The Combinatorial Decomposition of Stability Problems in Non-Statistical Classification Theory.

by

D. M. Jackson

L. J. White *

Technical Report

No. 70-87

December, 1970

Department of Computer Science
Cornell University
Ithaca, New York 14850

*Department of Computer and
Information Science
Ohio State University
Columbus, Ohio 43210

The Combinatorial Decomposition of Stability Problems In Non-Statistical Classification Theory

Chapter		Page
1	Introduction	1
2	The Definition of Stability	2
3	Algebraic Specification of the Error Matrix	4
4	Number of Tables	8
5	Algorithm for Computing $\phi^*(\underline{n})$ and Experimental Results	11
6	Monotonicity properties of $\phi^*(\underline{n})$	14
7	Convergence	19
8	The Diagonal Elements of an Error Matrix	23
9	The Behavior of $\overline{k}$ with respect to N, r and δ	28
10	Conclusions	30
	Acknowledgements	
	References	
	Tables I - VIII	

1. Introduction

Classification theory is concerned with the problem of resolving detailed information about the individuals of a population into generalizations about certain subsets of the population. Philosophically the problem is of great antiquity. The principles of classification are a necessary concomitant of any epistemological theory and provide the means by which general statements may be predicated of entities that comprise the real world. Classification theory accordingly comprises the techniques for constructing predictions and hypotheses from aggregations of individual facts relating to populations of objects.

Growing attention has been given to devising non-statistical classification algorithms for multivariate heterogeneous populations. Statistical techniques are avoided because they are either too costly or because an adequate statistical model for the population is not available. In a typical situation, a multivariate heterogeneous population of objects is given together with a characterization of each object in terms of observable attributes. The population is heterogeneous in the sense that it is conjectured that the population may be resolved into a number of intelligible classes. We assert that a class is intelligible if we are able to assign a name to the extension of the class using a terminology external to the process of construction, with the provision that the name specifies an intensional rule. Clearly, with the lack of a suitable underlying model for the population, an internal criterion for determining classes must suffice. The objective of the analysis is to determine a decomposition which will reveal the structure of the population and which will suggest further models for a more sophisticated treatment of the material (1).

2. The Definition of Stability

Non-statistical classification typically involves the use of similarity functions and classificatory criterion functions. A similarity function specifies a rule for computing the degree to which the characterizations of an arbitrary pair of objects conform to one another. Sokal and Sneath (2) have given a good account of the design of such functions. A criterion function specifies a rule for determining whether a particular collocation of objects is admissible as a class of the classification. Examples of specific criterion functions are given by Parker-Rhodes and Needham (3), Lerman (4), and Friedman and Rubin (5), among others. A group of objects is a class of the classifications if the criterion function is locally minimum, and constructing a class is equivalent to optimizing the criterion function locally over a set of vertices of a hypercube if the attributes are binary. The problem of optimizing a function globally over the vertices of a hypercube is encountered, for example, in hierarchical classification (Hartigan (6)), and is solved generally under the assumption that the solution is accessible in a sequence of steps, each locally optimal. Gleason (7) comments interestingly on problems of discrete optimization.

Before the results of a cluster analysis are accepted, it is appropriate to determine whether the results satisfy a number of necessary conditions which characterize, in part, the relationship between an adequate classification and the original population. One such condition is that of stability. Instability, or sensitivity of the classification to small errors in the observations, is evidence that the objects have been insufficiently or inadequately characterized. The condition may be written:

$$\text{Given } \varepsilon>0, \exists\, \delta(\varepsilon)>0 . \Psi(\mathcal{D}+\delta\mathcal{D},\mathcal{D})<\delta \Rightarrow \Psi(\pi(\mathcal{D}+\delta\mathcal{D}),\pi(\mathcal{D}))<\varepsilon \qquad (2.1)$$

where $\mathcal{D}$ is a set of data, $\delta\mathcal{D}$ is an alteration to $\mathcal{D}$, π is a classification algorithm and Ψ is a function which measures the degree to which a pair of classifications or a pair of data sets conform with one another. The condition is clearly not sufficient since it does not ensure the adequacy of a particular configuration of classes in generalizing the detailed inter-object similarities. Indeed, assessment of the sufficiency, as opposed to the self-consistency, of a classification is logically external to the construction of the classification and utilizes categories of information which are inaccessible at the time of construction.

This paper is concerned with providing a technique for facilitating the examination of stability. Since the multi-state or continuous attributes may be decomposed into binary attributes, the binary case is of sufficient generality to examine in detail. Subsequent analysis will provide expressions for the expectation ϕ^* of statistical functions ϕ (mean, variance, moments, factorial moments) of the similarity f between arbitrary pairs of objects (t_j,t_k) when the states of the N attributes are susceptible to error independently and equiprobably with an assignable probability r. A technique for computing the error matrix $E^{(N)}$ comprising the expectations of a given statistical function for each pair of objects will be given. The manner in which the error matrix is used to determine the stability of classifications generated by specific algorithms will not be given here. The quantification by Ψ of small changes in non-hierarchical classification has been discussed in detail elsewhere (Jackson (8)).

In a more general context, the problem may be regarded as one of comparing pairs of fixed length binary sequences, and is suggestive of the classical Matching Problem (Bernouilli (9), de Moivre (10)). ϕ^* may be regarded as the expectation of a generalized "match" between two arbitrary compositions of objects of two kinds, labelled zero and one, when there is a probability of misrecognizing the kind of an object. An account of methods of asymptotic approximation for large N, when the distribution is almost normal, is given by Barton (11).

As a matter of terminology, "classification" has been used here to mean "the construction of a set of classes". "Identification" is reserved for the alternative meaning of classification, namely, "the assignment of an object to an appropriate class of a given classification". The task of identification might also permit predicitions or inferences to be made about the identified object, and accordingly has sometimes been called "extrapolation". Thus, from a knowledge of a classification and the attributes of a particular object we might be able to infer (predict, extrapolate) that the object possesses further attributes, hitherto unobserved. Fine (12) has axiomatized the non-statistical continuous attribute case for a model in which the prime objective is extrapolation. In the model presented here, extrapolation to a set of (binary) attributes is envisaged.

3. <u>Algebraic Specification of the Error Matrix</u>

It may be shown by the central limit theorem that for large N, ϕ^* is approximately normal. However, we are concerned here with determining

ϕ^* with prescribed accuracy for intermediate values of N for which normality does not occur.

Suppose that a multivariate heterogeneous population consists of M objects $\{t_i\}$, and that the characterizations of these objects are provided by composition from a population of N binary attributes. Define $\mathcal{D}$ to be the set of characterizations of all objects in the population. Accordingly, the object characterization of t_j is given by the j-th. column $D_{.j}$ of D. $D_{ij}=1$ implies that t_j possesses attribute a_i, while $D_{ij}=0$ implies that it does not. Denote by a the (unordered) pair of attribute states (0,0) where for some attributes a_i, t_j has state 0 and t_k has state 0. Similarly, denote by x and d the paired combinations of states (1,0)V(0,1) and (1,1) respectively. For the pair of objects (t_j,t_k), let n_1,n_2,n_3 be numbers of attributes having the combinations of states a,x,d respectively. For brevity, denote (n_1,n_2,n_3) by $\underline{n}$. Let f_{jk} be the similarity between an arbitrary pair of objects t_j,t_k.

An error occurs if an attribute state s_1 is recorded as s_2. The characterizations $D_{.j}$ and $D_{.k}$ are scrambled into $D'_{.j}$ and $D'_{.k}$. Let m_1,m_2,m_3 and $\underline{m}$ be quantities, analogous to the n's, developed from the scrambled characterizations. The value of f_{jk} after scrambling is $f(\underline{m})$. Let P(u,v) be the probability of the transition $v\to u$, where u and v are each one of a,x,d. The transition probability matrix $\mathcal{P}$ is given by:

$$\mathcal{P} = \begin{bmatrix} s^2 & rs & r^2 \\ 2rs & r^2+s^2 & 2rs \\ r^2 & rs & s^2 \end{bmatrix} \qquad \text{where } s=1-r$$

It may be shown (Jackson (13)) that $\phi^*(\underline{n})$ is expressible in the following form, involving the product of three interdependent multinomial distributions:

$$\phi^*(\underline{n})=\sum_{I} \frac{\prod_i n_i}{\prod_{i,j} I(i,j)!} \cdot \prod_{i,j} P(i,j)^{I(i,j)} \cdot \phi(\underline{m},\underline{n}) \tag{3.1}$$

where the sum is taken over all tables, I, with non-negative integer entries and prescribed row sums $\underline{n}$. $\underline{m}$ is given by the column sums. Both row sums and column sums sum to N. Products are taken over (1,2,3). This form is equivalent to

$$\phi^*(\underline{n}) = s^{2N} \cdot \sum_{\underline{m}} \phi(\underline{m},\underline{n}).h(\underline{m},\underline{n}) \tag{3.2}$$

where

$$h(\underline{m},\underline{n})=s^{-2N}.\underline{\text{coeff. of}}\ \prod_{i=1}^{3} q_i^{m_i}\ \underline{\text{in}}\ \prod_{j=1}^{3} ((\sum_{i=1}^{3} q_i P(i,j))^{n_j})$$

and the q's are dummy variables.

