(8) M. D. Cooper, *Anal. Chem.*, **25**, 411 (1953).
(9) K. Goto, T. Komatsu, and T. Furukawa, *Anal. Chim. Acta*, **27**, 331 (1962).
(10) A. Okac and M. Barlusek, *Z. Anal. Chem.*, **178**, 198 (1960).
(11) E. B. Sandell, "Colorimetric Determination of Traces of Metals", 3th ed., Interscience, New York, 1959, p 83.
(12) E. M. Gates and G. H. Ellis, *J. Biol. Chem.*, **168**, 537 (1947).
(13) "Handbook of Analytical Chemistry", 1st ed., L. Meites, Ed., McGraw-Hill, New York, 1963, p 6–34.
(14) "Standard Methods for the Examination of Water and Wastewater", 12th ed., American Public Health Assoc., Inc., New York, 1965, pp 317–320.
(15) M. Nishimura and E. B. Sandell, *Anal. Chim. Acta*, **26**, 242 (1962).
(16) H. Barnes, *Analyst (London)*, **76**, 220 (1951).
(17) R. M. Rush and J. H. Yoe, *Anal. Chem.*, **26**, 1345 (1954).
(18) W. Berger and H. Elvers, *Z. Anal. Chem.*, **171**, 185 (1959).
(19) R. Püschel, *Z. Anal. Chem.*, **221**, 132 (1966).
(20) K. L. Cheng and R. H. Bray, *Anal. Chem.*, **27**, 782 (1955).
(21) D. Bettridge, P. K. Todd, Q. Fernando, and H. Freiser, *Anal. Chem.*, **35**, 729 (1963).
(22) V. Michaylova and L. Yuroukova, *Anal. Chim. Acta*, **68**, 73 (1974).
(23) S. Shibato in "Chelates in Analytical Chemistry", Vol. 4, H. A. Flaschka and A. J. Barnard, Jr., Ed., Marcel Dekker, New York, 1972, pp 1–232.
(24) R. G. Anderson and G. Nickless, *Analyst (London)*, **92**, 207 (1967).
(25) O. A. Tataev and L. G. Anisimova, *J. Anal. Chem., USSR*, **26**, 166 (1971).
(26) T. Yotsuyanagi, K. Goto, M. Nagayama, and K. Aomura, *Bunseki Kagaku*, **18**, 477 (1969); *Chem. Abstr.*, **71**, 56313n (1969).
(27) K. Ueda, Y. Yamamoto, and S. Ueda, *Nippon Kagaku Zasshi*, **90**, 903 (1969); *Chem. Abstr.*, **71**, 119327t (1969).
(28) D. Nonova and B. Evtimova, *Talanta*, **20**, 1347 (1973).
(29) M. Hniličková and L. Sommer, *Collect. Czech. Chem. Commun.*, **26**, 2189 (1961).
(30) M. Tanake, S. Funahashi, and K. Shirai, *Anal. Chim. Acta*, **39**, 437 (1967).
(31) M. Kitano and J. Ueda, *Nippon Kagaku Zasshi*, **91**, 983 (1970); *Chem. Abstr.*, **74**, 71352e (1971).
(32) H. Flaschka and A. M. Amin, *Mikrochim. Acta*, **1953**, 414.
(33) F. H. Pollard, P. Hanson, and W. J. Geary, *Anal. Chim. Acta*, **20**, 26 (1959).
(34) W. J. Geary, G. Nickless, and F. H. Pollard, *Anal. Chim. Acta*, **26**, 575 (1962).
(35) N. A. Lange, "Handbook of Chemistry", 10th ed., McGraw-Hill, New York, 1967, p 972.
(36) E. B. Sandell, "Colorimetric Determination of Traces of Metals", 3rd ed., Interscience, New York, 1959, p 612.
(37) T. Iwamoto, *Bunseki Kagaku*, **10**, 189 (1961) (as quoted in Ref. 23).
(38) S. T. Talipov, R. K. Dzhiyanbaeva, L. V. Chaprasova, and R. T. Gutnikova, *Nauchn. Tr. Tashkent, Gos. Univ.*, **263**, 72 (1964); *Chem. Abstr.*, **62**, 15425c (1965).
(39) G. Nakagawa and H. Wada, *Nippon Kagaku Zasshi*, **83**, 1098 (1962) (as quoted in Ref. 23).
(40) I. M. Klotz and W. C. Loh Ming, *J. Am. Chem. Soc.*, **75**, 4159 (1953).
(41) A. Ringbom, "Complexation in Analytical Chemistry", Interscience, New York, 1963, p 143.

RECEIVED for Review March 17, 1975. Accepted August 29, 1975. This work is part of a cooperative research program sponsored by Statens naturvetenskapliga forskningsråd (The Swedish Natural Science Research Council) and the National Science Foundation. The U.S.A. portion is under the direction of Professor Abraham Clearfield, Department of Chemistry, Ohio University, Athens, Ohio 45701.

