What's in a Patent? Information! But Can I Find It?†

STUART M. KABACK

Research Information and Analysis Unit, Analytical and Information Division, Exxon Research and Engineering Company, Linden, New Jersey 07036

Received December 5, 1983

The information in chemical patents is used for many purposes, including but not limited to patentability, validity, and infringement studies, state of the art reviews, and the monitoring of competitive technology. For some of these purposes, only the information in patent claims is germane; for others, the examples are most important. Information is sometimes presented in a highly specific fashion, while at other times it is quite generic. Searches similarly are sometimes aimed at specifics and other times at generics—and often a specific search must contend with generic patents and vice versa. Overlying these problems are the difficulties of dealing with pictorial and numerical information and with the context in which information is presented. This paper examines these manifold problems, taking a close look at the kinds of information needed by the users of patent information.

During 1982–1983 I participated, along with a number of colleagues at other oil companies, in a study comparing patent retrieval from several different databases. The study was carried out under the auspices of the American Petroleum Institute's Patent Task Force, and some results from the study have been published.[1] One of the most striking findings was how often a needed reference was present in a database but was virtually unretrievable, unless one retrieved a very large number of references and culled through all of them. This was not a startlingly new perception, of course, but one that served to focus my attention on the problem, hence the title of this paper, which examines some of the problem areas in indexing and retrieving information from patents.

Patents are in some respects quite different from other sources of technical information. Once upon a time, nearly 60 years ago, there was a man who opened Pandora's box, and the world has never been the same since. His name was Eugene A. Markush, and he—or perhaps we had better blame his patent attorney—managed to get a U.S. patent claiming the preparation of a substance with an indeterminate structure. The first claim of Markush's pioneering patent is shown in Figure 1.

This is a process claim, and we are probably more used to seeing Markush-type product claims, replete with such well-defined substituents as R, X, and Y, as shown in Figure 2.

The grouping is heterogeneous to say the least: water, cholesterol, dicholesteryl ether, or oxygen. I think I can safely say that such a Markush claim would never be allowed. But Figure 3 shows a real Markush structure, taken from a real patent publication. Note that, while a few of the possible substituents are well-defined, such as nitro and nitrile, most of them are open-ended generic species. This single structure encompasses more potential compounds than the entire CAS Registry. For legal purposes, it constitutes a report in the literature of such disparate compounds as those shown in Figure 4, whether or not they were actually synthesized.

Because of this legal significance, a great deal of effort has been devoted by database producers and information scientists—especially those concerned with the synthesis of new compounds—to developing systems that are capable of storing and retrieving the indeterminate structure descriptions of Markush-type disclosures. It seems sometimes that virtually all of the work being done on chemical information handling is devoted to structures in some way, either to specific structures as in the CAS Registry or to Markush structures in several systems.

Let there be no misunderstanding. I am certainly not trying in any way to downgrade the importance of this work. I do believe, though, that the preoccupation with structural information had led to the neglect of other important problems in handling chemical information.

This paper will look to some extent at Markush searching, but for the most part it will examine other types of searching, the sort of thing that my colleagues and I frequently look for and often have difficulty finding.

First of all, Markush problems are significant not only for the makers of new compounds but also for process-oriented people like the petrochemical industry. Consider a simple Markush situation. Figure 5 shows an etherification reaction with the product in terms of a Markush structure. A claim on this product would encompass the common octane-improver methyl *tert*-butyl ether (MTBE), but also methyl *tert*-amyl, methyl *tert*-octyl, etc. A process claim very much like this appears in a recent U.S. patent. If one is interested in manufacturing MTBE with a catalyst similar to the one used in this patent, one needs to know about this patent. It should be indexed in a way that shows that MTBE is potentially one of the products. But if we look at a portion of the indexing record for this patent in *Chemical Abstracts* (CA), shown in Figure 6, we find only one ether product covered, along with its olefin starting material, and that turns out to be the *tert*-amyl ether, which was the only one made in an example. Chemical Abstracts Service (CAS) would have indexed MTBE as a product if it had been claimed specifically, even without its having been synthesized—and this is an important advance in CAS coverage over the past few years—but there is not any specific claim, only a Markush formulation, and so MTBE is not indexed.