On substitution of $\underset{\sim}{P}$, these reduce to the following forms which will be used extensively in the subsequent analysis:

$$\phi^*(\underline{n})=s^{2N}. \sum_{i=0}^{2(n_1+n_3)+n_2} t^i \sum_{\substack{I\epsilon I^*(\underline{n}) \\ k(I)=i}} (1+t^2)^{I_{22}}.K_I.\phi(\underline{m},\underline{n}) \tag{3.3}$$

where i)

$I^*(\underline{n})$ is the set of all tables such that $\sum_{i=1}^{3} I_{ij}=n_j$

ii)

$k(I)=N-\underline{\text{trace}}(I)+I_{13}+I_{31}$ (Table constant)

iii)

$m_i= \sum_{j=1}^{3} I_{ij}$, $\sum_{i=1}^{3} m_i=N$, and $\sum_{j=1}^{3} n_j=N$

iv)

$K_{\underline{I}}=s^{m_2-I_{22}}.\prod_{i=1}^{3} M(n_i:I_{i1},I_{i2})$ using the usual multinomial

notation for $n_i!/I_{i1}!I_{i2}!(n_i-I_{i1}-I_{i2})!$

or $\quad \Phi^*(\underline{N})=s^{2N}.\sum_{\underline{m}} h(\underline{m},\underline{n}).\Phi(\underline{m},\underline{n})$ (3.4)

where i) $h(\underline{m},\underline{n})=H_{\underline{m}}(\{t^2,2t,1\}^{n_1}.\{t,1+t^2,t\}^{n_2}.\{1,2t,t^2\}^{n_3})$

ii) $\{x_1,x_2,x_3\}=(x_1,x_2,x_3).(q_1,q_2,q_3)$

iii) $H_{\underline{m}}(X)=$ <u>coeff.</u> <u>of</u> $\prod_{i=1}^{3} q_i^{m_i}$ <u>in</u> X where X is a polynomial in the

q_i's with non-negative exponents.

In both of the above forms for $\Phi^*(\underline{n})$, $t=r/s$ and $s=1-r$.

A final form may also be mentioned (Jackson 13): $\Phi^*(\underline{n})=\sum_{\underline{i}} a_{\underline{i}}.B_{\underline{i}}$ (3.5)

where i) $\Phi(\underline{m},\underline{n})=a_{\underline{i}}(\underline{n}).\underline{m}^{\underline{i}}$

ii) $D_{ij}=P(i,j).d/dP(i,j)$ is a formal differential operator.

iii) $B_{\underline{i}}=\prod_j((\sum_k D_{jk})^{i_j}).\prod_k((\sum_j P(j,k))^{n_k})$

It is anticipated that this expression may be expanded quite readily by a suitable algebra system (Collins (14),(15)) to give an analytic solution for each element of the error matrix. The form results from a method related to the symbolic method of MacMahon (16) and its extensions (Kaplansky and Riordan (17)) for the classical Matching Problem referred to in Section 2.

Since the motivation of this analysis is the question of the stability of classifications, chief attention will be confined to a particular Φ, namely $\Phi=|f(\underline{n})-f(\underline{m})|$. The relevance of this function to the question of stability has been discussed elsewhere (Jackson and White (18)). In addition, attention will be confined to the Jaccard similarity function $n_3/(n_2+n_3)$, although comparable results will hold for the other functions.

4. Number of Tables

The evaluation of $\phi^*(\underline{n})$ involves summation over all valid tables, I, such that the row sums are given by $\underline{n}$. Bell (19) has computed the number of matrices for the special case of generalized stochastic matrices, with $\underline{m}=\underline{n}=(n,n,n)$, to be $(n+1)(n+2)(n(n+3)+4)/8$. The number of possible vectors $\underline{n}$, for given N, is $M(N+2:2)$. The number, T(N) of valid tables is $M(N+8:8)$ and is the number of ways that N indistinguishable objects may be distributed among nine cells. T(N) grows on the order of N^8.

The following theorems provide a bound on the number of tables with table constant less than or equal to a given value. It will be seen in Section 5 that this bound may be used to estimate the reduction in time taken to compute $\phi^*(\underline{n})$, to prescribed accuracy, achieved by utilizing only tables with table constant less than or equal to a determinable value $\bar{k}$.

Theorem 4.1

Let $(T(\underline{n},k)$ be the number of tables, I, such that $k(I)=k$. Then an upper bound for $T(\underline{n},k)$ is:

$$T^*(\underline{n},k)=(k+2)(k+4)(k+6)(k^2+8k+10)/2.4.6.10 \qquad \text{if } k \text{ even}$$

$$\text{or} \quad (k+1)(k+3)(k+4)(k+5)(k+7)/2.4.6.10 \qquad \text{if } k \text{ odd}$$

$T(\underline{n},k)=T^*(\underline{n},k)$ when $k\leq\min(n_1,n_2,n_3)$.

Proof From (3.3), the table constant is given by $k(I)=N-\underline{\text{trace}}(I)+I_{13}+I_{31}$. To determine $T(\underline{n},k)$, a number of cases must be considered.

First consider the case:

$$k \leq \min(n_1, n_2, n_3) \tag{4.1}$$

Define $j=I_{13}+I_{31}$, so $k-2j=I_{12}+I_{21}+I_{23}+I_{32}$. There are $(j+1)$ non-negative integer solutions to the first equation and $M(k-2j+3:3)$ non-negative integer solutions to the second. The latter is the number of ways of distributing $(k-2j)$ indistinguishable objects among four cells. Then, denoting $T(\underline{n},k)$ by $T^*(\underline{n},k)$ in this case, we have

$$T^*(\underline{n},k) = \sum_{j=0}^{\lfloor \frac{1}{2}k \rfloor} (j+1).M(k-2j+3:3) \tag{4.2}$$

$$= (k+2)(k+4)(k+6)(k^2+8k+10)/2.4.6.10 \qquad \text{if } k \text{ even}$$

$$\text{or } (k+1)(k+3)(k+4)(k+5)(k+7)/2.4.6.10 \qquad \text{if } k \text{ odd}$$

where $\lfloor x \rfloor$ denotes the integer part of x. The remaining cases involve additional constraints on the rows of I and therefore result in fewer tables. (4.2) serves as an upper bound in the cases in which (4.1) does not apply.

Theorem 4.2

Let $R(N,\bar{k})$ be the proportion of the feasible tables with $k(I) \leq \bar{k}$ to the number of feasible tables with $k(I)$ unrestricted. Then, an upper bound on $R(N,\bar{k})$ is given by:

i) $R^*(N,\bar{k})=M(N+2:2).(M(N+8:8))^{-1}.(T^*(\underline{n},1)+T^*(\underline{n},2)+...+T^*(\underline{n},\bar{k}))$

ii) $R^*(N,\bar{k})$ is asympototic to $R(N,k)$ in N

Proof

i) We know from Theorem 4.1 that $T^*(\underline{n},k)$ is independent of $\underline{n}$, and is accordingly valid for all $\underline{n}$. There are $M(N+2:2)$ such vectors $\underline{n}$. The number of feasible tables for $k(I)$ unrestricted is $T(N)$, which is equal to $M(N+8:8)$. $T^*(\underline{n},k)$ is an upper bound on the number of tables with $k(I)=k$. Hence (i) follows.

(ii) We shall prove that an upper bound for the number T_I of tables incorrectly included in $T^*(\underline{n},k)$ may be made arbitrarily small compared with the number of tables to be considered for each value of k. Consider the three cases **a)** $n_2<k$; b) $n_1<k$; and c) $n_3<k$. One or more of these conditions must hold for invalid tables to occur.

