Kekule structure[15,16] and are not aromatic.

From the above paragraph, it may be concluded that a systematic complete pseudoreduction from peri-fused ring system to an acyclic polyenyne is beyond our grasp. Although a less ambitious task, as suggested in reference 2, is not precluded, nevertheless, because a chemically viable (regardless of how energetically undesirable) pathway from "reactant" to "product" is not evident, the isomorphisms shown in reference 2 for coronene and ovalene appear to be merely a convenient balancing of bonds and hydrogen atoms between the initial and final forms—rather than there existing an actual pathway that *could* achieve this conversion.

REFERENCES AND NOTES

(1) Dias, J. R. "A Periodic Table for Polycyclic Aromatic Hydrocarbons. 4. Isomer Enumeration of Polycyclic Conjugated Hydrocarbons. 2". *J. Chem. Inf. Comput. Sci.* **1984**, *24*, 124–135.
(2) Ibid., p 128.
(3) For purposes of this paper, "rings" will have the implication of being nondegenerate (i.e., a triple bond will not be viewed as a conjugated two-atom ring) and simply closed (i.e., the only Dewar benzene structures considered will have parallel bonds. Crossed bond structures will be disallowed).
(4) From a mathematical bias, this particular set of polycyclic aromatic hydrocarbons may be formulated by starting with a "stack" of benzene "modules" and successively annellating one module at a time to the previous structure in order to form the final compound.
(5) Balaban, A. T.; Harary, F. "Chemical Graphs—V. Enumeration and Proposed Nomenclature of Benzenoid Cata-condensed Polycyclic Aromatic Hydrocarbons". *Tetrahedron* **1968**, *24*, 2505–2516.
(6) International Union of Pure and Applied Chemistry *Nomenclature of Organic Chemistry*: *Section A*; Pergamon: Oxford, England, 1979; p 23.
(7) The reason for this particular choice, or more precisely the problem with any other choice, will be described later in this section.
(8) Although pentacovalent, and even hexacovalent, carbon compounds do exist,[9] they are rare in occurrence; therefore, in this proposed pseudoconversion scheme they have been purposefully excluded as an allowed product.
(9) Greenberg, A.; Liebman, J. F. *Strained Organic Molecules*; Academic: New York, 1978; pp 375–385.
(10) Within the constraint of a maximum of four bonds per atom, polyfilar[11] fused ring systems do not lend themselves to strictly linear names; i.e., there does not exist an 18-atom linear name for this starphene. The correlation of this nonlinearity with the fact that it is not possible for all three rings of a starphene to be in aromatic conjugation at the same time[12] may be more than coincidence; however, because this is a *pseudo*conversion, rather than an observed reaction, formulation of a mechanism for this has not been investigated.
(11) Taylor, F. L. "Enumerative Nomenclature for Organic Ring Systems". *Ind. Eng. Chem.* **1948**, *40*, 735.
(12) Clar, E. *The Aromatic Sextet*; Wiley: London, 1972; p 37.
(13) Elk, S. B. "A Nomenclature for Regular Tessellations and Its Application to Polycyclic Aromatic Hydrocarbons". *Math. Chem.* **1980**, *8*, 126.
(14) Klarner, D. A. "Cell Growth Problems". *Can. J. Math.* **1967**, *19*, 851–863.
(15) Clar, E. *The Aromatic Sextet*; Wiley: London, 1972; p 108.
(16) Ibid., p 110.

Heuristic Approaches to the Design of a Cybernetic Electroanalytical Instrument

HARI GUNASINGHAM

Department of Chemistry, National University of Singapore, Kent Ridge, Singapore 0511

Received November 6, 1985

The design of a "cybernetic" electroanalytical instrument requires consideration of both heuristic and deterministic procedures. The former brings to bear expert knowledge to obtain a solution at minimum cost. Heuristic knowledge is most conveniently represented in an expert system structure where the knowledge base is separated from the inference procedure. A system design based on PROLOG is described that obtains a solution by a heuristic depth-first search of an AND/OR graph.