There are two main techniques today for building files of chemical structures: full topological coding, exemplified by the CAS Registry, and fragmentation systems, exemplified by such systems as the Derwent chemical code, the IFI/Plenum Comprehensive Data Base system, the API's system of chemical aspects, and the Gremas code of Germany's IDC.[2,3] There are linear notation systems, notably those based on Wiswesser line notation (WLN), but they are beyond the scope of this discussion.

A file dealing with chemical patent information should be able to cope with Markush structures at both the input stage and the search stage. At input, we would want the file to be able to accept a Markush formulation in such a way as to show that the file contains each of the individual structures encompassed by that Markush, so that a search of the file would be able to locate any of the structures.

Topological files have trouble with Markush structures. IDC's topological file has some capability for registering Markush formulations with specifically defined substituents,

† Presented at the Herman Skolnik Award Symposium, Division of Chemical Information, American Chemical Society, Washington, DC, Aug 30, 1983.

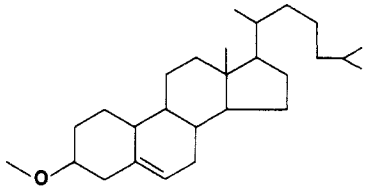
0095-2338/84/1624-0159$01.50/0

The process for the manufacture of dyes which comprises coupling with a *halogen-substituted pyrazolone*, a diazotized unsulfonated material selected from the *group consisting of aniline, homologues of aniline, and halogen-substituted products of aniline.*

Figure 1. First claim of Markush's U.S. Patent 1 506 316.

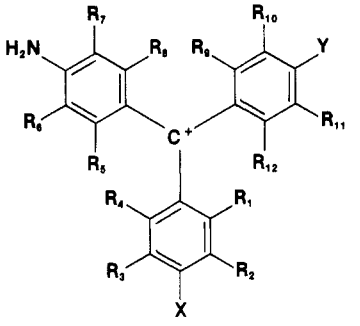
X—O—Y

X, Y are independently H or



or X and Y together are = O

Figure 2. Fantasy Markush-type structure.



A^-

Source:
DE 2,723,774

R_1-R_{12}	X, Y		
H	$-NR_{13}R_{14}$	X-R_2	
Alkyl		Y-R_{10}	may form
Haloalkyl		R_1-R_2	5-6 member
Halogen	R_{13}, R_{14}	R_5-R_6	ring
Alkoxy		R_9-R_{10}	
Alkylmercapto	H		
Alkylsulfonyl	Alkyl		
Alkoxysulfonyl	Haloalkyl		A^- is an anion
Alkylcarbonyl	Cyanoalkyl		
Alkoxycarbonyl	Cycloalkyl		
$-NO_2$	Aryl		
—CN	Aralkyl		

Figure 3. Real Markush-type structure.

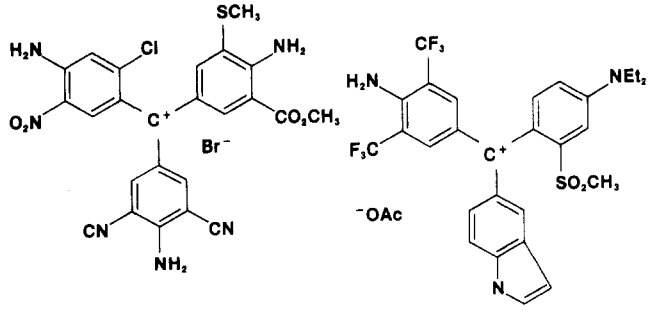
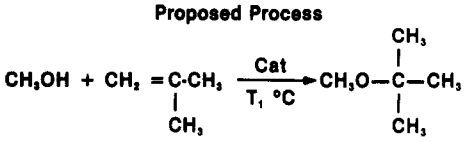


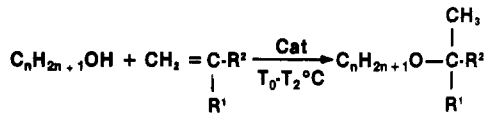
Figure 4. Members of a Markush family.

but it cannot handle generic ones; IDC leaves that to the Gremas code. To date, no topological system can fully cope with the registration of Markush structures, but important work is being done in this area, notably by the DARC group in France and by Professor Michael Lynch and his co-workers in the U.K.[4] CAS is also trying to solve this problem.