Case a) $n_2<k$ Allocate a_2 of the k objects to the second row. Invalid tables occur for $a_2=k-q$, when $0\le q\le k-n_2-1$, from $a_2\ge n_2+1$. There are $k-q+1$ different ways to allocate these objects to the second row. Let a_1 be the number of the q objects allocated to the first row. Then there remain

$$\sum_{a_1=0}^{q} (\lfloor a_1/2\rfloor+1)(\lfloor(q-a_1/2)\rfloor+1)$$ different ways of allocating the q

objects between the first and third rows. Thus, the total number of invalid tables is

$$\sum_{q=0}^{k-n_2-1} (k-q+1) \cdot \sum_{a_1=0}^{q} (\lfloor a_1/2\rfloor+1)(\lfloor q-a_1)/2\rfloor+1)$$
$$< (k+1)(k-n_2)^2(k-n_2+4)^2/16$$

Case b) $n_1<k$: By arguments similar to the above, the number of invalid tables is

$$\sum_{q=0}^{k-n_1-1} (\lfloor(k-q)/2\rfloor+1) \cdot \sum_{a_3=0}^{q} (\lfloor a_3/2\rfloor+1)(q-a_3+1)$$
$$< (k+2)(k-n_1)^3(k-n_1+2)/8$$

Case c) $n_3<k$: Identical to case (b) by symmetry of n_1 and n_3.

Summing the results for the three cases we have:

$$T_I < 2(k+2) \sum_{n_1=0}^{k-1} (k-n_I)^3 (k-n_1+2)/8 + (k+1) \sum_{n_2=0}^{k-1} (k-n_2)^2 (k-n_2+4)^2/16$$

$$< 5k^3 (k+1)(k+2)^2/16$$

Some of the tables have been counted more than once, but this simply weakens the upper bound on T_I. The number of tables with table constant k is bounded by $(N+1)(N+2)T^*(\underline{n},k)/2$ which includes the invalid tables. The ratio of the upper bound on T_I to this total may be made arbitrarily small for sufficiently large N since this ratio is equal to $2T_I/((N+1)(N+2)T^*(\underline{n},k))$

$$< \frac{300k}{(N+1(N+2)} \cdot \frac{(k+1)(k+2)}{k^2+8k+10} \cdot \frac{k}{(k+4)} \cdot \frac{k}{(k+6)} \qquad \text{if k even (from Theorem 4.1)}$$

$$\underline{\text{or}} \quad < \frac{300k}{(N+1)(N+2)} \cdot \frac{k+2}{k+3} \cdot \frac{k+2}{k+4} \cdot \frac{k}{k+5} \cdot \frac{k}{k+7} \qquad \text{if k odd (from Theorem 4.1)}$$

This proves the theorem.

5. Algorithm for computing $\phi^*(n)$ and experimental results.

The computation time required for calculating $\phi^*(\underline{n})$ depends on the number of tables considered, and the latter increases on the order of N^8. Much of the work in providing an efficient algorithm is centered around reducing the number of tables utilized while yielding $\phi^*(\underline{n})$ to prescribed accuracy. The algorithm has been implemented in FORTRAN and its salient steps are described briefly below.

The first step is to read the values of the error rate, r, the number of attributes N, and the stopping criterion δ. The value of r was fixed at 0.05 and δ at 10^{-3} for most of the experiments.

The second step is the initialization of the parameters which determine the tables. There are eight independent parameters, and these uniquely determine a table. The following eight constraints were used as bounds on the looping parameters:

$$
\begin{aligned}
&0 \le n_2 \le N \\
&0 \le n_1 \le N-n_2 \\
&0 \le I_{13} \le \min(n_1, \lfloor k/2 \rfloor) \\
&0 \le I_{12} \le \min(n_1-I_{13}, k-2I_{13}) \\
&0 \le I_{31} \le \min(n_3, \lfloor (k-2I_{13}-I_{12})/2 \rfloor \\
&0 \le I_{32} \le \min(n_3-I_{31}, k-2I_{13}-2I_{31}-I_{12}) \\
&0 \le I_{21} \le \min(n_2, k-2I_{13}-2I_{31}-I_{12}-I_{32}) \\
&0 \le I_{23} \le \min(n_2-I_{21}, k-2I_{13}-2I_{31}-I_{12}-I_{32}-I_{21})
\end{aligned}
\tag{5.1}
$$

The following quantities are then determined:

$$
\begin{aligned}
&n_3=N-n_1-n_2 && m_1=I_{11}+I_{21}+I_{31} \\
&I_{11}=n_1-I_{13}-I_{12} && m_2=I_{12}+I_{22}+I_{32} \\
&I_{33}=n_3-I_{31}-I_{32} && m_3=I_{13}+I_{23}+I_{33} \\
&I_{22}=n_2-I_{21}-I_{23}
\end{aligned}
\tag{5.2}
$$

The column sums are determined by these conditions. Accordingly, the rows of I are independent (uncoupled) and the validity of these conditions follows immediately. The conditions are necessary and sufficient for the existence of tables (generate all and only feasible tables)with specified k.

The third step is the determination of $\overline{k}$. The parameter k is incremented in an outer loop until, at some value $\overline{k}+1$, the admission of further tables contributes less than δ to the largest element of the error matrix. Determination of the largest element is given in Section 6. A proof that the convergence of this element controls the convergence of all other elements of the matrix, for the same value of $\bar{k}$, is given in Section 7.

The fourth step, given a unique table, is the computation of the contribution of that table to the summation ϕ^*. Logarithms are introduced to avoid scaling problems.

The fifth step is to increment and test the parameters. If the value of each parameter is less than its upper bound, a branch is executed to compute a new table.

Table I shows several elements of the matrix $E^{(10)}$ for successive values of $\bar{k}$ and r=0.05. It is seen that sufficient accuracy is given by $\bar{k}=5$. A proof of uniform convergence for all $\underline{n}$ for the same $\bar{k}$ is given in Section 6, and it is clear that $\phi^*(\underline{n})$ may be approximated by limiting the value of $k \leq \bar{k}$. The behavior of $\bar{k}$ with respect to N and r is given in Table II.

Table III indicates the reduction in computation time obtained by restricting k for several values of $\bar{k}$ and r=0.05. The first two columns give N and $\bar{k}$ respectively. Only tables with $k \leq \bar{k}$ are constructed. Column three gives the upper bound developed in Section 4 for the ratio of the number of valid tables with $k \leq \bar{k}$ to the total number of tables. Column five gives the computation time on the IBM 360/75, and column four the "time-saving ratio", or the ratio of the computation times with $k \leq \bar{k}$ and k unrestricted. Since the program was run in a multiprocessing environment these times include interrupts and wait states, and are therefore not reproducible. They serve as upper bounds on the amount of computation time required. The last line gives the computation time when k is not restricted. The ratio of valid tables provides a good estimate of the time-saving ratio, and is proposed for this use. Finally, column six gives an indication of the accuracy of the truncated calculation by specifying the error in $\phi^*(29,0,1)$ in the third significant figure.

For $E^{(10)}$ the single element corresponding to $\underline{n}=(99,0,1)$ has been computed. For $\bar{k}=21$ and $\bar{k}=22$, the estimated values of $\phi^*(\underline{n})$ are 0.883 and 0.884 respectively.

Sufficient accuracy is therefore obtained with $\bar{k}=21$. The time required to compute this element was approximately 0.15, so that the time required for the complete error matrix is approximately 15 mins. Without table truncation the time is estimated to be approximately three years.

Table IV shows $E^{(10)}$ for $r=0.05$, and the values of $\phi^*(\underline{n})$ are given in tabular form by specifying n_1 and n_3. If the matrix is presented in this way, the following characteristics emerge. These characteristics have also been observed for other values of N and r.

i) If $n_3 \neq 0$, then $\phi^*(\underline{n})$ is strictly monotonically increasing with n_1 and n_3. The largest values occur on the skew-diagonal.

ii) If $n_3=0$, then $\phi^*(\underline{n})$ is monotonic decreasing with n_1.

iii) The maximum $\phi^*(\underline{n})$ occurs when $\underline{n}=(9,0,1)$. In addition, if $n_1 \neq 10$ then $\phi^*(\underline{n})$ is strictly monotonically increasing with n_1 along the skew-diagonal. This is observed in general, the maximum occuring at $\phi^*(N-1,0,1)$.

6. Monotonicity properties of $\phi^*(\underline{n})$.

The monotocity of $\phi^*(\underline{n})$ with respect to n_1 and n_3 as independent variables has been suggested by the empirical evidence given in Section 5. The behavior of $\phi^*(\underline{n})$ is characterized by the following theorem.