Structural Interpretation of Spectra

N. A. B. Gray

King's College Research Centre, King's College, Cambridge, England, CB2 1ST

A computer program for the structural interpretation of spectra is described. The program uses a compound's formula, infrared, ultraviolet, proton resonance, and mass spectra. The program is based on the methods employed by spectroscopists as far as these can be identified from problem solving protocols. The program can process cyclic and acyclic polyfunctional molecules with molecular weights up to around 300 amu.

The acquisition of nuclear magnetic resonance, mass, infrared, and ultraviolet spectra is now commonly effected by laboratory computer systems. The resulting availability of computer compatible spectral data will enhance the value of computer assisted methods for identifying a compound's structure from its spectra. At least one file-search system using combined IR, UV, ^{1}H NMR, ^{13}C NMR, and mass spectra is under development (*1*); other established file-search systems have used either NMR (*2*) or IR (*3*) or mass spectra (*4, 5*) with the programs for processing mass spectral data being the most highly developed. In addition to compound identification by spectrum comparison methods, other processing methods are being developed including classification by pattern recognition techniques (*6, 7*) and structural inference by Artificial Intelligence (*8, 9*). Other empirical procedures for identifying structure from spectra by use of correlation tables have been proposed by Sasaki et al. (*10*) and by Beech et al. (*11*).

Like the Beech and Sasaki methods, the proposed method is purely empirical. The program attempts to simulate the problem solving processes employed by a spectroscopist as far as these processes can be identified from worked examples in undergraduate texts or from protocols recorded when students succeed in solving spectrum interpretation problems (*12*). Analysis of these protocols suggested a model based on the use of procedures to represent the properties and interactions of a small number of functional groups. Such procedures are used in an attempt to identify a unique set of fragments comprising the molecule and to determine how these fragments are bonded.

OVERVIEW OF METHOD

Figure 1 illustrates the spectral data used in the following sample problem solving protocol for $C_9H_{10}O_3$. "This compound must have five double bond equivalents. The infrared spectrum shows the characteristic ragged absorption of a carboxylic acid group in the region 3200–2500 cm^{-1} and the carbonyl stretch at 1720 cm^{-1}. There is one acidic proton at δ = 11.5 so there must be a $-CO_2H$ group. That leaves C_8H_9O and four double bond equivalents. There is a group of five protons in the aromatic region. That would fit with the UV which looks like a standard benzene and there are some bands in the IR at 1600 cm^{-1} and 1500 cm^{-1}. So it is probably $-C_6H_5$ which takes up all the double bonds leaving C_2H_4O. The aromatic protons at 7.2–6.8 are at comparatively high field, which suggests that the benzene is attached to an alkyl carbon or an oxygen (oxygen is more likely). That last oxygen must be an ether. The two $-CH_2-$ groups are bonded as shown by their splitting patterns. The one at 2.8 could also be bonded to the carboxylic acid; the other has a large shift of 4.2 so must have an oxygen bond. The compound must be $Ph-O-CH_2-CH_2-CO_2H$."

The student has solved this problem by considering first those groups for which the spectral evidence is most dis-

