GNF-SYSTEM FOR BACTERIA IDENTIFICATION

The GNF-System to identify a group of common bacteria in Taiwan was developed by Professor Tzai in the Yang Ming Medical College in Taipei, Taiwan. He used the results of 18 primary tests as criteria to select the best matching bacterium among 54 different but closely related bacteria. The primary test data table (Blocks 152 to 155) shows the percentages that a test may show a positive result on a bacterium. If two or more bacterium show similar test results, supplemental tests may have to be carried out to identify the bacterium positively.

The method to match a set of test results to the table is to do a clustering analysis in a 18 dimension space. The speci men showing the shortest distance from the test result can thus be identified as the culprit.

Among the 18 tests, three are negative to all 54 bacteria. Only 15 tests are meaningful and their results can be very conveniently codes into a 16 bit integer. It specifies a unique location in the 18 dimension space and its distance from any of the 54 known specimen can be readily calculated. The 5 closest specimen will be picked up and displayed on the CRT terminal for the operator to decide whether a positive identification has bee reached or supplemental tests are to be done.

Blocks 150 and 151 has the names of all the bacteria properly ordered. Blocks 152 to 155 contain the primary test data and Blocks 156 to 159 contain the supplemental test data. Only the primary test data are used in the present program.

150 LIST

(GNF-SYSTEM ORG-NAMES, CHT, 12-14-81)

EIKENELLA CORRODENS	ACIN. CALCOACETICUS-ANITRATUS
ACIN. CALCOACETICUS-HAEMOLYTICUS	ACIN. CALCOACETICUS-ALCALIGENES
ACIN. CALCOACETICUS-LWOFFII	ACHROMOBACTER XYLOSOXIDANS
ACHROMOBACTER (GROUP VD-1)	ACHROMOBACTER (GROUP VD-2)
ALCALIGENES FAECALIS	ALCALIGENES ODORANS
ALCALIGENES DENITRIFICANS	BORDETELLA BRONCHISEPTICA
GROUP IV-E	GROUP IV-C2
AGROBACTERIUM RDIOBACTER (VD-3)	GROUP-IIK-2 (PSEUDOMONAS-LIKE)
GROUP-VA-1 (PSEUDOMONAS-LIKE)	GROUP-VE-1 (PSEUDOMONAS-LIKE)
GROUP-VE-2 (PSEUDOMONAS-LIKE)	FLAVOBACTERIUM MENINGOSEPTICUM
FLAVOBACTERIUM ODORATUM	FLAVOBACTERIUM BREVE
GROUP-IIB (FLAVOBACTERIUM-LIKE)	GROUP-IIF (FLAVOBACTERIUM-LIKE)
GROUP-IIJ (FLAVOBACTERIUM-LIKE)	MORAXELLA LACUNATA
MORAXELLA NONLIQUEFACIENS	MORAXELLA OSLOENSIS
MORAXELLA PHENYLPYRUVICA	MORAXELLA ATLANTAE

151 LIST

(GNF-SYSTEM ORG-NAMES, CONT'D, CHT, 12-14-81)

MORAXELLA URETHRALIS	GROUP-MS
GROUP-M6	KINGELLA KINGAE
KINGELLA INDOLOGENES	KINGELLA DENITRIFICANS
PSEUDOMONAS AERUGINOSA	PSEUDOMONAS FLUORESCENS
PSEUDOMONAS PUTIDA	PSEUDOMONAS PSEUDOMALLEI
PSEUDOMONAS MALLEI	PSEUDOMONAS CEPACIA
PSEUDOMONAS STUTZERI	PSEUDOMONAS MENDOCINA
PSEUDOMONAS MALTOPHILIA	PSEUDOMONAS PUTREFACIENS
PSEUDOMONAS ALCALIGENES	PSEUDOMONAS PSEUDOALCALIGENES
PSEUDOMONAS DIMINUTA	PSEUDOMONAS VESICULARIS
PSEUDOMONAS ACIDOVORANS	PSEUDOMONAS TESTOSTERONI
PSEUDOMONAS PICKETTII	PSEUDOMONAS PAUCIMOBILIS

152 LIST

(GNF-SYSTEM PRIMARY-TEST DATA 1, CHT, 12-14-81)