INTRODUCTION

The rapid development of microelectronics technology has had a major impact on electroanalytical instrumentation. Although up to the late 1970s the major advancements have been in speed, accuracy, and sensitivity, we appear to be reaching the practical limits to further improvements in the sheer physical performance of the hardware. There is, for example, little point in improving sensitivity if interferences mask or distort the analytical signal. Rather, what is needed is better and more subtle qualitative information as well as new ways of assessing experimental data. This perception may be seen as the driving force for the future development of electroanalytical instrumentation.[1-3]

Recent work has sought to use computer-based techniques to improve the information content of electroanalysis. A number of software packages with limited inferential capabilities have been developed that enable the extraction and interpretation of information from the raw experimental data. In general, these packages employ rigorously defined algorithms based on statistical and computational methodologies to transform the raw experimental data. The transformed data are then compared with a theoretical model, and the conformity or deviation from the model guides the inference. With the advent of inexpensive personal computers, work in the area of software development has gained impetus.

Examples of analytical and inferential software include the following: use of a stored library of reaction schemes that could be automatically fitted to experimental data by computer developed by Harrison and Small;[4] automated analysis of data from a number of techniques by simplex fitting by Hanafay et al.;[5] pattern recognition for classifying Fourier-transformed cyclic voltammograms by Schachterle and Perone;[6] automatic classification of coulometry data with deviation-pattern recognition by Meites and Shia;[7] mechanistic classification with deviation-pattern recognition by Rusling et al.,[8-11] use of a computed function to assess electroanalytical data by Maloy;[12,13] automatic feature extraction from cyclic voltammograms by Speiser.[14] The above software were mainly applied to mechanistic analysis.

The computer can also be used to enhance the information content of electroanalysis by the intelligent control of the experiment itself. For example, complex potential waveforms and current-measurement protocols may be applied.[15] However, perhaps the most important use of the computer in experimental control lies in the fact that is can bring to bear a number of electroanalytical techniques to solve a problem.

0095-2338/86/1626-0130$01.50/0

Faulkner and co-workers[1-2,16] have described a microcomputer-based instrument that is able to intelligently coordinate a repetoire of electroanalytical techniques.

The ultimate vision of computerized electroanalysis lies in the bringing together of intelligent experimental control and automated inferential and decision-making capabilities so that the two may function in closed-loop interaction. Such a system has been proposed by Faulkner and co-workers as a "cybernetic potentiostat".[1-2]

HEURISTICS IN PROBLEM SOLVING

The inferential and control software that has been so far developed is amenable to algorithmic representation in that it is rigidly defined. Such processes may be termed deterministic. There is, however, another approach to assessing data and providing control that cannot be rigidly defined but is equally important. We are referring to the way the expert electroanalytical chemist functions. Experts usually apply the wealth of their knowledge of the subject (the expertise) to the problem at hand. Often, such expert knowledge is simply a rule-of-thumb with little theoretical basis. This approach to problem solving is heuristic.[17,18] The expert knowledge of an electroanalytical chemist is painstakingly obtained by empirical observation over the years or derived from other chemists, principally through the literature.

Conventional deterministic procedures are well-defined and posses little or no ambiguity. Heuristic procedures on the other hand may posses a high degree of uncertainty, because of uncertainty in the data in the first instance or the lack of an appropriate theory (either because there is no theory or because the theory is too complex to afford a solution without immense cost in computer resources).

