Dendral and Meta-Dendral - The Myth and the Reality

N.A.B. Gray,
Department of Computing Science,
University of Aberdeen†.

†Permanent address:
Department of Computing Science,
University of Wollongong,
PO Box 1144,
Wollongong NSW2500,
Australia.

Contents

1. The Dendral Project

The objective of the Dendral Project was the development of computer programs that would aid the structural organic chemist in the process of structure elucidation[1,2]. The chemical structure elucidation process does not involve numerical computations; instead, the process is concerned with symbol-manipulation tasks such as the modelling of symmetries in graphs (i.e. in mathematical representations of chemical structures), the analysis of the implications of known structural constraints on the possible combinations of particular substructures, and the processing needed to infer structural constraints from patterns observed in spectral data. The tie between the chemical researchers on the Dendral project and workers in Artificial Intelligence (AI) originated from the need to work on symbolic computations, and from a requirement for a method for encoding the chemists' rules and procedures for spectral interpretation and prediction.

The structure elucidation process, as practiced by natural products chemists, involves several steps. Chemical and spectroscopic data are interpreted to yield structural constraints, that is specifications of substructures that must either be present in, or absent from the chemical structure under investigation. All possible candidate structures consistent with these constraints must be identified. These candidate structures are then evaluated and additional discriminating experiments planned so that the correct structure may be determined. Finally, the proposed "correct" structure must be confirmed either by synthesis or, possibly, by X-ray crystallography. Various programs within the Dendral project were developed to assist in the first three stages – the interpretation of spectral and chemical data, the generation of candidate structures, and the evaluation of candidates and experiment planning. Other AI programs have been devised to assist in the processes of planning a synthesis as needed to confirm a proposed structure [3-5].

Many parts of the Dendral system are inherently purely algorithmic. For instance, an analysis of the stereochemistry of a chemical structure involves an iterative process for computing and investigating various symmetry groups of bonds and atoms. Other parts of the Dendral system are concerned with what is now regarded as the representation and use of knowledge.

The Dendral programs needed to represent structures and substructures – that is essentially factual knowledge concerning patterns of bonding relations that exist among atoms. The interpretation of spectral and chemical data requires either procedural knowledge

or rules that can suggest possible substructures that are consistent with observed spectral patterns.

Constraints on the presence and absence of substructures (as derived in a spectral interpretation "Planning" step) must be used to guide the processes that generate candidate structures in the second step of the structure elucidation process. Thus, it is necessary to have procedures that can perform a form of constrained search through a space of possible structures.

Finally, Dendral's evaluation of candidate structures is based in part on the prediction of detailed spectral properties and such prediction processes again require knowledge of how, for example, particular combinations of structural features cause specific fragmentation processes to occur in a mass spectrometer.

Although the Dendral project was always focussed on specialized domain-specific problem solving, the project soon established and subsequently maintained a more general AI research component. The earliest ties between Dendral and AI arose from a need for recursive list-processing languages. A recognition of the structure elucidation problem as an instance of AI's Plan-Generate-Test paradigm lead to the more substantial ties. The AI work focussed first on issues relating to the representation and use of domain knowledge in systems that could guide searches for appropriate candidate structures or could help in testing those candidates that were generated. Subsequent AI work was concerned with the capture and encoding of knowledge of the characteristic spectral properties of particular classes of compounds.

The AI workers associated with the Dendral project became concerned with issues of acquiring, representing, and using detailed domain-specific knowledge at a time when most AI work focussed on very general techniques of problem solving. The standard AI model of the time was resolution theorem proving – high performance AI programs were to be achieved using sophisticated inference systems based on theorem-provers, and little attention was given to the need to encode the domain knowledge that was to be used by the theorem prover.

The Heuristic Dendral program constituted an empirical demonstration of an alternative approach to problem solving – one that used detailed domain-specific knowledge to guide a relatively weak inference system. Heuristic Dendral illustrated how complex

real-world knowledge could be represented in terms of discrete situation→action rules processed by a rule-interpreter. Subsequently, the MetaDendral work showed how search methods could be used as a basis for programs that attempt theory formation and learning tasks. This focus on knowledge and simple search techniques became a model for much subsequent work in applied AI, and lead indirectly to the entire Knowledge-Based Systems industry of the 1980s.

The success of the Dendral project is now legendary, as can be illustrated by the quotations in Table 1. Dendral is presented as an AI success – a program that has "redefined the role of humans and computers in chemical research". Dendral's supposed success is cited as an example of AI applied to the real-world and used as an argument justifying wider attempts to apply AI techniques.

However, this popular perception of Dendral is misleading. While the Dendral program had some success as a model for applied AI, most of the results of the project are in the form of algorithms and programs that have essentially no AI content. In fact, Dendral provides a further illustration of the way in which projects change as the problems become better understood. Just as the task of integration changed from an AI task based on heuristics (in programs like SAINT and SIN) to an algorithmic process (as in programs like MACSYMA) so Dendral changed from a demonstration system using heuristics, search and simple rules to a practical system using algorithmic methods.

2. Origins

Dendral's origins date back to the early 1960s. From his work on the biochemical origins of life, Professor Lederberg was interested in the scope for chemical isomerism – that is in determining how many different compounds could exist with any given chemical composition. Since the 1870s, organic chemists have been working with a more or less correct model of molecular structures in terms of atoms, bonds and stereochemistry; but in all this time, the problem of chemical isomerism had never been tackled. In the 1930s, a small subset of the isomerism problem was addressed by Henze and Blair who used a "tree" representation of structures and determined the number of possible isomers for acyclic alkanes. However, the general isomerism problem was unsolved. In fact in the early 1960s, there was not even a way representing chemical structures uniquely, and chemical nomenclature was based largely on various essentially ad hoc schemes.

Lederberg introduced a new approach to the definition of (topological) structure based on trees and "vertex graphs" [1,2]. Lederberg's system was not simply a description of structure, it was inherently an algorithm for generating all structures of a given chemical composition. Used descriptively, Lederberg's system defined structures in terms of cyclic components (characterised as substituted vertex graphs) and acyclic parts (characterized as trees). A generator algorithm could create representations of all isomers of a given composition by partitioning a molecule's constituent atoms into cyclic and acyclic parts and subsequently building up, and joing together these parts (using a vertex graph generation algorithm for the cyclic parts, and a tree generation algorithm – the "Dendritic" algorithm – for the acyclic parts). In principle, these algorithms solved the problem of topological isomerism (stereoisomerism could be tackled once representations of topological structures had been created).

Lederberg's generation algorithms are recursive and consequently ill-suited to Fortran – the main scientific programming language of the period. Preliminary implementations of the algorithms were attempted in the naturally recursive Lisp language, and so a connection was established with AI researchers. The Dendral acyclic generator is basically a recursive combinatorial algorithm that finds all distinct partitionings of a given set of atoms (e.g. C and H atoms) into simple radicals such as $-CH_3$, $-CH_2-$, >CH- and >C<, and further finds all distinct ways in which these radicals can be combined into larger units such as ethyl-, propyl-, or butyl- radicals. Early versions of the Dendral (*Den* dritic *Al* gorithm) acyclic generator were running in Lisp in 1965. Results on the numbers of isomers of different empirical compositions were eventually published [6].

The "scope of chemical isomerism" is a rather esoteric topic. There are no practical applications that require information as to the total number of distinct ways of bonding together a particular set of atoms. However, Professor Lederberg had a second, more immediate application that also required computer assisted handling of chemical structural information. At this time NASA was planning the Viking missions that would land research probes on Mars. There was a possibility that Mars either still had, or at one time had had life forms; one of the tasks for the Viking probe was to perform chemical tests for molecules indicative of life. The kinds of molecules that might indicate life forms were expected to include organic acids and esters, ethers, alcohols, ketones and aldehydes, amines, amino acids, and possibly small peptides. Professor Lederberg had responsibility for planning the experimental procedures for data acquisition and the analysis procedures for data interpretation that would resolve questions concerning life on Mars.

Of the analytic methods then available, only mass spectrometry constituted a viable candidate for inclusion in the Viking lander. A mass spectrometer (MS) could be built that would operate under the conditions that were expected to be experienced by the Viking lander. This MS could be coupled to a simplified gas chromatography (GC) separation system and the combined GC/MS system used to obtain mass spectra from infinitesimally small amounts of individual compounds extracted from samples of Martian soil. The spectra acquired from the compounds were to be transmitted back to Earth for analysis.

Some of the compounds could be expected to be identifiable through the use file-search techniques in which their mass spectra would be matched with reference data characterizing known compounds. However, although more than five million compounds have been reported, the largest files available (even in the 1980s) contain only around 80000 mass spectra. It would obviously be unreasonable to presume that any Martian life forms would necessarily use the same compounds as identified in terrestial life forms and characterized by data in available reference files. Consequently, many of the spectra might require other methods of interpretation and analysis.

It is unusual to have to attempt structure elucidation solely from mass spectral data. Normally, structure elucidation is only attempted when there is sufficient sample for many complementary spectral techniques to be applied. Further, a chemist generally has considerable information concerning the origin of a sample and data on the structure of other compounds as previously found in similar samples – such information can considerably help in the processes of spectral interpretation. Lederberg's problem with his potential Martian samples was thus really novel, and he had to plan a new approach to structure elucidation.

By the 1960s, there had been a reasonable amount of theoretical work on organic mass spectrometry and the first texts were being published[7]. It was known that observed ions can be correlated with fragments of the original molecular structure that result from cleavage of individual acyclic bonds, or pairs of bonds in rings, or sets of three bonds in fused and bridged ring systems.The propensity for cleavage of different bonds was also partially understood, and it was often feasible to rationalize the observed mass spectral data of a known compound in terms of fragmentations of the bonds in its structure.

Other empirical analyses had established correlations between particular patterns of ions and intensities in a spectrum and the presence of paticular functionality in a structure. It was possible for an experienced spectroscopist to look at the spectrum of a simple

monofunctional molecule, guess what functional group was present, and use this guess to help identify the molecular ion and, consequently, establish the molecular composition.

Lederberg's method for analyzing mass spectral data had to be based on such empirical methods for the identification of composition and functionality, and on procedures for predicting the spectral patterns that would be expected for given structures. If the composition and functionality of an unknown could be established, generation algorithms could be used to create all of its possible isomeric forms.