0	0	0	0	1	1	1	1	1	1	90	99	1	1	1	99	1	1
0	0	0	0	71	58	1	1	1	1	1	1	1	99	1	1	1	1
0	0	0	0	80	18	1	1	1	1	1	1	1	99	1	1	1	1
0	0	0	0	90	12	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	11	21	1	1	1	1	1	1	1	1	1	1	1	1
0	0	0	0	94	1	1	99	4	3	1	99	1	92	4	96	71	1
0	0	0	0	72	99	1	99	99	13	1	99	1	99	99	99	1	1
0	0	0	0	99	99	1	99	99	13	1	99	1	99	99	99	1	1
0	0	0	0	95	1	1	99	1	1	1	99	1	1	1	48	1	1
0	0	0	0	99	2	1	99	1	1	1	99	1	1	1	1	1	1
0	0	0	0	90	26	1	99	7	1	1	99	1	1	1	99	99	1
0	0	0	0	95	99	1	99	27	33	1	99	1	1	1	99	1	1
0	0	0	0	10	99	1	90	99	1	1	99	1	1	1	99	67	1
0	0	0	0	95	99	1	99	1	1	1	99	1	1	1	1	1	1
0	0	0	0	99	99	1	99	99	1	1	99	1	99	99	99	1	9

153 LIST

(GNF-SYSTEM PRIMARY-TEST DATA 2, 12-14-81, CHT)

```
 0  0  0  0  1 99  1 32  9  1  1 99  1 99 99 30  1  1
 0  0  0  0 99 99  1 89  7  1  1 99  1 95 99 91 70  1
 0  0  0  0 99 51  1 97  6  1  1  1  1 99 99 79  1 85
 0  0  0  0 99 50  1 99 23  1  1  1  1 99 99  1  1  1
 0  0  0  0  1  4 99  1  1  1  1 99  1 99 99  1  1  1
 0  0  0  0  1 99  1  1 79  1  1 99  1  1  1  1  1  1
 0  0  0  0  1 10 99  1  1  1  1 99  1 50  1  1  1  1
 0  0  0  0  1  4 98  1 13  1  1 99  1 92 93 17  1  1
 0  0  0  0  1  1 99  1  3  1  1 99  1  1  1  1  1  1
 0  0  0  0  1 99 99  1  1  1  1 97  1  1  1  1  1  1
 0  0  0  0  1  1  1  1  1  1  1 99  1  1  1 99  1  1
 0  0  0  0  1  4  1  1  1  1  1 99  1  1  1 95  1  1
 0  0  0  0  1  1  1  1  1  1  1 99  1  1  1 25  1  1
 0  0  0  0  1 97  1  1 72  1  1 99  1  1  1 65  1  1
 0  0  0  0  1  1  1  1  1  1  1 99  1  1  1  1  1  1
```

154 LIST

(GNF-SYSTEM PRIMARY-TEST DATA 3, CHT, 12-14-81)

```
 0  0  0  0 99  1  1  1 23  1  1 99  1  1  1  1  6  1
 0  0  0  0  1  1  1  1  1  1  1 99  1  1  1  1  1  1
 0  0  0  0  1  1  1  1  1  1  1 99  1  1  1  3  1  1
 0  0  0 50  1  1  1  1  1  1  1 99  1 99  1  3  1  1
 0  0  0 99  1  1 99  1  1  1  1 99  1 99 99  1  1  1
 0  0  0  0  1  1  1  1  1  1  1 99  1 99 92 93 88  1
 0  0  0  0 96 12  1 96  1  3  1 99 95 97 92 40 50 99
 0  0  0  0 47 42  1 99  1  6  1 99  1 99 98 14  6 99
 0  0  0  0 78 51  1 99  1 15  1 99  1 99 98  1  1 98
 0  0  0  0 96 43  1 99  1  5  1 99  1 99 99 86 99 99
 0  0  0  0  1 17  1  1  1  5  1 67  1 99 99 50  1 99
 0  0  0  0 96 46  1 99  1  1 55 89  1 99 99 27  1  1
 0  0  0  0 50 16  1 99 39  5  1 99  1 99 91 66 99  2
 0  0  0  0 50 50  1 99 50  1  1 99  1 99 99 99 99 99
 0  0  0  0 81  1  1 99  1  1  1  2  1 99 99 41  1  1
```

155 LIST

(GNF-SYSTEM PRIMARY-TEST DATA 4, CHT, 12-14-81)

```
 0  0  0 99 55  7  1 99  1  8 99 99  1 99 66 89  1  1
 0  0  0  0 75 22  1 99 19  1  1 99  1  1  1 46  1  1
 0  0  0  0 63  1  1 99 21  1  1 99  1  1 99 94  1 18
 0  0  0  0 50  1  1 99  3  3  1 99  1  1  1  3  1  1
 0  0  0  0 50  1  1 99  1  1  1 99  1 32  5  5  1  1
 0  0  0  0 48  1  1 99  3  1  1 99  1  1 99 86  1  1
 0  0  0  0 48  1  1 99  1  1  1 99  1  1  1 83  1  1
 0  0  0  0 99 99  1 99 41  1  1 99  1 99 99 99 99  1
 0  0  0  0 60  1  1 85 21  1  1 92  1 99 99  1  1  1
```