Proposed Process



Patented Process



$_n$ = 1-3

$T_0 < T_1 < T_2$

R^1, R^2 = CH_3, C_2H_5, C_3H_7, C_4H_9

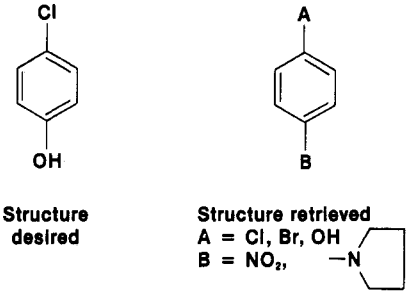
Figure 5. Petrochemical process in Markush form.

67-56-1, reactions, etherification of, by alkene (methanol) .....

513-35-9, etherification of methanol by ..... **(2-methyl-2-butene)**

994-05-8, prepn of (methyl tert-amyl ether)

Figure 6. CAS indexing of U.S. Patent 4 262 145 (partial).



Structure desired

Structure retrieved
A = Cl, Br, OH
B = NO_2, —N

Figure 7. Markush false drop.

But to this point, the only systems that can deal with Markush formulas are the fragmentation codes—and they deal with them with varying degrees of success. Fragmentation systems have the shortcoming of being unable to depict all of the details of how a given molecule is strung together. The more detail of a structure that remains undefined by the system, the more noise there will be at the retrieval stage. Todays operating fragmentation systems can frequently describe some of the details of molecular constitution—with the Gremas code probably the most effective—but they all have their limitations. And so, a fragmentation system, while it has the flexibility to cope with Markush formulations, tends to give relatively high noise levels at search time. Combinations of fragmentation systems with topological systems can improve the situation but do not fully solve the problem if the topological system cannot contain the structures implied by a Markush formulation.

At the searching stage, one of the most interesting developments has been the introduction of Markush searching of a file of definite structures by the Questel organization for its DARC file. For the most part, though, Markush searching is done with the fragmentation systems—and must be done in this way for files that contain Markush input.

As noted above, fragmentation systems give only a partial description of a molecule's constitution, a fact that contributes to noise in searching. Another factor contributing to noise is the fact that most systems cannot distinguish between sub-

stituents that are both present in a molecule and those that are merely alternatives.

This is illustrated in Figure 7, which shows a search for a chlorophenol. In most fragmentation systems, a search for a benzene ring substituted by chlorine and hydroxy will retrieve not only valid hits but also the structure shown on the right, in which Cl and OH are alternative substituents at the same position—if one is present, the other is not. Once again the Gremas system is capable of eliminating this type of false drop. Derwent can eliminate references to structures that *must* contain certain unwanted substituents. IFI can do this in the batch mode, but in its online system can do so only for relatively common compounds, not for uncommon compounds or Markush formulations. However, when the other substituents are only possibilities, such as the nitro group or pyrrolidine ring in this example, it is impossible for either system to eliminate them. Files of compounds continue to grow, and the more they grow the more important it becomes to eliminate searching noise. Systems that could cope with Markush alternatives of this sort would be more than welcome in the commercially available fragmentation files.

Let me touch on one other aspect of chemical substructures. Frequently, it is important to be able to search for a system that is modified by a compound with a given substructure. Let us say we want to find Ziegler catalysts that are modified by thiourea compounds. Assuming that we can adequately identify in the CASearch files the set of references on Ziegler catalysts, how do we intersect them with references on thioureas? Using the word thiourea will get some of the references, but what about those references with index terms beginning with the Registry Number for a thiourea but that never use the word thiourea in the indexing language?