Even if molecular composition and functionality are known, the numerous possibilities for different branching patterns in alkyl chains and for variations in the position of functional groups means that there can be hundreds, or even thousands, of possible candidate structures generated for quite simple empirical formulae. These various candidates can in principle be evaluated using procedures for spectral prediction. The bonds in a structure that are mostly likely to be cleaved can be identified, and consequently the masses of the most likely fragment ions can be calculated. Comparison of these predictions for ion patterns and the data actually recorded for an unknown should eliminate most candidate structures leaving just a few plausible structures that are consistent with all available data characterizing an unknown.

To make this approach possible, Lederberg required assistance from AI researchers – assistance in the development of ways of representing and using knowledge concerning both the semi-empirical correlations between spectral patterns and substructures, and the more exacting prescriptions of how a given structure should cleave.

3. Heuristic Dendral

The mid-1960s are associated with one of those waves of enthusiasm for theorem proving approaches that have enveloped AI researchers. The favoured model for problem solving was to express any problem in terms of a formal, mathematical model. Knowledge of a domain was first to be axiomatized; then each problem was to be expressed as a theorem to be proven by inference from the given axioms (more accurately, a negated statement of the problem had to be expressed as a theorem that then had to be disproved). The main focus of research interest was on the inference methods that could be employed by such theorem provers. Stanford's AI laboratory was pursuing long-term research into general theorem

proving methods, as well as carrying out work on robotics and vision, on problems associated with "understanding" natural language, on semantics of Lisp and so forth.

In addition to Stanford's main AI laboratory, there was smaller research group, the Heuristic Programming Project (HPP) group, interested in the application of weak AI techniques, such as heuristic state-space search, to real-world problems. It was researchers from this HPP group, lead by Feigenbaum and Buchanan, who cooperated with Lederberg on the problems of encoding and using chemical knowledge.

The Heuristic Dendral system was built up from the basic Dendral acyclic isomer generator. Dendral was extended by including a planning phase (the "Preliminary Inference Maker") that could perform a crude compound classification based on mass spectral data. Heuristic controls were incorporated in an attempt to filter the number of combinations of branched alkyl chains considered by the generator. An elaborate spectrum predictor was created to enable the final testing of candidates. Figure 1 provides an illustrative summary of the kinds of mass spectral processing performed in the various parts of early versions of the Heuristic Dendral program.

In the first stages of its development, the mass-spectral knowledge given to Dendral was bound closely into the code. The knowledge used in the Preliminary Inference Maker (PIM) was coded explicitly as a form of decision tree; an example tree is shown in Table 2 [8,9]. The PIM program worked through a hierarchy of compound classes (ketones, ethers, etc) verifying that the molecular composition was compatible with a possible classification and testing for specified ion patterns indicative of that class. A pattern of ions at specific masses (m/z values) was given for each class; if the pattern matched the observed spectral data, the Dendral generator would subsequently create structures incorporating the corresponding substructure.

Modifications of the Dendral generator were made in an attempt to exploit typical heuristic "state-space" search techniques from AI. The Dendral generator was now viewed as a "search" program. Dendral's "search" was really an exhaustive enumeration of all possible radicals that could be built with a set of available atoms and placed about a central atom or bond. If Dendral is treated as a state-space search, then states correspond to partially developed structures and operators correspond to the various (combinatorially generated) ways of partitioning the remaining atoms among those radicals that have still to be added to the partial structure. Goal-states correspond to complete structure representations that are

consistent with structural and spectral constraints.

If at a given stage, one could perform a test that established that a particular operator (i.e. a particular partitioning of atoms among radicals) could not yield a satisfactory solution, then it would be possible to "prune" away an entire branch of the generation tree and so reduce the size of the search space. The heuristic test added to the Dendral generator checked for consistency between radical composition and mass spectral data. Each radical considered by Dendral was attached to the partial structure by a single bond; cleavage of that bond would lead to two fragments whose masses could both be computed by the Dendral generator. A test required that at least one of these fragments was evidenced by an ion observed in the recorded mass spectrum. If the appropriate ions were not apparent, the "Heuristic" Dendral generator would discard that possible radical and all the structures that might have incorporated it. (These heuristic filters in the generator were abandoned in later versions of Heuristic Dendral – they involve assumptions about mass spectral processes that are not generally valid and so their use could lead to incorrect results.)

The Predictor program incorporated the most detailed mass spectral knowledge. The Predictor attempted to account for the relative intensities of the different fragment ions that could result from a given structure. The main part of the Predictor program was an iterative loop in which each bond of a structure was analyzed in turn to determine the expected intensities for the two fragment ions that would result from its cleavage.

This intensity analysis involved calculation of measures of the plausibility that a bond would break and of the relative likelihood of the charge being carried by each of the two fragments. The calculations involved numerous functions for assessing influence of neighbouring bonds, of the presence of hetero-atoms in fragments, of the relative sizes of fragments and so forth. Special case analyses were accorded to bonds in environments for which the general analysis proved inadequate. Thus, the knowledge of mass spectrometry in the Predictor program was initially coded in a conventional manner with tests, branches to special case code, and numerous functions. One the first reports on the problems of knowledge acquisition, and on possible methods of eliciting knowledge from an expert, concerns the difficulties of encoding a Lisp function that would constitute a correct description of ß-cleavage fragmentation processes of ketones [10].

When Heuristic Dendral was extended to handle additional classes of molecules such as ethers[11] and amines[12], the interpretation and spectrum prediction functions were

modified so that the system no longer relied solely on mass spectral data. With ethers it was found that mass spectral data did not provide sufficient discriminatory information to rank candidates reliably, so a proton nuclear magnetic resonance (nmr) spectrum predictor was added to the structure evaluation procedures. Proton nmr was subsequently used in the planning phase when Heuristic Dendral was extended to deal with amines. For acyclic monofunctional compounds, the possibilities for different types of branching in alkyl side chains constitute the major risk of a "combinatorial explosion" in the generator. Even a simple proton nmr interpreter could provide a count of the number of methyl groups in the molecule, and so could yield an extremely powerful constraint on the scope for branching in the alkyl chains.

These developments of Heuristic Dendral were accompanied by another more substantial change. Sections of the program, such as the Predictor, were rewritten into a form then described as "table driven programs"[10].

In early versions of the Predictor, mass spectral fragmentation behaviours described by chemists were encoded as individual Lisp functions. Thus, the programmer would invent a Lisp function, "VINYLIC", that contained special case coding to express the chemist's idea that "vinylic bonds, that is C-C single bonds adjacent to C=C double bonds, are generally not cleaved". The programmer's boolean VINYLIC function would involve hand crafted code to determine whether a C-C bond was adjacent to a C=C, and would be used in a special test that modified the plausibility of cleavage of that C-C bond.

The complexities of all such special case coding was severely limiting development of the system when it was perceived that, underlying all the special variations, there was a single form of processing. The chemists' descriptions of fragmentation processes were all simply definitions of substructures that had to be matched onto a given structure; if a matching was achieved, the plausibility of cleavage of a particular bond would be changed.

A "table-driven" version of the Predictor is illustrated in Figure 2. The chemists' definitions of factors affecting bond cleavages are represented using substructural templates; these templates define a bond environment and an associated change in break plausibility. (The term "table-driven" reflects the use of tabular record structures defining substructure connectivity and associated data .) A standard graph matching routine is used to find all possible matchings of each of the templates onto a structure whose spectrum is to be predicted. Each successful matching results in a change to a break plausibility. After the

matchings are complete, the spectrum prediction process converts from break plausibilities to predicted masses and intensities.

The first gain from the "table-driven" revolution was that the knowledge of the experts, the chemists, was now expressed in a language they understood – the language of chemical substructures (instead of Lisp code). Errors in the encoded knowledge could be more readily identified. Frequently, these errors were errors of omission – the chemist's description of a substructure would be overly general resulting in its use in an inappropriate context. The chemist could be shown a trace illustrating how the substructure had been matched with a structure. This trace would immediately reveal any overly general attempt at matching and so allow the chemist to identify additional constraints that were omitted from the original substructure description.The chemist could then use a structure editor to add a constraint, e.g. a requirement that a particular neighbour not be part of a π-system, to the existing substructure definition.

A beneficial side effect of the whole coding process was that the domain experts often found that their own understanding of compounds' fragmentation behaviour was enhanced by the need to formulate a description that could be used by the programs. Further, the descriptions formulated for the program were sometimes useful when explaining the same mass spectral processes to students.

A second gain from a tabular representation of domain knowledge is that this knowledge becomes data capable of being manipulated by other programs. Programs can in some cases check the consistency of different "chunks of knowledge" from these tables. Analysis of particular examples may allow for the automated expansion of the tables with additional data.

Finally, the tabular representations of Dendral's mass spectral knowledge were equivalent to situation→action rules, or production rules. The fact that a production rule representation of knowledge had again proved effective was encouraging. It helped to unify the AI research of the Dendral project with other work. Waterman had recently completed a study using production rules in a system that learnt heuristics[13]. Newell and Simon had conducted lengthy studies in cognitive psychology that showed that a production rule formalism accounted for many human problem solving skills[14].

4. Diversification

Heuristic Dendral worked[15]. It could infer substructural constraints from simple features in mass and proton nmr spectra. It could generate structures compatible with composition and substructural constraints. It could predict the spectra of generated candidates in reasonable detail, and could rank candidates according to some measure of compatibility between predicted and observed spectra.

However, the Heuristic Dendral program was no more than a demonstration, an amusing toy. Its structure generation algorithms were limited to acyclic structures; its spectral interpretation and prediction procedures were limited to monofunctional compounds. Now monofunctional acyclic compounds are important to chemists – after all, these compounds serve as the solvents in which synthetic reactions are performed or with which natural products are extracted – but they do not constitute a class where structure elucidation problems are common. The chemists who had been drawn into the project had been promised computer aids for practical structure elucidation problems, not merely special case toys for simple mass spectra. The practical chemists needed computer systems capable of handling the large, polyfunctional, polycyclic structures that are of biochemical interest.

Thus by about 1970/71, it was time to take stock of the work of the Dendral project and consider various possible development paths.These development paths were as illustrated in Figure 3.

For the programs to be useful in practical applications by chemists, the structure generation algorithms had to be capable of dealing with polycyclic, polyfunctional molecules. This extension to the generators was obviously a complex, largely mathematical task that could not be expected to achieve major results for some years; this work was continued as the "Dendral" project.

Work could continue on the application of rule-based systems to problems from particular well-defined subdomains of organic chemistry; such work could be expected to yield useful results in the short to medium term, and might also help identify automatic methods for acquiring and encoding the necessary domain-specific knowledge. This knowledge encoding/acquisition aspect particularly interested the AI researchers who wished to explore the processes of scientific theory formation[16]. The work on the use and formulation of knowledge, rules of mass spectra etc, formed the "MetaDendral" project.

