(10) Reingold, E. M.; Nievergelt, J.; Deo, N. "Combinatorial Algorithms: Theory and Practice"; Prentice-Hall: Englewood Cliffs, NJ, 1977; Chapter 4.
(11) Beech, G.; Jones, R. T.; Miller, K. "Structural Interpretation of Proton Magnetic Resonance Spectra by Computer: First-Order Spectra". *Anal. Chem.* **1974**, *46*, 714-718.
(12) Knuth, D. E. "The Art of Computer Programming"; Addison-Wesley: Reading, MA, 1968; Vol. 1, Chapter 1, p. 38.
(13) Harary, F. "Graph Theory"; Addison-Wesley: Reading, MA, 1969; Chapter 2.
(14) Shelley, C. A.; Munk, M. E. "An Approach to the Assignment of Canonical Connection Tables and Topological Symmetry Perception". *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 247-250.
(15) Farkas, M.; Munk, M. E., work in progress.

Monte Carlo Studies of the Classifications Made by Nonparametric Linear Discriminant Functions

TERRY R. STOUCH and PETER C. JURS*

The Pennsylvania State University, University Park, Pennsylvania 16802

Received August 28, 1984

Chance factors in pattern recognition studies utilizing nonparametric linear discriminant functions are examined. The relationship between complete linear separation of a data set and the dimensionality of the study is well-known. Also, due to the nature of the inequalities from which these numerical techniques are derived, 50% separation is always assured. This paper investigates the probability of achieving less than 100% but greater than 50% chance separations as a function of the dimensionality and class membership distribution. It is shown that the fraction of correct classifications due to chance factors increases dramatically as the dimensionality of the study increases. These results serve to redefine the level of expected chance classifications as a function of the number of observations, the dimensionality, and the class membership distributions. The results can be used to assess the classification results obtained with a given linear discriminant function.

The field of pattern recognition (PR) consists of techniques that are designed to classify numerical patterns.[1,2] When PR is applied to structure–activity relationship (SAR) studies, the patterns consist of measures of the compounds' physical and structural properties, and the classes are composed of compounds of like activity. These methods are useful when the activities of compounds are not known quantitatively. In such a case, the compounds can be assigned to a class of activity such as active, inactive, or very active. Such situations often occur when quantitative measures of activity are difficult, impractical, or impossible to obtain.

Discriminant generating methods of PR develop boundaries between the classes. The boundaries defined by linear discriminant functions (LDFs) are linear combinations of the variables and can be thought of as their corresponding geometrical forms: a line in two dimensions, a plane in three dimensions, and a hyperplane in higher dimensions. Algebraically, they are represented by eq 1, where the coefficients

$$f(\text{activity}) = a_0 + a_1x_1 + a_2x_2 + \ldots + a_dx_d \qquad (1)$$

of the variables (a_i's) comprise the discriminant and the variables (x_i's) are the measurements (descriptors) that describe the compounds and that comprise the patterns. d is the number of descriptors used in the discriminant. Parametric methods of discriminant generation use class statistics in order to generate discriminants; to do this, assumptions must be made concerning the distribution of the data. Nonparametric methods make no such assumptions and generate discriminants by analyzing each individual pattern. For this reason, many SAR PR studies have been performed with the nonparametric techniques.

It is well-known that, for a given number of compounds, the probability of fortuitously obtaining 100% complete and correct classification of the compounds increases as the number of features increases from 1 to the number of patterns in the study, N. This probability can be calculated with eq 2, where

$$P = 2\sum_{i=0}^{d} C_i^{N-1}/2^N \qquad (2)$$

$C_i^{N-1} = (N-1)!/[(N-1-i)!i!]$. N is the number of compounds, and d is the dimensionality, the number of descriptors. Figure 1 shows a plot of this relationship for 50 compounds. The only assumption that is made concerning the data is that it be in general position, that is, that no $d+1$ data points should be contained in a $d-1$ hyperplane. When a small number of descriptors is used to develop the LDF, the probability of achieving complete separation due to chance is small. As the number of descriptors approaches the number of compounds used in the study, however, the probability of such an occurrence increases. These classifications, while correct, are due only to artifacts of the mathematics governing the LDF-generating process. They are not due to any relationship between the compounds, and the resulting LDF will have no predictive ability beyond random guessing. This relationship has been known for some time and can be found in the literature.[1,2] Stuper and Jurs have shown that if the number of descriptors is kept below one-third the number of compounds used, the probability of complete separation due to chance can be kept low.[3] At the other extreme, it is intuitively obvious that for classes of equal size, 50% separation can be achieved by assigning all of the patterns in the study to one class. Often, however, LDFs are generated and used without reference to the levels of classification between these extremes. Often, random results are assigned the value of 50% correctly classified, and any result greater than 50% is considered to be due to the information contained within the descriptors and the explanation of at least a portion of the underlying SAR. This is contrary to the fact that the probability of 100% separation increases with an increasing number of descriptors. If the probability of 100% separation increases, then so too must the probability of less complete separations. This will have the effect of changing the level of classification that is considered