156 LIST

```
( GNF-SYSTEM SUPPLEMENTAL-TEST DATA 1, CHT, 12-14-81)
 1  1  1  1  1  1  1  1  1  1  1  1  1 50  1  0
99  1 67 99 99  1  1  1  1  1  3 99  0  1  0
93  1 67 99 99  1  1 98  1  1  4 99  0 99 99
 1  1  1  1 99  1  1 94  1  1  1 99  0 99 99
 1  1  1  1 99  1  1  8  1  1  1 98  0  1  0
 1  1  1 99 96  1  1  1  1  1 99 99  9  1  0
 1  1  1 99 50  1  1  1  1  1  1 99  9  1  0
 1 99 99 99 75 99  1  1  1  1  1 99  9  1  0
 1  1  1  1 67  1  1  1  1  6 27 98  9  1  3
 1  1  1  1 99  1  1  1  1  1 99 99  9  3  0
 1  1  1  1 67  1  1  1  1  1 99 99  9  3  0
 1  1  1  1  1  1  1  1  1  1  1 99  9  1 20
 1  1  1  1  1  1  1  1  1  1  1 43 91  1  0
 1  1  1  1 99  1  1  1  1  1  1 80  9  1  0
99 99 99 99 99 99  1  1  1  1  1 99  9  1 20
```

157 LIST

```
( GNF-SYSTEM SUPPLEMENT-TEST DATA 2, CHT, 12-14-81)
99  1 99 95  1 70 70  1 99  4  1 99  1  1  0
85  1 85 99 99  1  4  1  1  1  7 98  1 99  0
 1 99 99 99 99 99 22  1 99  1  1 99 71 44  0
 1 90 98 99 99  1 54  1 10  1  1 99  1  8  0
40 96 96  1  1 99  1  1 99 99  1 49  0 99  0
 1  1  1  1  1  1  1  1  1 93  1 98  0 99  0
 1  1 99  1  1  1  1  1  1 99  1 80  0 99  0
 3 10 93 40  1 99 99 99 35 99  1  1  0  1 12
 1  1  1  1  1  1  1  1  1  4  1  1  0 99  0
 1  1  1  1  1  1  1  1  1  1  1  1  0 99  0
 1  1  1  1  1  1  1  1  1  1  1  1  0 98  0
 1  1  1  1  1  1  1  1  1  1  1 10  0  1  0
 1  1  1  1 99  1  1  1  1  1  1 40  0  1  0
 1  1  1  1  1  1  1  1  1  1  1 81  0  1  0
 1  1  1  1  1  1  1  1  1  1  1 99  0  1  0
```

158 LIST

```
( GNF-SYSTEM SUPPLEMENTAL-TEST DATA 3, 12-14-81)
 1  1  1  1 99  1  1  1  1  1  1 99  0  1  0
 1  1  1  1  1  1  1  1  1  1  1 84  0  1  0
 1  1  1  1  1  1  1  1  1  1  1 50  0  1  0
 1  1 95  1  1  1  1  1  1  1  1 33  0  1 99
 1  1 99  1  1  1  1  1  1  1  1  1  0  1  0
 1  1  1  1  1  1  1 50  1  1  1  1  0  1  0
 1 69  4 82 96  1  1 60  3 11 38 99  1 51 43
16 92 41 98 99  1  5 92  2  1  1 99 71 99 17
17 22 24 97 99  1  7  1  2  1  4 99 71  1  1
99 99 99 86 99 57  1 86  1  1  1 99 71 99 53
99 83 99  1 99  1  1 50  1  1  1 99  0  1 50
98 99 96 98 99 64  1 43 80  1 64 93 71 70  6
 1 67 99 91 99  1 93 10  1  1  1 99  1  1  0
 1  1  1 99 99  1  1  1  1  1  1 99  1  1  0
99  1 99 54  2 99  1  1 95 99  1 99 71 99  0
```

ADJ,	Reduce the stack number to the range of 0-10 and compile it into the dictionary, building a table.
TABLE	Compile a screenful of data into the dictionary.
PRIMARY	The data table containing all the primary test data.
.NAME	Given a number on stack, pick the name of the bacterium in Blocks 150-1 and print it on CRT.