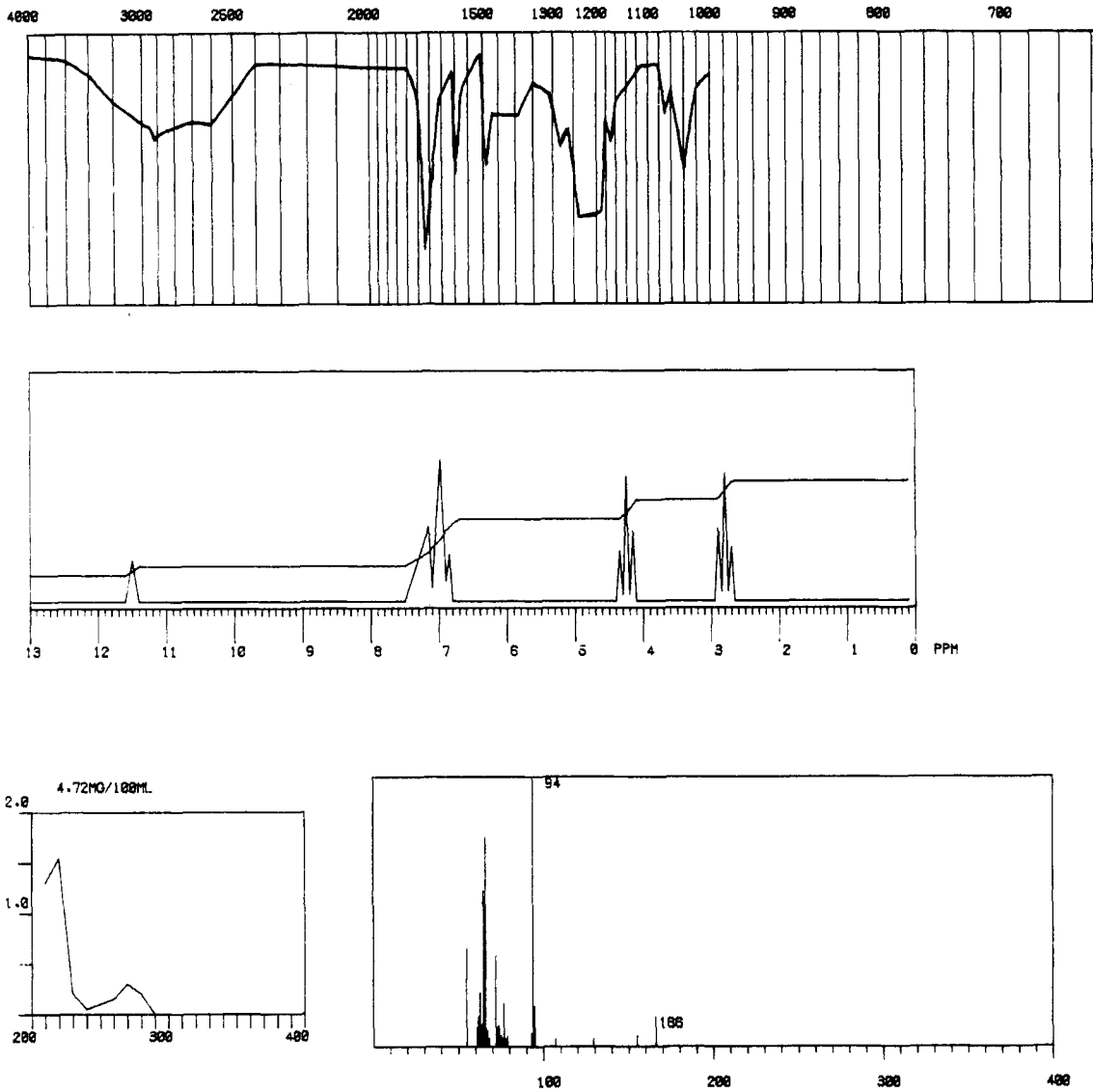


Figure 1. Spectral data used to illustrate the problem solving protocol

The plots were made from the crudely digitized data used by the program system. The original data were taken from Trost (*13*)

tinctive—such as the carboxylic acid group and the phenyl group. The assignment of atoms and double bond equivalents to identified groups simplifies subsequent processing steps. In this case, the student never considered alternative hypotheses such as the presence of a –CHO or a –CO–O–CO– groups—these possibilities having been eliminated by the assignment of the $-CO_2H$. This protocol shows how, as well as confirming the presence/absence of particular functional groups, the initial examination of the spectral data permits the imposition of constraints on the possible bonding of these groups. In simple cases, the constraints derived during the initial processing result in only a single possible molecular structure. More complex cases require the consideration of mass spectral evidence, or the use of addition rules for NMR shifts to determine how the constituent groups are combined in the molecule.

The program is closely modeled on the type of process implied by problem solving protocols such as that just described. The program is comprised of three simple modules. An input routine reads the molecular formula and spectral data, identifying centers of IR and UV absorption peaks (obtaining the extinction coefficients for the UV) and splitting the NMR spectrum into groups of peaks. The next routine—"Assign-fragments"—selects and applies those functional group procedures that are relevant to the analysis of a compound of the given composition. The functional group procedures are similar in form to the control structures in DENDRAL (*8*) or the "premise-action" rules of the MYCIN program for microbial therapy (*14*). Each begins with a set of conditional tests for spectral features, if the tests are satisfied the procedure will cause some fragment(s) to be assigned as a part of the molecule and will generate constraints on the possible bonding of the fragment(s). After a fragment has been assigned, it may be possible to simplify the list of procedures still to be applied by eliminating those which require atoms/double-bond-equivalents already committed. If the process of assigning fragments is successful, the result is a single set of fragments which together comprise the molecule. The final "Solve-structure" routine attempts to determine the connectivity matrix of these fragments. The entries in this connectivity matrix indicate the presence of single bonds between fragments (all multiply bonded groups such as >C=O or >C=C< are basic fragments). The connectivity matrix may be uniquely defined by the bonding constraints of the fragments. Otherwise, detailed processing is necessary to eliminate those extra possible bonds for certain fragments which are not consistent with the global molecular structure. (This simple form of connectivity matrix does not represent all structural details; for example, it does not dis-