0095-2338/85/1625-0045$01.50/0

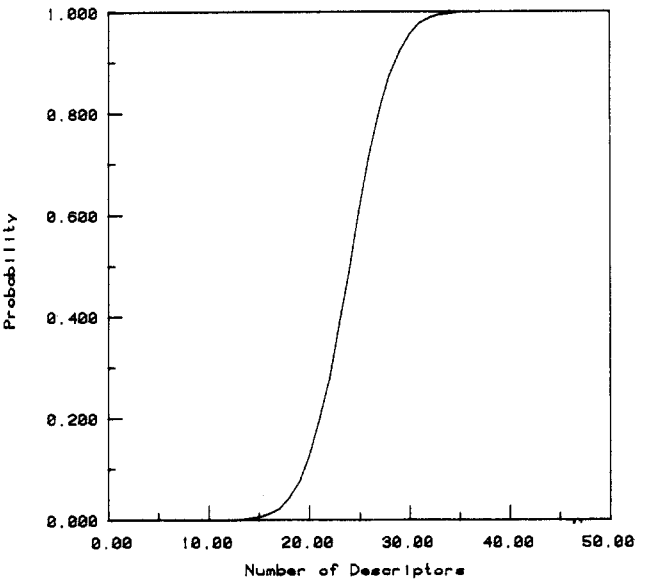

Figure 1. Probability of dichotomization (complete linear separation) of 50 patterns as a function of the dimensionality.

to be random. Random results will range between 50 and 100% as the number of descriptors increases from zero to the number of patterns, N. This effect is especially important if LDFs with less than 100% separatory ability are to be interpreted.

In order to properly evaluate the classification results that are obtained for a given number of compounds and a given number of descriptors, it is necessary to know what percentage of the classifications are due to chance. The studies reported here were performed in order to define random results as a function of the number of descriptors and the number of patterns. The effect of the distribution of the patterns between the classes was also examined.

METHODOLOGY

Monte Carlo techniques were used to simulate studies involving 26, 50, 100, and 200 compounds. These numbers were chosen because most studies found in the literature lie within this range and because they were of a size to be computationally convenient. For each study, pseudo random numbers were used as descriptors. These numbers were generated with the routine of Schrage,[4] which provides uniformly distributed numbers in the range of zero to one. Gaussian-distributed pseudo random numbers were developed by applying the Central Limit Theorem.[5] The random descriptors were uncorrelated, simple correlation coefficients were less than 0.4, and multiple correlation coefficients were less than 0.6 for "models" of up to 10 variables. Each descriptor was autoscaled to a mean of 0 and a standard deviation of 1. Sets of these random descriptors were submitted to the routines, which generated the LDFs that indicated the degree of separability of the data.

PR methods are capable of dealing with more than two classes; however, since most LDF SAR studies have been performed with only two and since multiclass problems can be reduced to a series of two-class problems, the studies reported here were restricted to two classes. Each study was started by random assignment of half of the patterns to class "one" and the other half to class "two". For the studies involving unequal class sizes, the required percentages of the total number of patterns were arbitrarily assigned to the two classes.

For each study, the number of descriptors in the sets was varied. The smallest sets were formed with only one. Many trials (from 10 to 50) were run for each number of descriptors in order to ascertain the range of classifications that a given dimensionality would support for a given number of patterns. Each trial was performed by obtaining the LDF that would yield the highest percentage classification for that data matrix within the constraints imposed on the iterative discriminant generating techniques. The number of trials performed for each dimensionality for each study varied somewhat depending on the study; fewer runs were made for the studies involving larger numbers of patterns since these consumed the most resources. The greatest number of trials were performed for the 50- and 100-pattern studies, whose sizes were representative of actual studies and of a size to allow fairly complete treatment.