Finally, Heuristic Dendral had demonstrated the utility of the production-rule formalism for encoding knowledge; there was interest in finding other practical domains where a similar encoding would also prove effective.

Rules for mass spectral prediction are complex. They require the matching of a subgraph representing chemical substructure onto a graph that represents the molecular structure. For each match found, the action component of the rule must be executed. This action component specifies bond cleavages, or cleavage plausibilities, that have to be interpreted in the context of a given match in order to derive the composition of the resulting ion. Thus, one requires a computationally sophisticated graph matcher to verify the premise of a rule and a quite specialized interpreter for the action part of a rule. Although these processes were elaborate, all that was really happening was that various attribute-value pairs of an object were being checked and, if certain conditions were satisfied, some weight was accorded to a particular expected result.

Heuristic Dendral made only rather limited use of situation→action rules. There were no real "chains of inference"; there was no need to apply one group of rules to establish the premise of some other rules. In the Predictor, there was a one step process – a substructure pattern (rule premise) was matched and spectral feature predicted (rule action). There was no a need to perform any search for applicable rules; all rules in the program's tables had to be applied to a structure (many of course might fail to match and so not "fire"). Single-step inference and exhaustive rule-matching were appropriate for Heuristic Dendral's Predictor and for the corresponding interpreter – but obviously more general mechanisms of rule usage can be devised.

Researchers in the HPP group sought applications of production-rule systems in problems of classification or diagnosis (as in PIM's spectrum interpretation) and also in tasks that involved predictions of system behaviour (as in the Predictor). The newer systems focussed on rule-usage. The rules themselves were simplified. A rule's premise became a statement of conjoined tests on the values of various fields (attributes) in a record structure that defined a particular object. The action parts of the rules were also simplified – actions simply updated the values in other fields of the object's record structure. Dendral's simple regimes of one-shot rule application and fixed classification trees were inappropriate to systems intended to demonstrate the power of knowledge-rich problem solving. HPP's more general systems utilized schemes in which chains of inference were built up dynamically through a search for those rules that could be applied in a given context and

which would establish conditions for other rules to fire.

For the Dendral researchers, improvement in the expert's own understanding of a problem domain and the potential for use of the expert's rules in the context of instruction were both incidental. For the HPP group, these aspects of the use of knowledge were of much more central concern. The HPP group envisaged other knowledge based systems providing users with explanations as to how problems were being solved.

The complexity of the structure elucidation problems addressed by Dendral, and the rather specialized domain-knowledge that they involved, reduced the effectiveness of Dendral as a demonstration of the knowledge-based approach to problem solving. (It has been suggested that no other AI worker ever read a complete Dendral paper). Practical demonstration of the knowledge-based approach to problem solving required a diagnostic task, whose complexity could be understood by those who were not specialists in the particular domain, that did involve fair amounts of specific knowledge and which occurred with sufficient frequency that it would be obviously advantageous to provide some automated assistance for human problem solvers.

Of the various projects at HPP in the early 70s, Shortliffe's project on medical diagnosis constituted the best domain for further development of knowledge based systems. Work on Shortliffe's project lead eventually to the Mycin system for medical diagnosis[17]. Mycin solves a relatively simple problem of diagnosis and therapy recommendation. Such a problem could have been "solved" through the use of some fixed classification scheme working in a data-directed manner from a comprehensive data record for a patient that would be filled in prior to the attempted diagnosis. HPP's focus on the representation and use of knowledge resulted in a more flexible approach, and it is with Mycin that we have the first demonstration of goal-directed search for rules, consultative-style problem solving, explanations of rules to assist both expert and normal user, and so forth. However, the Mycin program lies outside the scope of this review.

5. Dendral

The Dendral researchers saw their task as involving (1) the implementation of algorithms for exhaustive isomer (structure) generation, (2) the extension of these topological structure generators to include stereochemistry, (3) the provision of a mechanism for applying

structural constraints to the generator, and (4) the creation of systems that would help a chemist review large sets of candidate structures so as to identify additional discriminating experiments. The various structure generation systems developed by the Dendral group are noted in Table 3.

The Dendral group emphasized the importance of exhaustive, provably correct algorithms for isomer generation. The first proposed applications of the isomer generator were to be to the theoretical problem of the scope of chemical isomerism – the programs were to finally solve questions regarding the number of possible structural forms that could be built from a given set of atoms. However, exhaustiveness is also a prerequisite for a system that would satisfy the longer term objective of providing chemists with an aid for structure elucidation. For a structure to be "identified" on the basis of solely chemical and spectral evidence, it is necessary to show that no other structure is compatible with the evidence (even then, it is usually necessary for the proposed structure to be proven by some unambiguous process, such as synthesis). Therefore, all possible structures must be considered – and an exhaustive generator is required.

The first task of the Dendral researchers was to implement Lederberg's structure generation algorithms. Lederberg's approach involved combinatorial algorithms that found all possible distinct partitionings of atoms among cyclic and acyclic parts, and all possible divisions of the "cyclic" atoms among one or more separate ring systems. The forms of the acyclic parts could be explored using the Dendral generator. The cyclic parts were to be derived through an analysis of allowable vertex graphs.

Simple algorithms exist that can identify those vertex graphs that have to be used when generating cyclic structures for any given set of atoms of known degree. The forms of these reference vertex graphs can be obtained from a library. One of the first tasks of the Dendral researchers was to expand the pre-existing libraries of vertex graphs with those additional examples that would be needed by the structure generators[18]. (The various Dendral papers, like reference 18, that report on the development of, or results obtained from specialized algorithms were all published under the imprimatur *"Applications of Artificial Intelligence for Chemical Inference"* because this provided a useful identifier for related research. This header was used irrespective of the fact that most of these papers did not address any AI related issues.)

The reference vertex graphs obtained from a library have to be built up into

representations of complete cyclic parts of a molecule. This construction process involves an extensive sequence of "labelling" steps in which additional nodes are added to the edges of the vertex graph, and then atom names are attached to the nodes. The main difficulties in the structure generation process relate to problems of dealing with symmetries in the initial vertex graphs and in the partially labelled structures. The development of an exhaustive, irredundant generator required the development of new mathematical techniques for resolving the various labelling problems[19,20]. Developments eventually lead to the StrGen program for structure generation[21,22]. StrGen was a purely combinatorial program; all its atom partitioning and graph labelling steps reduce to the selection of n objects from some set of m objects. StrGen involved no heuristics, no search, no Artificial Intelligence.

The StrGen system was used to solve specialized problems relating to the scope of isomerism[23,24]; however, application to practical structure elucidation problems required further developments. StrGen could treat known constituent parts of a molecule as atomic entities or "superatoms". Thus, if it was known that a compound incorporated a carbonyl group, then the empirical composition given to StrGen could be changed by eliminating one oxygen and one carbon atom and adding a bivalent "carbonyl atom". The "structures" generated would not represent actual chemical structures. A second processing phase was required to derive representations of chemical structures. In this second "Embed" phase, the superatom parts would be "embedded" – that is expanded out to reveal their internal structure. The general embedding process is complex and, once again, involves detailed analyses of symmetries[25]. Depending on the ways in which superatoms and other structural constraints were defined, it was possible for duplicate structures to result from the embedding step. Consequently, a final processing step required conversion of structures to canonical form[26] and filtering of any duplicates that this revealed. Additional substructural constraints could be applied through the use of "graph-matching" techniques; for example, if a particular substructure was known not to be present in a molecule, then a representation of this structure could be matched with all the final generated structures and those possessing the substructure could be discarded.

Dendral's first structure elucidation system was Congen – the CONstrained GENerator[27]. The Congen system, illustrated in Figure 4, comprised an executive module that provided a convenient user-interface and which could invoke a number of other data-analysis and processing programs. The programs included a structure editor that allowed a chemist to define superatom parts, structure drawing routines, the StrGen

program for intermediate structure generation, and the Embed program that derived representations of the final structures using a combination of a version of Strgen's labeller (to embed superatom parts) and a canonicalization routine.

Congen did not attempt any interpretation of spectral data. The structural constraints used by the program had be derived by the chemist user. The chemist had to interpret available spectral and chemical data and define any inferred structural constraints either as superatom parts or as restrictions on the bonding of particular superatoms. Thus, Congen did not have any AI "planning" component. Further, Congen did not attempt to apply any heuristic filters to the output of either the Strgen or embedder programs (for example, it did not discard "strained structures" with fused small rings). There were no heuristics, no domain rules, no knowledge in Congen – the program was based entirely on provably correct, graph-theoretic, combinatorial algorithms.

This lack of heuristic filters was not a consequence of any problems in implementing such procedures – it was an essential feature of a structure elucidation system. The structure elucidation process requires that all candidates be considered and that candidates only be discarded if there exists definitive evidence against them. Many heuristics could have been defined and used to constrain a generator (e.g. "cyclopropene rings are extremely rare, so don't create them") but such usage was inappropriate – the unknown whose structure was being determined might be one of the few naturally occurring cyclopropenes.

Once final candidate structures had been produced by Congen they could be evaluated by means of a number of structure surveying routines. Spectrum prediction routines, analogous to Heuristic Dendral's Predictor function, could be used as part of this surveying process. Mass spectra could be predicted either by means of detailed class specific rules [28] or through a more generally applicable semi-empirical theory[29]. Carbon-13 nmr spectra could be predicted using a semi-empirical system that employed model substructures[30]. Candidate structures could be ranked according to a measure of the compatibility of observed and predicted spectra. However, it was not possible to exclude poorly ranked candidates – the semi-empirical theories used for spectrum prediction were too crude and could not be used in any "proof of structure" process. Structure ranking by spectrum prediction could at best help the chemist focus on the more plausible candidates.

Further development of the Congen program was mainly the work of Carhart. Carhart replaced the complex vertex graph analysis with an algorithm for generating

intermediate structures that was related to the structure generation procedures originally developed by Sasaki[31] and to the canonicalization procedures originally devised by Morgan[26]. The handling of symmetries in Carhart's algorithm was less sophisticated than the vertex graph approach but, overall, this new algorithm was more efficient in its processing of typical structural problems. Carhart's Congen-II became the service version of the Congen system as offered on the Sumex computer facility from about 1978. Nourse's Stereo module completed the structure generation process by generating all possible configurational stereoisomers of each candidate produced by the Congen system[32].