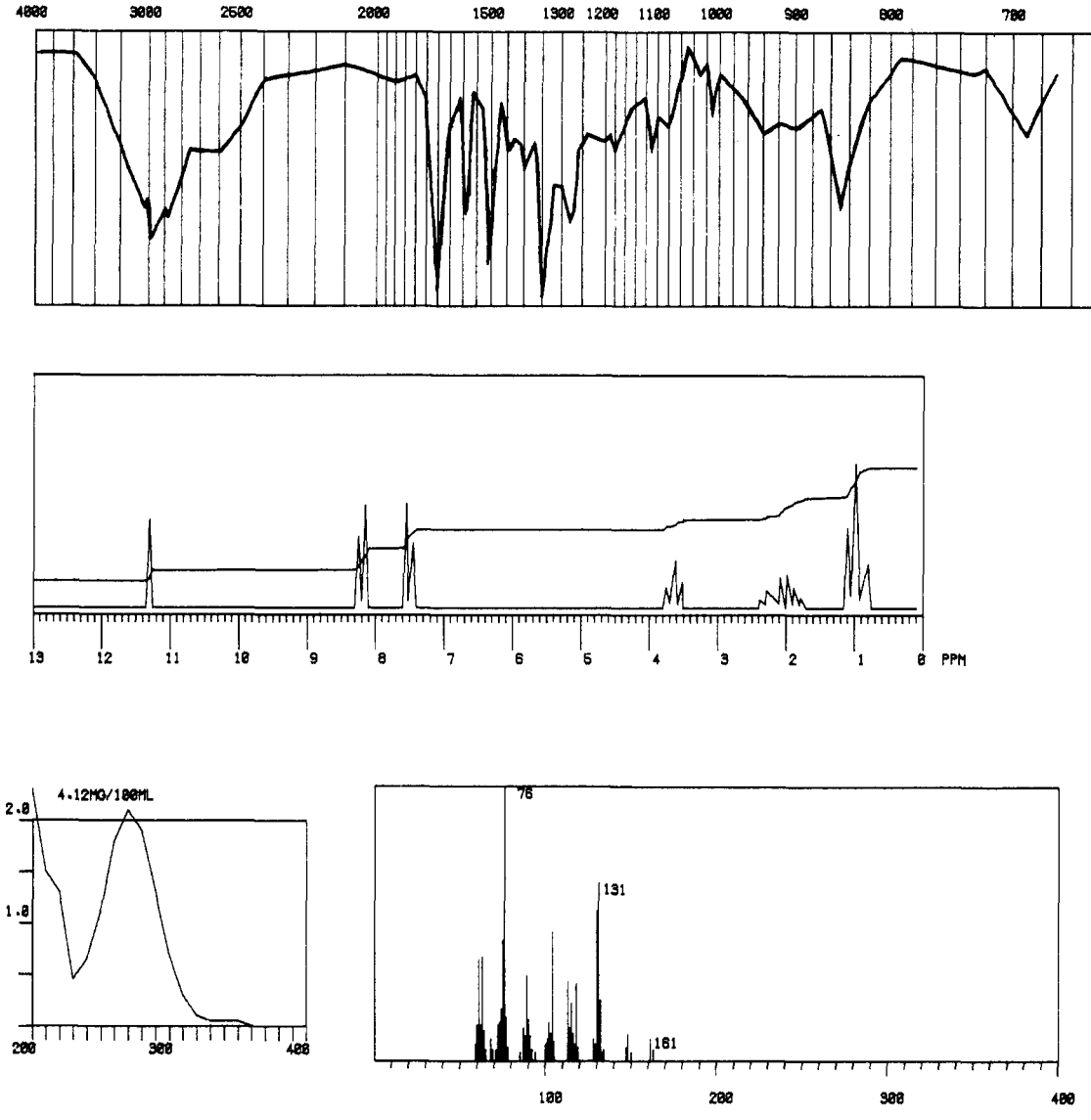


Figure 2. Spectra of the example compound used to illustrate the operation of the program system

tinguish the isomers for an asymmetric tetravalent carbon atom or even for >C=C<.)

The data used with this program have been taken from two published sets of problems for undergraduates (*13, 15*). The program requires the formula, IR, UV spectra, and the proton resonance spectrum including an integrator trace and may use the mass spectrum if this is available. The data in the NMR, UV, and IR traces is crudely digitized—the program allows for values of UV at intervals of 10 nm, IR at intervals of 10 cm^{-1}, and NMR in steps of of 0.01 ppm; in practice, the errors in reading IR and NMR data from the small reproduction spectra must be about ±10 cm^{-1} and ±0.03 ppm. (The limited accuracy of the NMR data prevents the program from identifying bonds by use of matching splitting constants.)

ASSIGNMENT OF FRAGMENTS

There are about 20 functional group procedures available to the "Assign-fragments" routine. The carboxylic acid procedure is typical of many of those for simple functional groups. This procedure has the form:

If: infrared (ragged, medium, 3200, 2500) and infrared (broad, strong, 1740, 1670) and proton 1 (13, 9)

Then: mark spectral features as accountable, create valence description for $-CO_2H$, assign atoms and double bonds.