There are many forms of nonparametric linear classifiers including the linear learning machine[6] and related least-squares techniques.[7,8] Each of these has its own characteristics, and several were used in tandem in order to achieve the "best" classification results. For each trial, a check was first made for complete separability with the linear learning machine, which works well for completely separable data. If complete separability was not found, the adaptive least-squares (ALS) method of Moriguchi[7] was used to find the best classification possible. In order to save computer time, this iterative technique was initiated with a LDF that was developed by a Bayes linear classifier algorithm.[2] This was chosen for its speed and efficiency at arriving at a LDF that, while normally not obtaining the highest possible classification results, served as an effective initialization for the ALS routine. The routines used and the order in which they were used were based on our past experiences with these algorithms. We chose the approach that would yield the highest possible classification success rate with the smallest consumption of computer time. Since many trials were run, some with quite large data matrixes, the conservation of computer time expended was a goal. In order to further minimize time and in order to automate these studies, the iterative routines were allowed to execute only a set number of iterations which was constant for all trials. For the ALS routine, this number was set at 800 after many initial trials at much higher numbers of iterations showed that classifications normally plateaued below this number. The linear learning machine was allowed to iterate several hundred times. When operating on separable sets, this algorithm normally converges in far fewer iterations. While this procedure saved time and effort, it may have resulted in less than optimal LDFs.

All the studies reported here were performed on a PRIME 750 computer with 2 megabytes of main memory with the PRIMOS revision 18.3 system software. The software used to perform the studies was the ADAPT system described by Stuper, Brugger, and Jurs.[6]

RESULTS

Figures 2–5 show the results for the studies using 26, 50, 100, and 200 patterns. Each point shows the percentage of correctly classified patterns obtained for each trial. For each study, the apparent separability of the data increased as the number of descriptors increased, even though there was no information contained within the "descriptors" and no possible relationship between the random patterns. These results were obtained by using uniformly distributed random numbers. Studies using normally distributed random numbers yielded essentially identical results.

The results of the four studies shown in Figures 2–5 can be compared to one another by dividing the number of descriptors used in each trial by the total number of patterns used in that study. This adjusts the range of the x values of each study to be within the range of 0–1. The data was replotted in Figure 6 with the newly scaled abscissa; the multiple trials at each dimensionality were averaged for clarity. All of the studies

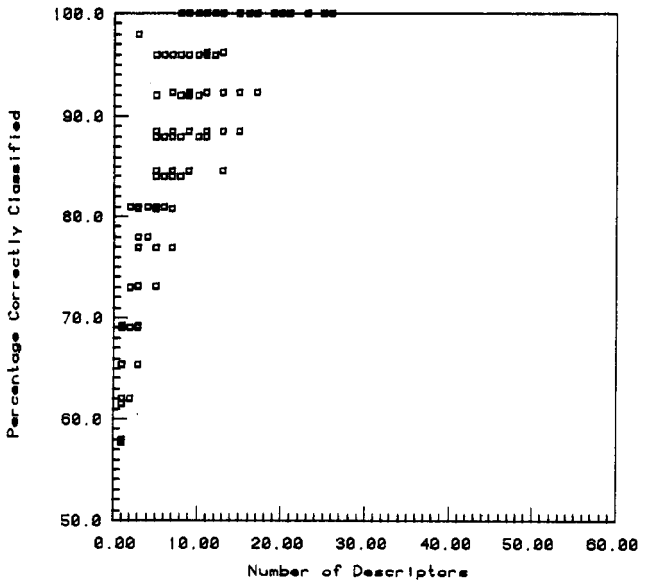

Figure 2. Plot of the percentage of correctly classified patterns vs. the dimensionality for a study of 26 random patterns.

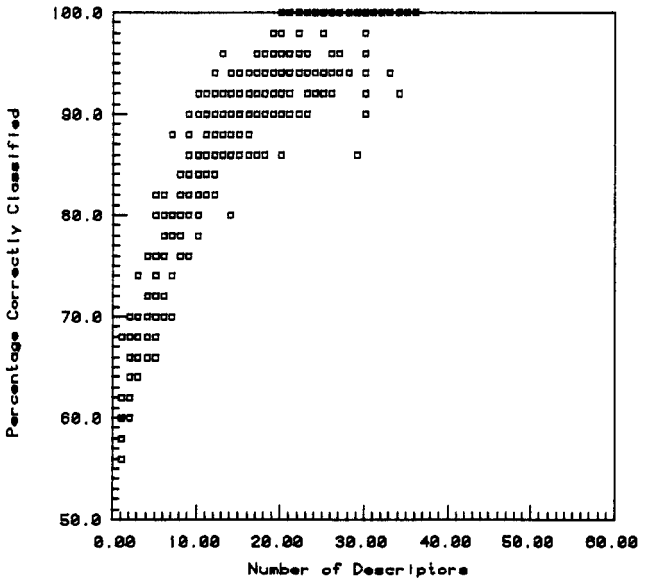

Figure 3. Plot of the percentage of correctly classified patterns vs. the dimensionality for a study of 50 random patterns.