The Congen approach of first generating intermediate structures containing superatoms and then embedding these superatoms was eventually replaced by Carhart's Genoa algorithm[33]. Frequently, it is difficult for a chemist to derive distinct superatom parts from the available information on overlapping substructures inferred from spectral or chemical data. Carhart's Genoa algorithm provided a much more direct way of assembling structures from large, possibly overlapping, substructural components. Once again, the Genoa system contains no heuristics, no state-space search, no knowledge, no AI; it is a purely algorithmic system based on a generalization of standard graph-matching procedures.

6. MetaDendral

In 1970/71, the Dendral structure-generator was known to constitute a long-term development project; in the meantime, practical, usable application programs were required. In order to justify further funding, the Dendral project had to be shown to be capable of delivering tools that could be used by chemists to help solve practical problems from the laboratory. The part of the Dendral project that lead eventually to the MetaDendral work on theory formation started as a search for immediate applications of a rule-based system that could express correlations between substructural features and spectral processes.

As shown in Table 4, the MetaDendral project developed the programs Planner, IntSum, and the RuleGen/RuleMod system. IntSum and Planner were special purpose tools devised to solve immediate problems for working chemists – these programs were basically concerned with the identification of characteristic mass spectral fragmentation behaviours as exhibited in the spectra recorded for many members of some class of molecules, and the use of information on molecular fragmentations in restricted types of structure identification

problem. The RuleGen/RuleMod system sought to take data, such as those produced by IntSum, and derive general rules describing mass spectral fragmentations; these rules would be similar in nature to those given by chemists to the original Heuristic Dendral program.

The Planner program[34] was intended to assist the chemist in the process of establishing the likely structure of a newly isolated compound in a known family through the interpretation of its mass spectral data. The program's operation is suggested by the scheme shown in Figure 5. Planner took as its inputs a definition of the common skeleton of the family of compounds, a set of rules characterizing the main fragmentation processes of that common skeleton, and composition and mass spectral data characterizing a particular unknown; the program's outputs were specifications of the probable nature and position of substituent groups.

Planner worked by calculating the compositions of the ions that would result from the specified skeletal break processes being applied to structures with particular assumed distributions of substituent atoms. If an ion of appropriate composition was found in the recorded spectrum, evidence would be accumulated towards establishing that particular assumed distribution of substituent atoms. Thus, in the example in Figure 5, process B can be applied under the assumption that both oxygens are in the leftmost part of the molecule, or that one oxygen is on either side of cleavage, or that both oxygens are attached to the three-carbon fragment that would be lost – these assumptions lead to predictions of ions at $C_{15}H_{18}O_2$, $C_{15}H_{18}O$, and $C_{15}H_{18}$. If $C_{15}H_{18}O$ was the only one of these ions observed, then this would constitute evidence that one -OH was attached to either C(15), or C(16), or C(17), and that the other -OH was bonded to one of C(1)-C(14) or C(18). Through such analyses of each break process and a subsequent weighting of accumulated evidence, Planner could frequently identify and localize the substituent groups of a compound.

Programs such as Planner are only useful if there are sets of class-specific rules for the particular compounds of interest. Such rules are derived through analysis of similarities in the spectra of several closely related, model structures. This analysis involves finding those break processes that seem generally applicable to molecules in a particular class and which account for most of the important ions in the observed spectra. The analysis process is basically a search – one has to explore some range of possible break processes and find those for which there is reasonable evidence among the available data.

There are some general constraints on the search. Thus, it is well known that bonds

incident on tetravalent carbons are more easily broken than those incident on bivalent carbons, cleavages of bonds ß to π-systems tend to be favoured, and so forth. Using such constraints, the chemist postulates a small set of processes, tries these out on the available example compounds and checks whether they lead to observed fragments, possibly modifies a process by allowing for further loss of hydrogens etc, and finally stops when a particular set of break processes can be used to rationalize the majority of the intense ions in the spectra of each of the available example compounds.

Humans rarely perform effectively at such tasks involving a search through large numbers of data points and combinations of different interpretations. In particular, humans are inclined to stop when one set of break processes is found – even though there may well be alternatives of at least equal plausibility that have still to be considered. Of course such searches are relatively readily programmed, and the computer can be relied on to perform an exhaustive and unbiased search. A natural extension to the Planner system was a program that would perform analyses of sets of mass spectra and would identify all possible break processes that were evidenced in the data.

This program – called IntSum for Interpretation and Summary program[35] – took as inputs a definition of a common skeleton, a set of constraints on the nature and complexity of the fragmentation processes that were to be considered, and a set of structure definitions and details of recorded mass spectra for those structures. IntSum identified all possible skeletal fragmentations that were consistent with the given constraints and for which there was evidence in the observed spectra. The final outputs from the IntSum program took the form of tables identifying the various possible processes and their supporting evidence. From these output data, chemists could infer general rules of mass spectral behaviour for subsequent use either in the Planner program or in more conventional methods of spectral processing.

The next program in the MetaDendral sequence (the RuleGen/RuleMod system [36]) was devised to accomplish this rule generation or "theory formation" task. It is the RuleGen/ RuleMod system that is commonly cited in claims that computers have discovered new theories for the physical sciences. The RuleGen/RuleMod system was devised originally for the analysis of mass spectral data; Mitchell subsequently extended the RuleGen system to the analysis of ^{13}C nmr data[37]. Mitchell's system generated rules that relate precise ^{13}C shift ranges to specific environments for the resonating atom; Mitchell's ^{13}C RuleGen program provides the clearer illustration of the mechanism for rule generation.

Heuristic search techniques can again be applied. Here, the states in the search space correspond to rules of the form Substructural-environment→Shift-range. States differ primarily in their substructural descriptions. Operators, that define state-state transitions, actually change the substructure descriptions – for example by adding detail to the specification of a neighbouring atom's chemical type, or number of substituent hydrogens, and so forth. The shift range that is to be associated with a newly created substructure can be found by searching a reference file of structures and their assigned spectra for all instances of that substructure, and recording the minimum and maximum shifts that have been observed for ^{13}C atoms in that specified substructural environment. Goal states can be characterized by various empirical criteria that define a useful prediction rule; for example, a rule may require a "sufficiently narrow" associated shift range and an "adequate number" of supporting instances in the reference data file.

Figure 6 illustrates RuleGen's search for ^{13}C prediction rules; these rules were developed from data on about 100 reference spectra of alkanes and alkyl amines. RuleGen began with a very general substructure and correspondingly weak prediction; the initial rule specifies that a carbon in an unrestricted substructural environment has to have a shift (or ∂-value) in the range $-\infty$ to $+\infty$. New states were created by the application of operators that added detail to the substructure descriptions of existing states; thus, the direction of search was from general to specific. As illustrated in Figure 6, RuleGen first added detail on the degree of the resonating atom (so distinguishing CH_3–, –CH_2–, –CH<, and >C< resonances) and then explored elaborations such as CH_3–C, CH_3–N, CH_3–X_{sec}– etc (these representing all carbon-bonded and nitrogen-bonded methyls and all methyls attached to any secondary center). The shift ranges assigned to these substructures were derived from the data characterizing the 100 alkanes and amines in the training set. Various heuristics determined whether (1) a state represented a rule of adequate quality for inclusion in the final rule set, or (2) a state represented an overly general rule that should be refined by the application of further operators, or (3) a state appeared to be on an unproductive path that could be abandoned.

MS-RuleGen's search for mass spectral rules was basically similar to the search for ^{13}C rules. However, the rules were inherently more complex and the substructure building operators were more diverse; both these factors act to enlarge the search space of possible rules. The search for rules was again from general to specific, starting with a seed-rule of the form X*X (interpreted as meaning "a bond between any pair of atoms is cleaved").

Operators could add atoms to this substructure, or add attribute values to existing atoms. First level elaborations of the seed rule included rules such as C*X (i.e. "cleave an atom between a carbon and any other atom"), XH_2-X ("cleave a bond between an atom with two hydrogens and another atom"), X_{sec}*X ("cleave a bond between a secondary atom and another atom"), and so forth. Evidence for a rule was found by matching these rule substructures against the substructural environments of those bonds cleaved in the fragmentation process that IntSum had identified.

Heuristics were again used to control the exploration of the space of rules defined by the substructure-elaboration operators. As in the case of ^{13}C data, the system frequently generated several apparently distinct rules that covered the same data points, and there were numerous partial overlaps among rules. A separate program, RuleMod, was needed to eliminate duplicates and merge overlapping rules.

The nature of mass spectral data limits the utility of the entire MS-RuleGen/RuleMod approach. The system can really only search for evidence for rules that apply to most members of a family of compounds. The result is a rather sparse set of rules; the rules rarely account for as much as 50% of the ion current. The rules are usually of low discriminatory power. They characterize fragmentations that are common to the class; but the rules don't always distinguish well among class members (and this discrimination task represents the only possible use of such rules).

In the original paper [36], it is noted that the MetaDendral program was "not discovering a new framework for mass spectrometry". Rather, the rules derived by the RuleGen/RuleMod system were compared with direct spectral⇔structural correlation rules such as were then being investigated by pattern recognition techniques. Contrasting their approach with the pattern recognition work, the authors of the MetaDendral papers emphasized their interest in providing the chemist both with explanations as to the origin of rules and with methods of constraining a search for rules. In contrast to this rather conservative presentation of the MetaDendral results in the original scientific literature, reports in the secondary literature and in AI journals have been enthusiastic. MetaDendral is described as having "created rules for sub-families of molecules for which none existed before" and the general tenor of reports is to the effect that MetaDendral has achieved some significant advance in the theory of mass spectrometry.

Given these enthusiastic reports of the MetaDendral results, it is worth trying to

assess the practical utility of MetaDendral's new rules or theories. One way of assessing the value of a new theory is to examine the frequency with which the work is subsequently cited. The evidence from the Science Citation Index confirms the supposition that the MetaDendral rules for keto-androstanes are of limited utility. The MetaDendral rules were cited in about twenty papers published in the next ten years; six of these citations were "continuity" references in subsequent Dendral project publications, seven citations were from general reviews along the lines of "Computers Techniques for Mass Spectroscopy", four citations were by authors of papers on alternative methods for interpreting mass spectra, and one citation was in a general review on mass spectral techniques for steroids. There was one reported attempt to use the keto-androstane rules – Gurst noted that they failed to predict a fragmentation uniquely characteristic of a particular subclass of diketo-androstanes [38] (this result is exactly as one would expect, the MetaDendral rules characterize those processes that are common to a class of compounds).