Both SDC and DIALOG provide the facility of dictionary files, from which we can identify a set of Registry Numbers and transfer those numbers automatically into the corresponding CASearch files. The problem is that a quick check of DIALOG's Chemname file shows over 3000 hits for the term thiourea. Neither SDC nor DIALOG can effectively transfer a list this long. What we need in the CASearch indexing record is not only the Registry Number, as important as it is, but also the name that it stands for. That would allow us to combine more effectively the terms that describe structural features with other bibliographic descriptors. A valid alternative, of course, would be the ability to transfer very long lists of Registry Numbers from the dictionary files to the bibliographic files and the ability to search those long lists effectively. The fragmentation systems are excellent at handling this type of problem.

What substances in a patent do we need information on? Clearly, the products of reactions, but what about starting materials and isolated intermediates? In some fields, starting materials are relatively unimportant, most especially in the pharmaceutical and agricultural chemical fields with their focus on new compound synthesis. But in a field such as petrochemicals, where there may be several attractive routes to the same end product, the reactants and intermediates can be of critical interest. If one is interested in synthesizing styrene by the side-chain alkylation of toluene with methanol, one is unlikely to want references on the route that involves coupling of toluene to stilbene, followed by disproportionation with ethylene. If the stilbene route is of interest, one may want to pick it up at the intermediate stage or the starting point. Whichever route is of concern, one is very unlikely to want the retrieval cluttered by all of the art on the traditional route, alkylation of benzene with ethylene followed by dehydrogenation. Sometimes we start off with no specific end product in mind, when we are looking for new uses for a given starting material, and just want to find what it has been converted to in the past. Most systems do deal reasonably well with starting materials, although the Derwent database did not begin indexing them until the middle of 1981, and then only for a limited number of specific compounds. We need indexing of all starting materials, common and uncommon, and not only of specifics but also of generics. IFI and API do this; hopefully, Derwent will get around to full coverage of starting materials.

Products. Starting materials. Active ingredients of formulations. Catalysts. Obviously, these need to be indexed. But what about supposedly unimportant auxiliaries, such as the lowly solvent? We all expect the solvent to be indexed if the patent teaches that a reaction must be run in a special solvent, but what about run-of-the-mill disclosures that a given solvent has been used in formulating a paint or a drug or in carrying out a reaction? Patents are used not only for legal purposes but also in the formulation of research and marketing strategies, and such strategies often need to know what solvents have been used in a given industry or what uses a given solvent might have.

These supposedly insignificant substances in a patent, that nobody tends to index, can turn out to be quite significant even for legal purposes. For example, perhaps a researcher has invented a process for carrying out a reaction in a Teflon-lined reactor, which gives results different from those obtained in steel equipment. An earlier patent may well have disclosed the use of such a reactor for this process, without remarking on any significant effect. One can of course retrieve all of the references on such a process, but they may be very voluminous. In my ideal world, every substance and material involved in a patent will be indexed, with an accompanying role indicator to describe its function. Unrealistic? Perhaps. Costly? Undoubtedly. But certainly beneficial to users of patent information.

Let us examine further the role indicator. Many of us are familiar with relatively simple role indicators, for products, reactants, and substances merely present. Such indicators are invaluable searching tools, allowing one to discriminate among cases in which A is converted to B, B is converted to A, or both A and B are produced from C. They are certainly easier to use than a Boolean operation in which one links a Registry Number in CASearch with a series of terms that might be used to describe preparation, synthesis, manufacture, or what have you, and it would help greatly if CAS would begin assigning role indicators intellectually. DIALOG of course produces algorithmically product and derivative role indicators in its CAS files, and this is certainly useful, but it is less than 100% effective. API, IFI/Plenum, and now Derwent have all had success with indexer-assigned role indicators. It would help if CAS followed suit.