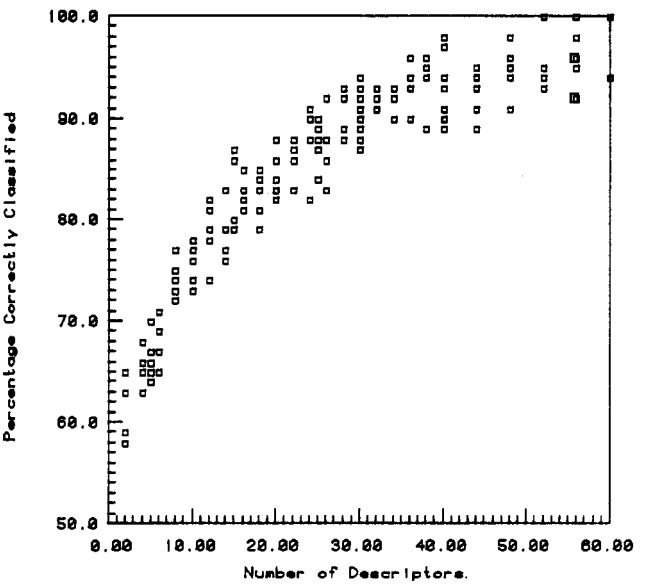

Figure 4. Plot of the percentage of correctly classified patterns vs. the dimensionality for a study of 100 random patterns.

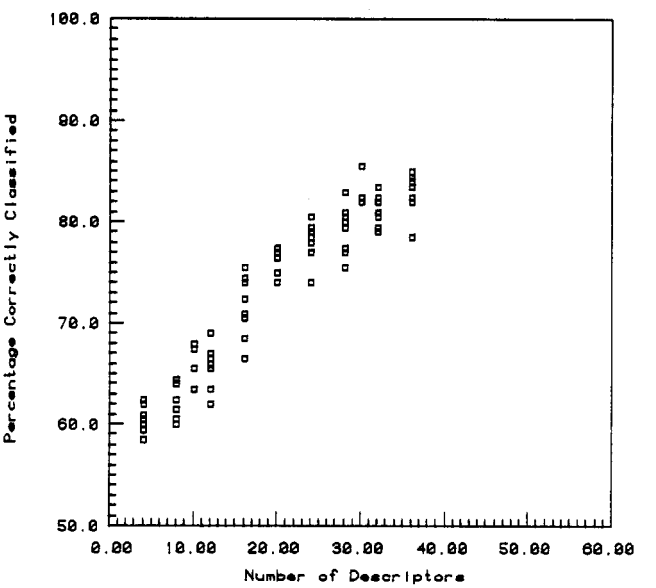

Figure 5. Plot of the percentage of correctly classified patterns vs. the dimensionality for a study of 200 random patterns. Only partial results were obtained due to computational constraints.

take on a common form. For any one value of the ratio of the number of descriptors to the number of patterns (d/N), the percentage classified is the same regardless of the number of patterns used in the study. While these studies cover the range of size that most PR studies would encounter, the agreement between them suggests that the results would apply regardless of the number of patterns used. This curve shows classification success rates well above 50% even for small numbers of descriptors. It also shows that the classifications are dependent on the number of descriptors used in developing the LDF. While the 1/3 ratio reported by Stuper and Jurs will keep the probability of achieving 100% separation low, at that ratio random classification results range around 90%. Even one descriptor for every 10 patterns yields random correction classifications of about 75%.

These results can be explained by considering eq 2, the equation for calculating the probability of 100% separation at any given dimensionality. The denominator is the number of ways that N patterns can be arranged into two classes. It represents the total number of dichotomies for N patterns, that is, every possible way that N patterns can be arranged into two classes. The numerator in the equation represents the number of linear dichotomies that can be obtained for a particular dimensionality and is derived by Nilsson.[1] This value is the number of LDFs that could be calculated for a given number of descriptors and a given number of patterns. As the dimensionality increases, the number of calculable LDFs increases also, and the probability of one of these being the particular dichotomy that was defined as the class assignments increases correspondingly. When the number of descriptors is equal to the number of patterns, all of the possible dichotomies can be obtained by a LDF. One of these, of course, has to be equivalent to the class assignments.