By way of example, the electroanalytical chemist knows from experience that in order to determine copper by anodic stripping voltammetry (ASV), a suitable stripping medium such as dilute nitric acid may be used. Of course he may have been able to deduce this from first principles based on the structure and stability of copper complexes; however, this would be extremely costly in computer resources. The problem would be compounded if a number of possible alternatives exist and where the correct choice depends on experimental constraints. For the above example, complications arise when copper has to be determined in the presence of excess zinc with a thin mercury film electrode.[19]

In regard to the cybernetic instrument proposed by Faulkner, there is much to be gained if such a system could employ heuristic procedures to augment the conventional deterministic ones. The approach to this problem has been given by the artificial intelligence field of knowledge-based or expert systems. Knowledge-based systems employ convenient structures for representing and using heuristics. Where they differ from conventional software systems is in the differentiation between the knowledge base and the inference procedure. The former contains the rules-of-thumb plus any data or facts that may be relevant to the problem domain. The latter guides the extraction of plausible inferences from the knowledge base. By keeping the two distinct, it is feasible to alter the knowledge base without affecting the inference procedure.[18,20]

The subject of this paper concerns the design of a knowledge-based system as a component part of a cybernetic electroanalytical instrument. Although the emphasis is on the design, specific examples are given that explore the role of heuristic procedures in the planning of experiments and the assessment of experimental data.

SYSTEM DESIGN

Figure 1 describes an implementation of a cybernetic system for electroanalysis that combines deterministic and heuristic

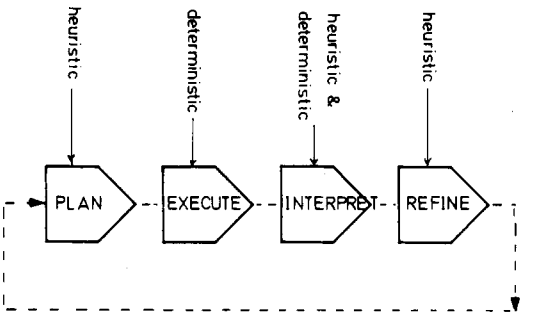


Figure 1. A cybernetic electroanalytical instrument is perceived as consisting of four phases. Apart from the execute phase, heuristic strategies may be applied to each of the phases.

procedures. The system consists of four phases.

The first phase implements the planning function. It is perceived as being largely heurisitic. The plan phase devises the experimental strategy, that is, the sequence of electroanalytical experiments that have to be performed, the data that have to be extracted, and the appropriate experimental conditions. The plane phase is crucial to the solution as any error would propagate into the remaining phases.

The execute phase is obviously deterministic. Included in this phase are the control procedures that execute the actual electroanalytical experiments. The computer may generate any conceivable potential waveform to be applied to the test cell and then acquire the current reponse. A repetoire of techniques in the way suggested by Faulkner may also be employed. The control programs also include those which implement the automation of sample handling (and here we may include the manipulation of robots) and peripheral control (e.g., disk drives, plotters, and display screen).

The interpret phase is seen as a combination of deterministic and heuristic procedures. The computer first seeks a solution by subjecting the raw experimental data to some transformation, which may be fitted to a theoretical model, as described earlier. Heuristic strategies may be employed to narrow the search for an appropriate model.[17] However, in the case of incomplete information or deviations from theory, other heuristic approaches may be employed to clarify or refine the interpretation. The success of an interpretation is rated by a certainty factor.

In the final refine phase, the system evaluates the success of the experiment and, depending on whether a satisfactory solution has been achieved, modifies the knowledge base, either by adding some new information or by modifying or deleting existing information. A satisfactory solution is determined by a weighted summation of the certainty factors derived in the interpret phase. In the event an unsatisfactory solution has been obtained, the system returns to the plan phase to see whether it could set up an improved experimental strategy based on the results of the previous experiment, which have now been encoded into the knowledge base.

In the above description, although our ideal is completely unattended operation, at every phase the user is able to interact with the system to guide it to a solution by posing constraints or by aiding in the interpretation. The system structure also affords automatic learning as each successful solution is added to the knowledge base and subsequent experiments may then have access to this solution.