Both the ^{13}C and MS versions of RuleGen are actually given a "theory" for the phenomena that they explore. In the case of ^{13}C RuleGen, this theory is a statement to the effect that shifts can be adequately explained in terms of topological models for substructures that can involve specifications of attributes (such as atom-type or number of hydrogens) on atoms up to two or three bonds from the resonating centre. This theory is built in; it is expressed through the forms allowed for substructures used in the rules and through the operators for modifying these substructures. The RuleGen programs can never transcend any limitations of the given theory; they can merely explore how well it can account for a given set of training data. The generated rules represent a particular instantion of the given theory as it applies to the specific training data.

The theoretical models given to MetaDendral RuleGen programs were in fact overly general, with significant constraints omitted. A spectral measurement (such as a ^{13}C shift) is a measure of how a particular subpart (a ^{13}C atom) is embedded in a whole structure. The properties of the subpart are going to be influenced most strongly by the character of its immediate neighbours, a second shell of neighbours will further influence the subpart's behaviour, other more remote neighbours may also influence the observed property to some measurable degree. There is an inherent hierarchy in these interactions; and, if the α-neighbors of that atom have not been fully specified, there is no point in seeking correlations between an atom's nmr-shift and parameters such as the number of γ-hydrogens.

If the given theory embodies all the appropriate physical constraints, ^{13}C rule formation becomes essentially deterministic. There is no need for heuristic search, nor for empirical evaluation functions and so forth. The stereochemical environments of each carbon atom in each reference structure can simply be encoded, algorithmically, out to some maximum bond-radius. Each combination of substructure code and associated chemical shift either becomes a new, specific "rule" (with a point-prediction for shift rather than a range), or the code/shift combination provides new data to update the shift range associated with an existing rule with the same substructure code. Further, the coding procedure can be arranged so that the code (substructure definition) is canonical at each bond-radius (or shell level). If two atoms are in identical environments out as far as three-bond radii, then their 3-shell codes will be identical. Each code/shift combination can be used to update more general rules using 3-shell, 2-shell, and 1-shell descriptions (of course, these general rules are associated with increasingly broad prediction ranges). The final Dendral programs for ^{13}C spectrum prediction and interpretation [30,39] were based "rules" created through such a deterministic approach. Once again, when the initial analysis is more complete, one moves from heuristics and search to algorithmic methods.

7. Lessons from the Dendral Project

The achievements of AI and in particular of Expert Systems are mainly presented through secondary literature – reviews, books, product presentations and so forth. These presentations are normally designed to popularise some technique, or to encourage the wider use of particular approaches to problem solving, or to sell a product. They use data on previous AI systems to provide supporting evidence for their claims. The Dendral project has figured prominently in such reports – and, for the most part, is presented in a totally inappopriate way.

Even in those Dendral/MetaDendral programs that did utilize rules for spectral interpretation and prediction, the AI-related code for creating and using rules represented only part (often only a small part) of the total system code. Most of the code of these systems was involved in the user interface, in structure and substructure editors, in specialized file system, in graph-matcher and canonicalization routines, and so forth. It is likely that the rule creation/interpretation code will only represent a small part of the total

code needed for problem solving in most complex domains. Systems as elaborate as the MetaDendral programs simply cannot be built from standard Expert System shells.

The Dendral programs of greatest practical value to chemists are the basic structure generators Congen and Genoa, and the Stereo module. Even these programs are of relatively limited use – they simply don't address a problem that arises with sufficient frequency as to really justify the use of computer aids. The chemical structural problems that do arise frequently (and that could utilise computer aids) are of the form *"What is the nature of this pollutant?"* – in these problems a rough classification of the compound (or mixture) as polynuclear aromatic, organo-phosphorous, chlorinated biphenyl etc is required, the data available to the classification procedure concern the origin of the sample and the results of a few quick physico-chemical tests. These problems are insufficiently constrained for use of structure generators like Congen, and in any case there is no interest in the specific structural form of the pollutant.

Complete structure elucidation problems, such as can be handled by Congen, are relatively infrequent. Most of the time of the typical individual graduate student in a natural products laboratory is spent on the isolation and re-identification of previously known compounds; the three or four new compounds encountered each year represent the main results of that individual's work, and the elucidation of their structure is the only intellectually stimulating part of the entire exercise. The natural products researcher is unlikely to hand over the only "fun" part of the work to a computer program, particularly as in most cases biochemical constraints on skeletal form are such as to make the structural problem readily soluble with no need to exhaustively consider hundreds of potential candidates. Further, structure elucidation ultimately requires proof – proof either through X-ray crystallography or by means of one of the newer skeleton-tracing nmr techniques. Identifications based on Dendral-like procedures of candidate generation and ranking do not constitute proof of correct structure elucidation and so do not constitute complete solutions to a structure elucidation problem.

Although available for many years through the networked Sumex-Aim facility, Congen was never widely used (when an attempt was made to discover its usage, only about eight applications citing its use could be traced). The two main review papers that present the Congen/Genoa programs as aids for structure elucidation use contrived examples based on previously completed and published structure eludication problems – there weren't any novel structural problems being solved by these programs that could have been used as

the basis for illustration in such review papers [40,41]. The Genoa/ Stereo program was made available through the company Molecular Design[42], but it represented only a minor and not very successful product line added to the repertoire of an already existing company. Claims of widespread use of the Dendral programs are simply without foundation. In any case, the programs Congen, Genoa, and Stereo have essentially no AI content and even if they had been widely used such usage would not constitute evidence for the successful application of AI techniques. (If Dendral is "one of the most successful AI programs ever" then god help the unsuccessful AI programs.)

Other misrepresentations of Dendral indicate how reports of the actual achievements of programs require much more careful phrasing. For example, original reports have compared the performance of graduate students and of the Heuristic Dendral program at the task of identifying a structure solely from composition and mass spectral data[10], and the correct results obtained from the structure generation algorithms of StrGen have been compared with the less successful attempts of post-doctoral students assigned similar structure-generation problems[21]. However these are not realistic tasks. A chemist normally interprets spectra in the context of considerable additional information concerning the origin and isolation of a compound. Structure generation problems, as actually encountered by chemists, involve small numbers of distinct multi-atom fragments and involve quite different conceptual problems from those associated with the generation of structures from large numbers of identical atoms. The lack of realism of these tasks is not stressed in the original reports – because these reports presume substantial shared background knowledge. Those familiar with the domain understand that the tasks reported represent only a small part of the total process of structure elucidation, and that the examples chosen have been selected to give an easily comprehended measure of the scope of problems that can be tackled by the programs. However, those who do not share this presumed background knowledge tend to take these measures of performance quite out of context – and so the Dendral programs come to be reported as having greater expertise than post-doctoral scientists.

The MetaDendral systems had little potential for real application. In principle, the Planner system could be used to help identify structures that had some novel pattern of substituents on a standard skeleton. However, Planner was limited by its requirement that the fragmentation processes of a skeleton would not be substantially altered by the presence of the substituents – while this requirement was met in the case of the simply substituted estrogens, it is not a requirement that would generally be satisfied. In any case, such

structural problems are much more readily solved through the use of other spectral techniques.

The RuleGen/RuleMod analysis of mass spectral data could not generate rules that would be of significant value in structure elucidation. The spectral⇔structural correlations that these programs could identify were for compounds in well defined chemical classes and so were irrelevant for general structure elucidation problems. Further because the selection method picked those processes for which there was general evidence among the available reference data on example compounds, the rules were usually of low discriminatory power. These rules could not really be used as the basis of ranking candidates even when one did get a structure elucidation problem involving compounds in the class for which the rules had been developed.

Claims that "Dendral surpasses all humans at its task and, as a consequence, has caused a redefinition of the roles of humans and machines in chemical research" are made in ignorance. Such claims attempt to establish a myth – a myth that may in the short term help to sell some techniques. But, such myths only obscure the real achievements of the Dendral project.

The Dendral project has real achievements. Many are esoteric. Dendral led to advances in combinatorial mathematics, particularly graph labelling techniques[20,21,25]. Lederberg's Vertex Graph method for describing molecular structures, combined with methods for embedding complex subgraphs within other graphs led first to systems that could help enumerate isomers and then to practical structure generation systems. Although other methods are now used to establish unique identifiers for chemical structures[26], Lederberg's scheme for classifying structures is still of some current research interest. Brown, Carhart, and Nourse invented and refined many algorithms for the representation of canonical forms for structures, for subgraph matching, for analysing symmetry, and for handling stereochemistry. Few of these algorithms have been made as accessible to other researchers as one might have wished; however, these algorithms continue to be used and developed by specialists in companies such as Molecular Design[42].

Possibly, the real achievement of Dendral was to help change the way people thought about computers and problem solving. The success of Heuristic Dendral, and its stepchild Mycin, encouraged people to look beyond numerical computation and data processing tasks. Dendral and Mycin showed how complex real-world classification and diagnostic problems

could be tackled provided that adequate domain-specific knowledge was given to a program that incorporated even a relatively simple inference procedure.

References

1. R.K. Lindsay, B.G. Buchanan, E.A. Feigenbaum, and J. Lederberg, *Applications of artificial intelligence for organic chemistry: The DENDRAL project*, McGraw-Hill, New York, 1980.

2. N.A.B. Gray, *Computer assisted structure elucidation*, Wiley-Interscience, New York, 1986.

3. E.J. Corey, A.K. Long, and S.D. Rubenstein, Computer-Assisted Analysis in Organic Synthesis, *Science*, 228 (1985) 408-418.

4. A.P. Johnson, Computer Aids to Synthesis Planning, *Chemistry in Britain*, 1985 (January) 59-67.

5. W.T. Wipke, G.I. Ouchi, and S. Krishnan, Simulation and Evaluation of Chemical Synthesis – SECS: An Application of Artificial Intelligence Techniques, *J. Artificial Intelligence*, 11 (1978) 173-193.

6. J. Lederberg, G.L. Sutherland, B.G. Buchanan, E.A. Feigenbaum, A.V. Robertson, A.M. Duffield, and C. Djerassi., Applications of artificial intelligence for chemical inference. I. The number of possible organic compounds: acyclic structures containing C, H, O, and N., *J. Am. Chem. Soc.*, 91 (1969) 2973-2976.

7. H. Budzikiewicz, C. Djerassi, and D.H. Williams, *Structure elucidation of natural products by Mass Spectrometry*, Holden-Day, San Francisco, 1964.

8. B.G. Buchanan, G.L. Sutherland, and E.A. Feigenbaum, Heuristic Dendral: A Program for Generating Explanatory Hypotheses in Organic Chemistry, in B. Meltzer and D. Michie (Editors), *Machine Intelligence, 4*, Edinburgh University Press, Edinburgh, 1969, pp. 209-254.