This reasoning can be extended to separations that are less than the 100% complete separation considered by eq 2. Only two of the total number of dichotomies will provide complete separation (the patterns can be classified as in either of the two classes interchangably and still be linearly separable). All of the dichotomies are capable of 50% separation. Once again, there is a middle ground to these extremes. There are dichotomies that will provide for the range of classification results. A simple example using four patterns will illustrate this. Figure 7 shows the 16 ways that four patterns can be

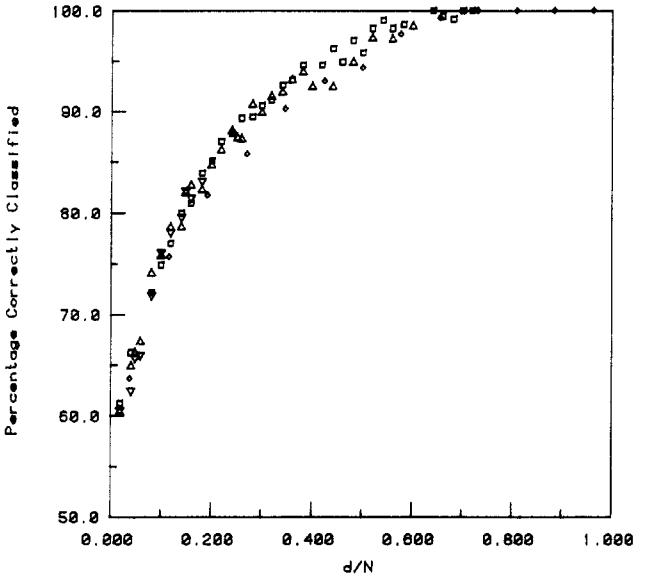

Figure 6. Results of scaling the abscissa values of Figures 4–7 by dividing each dimensionality by the total number of patterns used in the study: (diamonds) 26-pattern study; (squares) 50-pattern study; (triangles) 100-pattern study; (inverted triangles) 200-pattern study.

index	class 1	class 2	% correctly classified
1	1234	-	50
2	123	4	75
3	124	3	75
4	134	2	75
5	234	1	75
6	12	34	100
7	13	24	50
8	14	23	50
9	23	14	50
10	24	13	50
11	34	12	100
12	1	234	75
13	2	134	75
14	3	124	75
15	4	123	75
16	-	1234	50

Figure 7. Simple four-pattern, two-class problem showing all possible dichotomies and their corresponding degrees of separation for the arbirary class assignments.

arranged into two classes. If patterns one and two belong to class one and patterns three and four belong to class two, of the 16 possible dichotomies two will yield 100% classification, six give 50%, and eight give 75%. These classifications are also listed in Figure 7. This effect can be extrapolated to problems involving larger numbers of patterns; some fraction of the total number of dichotomies will account for any given degree of separation. A greater fraction of the total will allow lower separations while the more complete separations will be achievable by a smaller fraction of the total number of dichotomies.

These observations may be clarified by a different view of the data plotted in Figure 3. The number of occurrences of several degrees of separation (e.g., 70% correctly classified, 80% correctly classified, etc.) were summed for each dimensionality for the 50-pattern study. The fraction of the total number of occurrences of each degree of separation was plotted against the dimensionality as shown in Figure 8. For any dimensionality, the probability of achieving a given degree of separation can be obtained. For example, the probability of achieving 80% or better separation at two dimensions is very low. At seven dimensions, the probability is about half; the random classification results at seven descriptors will center

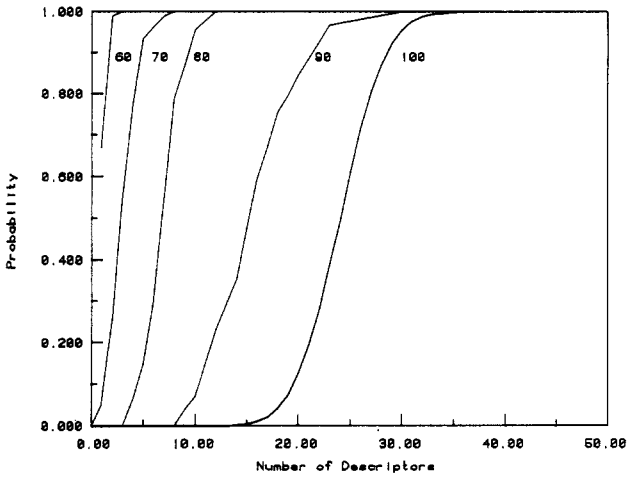

Figure 8. Data from Figure 5 replotted. The curves represent the sample cumulative probability functions for obtaining greater than or equal to the noted degree of separation.