But the roles product, reactant, and present are restrictive. What we really need are additional roles, which can designate solvents, catalysts, and other concepts. It is interesting to note that the Gremas system includes about 30 roles![3] Gremas is admittedly a costly system at the indexing stage, but it takes the bits and pieces of information available in documents and organizes them in a way that increases their information content. As our information files continue to grow, we are going to need such organized systems, or else we run the risk of drowning in seas of false drops. Is it a coincidence that the winners of the Skolnik award of the ACS Division of Chemical Information include Robert Fugmann of Hoechst, so responsible for the development of Gremas; Herman Skolnik himself, whose long history of accomplishments includes the concept of the multiterm index, aimed at increasing the information conveyed by a set of indexing terms;[5] or our present award winner Russell Rowlett? One of my earliest memories of Rowlett was his presentation to an earlier meeting of the

Division of Chemical Information of the then-new concept of subdivided headings in CA indexes, a simple concept that made it far easier to obtain needed information by organizing pieces of information in a meaningful way. Fugmann's papers on the occasion of his winning the Skolnik Award are particularly instructive.[6,7]

Let us look at some more of the things we need in the indexing of patents. There are times when one must search for something that is not present, where an ingredient is usually used in a given process, but one wants only those references where it is not there. The thing that is absent may be air, light, water, or copper—the list could go on and on. Remember that Ziegler discovered his catalysts by trying to find what elements other than nickel should be *excluded* from the system. Sometimes there are adequate terms to describe the missing substance, anhydrous, anaerobic, and so forth, but sometimes there are none. The API has an interesting method of dealing with this situation: linking the term NONE to the substance that is absent. An alternative technique might be to have a special role for a substance that is absent. Of course, the patentee must first call attention to its absence, but once that is done, it is important to be able to search for what is not there.

Let us look further at the link. Linking logic is often used in fragmentation systems to tie together the fragments of a given chemical substance, and in recent years, we have seen linking operators introduced in online systems to enable us to select terms that occur in the same context, such as a single CAS indexing phrase. Users of the API databases will be familiar with the way in which API allows other types of concept to be linked, such as the quantity NONE. But there are other ways in which linking logic might be applied. Consider a search for a hydrodesulfurization process in which a bottoms cut is recycled to the reactor. If one merely combines the concepts hydrodesulfurization, bottoms, and recycle, one will get a lot of useless information, because there is always some bottoms fraction produced and the hydrogen, if nothing else, is normally recycled. But if it were possible to link the term recycle to the specific cut that is recycled, then the information content of the indexing would be greatly enhanced, and one would get a cleaner search output with less noise.

In the petroleum refining industry there are many combination processes, in which a series of steps take place. Many patented combination processes may share the same steps, and an invention may involve changing the sequence of the steps. The art of Ziegler catalyst preparation is another in which the sequence of combining familiar catalyst ingredients is often the key inventive feature. There is a need for an indexing system that could distinguish the order in which things take place.

Then there is the problem of numerical values, not limited to patent information but more serious with patents than elsewhere, because patents invariably provide value ranges that may extend far beyond the real operating values. Reaction temperatures, for example—but are they needed for legal purposes or for research and development information? For example, perhaps we are interested in polymerization catalysts effective at temperatures above 150 °C. There are many patents that will disclose and even claim ranges such as 50–200 °C, but examination shows that all of the data were collected at 70–90 °C. These may be of importance legally, but they are not likely to interest the research and development people very much—they will be interested in those patents where the temperature exceeded or at least approached 150 °C. Nobody effectively handles things like reaction temperatures, pressures, or the like. There are systems that have indexing terms to deal with these subjects, but the ranges are generally of limited usefulness, and the application of the terms tends to be incomplete. And if we are dealing with a field in which there are 1000 patents describing a process run below a given temperature and 20 run above that temperature, we are dealing with a field where users need help from indexing and retrieval systems. They are not getting much so far.

Other numerical values can be equally important. Concentration is one—was the oxidant air, enriched air, or pure oxygen? Was the sulfuric acid dilute, or was it oleum? What about the molecular weight of polymers? It makes an enormous difference whether the polyethylene produced by a given catalyst is a liquid oligomer, a wax with 15 or 20 ethylene units, a standard polyethylene with a molecular weight of a few hundred thousand, or an ultrahigh molecular weight species with a molecular weight of over 10 million. Yet sometimes these vastly disparate species are indexed as the very same thing.