Theorem 6.1

For $t \in (0,t_0)$ for some t_0, then:

i) $\phi^*(\underline{n})$ is strictly monotonically increasing with respect to n_1 and n_3 as independent variables (if $n_3 \neq 0$).

ii) $\phi^*(n_1,0,n_3)$ is strictly monotonically increasing with respect to n_1 (if $n_3 \neq 0$).

iii) $\phi^*(N-1,0,0) > \phi^*(N,0,0)$

iv) $s^{-2N}.\phi^*(n_1,n_2,0)=t+0(t^2)$

Proof

i) From (3.3), $s^{-2N}.\phi^*(\underline{n})$ is a polynomial of degree 2N in t, and is therefore dominated by its constant and linear terms, for t sufficiently small, i.e. $t \in (0,t_0)$. Let $X=n_2+n_3\neq 0$. The case $n_2=n_3=0$ is excluded for special treatment since $f(\underline{n})$ is undefined at (N,0,0). The constant term is given by $\underline{n}=\underline{m}$, which gives zero contribution to $\phi^*(\underline{n})$. The linear term is given by the four tables I_1, I_2, I_3, I_4 whose column sums are $\underline{m}_1=(n_1+1,n_2-1,n_3)$, $\underline{m}_2=(n_1-1,n_2+1,n_3)$, $\underline{m}_3=(n_1,n_2+1,n_3-1)$ and $\underline{m}_4=(n_1,n_2-1,n_3+1)$ respectively. Thus:

$$s^{-2N}.\phi^*(\underline{n})=(2n_1n_3/(X(X+1))+2n_2n_3/(X(X-1))+n_3/X+1)t+0(t^2)$$

for $n_2\neq 0$, then:

$$s^{-2N}.(\phi^*(n_1,n_2-1,n_3+1)-\phi^*(n_1,n_2,n_3))$$

$$= (1/X).(2n_1/(X+1)+2(n_2+1)/(X-1)+1)t+0(t^2)$$

which is positive, so $\phi^*(\underline{n})$ is strictly montonically increasing wrt n_3. Now, assume that $n_1\neq 0$ and $n_3\neq 0$. Then:

$$s^{-2N}(\phi^*(n_1,n_2,n_3)-\phi^*(n_1-1,n_2+1,n_3))$$

$$= (2n_3/((X-1)X(X+1)(X+2))).(2n_1(X-1)+n_2(X+2)+X(X-1))t+0(t^2)$$

which is positive, so $\phi^*(\underline{n})$ is strictly montonically increasing wrt n_1.

ii) If $n_3\neq 0$, then:

$s^{-2N}\phi^*(n_1,0,n_3)=(2+(2n_1/(n_3+1)))t+0(t^2)$ which is positive so $\phi^*(n_1,0,n_3)$ is strictly monotonically increasing wrt n_1.

iii) $s^{-2N}\phi^*(N-1,0,1)=(N+1)t+0(t^2)$ and $\phi^*(N,0,0)=0(t^2)$ whence $\phi^*(N-1,0,1)>\phi^*(N,0,0)$.

iv) If $n_3=0$, valid tables which result in linear terms in t are given by the column sums $(n_1-1,n_2+1,0)$, $(n_1+1,n_2,0)$ and $(n_1,n_2-1,1)$. $\phi(\underline{m},\underline{n})$ is zero for the first two tables and is equal to $1/n_2$ for the third table. The value of K_I for this table is n_2. Accordingly, $s^{-2N}.\phi^*(n_1,n_2,0)=t+0(t^2)$.

The following theorem estimates the value of t_0 for certain elements.

Theorem 6.2

If $t\in(0,N^{-1/2})$ and $\phi(\underline{m},\underline{n})>0$ for all $\underline{m},\underline{n}$ then $\phi^*(n_1,0,n_3)$ is a monotonically increasing function of n_1.

Proof

From (3.4), we know that $\phi^*(n_1,0,n_3)-\phi^*(n_1-1,0,n_3+1)=$

$$s^{2N}.(1-t^2).\sum_{\underline{m}} \phi(\underline{m},\underline{n}).H_{\underline{m}}(\{-1,0,1\}\cdot\{1,2t,t^2\}^{n_3}.\{t^2,2t,1\}^{n_1-1}) \quad \text{(by}$$

extracting common factors)

$$=s^{2N-1}.(1-t^2).\sum_{\underline{m}} \phi(\underline{m},\underline{n}).t^{2m_1+m_2}.\sum_{\underline{j}} M(n_1-1:\underline{j})M(n_3+1:\underline{m}-\underline{j}).t^{2(j_3-j_1)}.$$

$(t^{-2}(m_3-j_3)-(m_1-j_1))/(n_3+1)$ (by replacing $H_{\underline{m}}(\{-1,0,1\}.(....))$ by $H_{m_1,m_2,m_3-1}-H_{m_1-1,m_2,m_3}.(....)$ and then removing common factors)

This is non-negative if $t^2\leq\min_{\underline{m}}((m_1-j_1)/(m_3-j_3)$ where $m_1>j_1,m_2\geq j_2$, $m_3>j_3$.

$=1/N$ and theorem follows.

We have assumed that m_1,m_3,m_1-j_1,m_3-j_3 are non-zero in certain cases. We may show by comparable analysis that the terms for which some of these are zero donot affect the conclusion. This is omitted for brevity.

We now derive another monotonicity relation which will be used in Section 7 to prove convergence. The following definitions are needed:

$$\left.\begin{aligned} &\text{Let } P_i(\underline{n}) = \sum_{\substack{I\in I^*(\underline{n})\\ k(I)=i}} K_I \\ &w_k(\underline{n}) = s^{2N}.\sum_{i=0}^{k} t^i.P_i(\underline{n}) \\ &\quad\text{(the partial sums of } w(\underline{n})\text{)} \\ &w(\underline{n}) = w_k(\underline{n}) \text{ at } k=2(n_1+n_3)+n_2 \end{aligned}\right\} \tag{6.1}$$

Comparing $w(\underline{n})$ with the probability weights for the distribution of $\phi(\underline{m},\underline{n})$ given in (3.3), we see that $w(\underline{n})\leq 1$, equality holding for $t=0$. The following theorem characterizes the monotonicity of $P_i(\underline{n})$ and $w_i(\underline{n})$.

<u>Theorem 6.3</u>

If $\underline{n}_A=(n_1,n_2,n_3)$ and $\underline{n}_B=(n_1+n_2,0,n_3)$, $n_2\neq 0$ then:

$$P_i(\underline{n}_B)>P_i(\underline{n}_A) \text{ and } w_i(\underline{n}_B)>w_i(\underline{n}_A)$$

<u>Proof</u>

From (6.1):

$$w(\underline{n}_B) - w(\underline{n}_A) = s^{2N}.\sum_{i=0}^{2N} t^i.(P_i(\underline{n}_B)-P_i(\underline{n}_A)) \tag{6.2}$$

where $P_i(\underline{n}_A)=0$ for $2(n_1+n_3)+n_2+1\leq i\leq 2N$.

The following parametrisation decomposes the summation over $I^*(\underline{n}_A)$ and $I^*(\underline{n}_B)$ so that summations common to $I^*(\underline{n}_A)$ and $I^*(\underline{n}_B)$ may be factored. Let $I\in I^*(\underline{n}_A)$ and $J\in I^*(\underline{n}_B)$ be arbitrary tables such that:

i) $k(I)=k(J)=k$

ii) $I_{3i}=J_{3i}$ for $i=1,2,3$.

iii) $I_{13}=J_{13}=a_1$

Then the tables I and J are given as follows:

$$\begin{array}{|ccc|c} * & b_2 & a_1 & n_1 \\ d_1 & * & d_2 & n_2 \\ I_{31} & I_{32} & * & n_3 \\ \hline \end{array} \qquad\qquad \begin{array}{|ccc|c} * & a_2 & a_1 & n_1+n_2 \\ 0 & * & 0 & 0 \\ I_{31} & I_{32} & * & n_3 \\ \hline \end{array}$$

$I \in I^*(\underline{n}_A)$ $\qquad\qquad\qquad$ $J \in I^*(\underline{n}_B)$

where the asterisk in the tables signifies an element which is determined by its neighbours. Since $k(I)=k(J)$ from (i), then:

(iv) $a_2=b_2+d_1+d_2$

For brevity, denote the parameters I_{31}, I_{32}, k, a_1 of $I^*(\underline{n}_A)$ by π_A and the parameters $I_{31}, I_{32}, k, a_1, b_2, d_1$ of $I^*(\underline{n}_B)$ by π_B. Denote the parameters b_2, d_1 by π_C. If π_A is given then J is uniquely determined. If π_B is given then I is uniquely determined. J is a valid table if:

(v) $a_1+a_2 \leq n_1+n_2$

(vi) $2I_{31}+I_{32} \leq k$

Now π_A induces a partition on $I^*(\underline{n}_B)$ since if J is valid, then $a_1+a_2>n_1+n_2$ (from (v)), whence $a_1+b_2+d_1+d_2>n_1+n_2$ (from (iv)) so I is valid Thus, from (6.2);

$$w(\underline{n}_B)-w(\underline{n}_A)=s^{2N}.\sum_{\pi_A} t^k.\Big(\sum_{\substack{\pi_C \\ I\in I^*(\underline{n}_B)}} K_I - K_J\Big) \qquad (6.3)$$

Now $\displaystyle\sum_{\substack{\pi_C \\ I\in I^*(\underline{n}_B) \\ \pi_A \text{ fixed}}} K_I \leq \sum_{\substack{b_2=0 \\ \pi_A \text{fixed}}}^{a_2} \; \sum_{d_1=0}^{a_2-b_2} K_I$ (from (iv); since the latter includes invalid tables)

$$= Q.\sum_{b_2=0}^{a_2} 2^{b_2}.M(n_1:a_1,b_2).\sum_{d_1=0}^{a_2-b_2} M(n_2:d_1,a_2-b_2-d_1)$$

(where $Q=2^{I_{32}}.M(n_3:I_{31},I_{32})$ and K_I is determined from (3.3(iv)).

$=Q2^{a_2}.M(n_1:a_1).M(n_1+n_2-a_1:a_2)$ (by the Vandermonde Convolution Theorem)

$<Q2^{a_2}.M(n_1+n_2:a_1).M(n_1+n_2-a_1:a_2) = K_J$ (by rearrangement)

Thus, $P_i(\underline{n}_B)>P_i(\underline{n}_A)$ whence $w_i(\underline{n}_B)>w_i(\underline{n}_A)$ from (6.3). This proves the theorem.

7. Convergence

If f(N,0,0) is defined to be unity, then this is the largest element in the matrix. The tables of results, however, have been computed with f(N,0,0)=0. (N,0,0) occurs only when the two binary sequences consist of zeroes alone. For a proper characterization of the objects, this case will not occur.

The following theorem confirms convergence of the algorithm for the special case $\phi^*(N,0,0)$. This is the simplest case and allows extension in Theorems 7.2 and 7.3 to arbitrary $\underline{n}$. Theorem 7.1 specifies for the numerical algorithm the summands which may be neglected at a cost of the most ϵ in the truncation error. We shall first need the following lemma.

Lemma 7.1 If $P_k = \sum_{\substack{I\ I^*(N,0,0)\\ k(I)=k}} K_I$ and $R_{k+1}=P_{k+1}/P_k$ then:

(i) R_k is monotonic decreasing in k and P_k is concave unimodal in k.

(ii) P_k has a maximum at k=N and is symmetrical about k=N.

Proof

If $k \leq N$ then $P_k = \sum_{a=0}^{\lfloor k/2 \rfloor} T(a,k)$ and if $k > N$ then $P_k = \sum_{a=k-N}^{\lfloor k/2 \rfloor} T(a,k)$

where
$$T(a,k) = 2^{k-2a}.M(N:k-2a,a). \qquad (7.1)$$

The summation bounds for $k>N$ are justified as follows:

$I_{12} = k-2a \geq 0$ (from (5.1)) so $a \leq \lfloor k/2 \rfloor$. Also, for $k>N, I_{11} = N - I_{12} - I_{13} \geq 0$

so $N-(k-2a)-a \geq 0$ whence $a \geq k-N$.

Suppose that k is even. Then $P_k = \sum_{a=0}^{k/2} T(a,k)$ and $P_{k+1} = \sum_{a=0}^{k/2} b(N,k,a).T(a,k$

where $b(N,k,a) = 2(N-k+a)/(k+1-a)$. Now $b(N,k,a)$ is strictly montonically decreasing in k, so P_k is also. It follows that P_k is concave unimodal with k.

Suppose that k is odd.

Then,
$$P_k = \sum_{a=0}^{(k-1)/2} T(a,k) \text{ and } P_{k+1} = \sum_{a=0}^{(k-1)/2} b(N,k,a).T(a,k) + 2.T(k-1/2\ ),k)/(k+1).$$

The isolated term combines with the term for a= (k-1)/2 to form a coefficient $b(N,k,(k-1)/2) + 1/(k+1)$ which is monotonic decreasing in k. It follows that P_k is concave unimodal with k. (ii) follows directly from the expressions for P_k given above, by change of variable.

Theorem 7.1 Suppose that ϕ^* is non-negative and bounded (without loss of generality assume $0 \leq \phi \leq 1$). Let ϕ_i^* be the partial sum of $\phi^*(N,0,0)$ for $k(I) \leq i$.

Then $\phi^*_{\bar{k}} - \phi^*_{\bar{k}-1} < \delta$ and $t^{\bar{k}}.P_{\bar{k}+1} < t^{\bar{k}}.P_{\bar{k}} \Rightarrow \phi^*_{2N} - \phi^*_{\bar{k}} < \varepsilon$

where $\varepsilon=\begin{cases}\delta(N-\bar{k}+(1-t)^{-1}) & \bar{k}<N\\ \delta t/(1-t) & \bar{k}\geq N\end{cases}$

<u>Proof</u> Let $P'_k=\sum_{\substack{I\in I^*(N,0,0)\\ k(I)=k}} K_I.\phi_I$. Then from the form of ϕ^* given in (3.3) we have $K_I=2^{I_{12}}.M(N:I_{11},I_{12})$ and $k(I)=I_{12}+2I_{13}$. Now $P'_k\leq P_k$ as $0\leq\phi\leq 1$.

Since P_k is concave unimodal in k (from Lemma 7.1) then $t^k.P_k$ is concave unimodal for t<1.

Let k_0 be the smallest k such that $t^{k+1}.P_{k+1}<t^k.P_k$. k_0 exists since P_k is unimodal.

k_0 must be the smallest k such that $t.R_k<1$. (7.2)

Then for any $k>k_0$:

$$t^{k+1}.P_{k+1}<t^k.P_k.$$

For $\bar{k}<N$: $\phi^*_{2N}-\phi^*_{\bar{k}} = s^{2N}.\sum_{k=\bar{k}+1}^{2N} t^k.P'_k$

$$\leq s^{2N}.\sum_{k=\bar{k}+1}^{N} t^k.P_k+s^{2N}.t^N\sum_{i=1}^{N} t^i.P_{N+i}$$

(by symmetry of P_k (from Lemma 7.1, and $P'_k\leq P_k$)

$<\delta(N-\bar{k}+(1-t)^{-1})$

For $\bar{k} \geq N$: $\phi^*_{2N} - \phi^*_{\bar{k}} = s^{2N} \sum_{k=\bar{k}+1}^{2N} t^k . P'_k < s^{2N} . \sum_{k=\bar{k}+1}^{2N} t^k . P_{\bar{k}}$

(since $\bar{k} \geq N, t<1, P_k$ monotonic decreasing for $k \geq N$)

$< \delta\ t/(1-t)$

This completes the proof.

Convergence is now shown for any $\phi^*(\underline{n})$ in the matrix $E^{(N)}$, for the values of $\bar{k}$ and δ specified in Theorem 7.1, with truncation error less than the ε of Theorem 7.1.

<u>Theorem 7.2</u> If $0 \leq \phi \leq 1$ and ϕ^*_i is the partial sum of $\phi^*(n_1,0,n_3)$ for $k(I) \leq i$, then for $\bar{k}$, ε and δ defined in Theorem 7.1: $\phi^*_{2N} - \phi^*_{\bar{k}} < \varepsilon$

<u>Proof</u> From the form of ϕ^* given in 3.3:

$$\sum_{i=0}^{2N} t^i . P_i(n_1,0,n_3) = (1+t)^{2N} \quad \text{(weights of the probability distribution)}$$

where

$$P_i(\underline{n}) = \sum_{\substack{I \varepsilon I^*(\underline{n}) \\ k(I)=i}} K_I$$

Since this is an identity for all t, and

K_I is independent of t, then:

$$P_i(n_1,0,n_3) = M(2N:i) \qquad (7.3)$$

which is independent of n_1 and n_3.