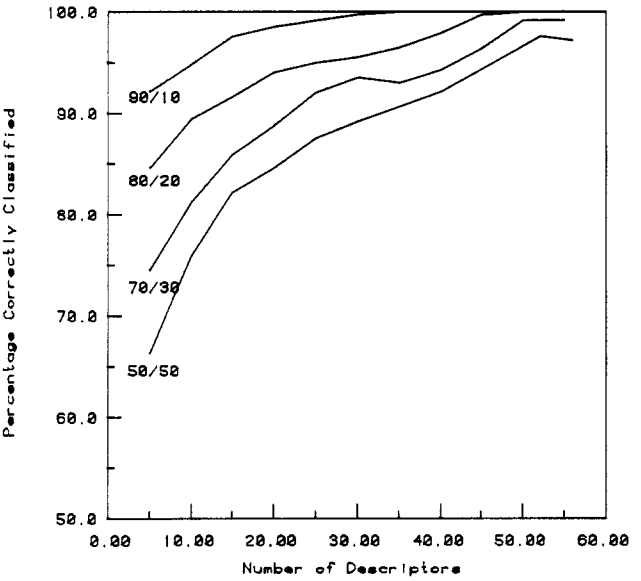

Figure 9. Plot of the percentage of correctly classified patterns vs. the number of random descriptors for a 100-pattern study where the patterns were unequally distributed between the classes.

around this value. Above 15 descriptors, the probability of achieving greater than this degree of separation is unity, and any random classifications will be greater than 80%. While only a few degrees of separation are plotted in this figure for clarity, the others could also be plotted, and the range of classification values at a particular dimensionality could be determined by the curves that cross that abscissa value.

All of the results discussed to this point were obtained by using equal class sizes, with half of the patterns assigned to one class and the other half assigned to the second class. Unequal class sizes serve to increase the random classifications that have been examined so far. At zero descriptors, random chance will account for greater than the 50% separation achievable for equal class sizes. These classifications will be determined by the percentage of the total number of patterns that the largest class represents. For example, if a study is conducted with 100 patterns that are divided 70/30 between the two classes, 70% of the patterns can be correctly classified by assuming that all of the patterns belong to the larger class.

Figure 9 shows the results of studies done for 100 patterns with varying distributions of the patterns between the two classes. These studies were conducted as were those for even class sizes with multiple trials (greater than nine) at several dimensionalities; the trials were averaged before being plotted

for clarity. The results can be compared with those for equal class sizes which were also plotted. It is clear that, as the class sizes become more unequal, the expected random results at a particular ratio of d/N increase.

These random classifications due to the dimensionality of the study should not be confused with the prediction results based on random guessing with a knowledge of the class distributions. This relationship has been noted in several PR studies and is represented in eq 3. In this equation, p stands

$$\text{percentage classification} = (p^2 + q^2) \times 100 \qquad (3)$$

for the fraction of the patterns contained in one class, and q is the fraction of patterns in the other.[9] For example, if a population was known to consist of 80% of class one and 20% of class two, then 68% could be guessed correctly if 80% of the guesses were for class one and 20% were for class two. This "zero-dimensional" prediction with prior knowledge of the class distributions is not equivalent to the classification results achievable by a LDF-generating routine since the PR routine is provided beforehand with the class memberships. The PR routines use this information to develop the LDFs on which the classification results are based.

As noted under Methodology, if the results presented here err, they err conservatively. Except in cases of complete separation, they would represent the low end of random classifications. The routines used to obtain these results are iterative and computationally intensive. They were run under conditions that experience dictated were likely to find the highest possible classifications in the amount of time that was alloted for each run. It is possible that higher classifications could have been obtained if the routines were differently configured or if they were given more time to execute.

CONCLUSIONS

The studies reported here have shown that chance classifications encountered in pattern recognition studies employing nonparametric linear discriminant functions are greater than previously thought. While random results were previously considered to be the 50% that would be obtained by random guessing with no prior knowledge of the class distributions, we have shown that random results are, in reality, far above this value and are a function of the ratio of the number of descriptors employed to the number of patterns involved. A LDF that is capable of 90% separation of a set of compounds may look as if it contains a great deal of information concerning the relationship between the patterns involved. However, if 15 descriptors were used to describe 50 compounds, this could easily be a random result.