DISTANCE	(sum addr1 addr2 count ---) Two vectors of count length are stored in addr1 and addr2. Subtract each pair of elements and sum the squares of the differences in the location of sum.
ADDR	From the specimen number, find the first entry in the data table.
DISTS	The array containing 54 distances as calculated.
CLRDIST	Initialize the DISTS array to zero.
CLUSTER	Given a specimen number, calculate its distances from all the other specimen. This shows how much they are different in this multidimensional space.
PRINT	Print the contents of the DISTS array.

```
( TEST DATA TABLES, CHT, 12-14-81)
: ADJ,   5 + 10 / , ;
: TABLE ( SCREEN --- )
     999 SWAP LOAD   BEGIN DUP 999 - IF ADJ, AGAIN DROP ;

VARIABLE PRIMARY
155 TABLE 154 TABLE 153 TABLE 152 TABLE

: .NAME ( N --- )   30 /MOD 150 + BLOCK SWAP 2+ 32 * +
      32 3 SPACES TYPE ;
```

```
( P@, S@, CHT, 12-15-81)
: DISTANCE ( ADDR1 ADDR2 ADDR3 COUNT --- )    2*
     OVER + SWAP DO DUP @ I @ - DUP * >R OVER R> SWAP +!
     2+   2 +LOOP 2DROP ;
: ADDR  ( ROW --- ADDR )
     53 SWAP - 36 *   PRIMARY 2+ + ;
VARIABLE DISTS  216 ALLOT
: CLRDIST   DISTS 216 ERASE ;
: CLUSTER  ( ROW --- )                    CLRDIST
     54 0 DO  DISTS I 4 * +   OVER ADDR I ADDR 18 DISTANCE
     LOOP DROP ;
: PRINT   CLUSTER 54 0 DO I 10 MOD 0= IF CR I 5 U.R THEN
      DISTS I 4 * + @  5 U.R LOOP ;
```

```
( BUBLE SORT, CHT, 1-12-82)
: ORDER   54 0 DO I 4 * DISTS +  2+  I SWAP
          ! LOOP ;

: +DATA   ( N --- ADDR )   2* 2* DISTS + ;
: >DATA   ( N1 N2 --- F)   SWAP +DATA @   SWAP +DATA @   < ;
: DX      ( N1 N2 --- , EXCHANGE D1 WITH D2)
          OVER +DATA 2@   2SWAP DUP +DATA 2@
          2SWAP >R +DATA 2!   R> +DATA 2! ;

: BUBLE   53 0 DO   54 I 1+ DO   I J >DATA
          IF I J DX THEN   LOOP LOOP ;
: 5PRINT  5 0 DO I +DATA 2@ ( DUP 10 < IF 2DROP LEAVE THEN )
          CR SWAP DUP 1+ 3 U.R .NAME 5 U.R LOOP ;
: SEARCH  1- CLUSTER ORDER BUBLE 5PRINT ;
```

ORDER	Append a ordered number to each entry in the DISTS array for the sorting purposes.
+DATA	Get the address of an DISTS entry.
>DATA	Return true if data at n1 is less than data at n2.
DX	Exchange 4 bytes of data at n1 with those at n2. Both the distance and the order tag are exchanged.
BUBBLE	Bubble sort the 54 distances in DISTS.
5PRINT	Print the name and the distance data in the top 5 entries in DISTS, after the sorting.
SEARCH	Print the 5 closestly related specimen of a specimen indicated by the stack number.

CODES	An array holding the decoded test data of a sample to be analyzed.
DECODE	Given the code representing the test results, bits set indicating positive result, fill CODES table with 10's and 0's.
COMPARE	Calculate the distance between the sample and all the 54 specimen. Fill the DISTS table.
SEARCH	Given the test code, identify and print out the closes five specimen on the CRT.

```
( COMPARE, CHT, 1-12-82 )
165 LOAD 166 LOAD 167 LOAD
VARIABLE CODES  108 ALLOT
: DECODE ( N --- )
     18 0 DO 2 /MOD SWAP IF 10 ELSE 0 THEN
     I 2* CODES + ! LOOP ;
: COMPARE     CLRDIST
     54 0 DO  DISTS I 4 * +   CODES I ADDR 18 DISTANCE
     LOOP DROP ;

: SEARCH ( N --- )     DECIMAL
      DECODE COMPARE ORDER BUBLE 5PRINT
       OCTAL  ;
```