Then, using δ and $\bar{k}$ defined in Theorem 7.1, we have:

$$\phi^*_{2N}(n_1.0,n_3) - \phi^*_{\bar{k}}(n_1,0,n_3) \leq s^{2N} . \sum_{i=\bar{k}+1}^{2N} t^i . P_i(n_1,0,n_3)$$

$$\leq s^{2N} . \sum_{i=\bar{k}+1}^{2N} t^i . P_i(N,0,0) < \varepsilon \text{ from Theorem 7.1}$$

This completes the proof.

The following theorem demonstrates convergence for $\phi^*(\underline{n})$. However, for the $\bar{k}$ and δ of Theorem 7.1, the truncation error is larger than the ε given in Theorem 7.1.

Theorem 7.3 If $0\le\phi\le 1$, and ϕ^*_i is the partial sum of $\phi^*(n_1,n_2,n_3)$ for $k(I)\le i$, then for $\bar{k}$, ε, and δ defined in Theorem 7.1:

$$\phi^*_{2N}-\phi^*_{\bar{k}} < \varepsilon' \text{ where } \varepsilon' = (1+t^2)^{n_2}\varepsilon.$$

Proof From the form of $\phi^*(\underline{n})$ given in (3.3) we have:

$$\phi^*_{2N}-\phi^*_{\bar{k}} \le s^{2N}.(1+t^2)^{n_2}.\sum_{i=\bar{k}+1}^{2N} t^i.P_i(n_1,n_2,n_3) \qquad (\text{since } I_{22}\le n_2)$$

$$\le s^{2N}.(1+t^2)^{n_2}.\sum_{i=\bar{k}+1}^{2N} t^i.P_i(n_1+n_2,0,n_3) \qquad (\text{by Theorem 6.3})$$

$$= s^{2N}.(1+t^2)^{n_2}.\sum_{i=\bar{k}+1}^{2N} t^i.P_i(N,0,0) \qquad (\text{since } P_i(n_1,0,n_3) = M(2N:i), \text{from } (7.3)).$$

$$< (1+t^2)^{n_2}.\varepsilon \qquad (\text{from Theorem 7.1})$$

This completes the proof.

8. The diagonal elements of an error matrix

A number of interesting results may be proved for the diagonal elements of an error matrix. The diagonal elements are given by $\underline{n}=(n_1,0,n_3)$. We know from Theorem 6.1 that diagonal elements dominate elements in the same row or column and that they are monotonic in n_1 and n_3.

<u>Theorem 8.1</u> If $\phi(\underline{m},\underline{n})$ is independent of $\underline{n}$ for $\underline{n}=(n_1,0,n_3)$ then $\phi^*(n_1,0,n_3:t)=\phi^*(n_3,0,n_1:t^{-1})$

<u>Proof</u> From the form of ϕ^* given in (3.4):

$$\phi^*(n_1,0,n_3)=s^{2N}.\sum_{\underline{m}}\phi(\underline{m},\underline{n}).H_{\underline{m}}(\{t^2,2t,1\}^{n_1}.\{1,2t,t^2\}^{n_3})$$

$$=(t/(1+t)).\sum_{\underline{m}}\phi(\underline{m},\underline{n}).H_{\underline{m}}(\{t^{-2},2t^{-1},1\}^{n_1}.\{1,2t^{-1},t^{-2}\}^{n_3})$$

$$=\phi(n_3,0,n_1;t^{-1})\text{ (since }\phi(\underline{m},\underline{n})\text{ is independent of }\underline{n})$$

This completes the proof.

A stronger form of this may be stated for $h(\underline{m},\underline{n})$ and any (n_1,n_2,n_3), namely $h(\underline{m}(n_1,n_2,n_3);t)=h(\underline{m}(n_3,n_2,n_1);t^{-1})$. The proof of this is analogous to the above.

The following identities hold:

<u>Theorem 8.2</u>

i) If f is the Jaccard similarity function then:

$$\sum_{\underline{n}} M(N:n_1).\phi^*(n_1,0,n_3)=(4t/(1+4t+t^2)).2^N$$

ii) If f is the Jaccard similarity function then:

$$\sum_{\underline{n}} M(N:n_1).(-1)^{n_1}.\phi(n_1,0,n_3)=0$$

<u>Proof</u>

i) From the form of $\phi^*(\underline{n})$ given in (3.4) we have

$$\sum_{\underline{m}} M(N:n_1).\phi^*(n_1,0,n_3)=s^{2N}\sum_{\underline{m}}(m_2/(m_2+m_3)).H_{\underline{m}}\sum_{\underline{n}} M(N:n_1).\{t^2,2t,1\}^{n_1}\{1,2t,t^2\}$$

$$=s^{2N}.\sum_{\underline{m}}(m_2/(m_2+m_3)).H_{\underline{m}}\{1+t^2,4t,1+t^2\}^N$$

$$=s^{2N}.(1+t^2)^{N-1}.4t.\sum_{m_1=0}^{N} M(N:n_1).\sum_{m_2=0}^{N-m_1-1} M(N-m_1-1:m_2)(4t/(1+t^2))^{m_2}$$

$$=(4t/(1+4t+t^2).2^N$$

ii) follows since the left hand side of the identity is equal to

$\sum_{\underline{m}}(m_2/(m_2+m_3)).H_{\underline{m}}(\{1,0,-1\}^N).(1-t^2)^N.s^{2N}=0$, since if $m_2\neq 0$ then $H_{\underline{m}}(\{1,0,-1\}^N)=0$

Analogous results may be obtained by the same methods for other similarity functions. Particularly simple expressions may be found for a number of the error matrix elements. These are presented in the following theorem.

Theorem 8.3 If f is the Jaccard similarity function then:

i) a) $\phi^*(N,0,0)=(t/(2+t)).(1-s^{2N})$

b) $\phi^*(N,0,0)\to t/(2+t)$ as $N\to\infty$

ii) a) $\phi^*(N-1,0,1)=(2/(2+t)).(1-t^2.s^{2N}-((1-t^2)/(Nt(2+t))).(1-s^{2N}))$

b) $\phi^*(N-1,0,1)\to 2/(2+t)$ as $N\to\infty$

Proof Straightforwardly, from the form for ϕ^* given in (3.4).

Theorem 8.1 may be applied to Theorem 8.3 (ii)(a) to give a closed form expression for $\phi^*(1,0,N-1)$. Summation techniques used in Theorem 8.3 may be used to show that diagonal elements $\phi^*(n_1,0,n_3)$ may be computed by an algorithm of order at most N^2.

This result is incorporated into Theorem 8.4. To prove it, we need the following lemma:

<u>Lemma 8.1</u> If $g(\underline{a},N,A,t) = s^{2N} \sum_{\underline{m}} \phi(\underline{m}(n_1,0,n_3)).H_{\underline{m}-\underline{a}}\{t^2,2t,1\}^{N-A}$

where $a_1+a_2+a_3=A, m_1+m_2+m_3=N$ and $m_1 \geq a_1, i=1,2,3$ then $g=L.a_2+K$

where L and K are functions of N,m_1,a_1 and t, and involve sums over at most N terms.

<u>Proof</u> Consider the operator $\sum_{\underline{m}} \equiv \sum_{m_1} \sum_{m_2}$

Now $m_1 \geq a_1$ and $m_2 \geq a_2$, giving the lower limits of the sums. Also $m_3 \geq a_3$ so $N-A-m_1-m_2+a_1+a_2 \geq 0$ whence $m_2 \leq N-A-m_1+a_1+a_2$. This is the upper limit for the summation over m_2. To be valid, it must be greater than the lower limit, whence $m_1 \leq N-A+a_1$. This is the upper limit for m_1. Thus:

$$\sum_{\underline{m}} \equiv \sum_{m_1=a_1}^{N-A+a_1} \sum_{m_2=a_2}^{N-A-m_1+a_1+a_2}$$

Now $H_{\underline{m}-\underline{a}}\{t^2,2t,1\}^{N-A} = M(N-A:m_1-a_1,m_2-a_2).(t^2)^{m_1-a_1}.(2t)^{m_2-a_2}$

$$\text{so } g=s^{2N}.\sum_{m_1=0}^{N-A} (M(N-A:m_1)/(N-m_1-m_2)).t^{2m_1} \sum_{m_2=0}^{N-A-m_1} (m_2+a_2).M(N-A-m_1:m_2).(2t)^{m_2}$$

(by change of variables $m'_i=m_i-a_i$ and rewriting m'_i as m_i)

$=L.a_2+K$

$$\text{where } L=s^{2N}.\sum_{m_1=0}^{N-A} (M(N-A:m_1)/(N-m_1-a_1)).t^{2m_1}.(1+2t)^{N-A-m_1}$$

and $K=(2t/(1+2t)).((1+t)^{-2A}-(A-a_1).L)$

Clearly, L and K are functions of N,A,m_1,a_1 and t, and contain irreducibl sums over at most N terms. This completes the proof.