These results should not discourage the use of these PR techniques, but it should encourage care in their use. Any numerical technique has limitations. A knowledge of these limitations is essential for the proper application of the techniques. These results will have little bearing on large studies involving hundreds of compounds where it is unlikely that the value of d/N will become very large due to computational constraints. In a case where this ratio is small, high separations can be interpreted with confidence. However, for small studies, random classifications could become large and could be misinterpreted. This is especially true if the class sizes are unequal. In such a case, even a few descriptors may fortuitously provide high classifications. This situation could easily arise in some SAR studies where the number of compounds from which to derive a relationship is small.

Of course, this does not mean that a high-dimensional relationship might not exist between the compounds in the worklist. It does mean, however, that if the number of compounds is small, then a randomly selected LDF with high separatory powers would be more likely to be obtained than one which actually explained the relationship. This is true since many random LDFs with high separatory ability would be possible and only a few LDFs would actually explain the SAR of interest.

The 1/3 ratio may avoid random, 100% separations; however, it may be unrealistically high if separations less than 100% are to be interpreted. Unfortunately, Figure 6 shows no clear delimiters of what a "safe" ratio may be. Probably the best approach to a PR LDF study is a knowledge of this curve and its implications. Certainly, if the ratio of d/N is kept low, then the chance of obtaining high random classifications will also be kept low.

Several other steps can also be taken to surmount this problem. First, more compounds could be added to the study. This would effectively decrease the ratio of d/N. As the number of patterns is increased, the random classifications due to a given dimensionality will decrease. This, however, is often impractical due to the limited amount of available data. Another approach would be to check the predictive ability of a LDF by external prediction. Compounds that are not involved in developing the LDF are "external" to its development. If their corresponding patterns are of the same type as those of those compounds used for the LDF development and if the LDF does explain some relationship between the "internal" compounds, then this LDF should be effective in correctly predicting the classes of the new compounds.

In addition, measures of the clustering of the individual classes may be valuable. The application of LDFs assumes some clustering of the data in the descriptor space; as such, the discriminants serve as a data reduction tool. Discriminants are appealing because they can separate two classes even if those classes are diffuse or irregularly shaped or consist of several clusters. In such cases, intraclass clustering may not be obvious; however, absence of any evidence of such clustering may be reason for suspecting an LDF's classification results, especially if they are not considerably higher than those shown in Figure 6.

Lastly, a problem that requires a large number of descriptors in order to explain the activities may be unrealistically complex. The problem might consist of smaller, mechanistically different subgroups of compounds. If such were the case, it might be unrealistic to include these subgroups in the same study.

As noted under Methodology, the data used in the studies presented here were artifically generated. There was no experimental noise, and the random descriptors were continuous, largely uncorrelated, and contained no omitted values. In these ways, they are unlike much of the "real" data to which these PR methods are often applied. We are currently investigating the effects of such nonideal data; the studies presented here, however, can serve as a comparison to those other works that deal with similar effects and that use similar data.[3,10]

As a final note, we would like to contrast this work with that of Topliss and Edwards.[11] They studied the effect of screening large numbers of variables with multiple linear regression analysis and showed that if the number of screened variables becomes large, the probability of achieving a random high correlation becomes large, also. We are currently investigating this effect in PR LDF studies; however, this study does not involve a screening of variables. This study is concerned only with the random classifications due to mathematical artifacts that are a result of the ratio of the number of descriptors to the number of observations. The parallel concern in regression studies would be the effect on the correlation coefficient as the number of variables was increased.

ACKNOWLEDGMENT

This research was supported by the U.S. Environmental Protection Agency under Cooperative Research Agreement CR 807531. The contents do not necessarily reflect the views

of the Agency, and no official endorsement should be inferred. The PRIME 750 computer used in these studies was purchased, in part, with the support of the National Science Foundation.

REFERENCES AND NOTES

(1) Nilsson, N. J. "Learning Machines"; McGraw-Hill: New York, 1965.
(2) Tou, J. T.; Gonzalez, R. C. "Pattern Recognition Principles"; Addison-Wesley: Reading, MA, 1974.
(3) Stuper, A. J.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1976**, *4*, 238.
(4) Schrage, Linus *Assoc. Comput. Mach. Trans. Math. Software* **1979**, *5*, 132.
(5) Muller, Mervin E. *J. Assoc. Comput. Mach.* **1959**, *July*, 376.
(6) Stuper, A. J.; Brugger, W. E.; Jurs, P. C. "Computer Assisted Studies of Chemical Structure and Biological Function"; Wiley-Interscience: New York, 1979.
(7) Moriguchi, Ikuo; Komatsu, Katsuichiro; Matsushita, Yasuo *J. Med. Chem.* **1980**, *23*, 20.
(8) Pietrantonio, Lucio; Jurs, Peter C. *Pattern Recognition* **1972**, *4*, 391.
(9) Fleiss, J. L. "Statistical Methods for Rates and Proportions"; Wiley: New York, 1973.
(10) Whalen-Pedersen, E. K.; Jurs, P. C. *J. Chem. Inf. Comput. Sci.* **1979**, *19*, 264.
(11) Topliss, John G.; Edwards, Robert P. *J. Med. Chem.* **1979**, *22*, 1238-1244.