PROLOG

LISP and PROLOG are the two main languages that have been used for building knowledge-based systems. Whereas LISP has been long popular in America, PROLOG has only recently gained attention, primarily because of the Japanese fifth-generation artificial intelligence project.[21] We have chosen to implement the knowledge-based system in PROLOG

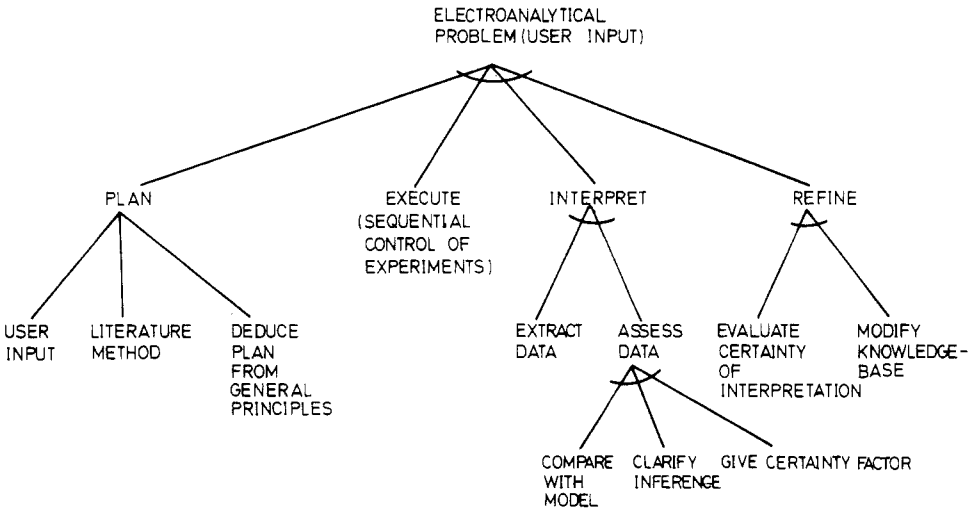


Figure 2. System implementation of the four phases of a cybernetic electroanalytical instrument as a heuristic search of an AND/OR graph.

because of its simplicity of use and elegant structure based on first-order predicate logic.[22-24] For the sake of clarity in the explanation of the present system implementation, essential features of the language are presented here.

A knowledge-based system essentially consists of rules and facts concerning the applications domain and a control structure to deduce inferences from the knowledge base. PROLOG efficiently represents rules and facts as relations. Also, programming simply involves declaring the knowledge base.

"Thiocyanate forms a stable complex with mercury" is a fact. This may be represented in PROLOG as

(i) stable-complex (thiocyanate,mercury)

We can add more facts to the knowledge base:

(ii) stable-complex (cyanide,mercury)

(iii) stable-complex (chloride,mercury)

(iv) stable-complex (thiocyanate,mercury)

(v) soluble-complex (nitrate,mercury)

(vi) insoluble-complex (chloride,mercury)

Fact vi states that thiocyanate forms a soluble complex with mercury.

A rule may have the form

X is a suitable stripping matrix for Y if X forms a stable complex with Y and X forms a soluble complex with Y

where X and Y are variables. In PROLOG, this may be represented by

(vii) stripping-matrix (X,Y) :- stable-complex (X,Y),
soluble-complex (X,Y)

Generating Inferences with PROLOG. The knowledge base that has been declared above does not commit itself to an inference procedure. This occurs automatically when a query (the problem to be solved) is presented to the PROLOG interpreter, which then searches through the knowledge base and, by a method of pattern matching, deduces a solution. The control structure for deducing an inference may thus be altered by altering the information encoded in the knowledge base. One of the features PROLOG is that it enables automatic backtracking when a search for a solution along a particular line fails.[23,24]

In the selection of a suitable medium for anodic stripping voltammetry, a number of practical constraints have to be satisfied such as the solubility and stability of the complex between the metal to be determined and the stripping medium. For the case of mercury, facts i–vii and rule viii represent the knowledge base from which a solution may be deduced.