<u>Theorem 8.4</u> $\phi^*(n_1,0,n_3)$ may be computed by an algorithm of order N^2.

Proof

From the form of $\phi^*(\underline{n})$ given in (3.4):

$$\phi^*(n_1,0,n_3)=s^{2N}.\sum_{\underline{m}} (m_2/(m_2+m_3)).\ \sum_{\underline{a}} (H_{\underline{m}-\underline{a}}\{t^2,2t,1\}^{N-n_3})(H_{\underline{a}}\{1,2t,t^2\}^{n_3})$$

(where $a_1+a_2+a_3=n_3, m_1+m_2+m_3=N$ and $m_i \geq a_i, i=1,2,3.$)

$$= \sum_{\underline{a}} H_{\underline{a}}\{1,2t,t^2\}^{n_3}.(L.a_2+K) \qquad \text{(from Lemma 8.1)}$$

$$= \sum_{a_1=0}^{n_3} M(n_3:a_1).(t^2+2t)^{n_3-a_1}.(K+(2t/(1+2t)).(n_3-a_1).L)$$

Now K and L are functions of a_1, from Lemma 8.1, and contain irreducible sums over at most N terms. The summation over a_1 has at most N terms. The irreducible sums therefore contain at most N^2 terms to within a multiplicative constant. Accordingly, there exists by construction an algorithm of order N^2 for computing $\phi^*(n_1,0,n_3)$. This completes the proof.

Theorem 8.5 If t=1 then

$$\phi^*(n_1,0,n_3)=(2/3).(1-(\tfrac{1}{2})^{2N})$$

Proof

From the form of $\phi^*(\underline{n})$ given in (3.4), we have:

$$\phi^*(n_1,0,n_3)=(\tfrac{1}{2})^{2N}.\sum_{\underline{m}} (m_2/(N-m_1).H_{\underline{m}}(\{1,2,1\}^{n_1}.\{1,2,1\}^{n_3}) \quad \text{if } t=1$$

$$=(\tfrac{1}{2})^{2N}.\sum_{m_1=0}^{N-1} M(N:m_1) \sum_{m_2=0}^{N-m_1-1} M(N-m_1-1:m_2).2^{m_2}$$

The upper limit of N-1 excludes the undefined value of $\phi(\underline{m},\underline{n})$)

$$=(\tfrac{1}{2})^{2N}.2.\sum_{m_1=0}^{N-1} M(N:m_1)\,3^{N-m_1-1}$$

$= (2/3)(1-(\frac{1}{2})^{2N})$, which completes the proof.

This has been used for validating the program for computing $\phi^*(\underline{n})$. Table V shows $\phi^*(n_1,0,n_3)$ computed for $N=10, t=1$. With $\delta=10^{-3}$, $\bar{k}$ had the value 16.

9. The behavior of $\bar{k}$ with respect to N, r and δ

We first obtain an improved estimate for the truncation error ε. From Section 7 we know that $T(a+1,k)/T(a,k)$ is monotonically decreasing so $T(a,k)$ is unimodal in a. The dominant term of P_k is given by the smallest a such that $T(a,k)\geq T(a+1,k)$. I.e. $4(a+1)N+6a+4\geq k(k+3)$. Suppose that a' is such a value and that $\bar{a}=a'+1$. Corresponding terms of P_k and P_{k+1} are related by the factor $b(N,k,a)$, so R_k may be approximated by $\bar{R}_k=b(N,k,\bar{a})=(4N(N-k)+6N+(k-1)(k-2))/(k+1)(2N-k+1)$. Table VI gives several values of $\bar{R}_k$ for $N=10$, and $k\leq 10$, and shows that R_k is monotonic decreasing in k, as proved in Lemma 7.1.

$\bar{R}_k$ can be used to obtain a larger value of $\delta(\varepsilon)$ than provided by Theorem 7.1. Clearly, δ is related inversely to computation time. We show that ε and δ are of the same order of magnitude.

Now

$$\sum_{k=\bar{k}+1}^{N} t^k.P_k \quad <\delta t.\bar{R}_{\bar{k}}/(1-t.\bar{R}_{\bar{k}}) \quad \text{(by monotonicity of } R_k\text{; Lemma 7.1)}$$

Also

$$t^N.\sum_{k=1}^{N} t^k.P_{N-k} < \delta\, t^2/(1-t^2)+\delta t^{2(N-\bar{k})}/(1-t)$$

Thus

$$\varepsilon/\delta = t.\bar{R}_{\bar{k}}/(1-\bar{R}_{\bar{k}}.t)+t^2/(1-t^2)+t^{2(N-\bar{k})}/(1-t) \qquad \text{if } \bar{k}<N$$

$$= t/(1-t) \qquad \text{if } \bar{k}\geq N$$

Table VII gives an improved truncation error estimate over the one given by Theorem 7.1. ε and δ are of the same order of magnitude. The truncation error is therefore estimated by the last summand of ϕ^* which has been included.

Now, from (7.2), k_0 is the smallest k such that $t.R_k<1$. Substituting the approximation $\bar{R}_{k_0}$ into $\bar{R}_{k_0}.t<1$, simplifying and solving the resulting quadratic for its smallest root, we have: $k_0<2Nt/(1+t)$ (9.1)

To determine an upper bound for $\bar{k}$, consider the value of $\phi^*(N,0,0)$. Now $\phi^*_{\bar{k}+1}(N,0,0)-\phi^*_{\bar{k}}(N,0,0)<\delta$ by the definition of $\bar{k}$, whence $s^{2N}.t^{\bar{k}}.P_{\bar{k}}(N,0,0)<\delta$. Now from (7.1) $P_{\bar{k}}(N,0,0)=\sum_{i=0}^{\lfloor\frac{1}{2}\rfloor} T(i,\bar{k})$. The largest element in the sum occurs for $i'=(\bar{k}(\bar{k}+3)-4(N+1))/(4N+6)$. Thus, the upper bound is given by

$$(\tfrac{1}{2}\bar{k}+1).s^{2N}.t^{\bar{k}}.T(i',\bar{k})<\delta \text{ and } t.P_{\bar{k}+1}(N,0,0)<P_{\bar{k}}(N,0,0).$$

Substituting for $T(i',\bar{k})$ from (7.1) and approximating asymptotically to the factorial by Stirling's approximation, we have:

$$s^{2N}.\bar{k}.((2N-\bar{k})/\bar{k})^{\bar{k}}.t^{\bar{k}}.(2N/(2N-\bar{k}))^{2N}<\delta \qquad (9.2)$$

Table VIII gives the lower (9.1) and the upper (9.2) bounds for $\bar{k}$, along with the experimentally derived values for $\bar{k}$, as N and r vary, and δ=0.001.

10. Conclusions

The algorithm which has been examined has been verified for the Jaccard similarity function. Uniform convergence of the method, however, holds for any statistical function ϕ and similarity function f such that $\phi(\underline{m},\underline{n})$ is bounded and non-negative. All elements of an error matrix converge for the same value of $\bar{k}$. Moreover, it is seen from the computational results that $\bar{k} \ll 2N$. Particular kinds of results, for example monotonicity properties and closed forms, obtained for the Jaccard function are derivable for other similarity functions.

A conservative upper bound on the order of the algorithm is $N^2\bar{k}^6$. Experimental evidence shows that the actual order of the algorithm is considerably less than this. A better estimate of order is under investigation. For a given set of data, not all of the $\underline{n}$ need to occur, and additional economy may be achieved by computing only the elements of the matrix which are required.

The analysis has provided a basis for determining the stability of classifications of objects described by binary attributes which are susceptible to error independently and equiprobably. The analysis is particularly valuable when the number of attributes is such that the asymptotic normality of ϕ^* does not hold.

Acknowledgement

This work was supported in part by the National Science Foundation Grants GN-27224 and GN-534.1.