There are other shortcomings in polymer indexing. It is clear that random styrene–butadiene copolymer rubber is very different from Shell's Kraton block styrene–butadiene–styrene copolymer, which is in turn very different from a high-impact polystyrene graft onto polybutadiene. Similarly, an ethylene–propylene copolymer elastomer is very different from a block propylene–ethylene copolymer, a polypropylene containing a few percent of random ethylene units, or a polyethylene containing a few percent of random propylene units. But in each case, these disparate materials are grouped under a single Registry Number. Is it too late to reverse this policy and make polymer registration more meaningful than it is today?

Another technology not well served by present registration or indexing policies is that of synthetic zeolites. We are faced today with a rapidly growing galaxy of these substances, aluminosilicates and their isomorphously substituted analogues, extremely important as catalysts, catalyst carriers, selective adsorbents, etc. Their chemical formulas are inexact, and they are usually defined in terms of their X-ray diffraction patterns. Their names are determined not by rigorous nomenclature but are rather conferred by their creators: an alphabet soup of Roman and Greek letters, ZSMs, and what have you. These different zeolites differ very greatly in their catalytic and adsorptive properties, and the differences can be greatly affected by cation substitution. It is imperative that all database producers do a thorough job of identifying the zeolites that are involved in both patent and literature references, something that is being done in a less-than-complete fashion today.

Chemical engineering often gets short shrift in abstracting and indexing. Sometimes the results of a chemical process can be very dependent on the engineering. Consider, for example, a polymerization reaction carried out in a stirred autoclave reactor and the same polymerization run in a tubular reactor. There is an enormous difference in the molecular weight distribution that is obtained. If the product is a copolymer, there will be great differences in comonomer composition and distribution. Yet it is very difficult to pinpoint patents that specify the use of a tubular reactor. It is rarely indexed and may even be omitted in abstracts, but at least the files with searchable abstracts give us a better chance. It is important for database producers to concentrate on more than the straightforward chemistry.

Another engineering situation brings up a second retrieval problem. There are many ways of carrying out polymer grafting processes. One of them is by grafting in a screw extruder. Our first obstacle is the extruder itself: will its use be indexed or won't it? But if we try to carry out a search on extruder grafting, we will get lots of noise, because we will get many references in which a graft polymer prepared by whatever method is later processed in an extruder. The terms grafting and extruder are floating around in the indexing

record and produce the undesired hit. This sort of noise could be eliminated by more compartmentalization of an indexing record—whether by the use of multiple indexing phrases that are kept as distinct subrecords or by the use of link operators to tie terms together.

I touched earlier on the need to get at information in patents that might seem to be trivial but turns out to be very useful. Patents can be rich sources of information on things beyond what they claim. Many a patent specification contains extended discussion of prior art, sometimes describing entirely different ways of doing things and sometimes describing closely related ones. Often a patent will disclose that its invention is important because of something that happens three or four steps later. For example, there is a patent that covers an improved process for step one of a four-step synthesis of a known commodity and that discloses the entire four-step synthesis as known prior art. This was the first public disclosure of the company's interest in this process and was published 5.5 years before their first commercial announcement of the process.

One could not expect any database producer to index the ultimate product for this patent. How could one find it, then, if one were interested in that product? One possibility is by having available full-text search of the patent specification, and such a system is currently available for U.S. patents since 1975 from Mead Data Central. But I am pessimistic about the value of such systems, because extraneous material is often introduced in patent specifications and this material is bound to lead to extra false drops. Admittedly, there are times when full-text search will be the only way to find a buried disclosure, but I think those times will be few and far between.