The premise part of this procedure checks for the -OH absorption and carbonyl stretch absorption of the carboxylic acid and tries to find a singlet resonance with shift between 13 and 9 ppm in the proton magnetic resonance spectrum. (The program does not include any provision for solvent effects that eliminate an acid proton resonance and so cannot process such data.) If the spectra satisfy the tests characterizing the $-CO_2H$ group, the action part of the rule is executed—a "valence description" (see below) is defined and the group is entered on the list of assigned fragments.

The procedures for aryl groups and alkanes/alkenes are more elaborate. Fused ring and heteroaromatics are not currently processed. The approach to handling other polyphenyl molecules is purely empirical and while it is usually correct for student problems, it is not generally valid. As yet the program has no model of the patterns of peaks in aryl group resonances and simply uses the number of protons assigned to aryl groups. If appropriate, the program attempts to identify the substitution pattern of an aromatic ring from the 900 cm^{-1} to 600 cm^{-1} region of the IR spectrum. The processing of alkane/alkene groups normally requires the program to make use of the evidence concerning other functional groups and their possible bonding as well as the immediate evidence of the proton resonance spectrum. For example, a singlet two proton resonance at 5.9 ppm might be due to a $-CH_2-$ or to two equivalent

>CH– groups; the program has to be able to determine that a $-CH_2-$ at 5.9 ppm requires two oxy links and then to be able to check if this requirement can be satisfied by other fragments in the molecule.

A fragment has two "valence descriptions" for each of its valences; the first of these descriptions defines the type of atom offering the valence and the second specifies the allowed matching valences on other fragments. These valence description pairs can be used to impose basic chemical constraints; thus, peroxide links can never be generated if oxy-link atoms do not possess the patterns that allow bonds to other oxy-links. In this role, the valence descriptions achieve a result similar to the "BADLIST" of the DENDRAL algorithm (*8*). In addition to these basic chemical constraints, the valence descriptions define compound specific constraints that can be identified from the spectral data. For example, a carbonyl group with IR absorption at about 1720 cm^{-1} might correspond to a dialkyl ketone or an aromatic ester, etc., but could not be a diaryl ketone; the bit patterns that are assigned to the two wanted valences of this carbonyl are arranged so that alkyl-ketone/ester/etc. structures are permitted but a diaryl ketone structure cannot be generated.

The steps executed by the program can be illustrated by tracing its analysis of the spectral data shown in Figure 2. The program first established the presence of a $-CO_2H$ group by use of the –OH and >C=O stretching absorptions and the acidic proton at 11.3 ppm. The methyl group (with resonance at 0.99 ppm) was identified and defined so that it had to be bonded to an alkyl group; the evidence for a bond to a $-CH_2-$ (from the splitting pattern) may be used to resolve any ambiguities. An aryl group was assigned to account for the four proton aromatic group at 8.2 and 7.5 ppm; examination of the infrared suggested p-C_6H_4. Hypotheses of amine, amide, and cyano groups were all eliminated before the nitro group was assigned on the basis of the rather weak evidence of an IR band in the fingerprint region. The remaining >CH– and $-CH_2-$ groups were readily identified, with bonding constraints derived from the resonance shift values.

SOLVE STRUCTURE PROCEDURE

The solve structure procedure generates an initial connectivity matrix containing entries for all possible bonds. The possible bonds are identified by working through pairs of fragments, checking for matching valence patterns. This initial matrix of possible bonds implicitly represents all structures that can be generated from the set of assigned fragments. Usually, for at least some of the fragments, there will be more possible bonds than the true valence requirements of these fragments. To identify a compound, the solve structure procedure must simplify the initial connectivity matrix by eliminating those possible bonds which do not correspond to real bonds in the molecule.

Some of the simplifications can be made for topological reasons; thus, a possible bond can be rejected if its formation would cause the molecular graph to be split into disconnected subgraphs. Once some of the bonds have been established, it becomes possible to re-use the spectra to get more precise constraints. The most important processing of this type is the use of Shoolery additive rules for determining the proton NMR shifts for $-CH_2-$ and >CH– groups. The program uses the conventional additive formulas for computing shifts due to different alpha substituents (some allowance being made for beta substituents). In many cases, comparison of an alykyl group's true shift with the shifts computed for the different substitutions enables some of the possible bonds to be eliminated. Except for methyl group splittings, the splitting patterns of alkyl protons are not used. Even for the simplest molecules processed, the spectra in this problem set rarely show straightforward first-order splitting patterns. In many cases, steric effects result in nonequivalent methylene protons and complex patterns and, commonly, the resonances of different >CH– and $-CH_2-$ groups just overlap and the patterns are smeared into a large broadened absorption.