9. A.M. Duffield, A.V. Robertson, C. Djerassi, B.G. Buchanan, G.L. Sutherland, E.A. Feigenbaum, and J. Lederberg, Applications of artificial intelligence for chemical inference. II. Interpretation of low resolution mass spectra of ketones. *J. Am. Chem. Soc.*, 91 (1969) 2977-2981.

10. B.G. Buchanan, G.L. Sutherland, and E.A. Feigenbaum, Rediscovery some Problems of AI in the Context of Organic Chemistry", in B. Meltzer and D. Michie (Editors), *Machine Intelligence, 5*, Edinburgh University Press, Edinburgh, 1969, pp. 253-280.

11. G. Schroll, A.M. Duffield, C. Djerassi, B.G. Buchanan, G.L. Sutherland, E.A. Feigenbaum, and J. Lederberg, Applications of artificial intelligence for chemical inference. III. Aliphatic ethers diagnosed by their low resolution mass spectra and NMR data, *J. Am. Chem. Soc.*, 91 (1969) 7440-7445.

12. A. Buchs, A.M. Duffield, G. Schroll, C. Djerassi, A.B. Delfino, B.G. Buchanan, G.L. Sutherland, E.A. Feigenbaum, and J. Lederberg, Applications of artificial

intelligence for chemical inference. IV. Saturated amines diagnosed by their low resolution mass spectra and NMR data, *J. Am. Chem. Soc.*, 92 (1970) 6831-6838.

13. D.A. Waterman, Machine learning of heuristics, Ph.D. thesis, Department of Computer Science, Stanford University, 1968.

14. A. Newell and H.A. Simon, *Human Problem Solving*, Prentice-Hall, Englewood Cliffs, 1972.

15. A. Buchs, A.B. Delfino, A.M. Duffield, C. Djerassi, B.G. Buchanan, E.A. Feigenbaum, and J. Lederberg, Applications of artificial intelligence for chemical inference. VI. Approach to a general method of interpreting low resolution mass spectra with a computer, *Helvetica Chimica Acta*, 53 (1970) 1394.

16. B.G. Buchanan, E.A. Feigenbaum, and J. Lederberg, A heuristic programming study of theory formation in science, in *Proceedings of the 2nd International Joint Conference on Artificial Intelligence*, September 1971, W. Kaufmann Inc, Los Altos, pp.40-50.

17. B.G. Buchanan and E.H. Shortliffe (Editors), *Rule Based Expert Systems: The Mycin Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, 1984.

18. R.E. Carhart, D.H. Smith, H. Brown, and N.S. Sridharan, Applications of artificial intelligence for chemical inference. XVI. Computer generation of vertex graphs and ring systems, *J. Chem. Inf. Comp. Sci.*, 15 (1975) 124-130.

19. H. Brown, L. Masinter, and L. Hjelmeland, Constructive graph labeling using double cosets, *Discrete Mathematics*, 7 (1974) 1-30.

20. H. Brown and L. Masinter, An algorithm for the construction of the graphs of organic molecules, *Discrete Mathematics*, 8 (1874), 227-244.

21. L. Masinter, N.S. Sridharan, R.E. Carhart, and D.H. Smith, Applications of artificial intelligence for chemical inference. XII. Exhaustive generation of cyclic and acyclic isomers, *J. Am. Chem. Soc.*, 96 (1974) 7702-7714.

22. L. Masinter, N.S. Sridharan, R.E. Carhart, and D.H. Smith, Applications of artificial intelligence for chemical inference. XIII. Labeling of objects having symmetry, *J. Am. Chem. Soc.*, 96 (1974) 7714-7723.

23. D.H. Smith, Applications of artificial intelligence for chemical inference. XV. Constructive graph labeling applied to chemical problems. Chlorinated hydrocarbons. *Analytical Chemistry*, 47 (1975) 1176-1179.

24. D.H. Smith, Applications of artificial intelligence for chemical inference. XVII. The scope of structural isomerism, *J. Chem. Inf. Comp. Sci.*, 15 (1975) 203-207.

25. H. Brown, Molecular Structure Elucidation. 3., *SIAM J. Appl. Math.*, 32 (1977) 534-551.

26. H.L. Morgan, The generation of a unique machine descriptor for chemical structure – A technique developed at Chemical Abstracts Service, *J. Chem. Doc.*, 5 (1965) 107-113.

27. R.E. Carhart, D.H. Smith, H. Brown, and C. Djerassi, Applications of artificial intelligence for chemical inference. XVII. An approach to computer-assisted elucidation

of molecular structure, *J. Am. Chem. Soc.*, 97 (1975) 5755-5762.

28. A. Lavanchy, T. Varkony, D.H. Smith, W.C. White, R.E. Carhart, B.G. Buchanan, and C. Djerassi, Rule-based mass spectrum prediction and ranking. Applications to structure elucidation of novel marine sterols, *Org. Mass. Spectrom.*, 15 (1980) 355-366.

29. N.A.B. Gray, R.E. Carhart, A. Lavanchy, D.H. Smith, T. Varkony, B.G. Buchanan, W.C. White, and L. Creary, Computerized mass spectrum prediction and ranking, *Anal. Chem.*, 52 (1980) 1095-1102.

30. N.A.B. Gray, C.W. Crandell, J.G. Nourse, D.H. Smith, M.L. Dageford, and C. Djerassi, Applications of artificial intelligence for chemical inference. XXXIV. Computer-assisted structural interpretation of C-13 spectral data, *J. Org. Chem.*, 46 (1981) 703-715.

31. Y. Kudo and S.I. Sasaki, Principle for exhaustive enumeration of unique structures consistent with structural information, *J. Chem. Inf. Comp. Sci.*, 16 (1976) 43-49.

32. J.G. Nourse, D.H. Smith, R.E. Carhart, and C. Djerassi, Computer assisted elucidation of molecular structure with stereochemistry, *J. Am. Chem. Soc.*, 102 (1980) 6289-6295.

33. R.E. Carhart, D.H. Smith, N.A.B. Gray, J.G. Nourse, and C. Djerassi, GENOA: A computer program for structure elucidation utilizing overlapping and alternative substructures, *J. Org. Chem.*, 46 (1981) 1708-1718.

34. D.H. Smith, B.G. Buchanan, R.S. Engelmore, A.M. Duffield, A. Yeo, E.A. Feigenbaum, J. Lederberg, and C. Djerassi, Applications of artificial intelligence for chemical inference. VIII. An approach to the computer interpretation of the high resolution mass spectra of complex molecules. Structure elucidation of estrogenic steroids, *J. Am. Chem. Soc.*, 94 (1972) 5962-5973.

35. D.H. Smith, B.G. Buchanan, W.C. White, E.A. Feigenbaum, C. Djerassi, and J. Lederberg, Applications of artificial intelligence for chemical inference. X. Intsum: a data interpretation program as applied to the collected mass spectra of estrogenic steroids, *Tetrahedron*, 29 (1973) 3117-3134.

36. B.G. Buchanan, D.H. Smith, W.C. White, R. Gritter, E.A. Feigenbaum, J. Lederberg, and C. Djerassi, Applications of artificial intelligence for chemical inference. XXII. Automatic rule formation in mass spectrometry by means of the MetaDendral program, *J. Am. Chem. Soc.*, 96 (1976) 6168-6178.

37. T.M. Mitchell and G.M. Schwenzer, Applications of artificial intelligence for chemical inference. XXV. A computer program for automated empirical 13C nmr rule formation, *Org. Magn. Reson.*, 11 (1978) 378-384.

38. J.E. Gurst and A.K. Schrock, Mass spectra of Androstane-7,17-diones, *J. Org. Chem.*, 45 (1980) 4062.

39. N.A.B. Gray, Applications of artificial intelligence for organic chemistry. Analysis of C-13 spectra, *J. Artificial Intelligence*, 22 (1984) 1-21.

40. C. Djerassi, D.H. Smith, and T.H. Varkony, A Novel Role of Computers in the Natural Products Field, *Naturwissenschaften*, 66 (1979), 9-21.

41. C. Djerassi, D.H. Smith, C.W. Crandell, N.A.B. Gray, J.G. Nourse, and M.R. Lindley, The Dendral Project: Computational AIds to Natural Products Structure Elucidation, *Pure. Appl. Chem*, 54 (1982), 2425-2442.

42. Molecular Design Ltd, 2132 Farallon Drive, San Leandro, CA94577.

Table 1.

Some assessments of Dendral as an Artificial Intelligence system

The first expert system, produced at Stanford and called Dendral, was developed by a professor of computer science E. Feigenbaum and nobel laureate J. Lederberg. Dendral attempted to capture Lederberg's expertise in analyzing organic compounds from mass spectrometry. Issued under a Stanford licence and one of the most successful AI programs ever, Dendral can be found in many chemistry laboratories.
J. Shurkin, IEEE Expert, Winter '86, p. 10.

Meta-Dendral, the first automated scientific discovery program, discovered previously unknown rules of mass spectrometry.
M. Walker, IEEE Expert, Spring '87, p. 69.

Dendral surpasses all humans at its task and, as a consequence, has caused a redefinition of the roles of humans and machines in chemical research.
F. Hayes-Roth, D.A. Waterman, and D.B. Lenat, Building Expert Systems, Addison-Wesley, 1983, p. 9.

Dendral is a success story. The results derived from its use are cited in over 50 scientific papers, which attests not only to its usefulness but also its scientific credentials. It is in regular and routine use. The number of its users was expanding so rapidly that in 1983 a separate company was set up for its distribution and continued enhancement.
P.S. Steel, "Expert Systems – A Practical Introduction", MacMillan, 1985.

An expert system can excel its human counterparts in terms of speed, completeness of knowledge, thoroughness, and its lack of human frailty. Existing expert systems such as DENDRAL and MYCIN have accomplished remarkable performance. The field of artificial intelligence is poised on the verge of a tremendous expansion of such systems into the marketplace. ART, the Automated Reasoning Tool, is literally a tool for building expert systems.
B.D. Clayton, "Inference ART Programming Primer", Inference Corporation, 1985.

Table 2.