References

1. Nagy, G. State of the art in pattern recognition. Proceedings of IEEE, 56, 1968,pp 836-862.

2. Sokal, R.R. and Sneath, P.H.A. Principles of Numerical Taxonomy. Freeman, San Francisco, 1963.

3. Parker-Rhodes, A.F. and Needham, R.M. The Theory of Clumps. Report No. 126. Cambridge Language Research Unit, Cambridge, England, 1960.

4. Lerman, I.C. On two criteria of classification. In: Numerical Taxonomy (Cole, A.J. ed) Academic Press, 1969, pp 114-127.

5. Friedman, H.P., and Rubin,J. On some invariant criteria for grouping data. Journal of the American Statistical Association, 62, 1967,pp 1159-1178.

6. Hartigan, J.A. Representation of similarity matrices by trees. Journal of the American Statistical Association, 62, 1967, pp 1140-1

7. Gleason, A.M. A search problem in the n-cube. Proceedings of Sympos in Applied Mathematics, 10, Combinatorial Analysis, 1960, pp 175-178 Bellman, R. and Hall, H. eds)

8. Jackson, D.M. Comparison of Classifications. In: Numerical Taxonomy, (Cole, A.J. Ed) Academic Press, 1969, pp 91-111.

9. Bernouilli, N. Montmort's "Essay d'Analyse sur les Jeux des Hazards". 2nd ed, Paris, 1714.

10. De Moivre, A. Doctrine of Chances, 1st ed, London, 1718.

11. Barton, D.E. The matching distributions: Poisson limiting forms and derived methods of approximation. Journal of the Royal Statistical Society, (B), 20, 1958,pp73-92.

12. Fine, T. Extrapolation when very little is known about the source. Information and Control, 16, 1970, pp331-359.

13. Jackson, D.M. An error analysis for functions of qualitative attributes with application to information retrieval. Proceedings of the symposium on Computer and Information Science, (Tou,J.T. ed) ,3,1969, (in press).

14. Collins, G.E. PM A system for polynomial manipulation. Comm. ACM, 9, 1966, pp. 578-589.

15. Collins, G.E. The SAC-1 system: An introduction and survey, To be presented at the Second Symposium on Symbolic and Algebraic Manipulation, Los Angeles, March 23-25, 1971.

16. McMahon, P.A. Combinatory Analysis, Vol. 1, Ch. 3, Cambridge University Press, 1915.

17. Kaplansky, I. and Riordan, J. Multiple matching and runs by the symbolic method. Annals of Mathematical Statistics, 16, 1945, pp. 272-277.

18. Jackson, D.M. and White, L.J. The weakening of taxonomic inferences by homological errors. Mathematical Biosciences, (in press), 1970.

19. Bell, A.G. Partitioning the integers in n dimensions, Computer Journal, 13, 1970, pp. 278-283.

n_1	n_3	k=1	k=3	k=5	k=6	k=20(exact)
0	10	0.038	0.090	0.095	0.095	0.095
0	6	0.036	0.062	0,064	0,064	0,064
0	0	0.019	0.047	0.050	0.050	0.050
2	8	0.046	0.108	0.114	0.114	0.114
2	4	0.038	0.067	0.069	0.069	0.069
2	0	0.019	0.047	0.049	0.049	0.049
4	6	0.059	0.135	0.143	0.143	0.143
4	0	0.019	0.046	0.048	0.048	0.048
6	4	0.083	0.187	0.190	0.190	0.190
6	0	0.019	0.045	0.047	0.047	0.047
8	2	0.138	0.278	0.289	0.289	0.289
8	0	0.019	0.042	0.044	0.044	0.044
9	1	0.208	0.325	0.397	0.397	0.397(largest element
10	0	0.000	0.015	0.016	0.016	0.016

TABLE I

Values of $\phi^*_{\underline{k}}(\underline{n})$ for N=10, r=0.05 and several values of $\bar{k}$.

r	N	$\overline{k}$	r	N	$\overline{k}$
0.05	10	5	0.20	10	10
	20	7		20	16
	30	9		30	21
	40	11		40	27
	50	12		50	32
0.10	10	7	0.25	10	11
	20	10		20	18
	30	14		30	25
	40	17		40	31
	50	20		50	38
0.15	10	8	0.30	10	12
	20	13		20	21
	30	18		30	28
	40	22		40	36
	50	26		50	43

TABLE II

Behavior of $\overline{k}$ with respect to r and N for several values of r and N. Values of $\overline{k}$ computed with truncation error of 10^{-3} in ϕ*.

N	$\overline{k}$	$R^*(N,\overline{k})/T(N)$	Predicted time-saving ratio	Actual running time (seconds)	Error in third significant figure
30	2	0.0002	0.0003	3.93	460
	3	0.0004	0.0006	6.64	286
	4	0.0010	0.0011	12.68	149
	5	0.0021	0.0019	22.02	66
	6	0.0040	0.0033	37.32	25
	7	0.0072	0.0053	60.10	8
	8	0.0121	0.0084	96.11	2
	9	0.0195	0.0123	140.24	Converged
	60		1.0	3.18 hrs (estimated)	$\phi^*=0.688$

TABLE III

Computation time estimates for $\phi^*(29,0,1)$, with $N=30, r=0.05$, for several $\overline{k}$.

n_1 \ n_3	0	1	2	3	4	5	6	7	8	9	10
0	.050	.052	.054	.055	.057	.060	.064	.069	.076	.085	.095
1	.050	.053	.055	.058	.062	.067	.073	.081	.091	.104	
2	.049	.053	.058	.063	.069	.076	.087	.099	.114		
3	.049	.055	.062	.070	.079	.092	.108	.127			
4	.048	.058	.069	.082	.098	.118	.143				
5	.047	.063	.081	.104	.131	.163					
6	.047	.074	.105	.145	.190						
7	.046	.097	.158	.229							
8	.044	.158	.289								
9	.040	.397									
10	.016										

TABLE IV

The error matrix $E^{(10)}$ for $\phi=|f(\underline{n}) - f(\underline{m})|$, $r=0.05$ and truncation error 10^{-3}.

n_1	$\phi^*(n_1,0,n_3)$	n_1	$\phi^*(n_1,0,n_3)$
9	0.6665	4	0.6661
8	0.6663	3	0.6660
7	0.6662	2	0.6659
6	0.6662	1	0.6657
5	0.6661	0	0.6655

Exact value (by Theorem 8.5) of $\phi^*(n_1,0,n_3)$ is $2/3+10^{-6}$.

TABLE V

Computed values of $\phi^*(n_1,0,n_3)$ for $t=1, n=10$, and $\delta=10^{-3}$.

k	Actual R_k	$\overline{R}_k$
1	9.50	10.5
2	6.00	6.67
3	4.24	4.76
4	3.21	3.60
5	2.50	2.84
6	2.00	2.29
7	1.63	1.88
8	1.33	1.56
9	1.10	1.30

TABLE VI

R_k and estimates for $N=10$ and $k\leq 10$

r	t	N	$\bar{k}$	ε/δ	$t.\bar{R}_{\bar{k}}$
0.05	0.053	20	11	0.166	0.14
		50	26	0.180	0.15
0.20	0.25	20	12	1.42	0.58
		50	33	1.11	0.51
0.30	0.43	20	21	0.75*	
		50	44	2.29	0.67
0.40	0.67	20	25	2*	
		40	45	2*	

TABLE VII

Improved truncation error estimate.

*since $\bar{k}>N$, the second form of the bound for ε applies.

r	t	N	$\bar{k}$	Lower bound for $\bar{k}$	Upper bound for $\bar{k}$
0.05	0.053	10	5	1.0	8
		20	7	2.0	11
		30	9	3.0	13
		40	11	4.0	15
		50	12	5.0	18
0.10	0.11	10	7	2.0	10
		20	10	4.0	15
		30	14	5.9	19
		40	17	7.9	23
		50	20	9.9	26
0.15	0.18	10	8	3.1	12
		20	13	6.1	18
		30	18	9.2	24
		40	22	12.2	29
		50	26	15.3	34
0.20	0.25	10	10	4.0	13
		20	16	8.0	21
		30	21	12.0	28
		40	27	16.0	34
		50	32	20.0	41
0.25	0.33	10	11	5.0	15
		20	18	9.9	24
		30	25	14.9	32
		40	31	1-.8	40
		50	38	22.3	47
0.30	0.43	10	12	6.0	16
		20	21	12.0	26
		30	28	19.1	36
		40	36	24.1	44
		50	38	30.1	53

TABLE VIII

Upper and lower bounds of $\bar{k}$ as r and N vary with δ=0.001.