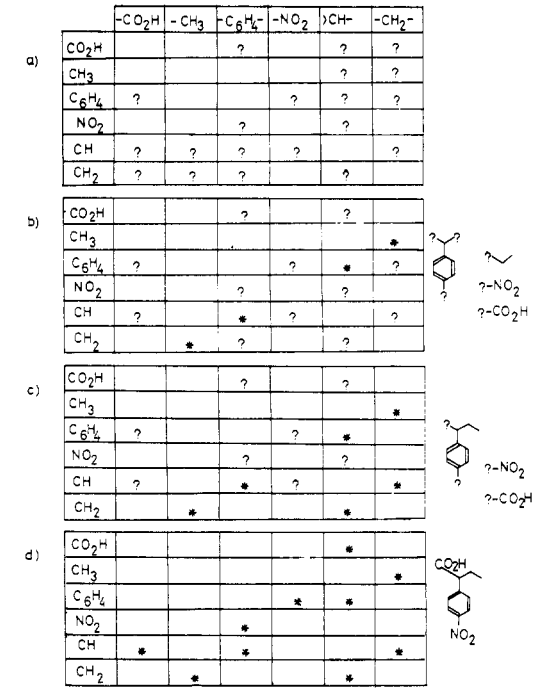
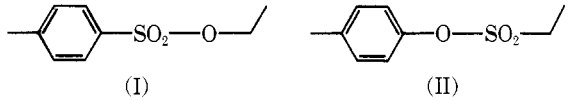


Figure 3. These different connectivity matrices illustrate the steps executed by the "solve structure" routine when combining the fragments assigned for the spectra in Figure 2

Question marks indicate possible bonds between pairs of fragments. Asterisks are used to denote bonds whose presence has been proven. Matrix *b* is derived by using the fact that the methyl group must have two α hydrogens. The next simplification to get *c* follows because the methine group must be bonded to some alkyl group. The final connectivity matrix representing the molecular structure is obtained by consideration of possible ^1H NMR shifts for the methine proton

The mass spectrum may be used at this stage in the processing of a molecule. The current program has inadequate rules for predicting mass spectra. A particular bond may be favored if intense ions can be found in the spectrum that can be associated with cleavages of that bond. This is sufficient to enable the program to select structure I as correct when given the choice of I or II.



(I) (II)

Relatively few cases as simple as this arise and so the mass spectrum is not really utilized. The adoption of some more elaborate mass spectrum processing procedure is one possible area for program extension.

Other procedures for simplifying the connectivity matrix are more empirical. Each fragment has associated with it some category data which can show if a set of fragments are thought to be equivalent—e.g., a four proton group resonating at 2.2 ppm in the NMR is assigned to two $-CH_2-$ groups in the same category. Fragments that are in the same category probably have similar bond patterns. Such a requirement can sometimes be used to restrict the possible choice of bonds for "equivalent" fragments. Another of the weak heuristics incorporated in the program is a procedure which causes favoring of five- or six-membered rings as opposed to larger ring systems.

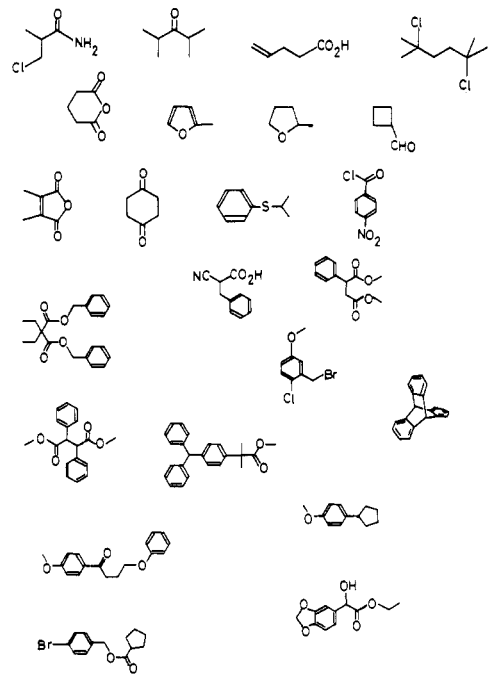


Figure 4. Examples of compounds correctly processed by the program

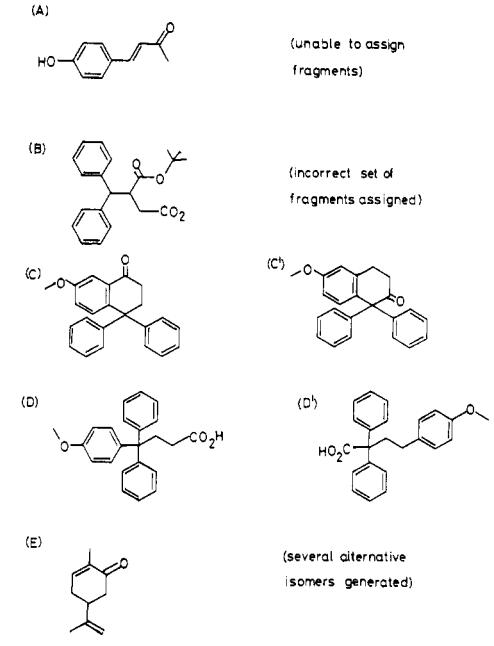


Figure 5. Examples of structures which are not processed correctly by the current program