There is a good and logical way to deal with disclosures like this, however. In a high-quality patent abstract, of the sort that Derwent has produced for many years, background material of this sort if generally included. In fact, Derwent did include in its abstract of this patent the disclosure of the ultimate purpose of the process. Searchable abstracts are valuable additions to our arsenal of information tools. IFI has them. Derwent has them for recent years and is preparing to add searchable abstracts to its backfile. API originally thought it did not need them because of its controlled indexing language but has now included them for a number of years and has found them quite valuable. Searchable abstracts are needed for CAS as well. Further, they should be available through all vendors of CASearch files, so that users can utilize whichever CASearch version is most appropriate for the search at hand.

The reason that searchable abstracts are so valuable is that the ideas expressed in them have extra value: they have been selected by a document analyst for their importance in relation to the document. The claims of a patent are similarly added-value terms because of their legal significance, which is why I am so enthusiastic about the Derwent–SDC U.S. patent files, which make searchable the full claims of U.S. patents. Patent examples would probably also be valuable as full-text searchable files. Recently, somebody spoke of the continuing importance of the abstract as an information filter, because it selects and highlights the most significant information in a document. It is no coincidence that that somebody was Russell Rowlett, delivering the 1981 Miles Conrad Award Address to the National Federation of Abstracting and Information Services.

One thing about abstracts, though. An abstractor must not try to be a creative writer. He or she must strive wherever possible to use the most standard terminology, to put things into idiomatic language. This is especially important with some of the abstracts in semi-English that come out of Japan into the Derwent system. Stilted language is not ideal in a printed product, of course, but it becomes a serious handicap when it becomes input to a searchable database. A human scanning an unidiomatic abstract can compensate; a computer cannot.

I have asked for a great deal from the database producers, without even touching on the subject of reaction indexing. Certainly, there are more individual problem areas I could mention, but there is just one more thing I would like to propose: the idea of patent offices, especially the U.S. office, requiring the use of CAS Registry Numbers to describe substances in patent specifications. This could enable all producers of patent databases to include the most reliable means of identifying specific substances, eliminating considerable indexing effort, avoiding nomenclature ambiguities, and cutting error rates. There are clearly obstacles to the establishment of such a policy, but I can envision some sort of cooperative arrangement with CAS that could make the policy work. It would benefit everyone involved in patent information.

I have proposed many things in this paper, some relatively simple and some costly to implement. Clearly, one could not expect the same costly steps to be taken by all, or several, of the patent database producers. Indeed, one thinks often of the several patent information services that we subscribe to that often process the same information in their own special ways. How much of this is wasteful, and how much produces complementary information that benefits us? In an *Online* article that I referred to at the outset, I fantasized about a master file of patent information, incorporating the inputs of CAS, Derwent, API, and IFI.[1] Could we take that fantasizing a step further, with each of the mythical participants being responsible for its own area of strength, upgrading the quality of their input perhaps in some of the ways I have suggested here? If such a Utopia could ever be created, we might yet be able to provide an extraordinary resource for the users of patent information.

REFERENCES AND NOTES

(1) Kaback, S. M. "Online Patent Searching: The Realities". *Online (Weston, Conn.)* **1983**, *7*, 22–31.
(2) Rössler, S.; Kolb, A. "The GREMAS System, an Integral Part of the IDC System for Chemical Documentation". *J. Chem. Doc.* **1970**, *10*, 128–134.
(3) Fugmann, R. "The IDC System". In "Chemical Information Systems"; Ash, J. E.; Hyde, E., Eds.; Ellis Horwood: Chichester, England, 1975; pp 195–226.
(4) Welford, S. M.; Lynch, M. F.; Barnard, J. M. "Towards Simplified Access to Chemical Structure Information in the Patent Literature". *J. Inf. Sci.* **1983**, *6*, 3–10.
(5) Skolnik, H. "The Multiterm Index: A New Concept in Information Storage and Retrieval". *J. Chem. Doc.* **1970**, *10*, 81–84.
(6) Fugmann, R. "Introduction to the Symposium on the Employment of Grammar in Indexing Languages". *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 117–118.
(7) Fugmann, R. "Role of Theory in Chemical Information Systems". *J. Chem. Inf. Comput. Sci.* **1982**, *22*, 118–125.