Substructure Searching of Heterocycles by Computer Generation of Potential Aliphatic Precursors

RICHARD L. M. SYNGE

School of Chemical Sciences, University of East Anglia, Norwich, NR4 7TJ, England

Received May 8, 1984

Heterocyclic structures, in natural and synthetic organic compounds, are mostly formed by cyclizations of aliphatic precursors. Computer programs have been written that notionally reverse this process. By breaking, in turn, one of the heterobonds in each ring of each heterocyclic region of a molecule, a permuted set of aliphatic tree structures is generated. Exhaustive search of the "bonded-atom" strings present in such a set of trees can reveal unsuspected structural and biosynthetic aspects of a molecule. Some current methods of substructure searching, by over-emphasizing cyclic structures, fail to detect these relationships.

INTRODUCTION

Heterocyclic structures tend to arise, both in nature and in the laboratory, by cyclizations of aliphatic compounds. It was recognition of this general principle that gave Robert Robinson such great insight into structures and laboratory syntheses of natural products.[1]

Much of what Robinson grasped intuitively has since been confirmed (in a general way) by biosynthetic experiments using isotopic tracers. Conversely (and using computer programs rather than experiment), it should be possible to "disinter", from heterocyclic structures, some of their hypothetical aliphatic precursors, thus throwing light on analogies in the biosynthetic pathways that have produced them and, perhaps, on their pharmacological properties. This paper describes one possible approach to this goal. This may also turn out helpful as a cheap (and, in some respects, impressionistic) aid in substructure searches of not-too-large files. It may also have its uses in tracking down compounds that fall within "generic" (Markush) claims in chemical patents.

In general, chemical nomenclatures and notations have been based, almost obsessively, on cyclic structures. This has been an inevitable consequence of the classificatory system developed in *Beilstein*. It is interesting that Prager and Jacobson, in their preface to the still-current 4th edition,[2] emphasized that *different* systematic arrangements were needed for particular purposes. Such a need is manifest in terpenoid chemistry, where the cyclizations are of less interest than the aliphatic skeletons that may or may not become cyclized. Randić and Wilkins[3] have developed interesting algorithms for searching such structures. Their approach, like that of this work, attaches equal emphasis both to cyclic and to aliphatic structures.

Graph theory[4] shows that, if *one* interatomic bond in each ring of a cyclic compound is broken, an aliphatic tree structure will result, without disconnection of any part of the molecule. In devising the present approach, heterocyclic moieties (or "domains") of each molecule have been treated separately, where a carbocyclic structure comes between them, and carbocyclic structures have been excluded from the manipulations. In each heterocyclic "domain", the ring bonds in which heteroatoms participate are broken in turn, so that all possible break permutations give rise to aliphatic trees (see Figure 1).

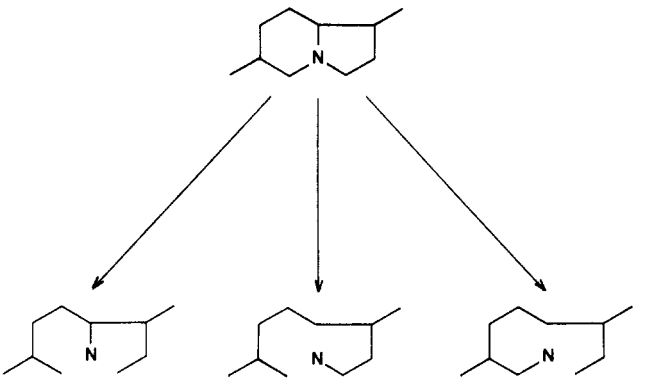

Figure 1. Formulas (lower) show the three heterobond break permutations by which the (upper) bicyclic heterocyclic compound is aliphatized.

Heterobond breakage in symmetrical compounds will yield the same aliphatic product more than once. However, such symmetry is rare among organic compounds, so nothing has been done to curtail resulting redundancies. Stereochemistry has been ignored throughout. Bond breakage is notionally by hydrogenolysis—thus, the structures produced in Figure 1 are all aliphatic primary amines. Notionally too, unsaturated

0095-2338/85/1625-0050$01.50/0