The operation of these simplifying procedures can be illustrated by the processing of the set of fragments assigned to the spectra of Figure 2. The initial connectivity matrix of these fragments is shown in Figure 3*a*. All the groups have more possible bonds than actual valences. The first simplification can be made using the fact that the methyl group requires two α hydrogens so establishing the $-CH_2-CH_3$ link and incidentally eliminating the possible $-CH_2-CO_2H$ bond and forcing a bond between the methine and aryl groups. The simplified matrix is shown in Figure 3*b*. The >CH– requires one alkyl bond so the structure >CH–CH_2-CH_3 is forced. Finally, use of NMR shifts identifies the bond between the methine group and the carboxylic acid group. The program prints the final connectivity matrix of Figure 3*d* as the structure identification. Several other data processing systems can transform such data into more conventional structural diagrams (*16*); such output transforms were not considered sufficiently important to be incorporated into the program at this stage.

RESULTS

The program has been implemented on an IBM 370/165 in the ALGOLW programming language and runs as a standard off-line job requiring 240kbyte. A file of eighty-five example compounds can be processed in about 20 seconds; two thirds of the processing time being simply for input and output. Figure 4 shows the structures of some of the compounds for which the program operates correctly. While some of these structures might represent quite difficult interpretation problems for students, none would tax an experienced spectroscopist.

The various examples for which the program has failed are generally more interesting. Structure A from Figure 5 illustrates a deficiency in the basic processing of the proton resonance data. This processing does not properly consider multiplet peak patterns, instead relying largely on the integrator trace and limits on the maximum values allowed for coupling constants. Structure A has a complex spectrum with a multiplet due to the –CH=CH– link overlapped with the resonances associated with the aromatic protons and the resonance of the hydroxy proton. For a spectroscopist, the different groups are easy to distinguish because of the regularity of peak patterns. While it is relatively simple to abstract the conscious logical steps of spectrum interpretation from student protocols, it is not easy to identify the operations used in this type of pattern processing. A realistic data interpretation system would require either a routine which groups NMR resonance peaks by some detailed pattern processing or would have to be provided with the groupings by the spectroscopist.

In a number of cases (such as example B), the program assigns an incorrect set of fragments. Such errors are due in part to the basic protocols being inadequate for complex molecules and also result from oversimplifications made when identifying and encoding the operations implicit in the protocols. The use of ^{13}C NMR would permit the processing of structures such as B where the program failed to detect the carbonyl group of the ester because its presence is masked by the acid etc. Simple processing of the ^{13}C NMR spectrum would allow the program to derive constraints on the number of carbonyls, alkene/aromatic carbons, etc., by counting the number of resonances in different regions of ^{13}C shift values.

Even if all the fragments have been correctly assigned, the "solve structure" procedure may not be able to identify the true compound. Examples C and D of Figure 5 illustrate cases where incorrect structures were derived. When attempting to select a structure, the program puts most emphasis on the evidence from the proton NMR. For C, this led the program to derive the structure C′ which is inconsistent with the infrared spectrum. C′ was selected because the recorded shifts for the methylene protons were larger than the additive rules predicted for structure C. Similarly, structure D′ was generated for D, but, in this case, if a mass spectrum had been available, the true structure would have been selected. The "solve structure" routine also fails to determine correct structures if its spectrum processing routines cannot resolve all the ambiguities between possible bonds. In this case, the program can either print the partially processed connectivity matrix (with question marks indicating unresolved alternative bonds) or can systematically generate the possible isomers. Thus, for example E of Figure 5, the program generated three isomeric types including the correct structure (this process does not allow for isomerism such as that due to different

substitution arrangements on the >C=CH- fragment). Currently, the isomer generation process does not attempt to rank alternatives on some calculated merit.

CONCLUDING COMMENTS

The main limitations of the program are in the modeled protocols, especially in those for the analysis of proton NMR data. The spectrum interpretation protocols used to develop the program were obtained from relatively inexperienced students and a limited number of worked examples in texts. The use of protocols recording the interpretation steps of an experienced spectroscopist might enhance the program's performance a little for aromatic and other unsaturated compounds. It is probable that an experienced spectroscopist combines the fragment assignment and structure building steps to a much greater extent than is allowed for in the current program. Consequently, the use of more sophisticated interpretation protocols might necessitate some reorganization of the program. An alternative to the protocol approach for eliciting spectral knowledge from experts has been discussed by Buchanan et al. (*17*). Because programs based on human problem solving follow a fairly conventional interpretation process, a trace of the processing of a particular compound can be understood by the spectroscopist. When interpretation errors occur, it is possible to identify the point at which the model for data interpretation has proved inadequate. The fact that the causes of interpretation errors can be localized, and then presumably corrected, is an advantage that is not shared by statistical pattern classification systems.