Decision tree for classification of the spectra of ketones as used in the Heuristic Dendral program

| Ketone –
| Find two fragment ions, A and B, such that the sum of their masses is 28 atomic mass units (amu) greater than the molecular weight M.
| Confirm the presence of ions at A-28amu and B-28amu
| | Methyl-ketone3 –
| | Require intense ions at m/z 43 and m/z 58.
| | Require ions at M-43 and M-15
| | | Generate structures incorporating
| | | CH_3-(C=O)-CH_2-C(<)-CH<
| | Ethyl-ketone3 –
| | Require intense ions at m/z 57 and m/z 72.
| | Require intense ions at M-29 and M-57
| | | Generate structures incorporating
| | | CH_3-CH_2-(C=O)-CH_2-C(<)-CH<
| | NPropyl-ketone3 –
| | Require intense ions at m/z 43, 71, and 86
| | Require an ion m/z 58
| | | Generate structures incorporating
| | | CH_3-CH_2-CH_2-(C=O)-CH_2-C(<)-CH<
| | IsoPropyl-ketone3 –
| | Require intense ions at m/z 43, 71, and 86
| | Require absence of any ion at m/z 58
| | | Generate structures incorporating
| | | $(CH_3)_2$>CH-(C=O)-CH_2-C(<)-CH<
| | Generate structures with >C=O but prohibit any occurrences of
| | methyl-ketone, ethyl-ketone or propyl-ketone substructures
| Ether
| Verify absence of ions at M-17 and M-18
| ...
| ...

Table 3.

Structure generation programs developed for the Dendral project

Dendral algorithm:

Algorithm for generating acyclic, "tree-structured" graphs; as used in Heuristic Dendral

StrGen program:

Isomer generating system employing atom-partitioning, vertex graph catalogues, labelling algorithms, and the Dendral algorithm.

Congen program:

Candidate structure generation system based on concept of superatoms, use of the StrGen system for intermediate structure generation, and additional labelling procedures for embedding of superatoms.

Congen-II:

Congen - but using a more practical, if mathematically less sophisticated algorithm for the generation of intermediate structures.

Stereo:

Module for the constrained generation of stereoisomers for candidate structures.

Genoa:

Structure generation algorithm based on a generalization of standard graph matching techniques.

Table 4. The MetaDendral programs

Planner:

A program for the "interpretation" of high resolution mass spectral data of a compound incorporating a known skeleton and substituent groups whose presence will not affect the fragmentations of the skeleton.

Planner employs rules, developed by a chemist, that characterize the fragmentation processes of the skeleton.

Planner can derive information concerning the placings of substituents on the known skeleton.

IntSum:

A program for analyzing a compound's high resolution mass spectrum in terms of possible fragmentations of its (given) structure.

IntSum identifies all those fragmentations of a structure for which there is evidence in its spectrum, and can summarize data on related compounds in terms of equivalent fragmentations of a common skeleton.

IntSum provides the information required by a chemist who is formulating rules for the prediction of mass spectra, or rules for use in an interpretation program such as Planner.

RuleGen/RuleMod (mass spectral version):

RuleGen automates the process whereby general rules (mass spectral "break processes") are inferred from descriptive data such as produced by IntSum

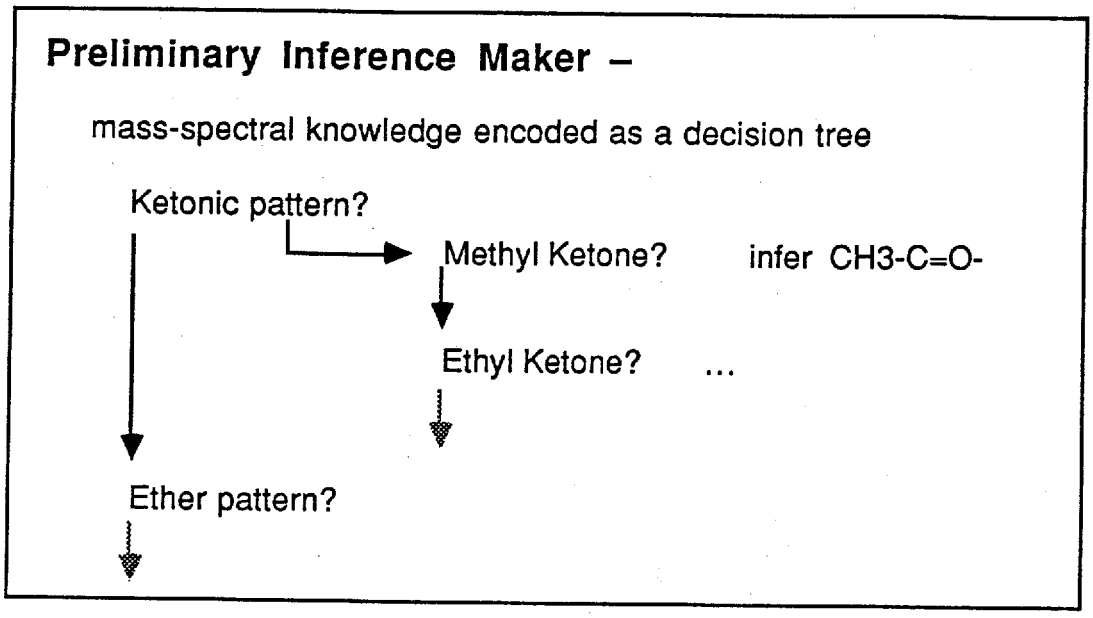
RuleMod refines the rules generated using the RuleGen procedure.

RuleGen (Carbon-13 nmr version):

A version of the RuleGen algorithm used to derive rules for predicting Carbon-13 nmr shifts of atoms in specified substructural environments.

Figure 1.

Mass spectral processing techniques used in early versions of the Heuristic Dendral program



Dendral generator

Heuristic filter – check for some spectral evidence for each radical considered when building up structure

Predictor

Analyze molecular ion and class specific fragmentations.

For each bond in structure

test for special cases

MethylP ...

Vinylic ...

.

.

otherwise: compute plausiblity of fragmentation using ...

calculate ion intensities by combining fragmentation plausibility ionization probability

Figure 2.

A "table-driven" spectrum predictor

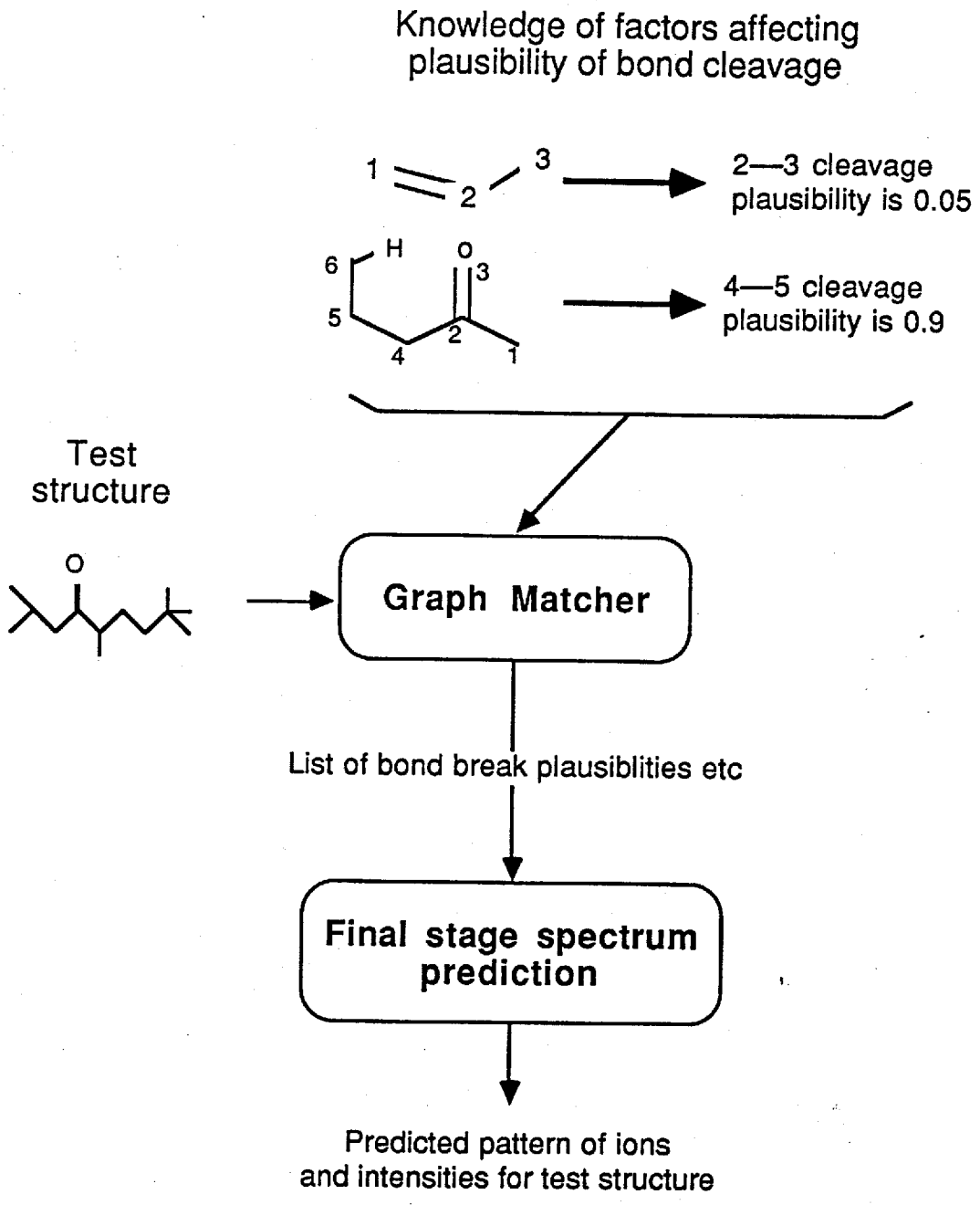


Figure 3.