The PROLOG query

?- stable-complex (X,mercury)

asks which matrices form stable complexes with mercury. PROLOG arrives at the answer by searching through the knowledge base to find instances of the above predication. It finds the answers

X = thiocyanate

X = chloride

X = bromide

A more complex query requires invocation of rule viii:

?- stripping-matrix (X,mercury)

PROLOG answers this query by checking if the assertions on the right-hand side of the rule are satisfied; that is, it must instantiate values of X for the assertions stable-complex (X,mercury) and soluble-complex (X,mercury). It returns with the answer

X = thiocyanate

SYSTEM IMPLEMENTATION

The first step in implementing any heuristic procedure is to define the problem as a state-space search. Problem reduction is an efficient way of doing this.[18,24] The approach is to decompose the problem into increasingly smaller subproblems, which can be more easily solved. The result is an AND/OR graph. Figure 2 shows an AND/OR graph that implements the four phases of the cybernetic instrument. The solution to this graph is obtained by a depth-first search having a left-to-righ hierarchy.

In response to a user query (the problem to be solved), PROLOG invokes the pattern-matching procedure, which results in the search of the AND/OR graph. Along each branch of the graph the search is highly selective and depends on the results of searches along previous branches. The user may also pose constraints to limit the search. In the event of failure of a search along a particular branch, PROLOG's in-built backtracking facility enables the system to automatically seek a solution with a modified experimental strategy.

In general, the solution to any plan problem has three alternatives, and the search of these is done in a best first fashion. First, the user may propose the complete plan overiding the

system. Second, the system can extract an appropriate plan from the knowledge base that matches the application at hand. The plan would be a literature method in that it has been already validated and recorded in some scientific paper (and is generally accepted to be a correct solution). And finally, the system can actually deduce a plan from general principles. In all three, the plan must select the appropriate electroanalytical technique(s), select the optimal experimental conditions, and decide on what data interpretation would be necessary.

It should be noted that PROLOG only explicitly invokes that part of the AND/OR graph it wishes to search. For example, suppose the user wants to determine some heavy metals in biological samples. He inputs the list of metals to be determined and any relevant constraints such as sample origin and suspected concentration ranges. If the user does not wish to overide the system and propose a plan, the system first searches for a literature method. A suitable method may one due to Adeloju et al.,[25] who proposed a simple multielement approach for the sequential determination of heavy metals by polarography, stripping voltammetry, and adsorption voltammetry. In the event the first two options had failed, PROLOG would backtrack to the last branch and plan a separation from general principles.

For this example, the devising of a plan from general principles requires a heuristic selection strategy: First, the system selects a technique appropriate for the maximum number of elements. Then, the system selects a technique for the maximum number of the remaining elements. It goes on in this way until all the elements have an appropriate technique. In selecting the technique, the system must also determine the optimal experimental conditions such as the electrolyte solution, bearing in mind sensitivity and selectivity.

Clearly, the plan obtained from the literature would be more efficient than one deduced from general principles. In the present example, Adeloju's literature method enables the analysis of all the elements in a single sample.[25] The more general deduced method being less expert would generate a plan that would require multiple samples.

The search along the execute phase branch is deterministic and always succeeds. In contrast with the heuristic procedures that are implemented in PROLOG, deterministic procedures are most conveniently represented in procedural languages such as BASIC and Assembly language. An interface between the two language systems is thus required. This may be done by running the procedural languages on one microcomputer and PROLOG on another. Communication can then be afforded through a serial link. A detailed description of the implementation of a loosely coupled PROLOG interface is given in a separate paper.[26]

Many of the deterministic techniques for interpreting electroanalytical data involve an exhaustive search for a plausible solution. For example, Rusling has proposed the search of a binary tree.[8-10] And in the approach of Harrison[3] and Hanafay et al.,[4] the fitting of data to the appropriate mechanism requires a search of all mechanisms in the library. Such exhaustive searches are costly in computer time and may be prohibitive if the number of possible solutions increases significantly. Heuristic search procedures can, however, narrow the search by posing contraints derived from the knowledge base.