The manipulations on the connectivity matrix by the "solve structure" routine correspond to a search strategy in which partial hypotheses are tested as they are generated. Such a strategy was one of three considered for the DENDRAL system (*17*). It is not an effective strategy if only mass spectral data are available (hence DENDRAL's approach of generating and then testing all hypotheses in some subspace) but it has been used in the AI program for processing ^{13}C NMR spectra (*18*). The normal "tree" search operations of AI programs are implicitly mapped through the manipulations of the connectivity matrix which serves primarily as a device for causing the effects of any local bond selections/deletions to be rapidly propagated throughout the graph of possible molecular connections.

In one respect, the program does not model the interpretation processes of practical chemists. Unlike students who must interpret the spectra of arbitrary compounds, a working chemist will be able to use his knowledge of the origin of a compound as an aid to the interpretation of its spectra. The current offline program could be extended by adding an option for reading in a list of fragments that the chemist knows to be present [cf., GOODLIST in the DENDRAL system (*8*)]. The additional input routines would have to be capable of identifying those spectral features due to the predefined fragments and of deriving their bonding constraints. A better solution might be the incorporation of the data processing algorithms in an interactive program for a laboratory data system. This approach would allow the chemist to provide a partially interpreted proton resonance spectrum and could permit the definition of partial structure either before "Assign-fragments" or at the point where the initial connectivity matrix is generated. As the performance of the current program is, at best, equal to that of an average chemist, there is little scope for immediate exploitation in practical analytical chemistry but an alternative possible area of application is in Computer Assisted Instruction.

LITERATURE CITED

(1) P. R. Naegli and J. T. Clerc, *Anal. Chem.*, **46**, 739A (1974).
(2) R. J. Feldmann and S. R. Heller, *J. Chem. Educ.*, **49**, 291 (1972).
(3) E. C. Penski, D. A. Padowski, and J. B. Bouck, *Anal. Chem.*, **46**, 955 (1974).
(4) S. R. Heller, in "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann, and E. Hyde, Ed., Wiley-Interscience, New York, 1974, p 175.
(5) Kain-Sze Kwok, R. Venkataraghavan, and F. W. McLafferty, *J. Am. Chem. Soc.*, **95**, 4185 (1973).
(6) B. R. Kowalski, in "Computers in Chemical and Biochemical Research", **2**, C. E. Klopfenstein and C. L. Wilkins, Ed., Academic Press, New York, N.Y., 1974.
(7) P. C. Jurs and T. L. Isenhour, "Chemical Applications of Pattern Recognition", Wiley-Interscience, New York, N.Y., 1975.
(8) B. G. Buchanan, A. M. Duffield and A. V. Robertson, in "Mass Spectrometry—Techniques and Applications", G. W. A. Milne, Ed., Wiley, New York, N.Y., 1971, p 121.
(9) D. H. Smith, B. G. Buchanan, R. S. Engelmore, A. M. Duffield, A. Yeo, E. A. Feigenbaum, J. Lederberg, and C. Djerassi, *J. Am. Chem. Soc.*, **94**, 5962 (1972).
(10) S. Sasaki, Y. Kudo, S. Ochiai, and H. Abe, *Mikrochim. Acta*, 726 (1971).
(11) G. Beech, R. T. Jones, and K. Miller, *Anal. Chem.*, **46**, 714 (1974).
(12) A. Newell and H. Simon, "Human Problem Solving", Prentice Hall, Englewood Cliffs, N.J., 1972.
(13) B. Trost, "Problems in Spectroscopy", W. A. Benjamin, New York, N.Y., 1967.
(14) E. H. Shortcliffe, Stanford Artificial Intelligence Laboratory, Memo AIM-251,STAN-CC-74-465, Stanford University, Stanford, Calif., 1974.
(15) D. H. Williams and I. Fleming, "Spectroscopic Problems in Organic Chemistry", McGraw-Hill, London, 1967.
(16) R. J. Feldmann, in "Computer Representation and Manipulation of Chemical Information", W. T. Wipke, S. R. Heller, R. J. Feldmann, and E. Hyde, Ed., Wiley-Interscience, New York, 1974, p 55.
(17) B. G. Buchanan, G. L. Sutherland, and E. A. Feigenbaum, in "Machine Intelligence 5", B. Meltzer and D. Michie, Ed., Edinburgh University Press, 1969, p 253.
(18) R. E. Carhart and C. Djerassi, *J. Chem. Soc., Perkin Trans. 2*, 1753 (1973).

RECEIVED for review May 19, 1975. Accepted August 28, 1975.