Diverse developments from the initial Heuristic Dendral Project

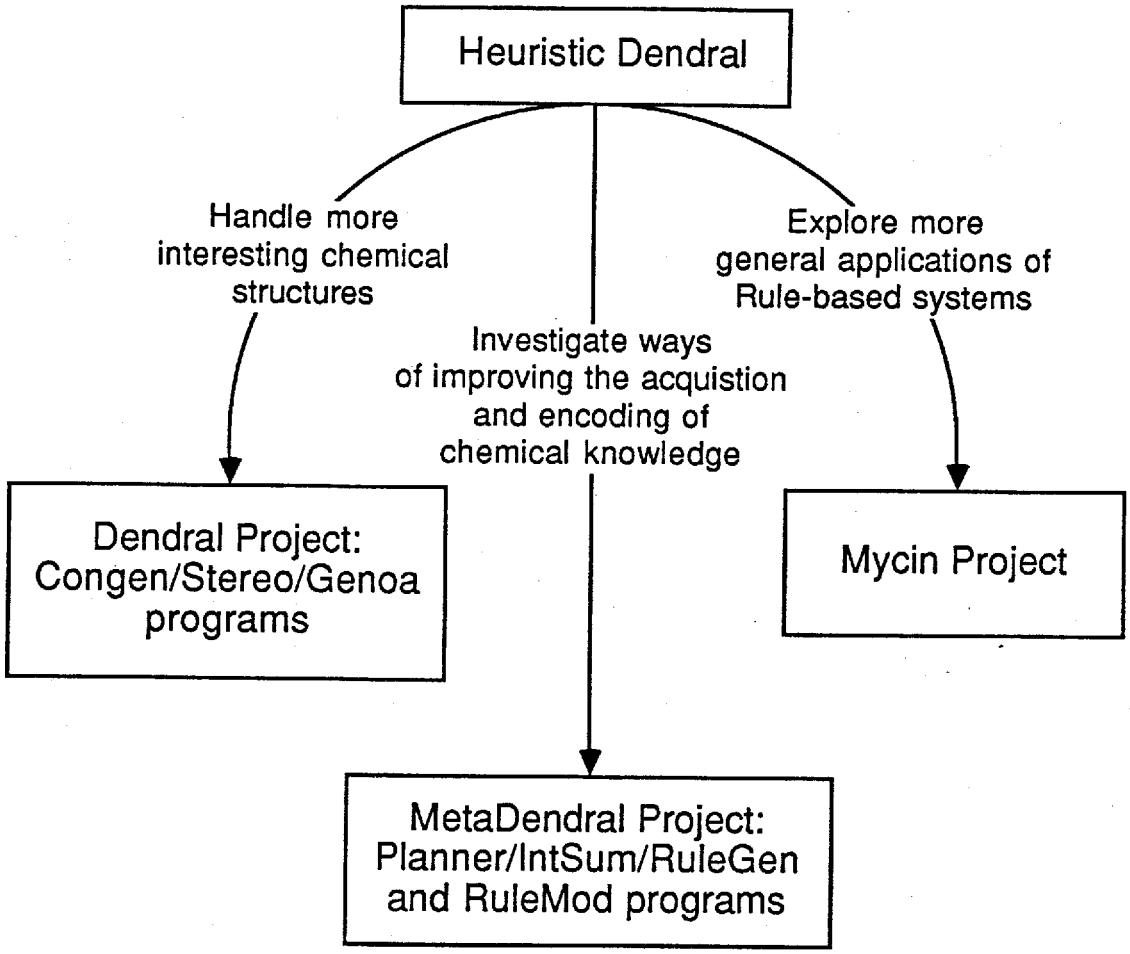


Figure 4.

In the Congen system, the human user performed the data interpretation tasks and the program handled the combinatorics of structure generation.

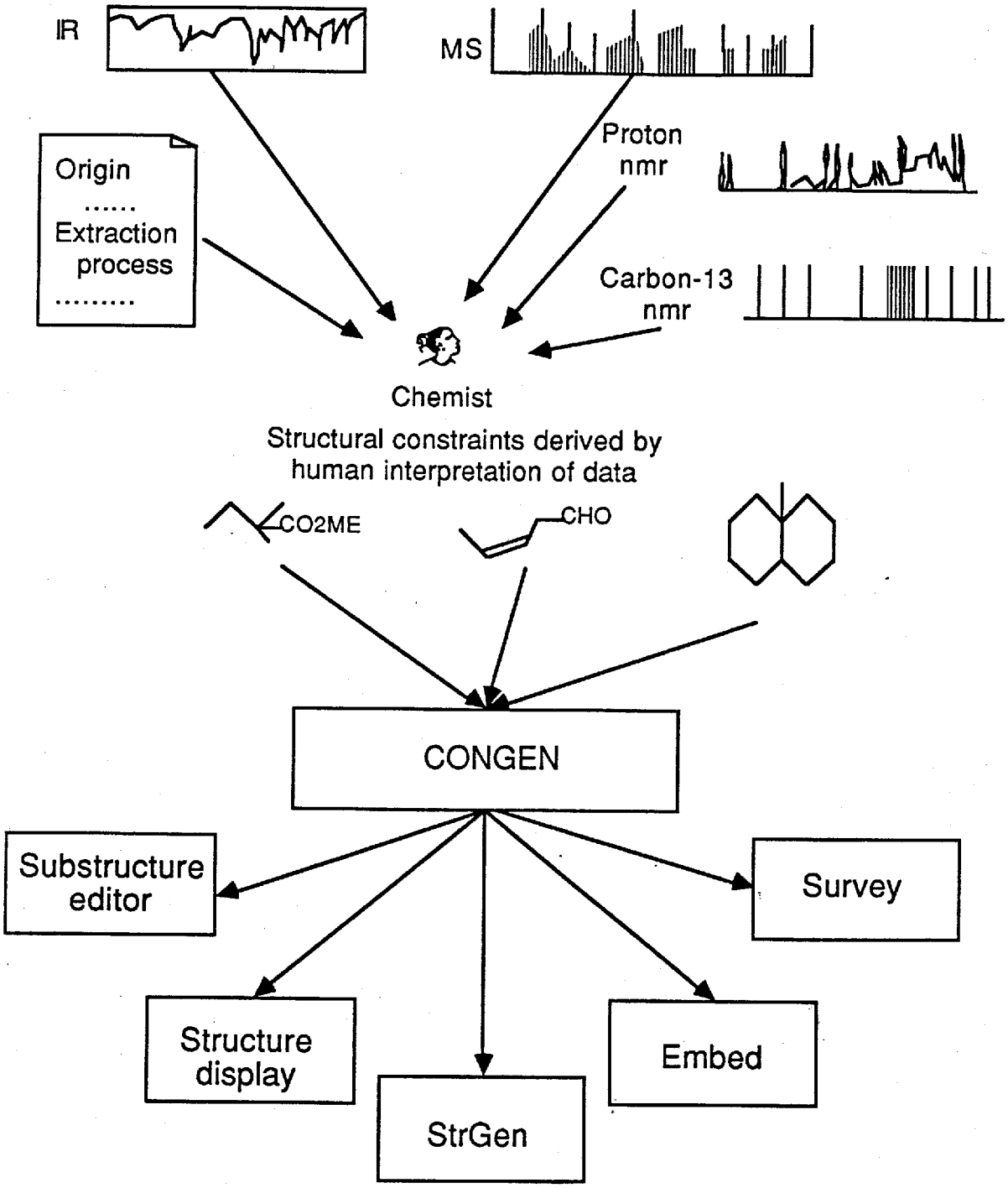


Figure 5.

The Planner program assisted in the interpretation of mass spectra of compounds from a known class.

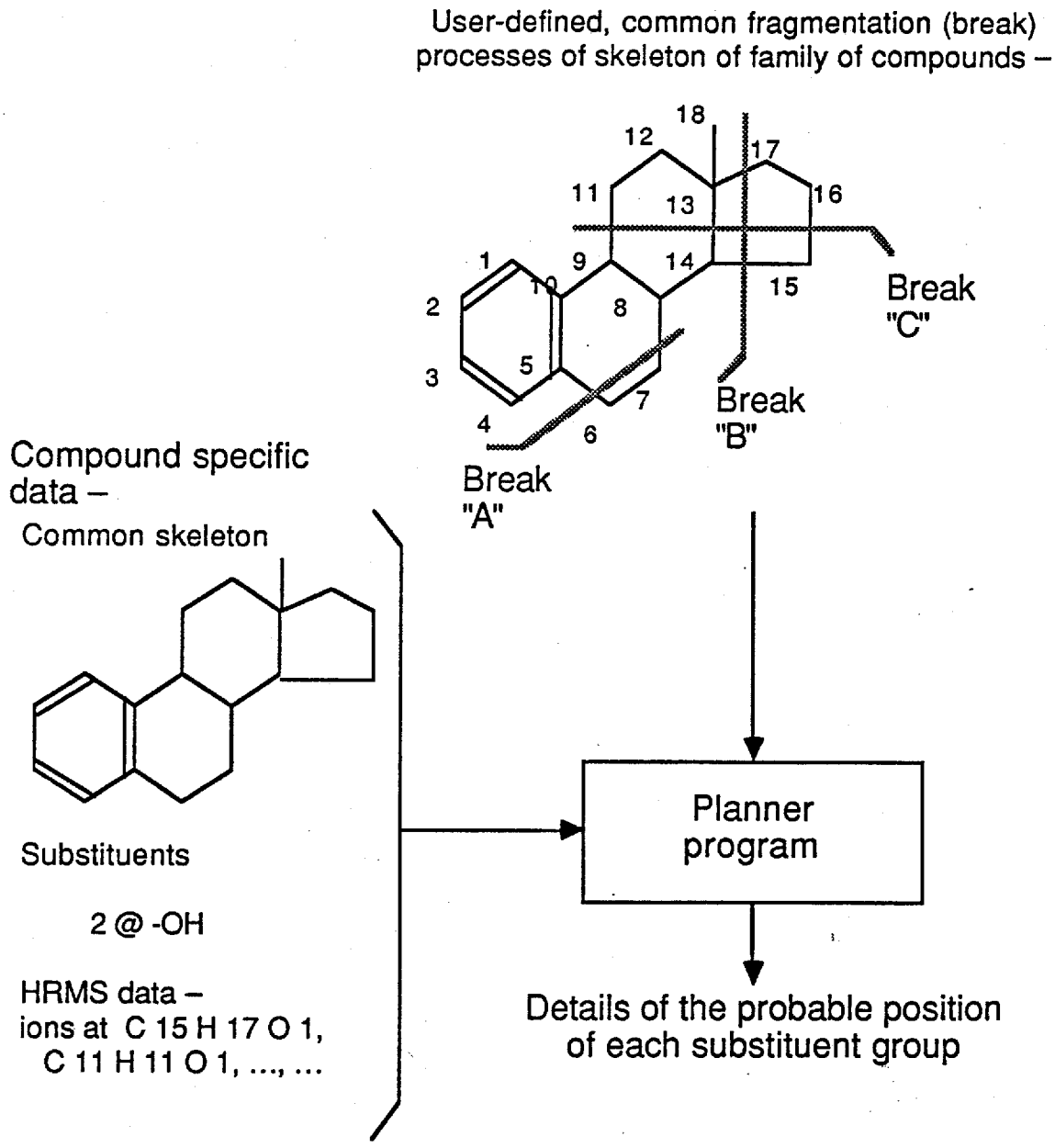


Figure 6. RuleGen's search for Carbon-13 nmr prediction rules.