In the solution to the problem of data interpretation, the treated experimental data are first subject to comparison with a theoretical model by a heuristic search of possible solutions. On failure of this search (or if there is a high degree of uncertainty in the inference), other heuristic strategies may be brought to bear. For instance, empirical evidence may be employed to clarify or increase the certainty of an inference. Much benefit may also be derived by the use of analogy as guide to the solution. For example, an analogous treatment arises in the analysis of the stripping voltammetry of bromide on lead and silver.[27]

In the final refine branch, the search examines whether a satisfactory solution has been obtained. The assessment is based on a summation of the weighted certainty factors derived from the interpret phase. If the interpretation is inconclusive, then the system returns to the plan phase to examine a new strategy. On the other hand, if the certainty of the solution is high, the refine phase can add it to the knowledge base. A powerful feature that may be employed in the refine phase is the use of induction to derive rules where we start with specific examples that can be generalized to a higher level rule.

It should be noted that the control structure of PROLOG naturally implements a depth-first search of the AND/OR graph. However, it is feasible to implement a breadth-first search.[17,24] The advantage of the latter is that an initial experimental strategy may be tested. For example, choice between the use of differential pulse polarography (DPP) and anodic stripping voltammetry (ASV) may depend on the concentration ratio of the metals to be determined.[25] On the basis of a user input on an assumed concentration, the plan phase could suggest one or the other technique for each metal. On execution of the initial experiment, however, it may be found that the initial assumption was incorrect, in which case a fresh plan is generated. The advantage of the breadth-first search is that it affords flexibility. The disadvantage is that it requires large computing resources and must be explicitly implemented in PROLOG.

In the present implementation, the refine phase is the least developed of the four phases, and the plan and execute phases are the most developed. However, this is an acceptable feature of building knowledge-based systems where the development is an incremental process that usually takes many man years. The present system design in PROLOG is, however, adequate for even a large knowledge base. The crucial work that remains to be done is the actual knowledge acquisition.

REFERENCES AND NOTES

(1) He, P.; Avery, J. P.; Faulkner, L. R. "Cybernetic Control of an Electrochemical Repertoire". *Anal. Chem.* **1982**, *54*, 1313A–1326A.
(2) He, P.; Faulkner, L. R. "Digital Computers in Electrochemistry". *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 275–282.
(3) Gunasingham, H.; Ang, K. P. "Voltammetry as a Model for Teaching Chemical Instrumentation". *J. Chem. Educ.* **1985**, *62*, 611–612.
(4) Harrison, J. A.; Small, C. E. "The Automation of Electrode Kinetic Measurements. I. The Instrumentation and the Fitting of the Data Using a Library of Reaction Schemes". *Electrochim. Acta* **1980**, *25*, 447–452.
(5) Hanafay, M. K.; Scott, R. L.; Ridgway, T. H.; Reilly, C. N. "Analysis of Electrochemical Mechanisms by Finite Differential Simulation and Simplex Fitting of Double Potentiostat Current, Charge and Absorbance Responses". *Anal. Chem.* **1978**, *50*, 116–137.
(6) Scharchterle, S. D.; Perone, S. P. "Recognition of Voltammetric Data by Computerized Pattern Recognition" *Anal. Chem.* **1981**, *53*, 1672–1678.
(7) Meites, L.; Shia, G. A. In *Chemometrics*; Kowalski, B. R., Ed.; American Chemical Society: Washington, DC, 1977; pp 127–152.
(8) Rusling, J. F. "Computerised Interpretation of Electrochemical Data Using Deviation-Pattern Recognition". *TrAC, Trends Anal. Chem. (Pers. Ed.)* **1984**, *3*, 91–94.
(9) Rusling, J. F. "Computerized Method for Mechanistic Classification of One-Electron Potentiostatic Current–Potential Curves". *Anal. Chem.* **1983**, *55*, 1713–1718.
(10) Rusling, J. F. "Applications of a Computerized Method for Mechanistic Classification of One-Electron Potentiostatic Current–Potential Curves". *Anal. Chem.* **1983**, *55*, 1719–1723.
(11) Rusling, J. F.; Conners, T. F. "Determination of Rate Constants of Pseudo-First-Order Electrocatalytic Reactions from Overlapped Voltammetric Data". *Anal. Chem.* **1983**, *55*, 776–781.
(12) Therdteppitak, A.; Maloy, J. T. "Digital Electrochemical Transient Analysis: Diagnosis of an Elementary Reaction Mechanism as an Illustration of a Data Acquisition Strategy". *Anal. Chem.* **1984**, *56*, 2592–2594.
(13) Therdteppitak, A.; Maloy, J. T. "Instrumental Interpretation of Current Transients: A Harbinger of an Artificial Intelligence for Electroanalysis". *1985 Electroanalytical Symposium*; BAS Press:

West Lafayette, IN, 1985; Abstr. 62.
(14) Speiser, B. "Multiparameter Estimation: Extraction of Information from Cyclic Voltammograms". *Anal. Chem.* **1985**, *57*, 1390-1397.
(15) Borman, S. A. "New Electroanalytical Pulse Techniques". *Anal. Chem.* **1982**, *54*, 698A-705A.
(16) Eklund, J. A.; Faulkner, L. R. "Pursuing Major Conclusions: The Intelligent Director of an Electrochemical Repertoire". *1985 Electroanalytical Symposium*; BAS Press: West Lafayette, IN, 1985; Abstr. 64.
(17) Georgeff, M. P. "Strategies in Heuristic Search". *Artificial Intelligence* **1983**, *20*, 313-425.
(18) Barr, A.; Feigenbaum, E. *Handbook of Artificial Intelligence*; William Kaufman: Los Altos, CA, 1982; Vol. 1.
(19) H. Gunasingham, K. P. Ang, and C. C. Ngo, in preparation.
(20) Nau, D. S. "Expert Computer Systems". *Computer* **1983**, *15*, 63-84.
(21) Fuchi, K. "The Direction the FGCS Project Will Take". *New Generation Comput.* **1983**, *1*, 3-9.
(22) Gunasingham, H.; Srinivasan, B.; Ananda, A. L. "Design of an Expert System for Planning HPLC Separations". *Anal. Chim. Acta*, in press.
(23) Dahl, V. "Logic Programming as a Representation of Knowledge". *Computer* **1983**, *16*, 106-111.
(24) Rich, E. *Artificial Intelligence*; McGraw Hill: New York, 1983.
(25) Adeloju, S. B.; Bond, A. M.; Briggs, M. H. "Multielement Determination in Biological Materials by Differential Pulse Voltammetry". *Anal. Chem.* **1985**, *57*, 1386-1390.
(26) A. L. Ananda and H. Gunasingham, in preparation.
(27) Bard, A. J.; Faulkner, L. R. *Electrochemical Methods*; Wiley: New York, 1980.

Molecular ID Numbers: By Design†

MILAN RANDIĆ*

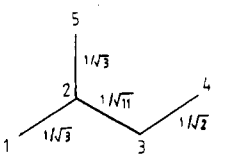
Department of Mathematics and Computer Science, Drake University, Des Moines, Iowa 50311, and Ames Laboratory, Iowa State University, Ames, Iowa 50011

Received November 27, 1985

The paper improves the discrimination ability of ID (identification) numbers by making use of a new set of weights for bonds, based on prime numbers.

