最后加“送毕”。

例如: 扩展 F3 7324 1025 4113 5416;  
1005 4432 1544 1233;  
5004 1772 0;  
4324 1750 2642 1344;  
5005 4344 2543 1003;  
5142 1743 2; 送毕

将十进制特征表化为八进制数的优点在于减少穿孔的工作量。对同一特征表而言,可减少4.6倍工作量。而且便于书写及数值检查,从而大大减少人为的错误几率。

#### 四、纸带穿孔

方法详见中国科学院计算技术研究所和计算中心编:《013 BCY 上机手册》。

#### 五、上机操作

步骤详见《013 BCY 上机手册》。

#### 六、输出打印格式

##### 1. 十进制特征表:

以“0”、“1”形式连续输出(如有  $nc$  项,则该项的位置以“3”表示),直至输入的  $M \times N$  个数据全部打印完毕为止。

##### 2. Ssm 值表:

取三位有效数字,省去小数点及百分比号,第一株菌的头一个 Ssm 为“0”,第二株菌的头两个 Ssm 为“0”,依此类推,一直打印到  $(M - 1)$  株菌(该菌有  $M - 1$  个 Ssm 值打印成“0”),最后打印的一个 Ssm 值,应是  $(M - 1)$  株菌之间诸 O. T. U. 相比的结果。

##### 3. 按 Ssm 大小值排列:

输出格式为 \* BDS = Ssm, P = 第 M 行, Q = 第 N 列,直至 Ssm 三角矩阵全部印完为止。

4. 菌株编号: 将已形成单连锁群的菌株,按连锁水平的顺序,自大而小逐级打印本水平内可划分簇群的数目及各簇群内包括的菌株编号。

最后,打印输出形成单连锁群后菌株的 Ssm 值表,其格式同 2。

#### 参 考 文 献

- [1] Sokal, R. R. AND P. H. A. Sneath: *Principles of Numeric Taxonomy*, W. H. Freeman & Co., 1963, pp. 175—185.
- [2] Lockhart, W. R. and J. Liston: *Methods for Numerical Taxonomy*, American Society for Microbiology, Maryland, 1970, pp. 27—52.
- [3] Sneath, P. H. A.: *Methods in Microbiology* (ed. by Norris, J. R. and D. W. Ribbons), Academic Press, London and New York, 1972, pp. 29—98.
- [4] 徐浩: 微生物学通报, 1(3): 32—35, 1974.
- [5] 徐浩: 微生物学报, 15(1): 34—36, 1975.
- [6] Skerman, V. B. D.: *A Guide to the Identification of the Genera of Bacteria*, 2nd ed., 蔡妙英等译:《细菌属的鉴定指导》, 科学出版社, 北京, 1978, 第 3—13 页。