Recently, I proposed a novel structural index—called molecular ID number—as a potentially useful label for molecular skeletons.[1] The index was a result of developing characterizations of complex molecules with weighted paths.[2] It proved valuable in a nonempirical approach to structure–activity. For example, use of the index has made it possible to classify a dozen anticholinergic compounds among some 40 therapeutically useful drugs exhibiting other activities.[3] When ID values were derived for all alkanes up to, and including, undecanes (in all, over 300 acyclic structures), it was observed that in no case did a duplicate numerical value occur. This suggested that this parameter, which is the total number of all suitably weighted paths (vide infra), is a *potential* structure discriminator. One can contrast this index to other graph theoretical (frequently referred to somewhat imprecisely as topological) indices of limiting discriminatory power: Hosoya's Z index,[4] the connectivity index,[5] and Balaban's J index,[6] which have been recently examined in a comparative study.[7] The uniqueness of ID was neither claimed nor was easy to prove (as is true for many other schemes related to graph isomorphism problem). Usually, the lack of uniqueness can be established by finding a counterexample. The report on molecular ID numbers[1] ended with an invitation to search for a counterexamples. A systematic search can provide insight into the number of counterexamples over a certain field size, an important information in evaluating practial use of such discriminators. While in mathematics interest in a conjecture immediately collapses as a single counterexample is found, in *applied* fields this is not necessarily the case. For example, many cluster analyses frequently do not suggest a single compound as candidate.[8] Thus, despite the unresolved issue of uniqueness, molecular ID numbers remain of considerable practical potential. Figueras[9] upgraded the existing ALL PATH program, which enumerates paths of different length in a graph, written in BASIC[10] by offering a much faster (turbo) Pascal version and at the same time extended the considerations to incorporate heteroatoms as distinctive items. It appeared that the ID numbers—arrived at by accident—have desirable features and promise in differentiating a large body of chemical compounds.

Table I. Bond Weights Based on the First Nine Prime Numbers and Enumeration of All Paths in 2-Methylbutane Using Prime Number Weights $1/[P_i]^{1/2}$ [a]



bond type	weight
(1, 2)	$1/2^{1/2}$
(1, 3)	$1/3^{1/2}$
(1, 4)	$1/5^{1/2}$
(2, 2)	$1/7^{1/2}$
(2, 3)	$1/11^{1/2}$
(2, 4)	$1/13^{1/2}$
(3, 3)	$1/17^{1/2}$
(3, 4)	$1/19^{1/2}$
(4, 4)	$1/23^{1/2}$

atom	p_1	p_2	p_3
1, 5	$1/3^{1/2}$	$1/(3^{1/2}\cdot3^{1/2}) + 1/(3^{1/2}\cdot11^{1/2})$	$1/(3^{1/2}\cdot11^{1/2}\cdot2^{1/2})$
2	$2(1/3^{1/2}) + 1/11^{1/2}$	$1/(11^{1/2}\cdot2^{1/2})$	
3	$1/2^{1/2} + 1/11^{1/2}$	$2[1/(11^{1/2}\cdot3^{1/2})]$	
4	$1/2^{1/2}$	$1/(2^{1/2}\cdot11^{1/2})$	$2[1/(2^{1/2}\cdot11^{1/2}\cdot3^{1/2})]$

$P_1 = 1/2^{1/2} + 2/3^{1/2} + 1/11^{1/2} = 2.163\,318$
$P_2 = 1/(2^{1/2}\cdot11^{1/2}) + 2/(3^{1/2}\cdot11^{1/2}) + 1/3^{1/2} = 0.894\,689$
$P_3 = 2/(2^{1/2}\cdot3^{1/2}\cdot11^{1/2}) = 0.246\,182$
$ID = P_1 + P_2 + P_3 + N = 8.304\,191$

[a] Bond type case (1, 1) is unimportant because it concerns only ethane.

However, Szymanski, Müller, Knop, and Trinajstić[11] undertook to systematically examine all acyclic structures beyond undecanes up to $n = 20$ carbon atoms using their very efficient program of generating trees.[12] In the field of 618 050 (i.e., well over half a million) structures, they found 124 pairs of nonisomorphic alkanes and a triple having exactly the same ID number (some of these are shown in Figure 1). The smallest pair has 15 vertices, two graphs among 13 476 different possible alkanes with 15 carbon atoms![13] There is one pair among alkanes having 16 carbon atoms (i.e., among 10 359); there are four pairs of alkanes having 17 carbon atoms (i.e., among 60 529); etc. All the computations were performed

† Ames Laboratory is operated for the U.S. Department of Energy by Iowa State University under Contract W-7405-Eng-82. This work was supported by the office of R. S. Hansen, Director.

* Address correspondence to the author at Drake University.

