for G2 matches the query HY.

CONCLUSION

For many years, searchers have been asking CAS to offer access to the generic and prophetic chemistry represented by Markush structures in patents, i.e., to those chemical substances that were not specifically prepared or claimed and therefore are not in files of specific substances such as the REGISTRY File. MARPAT now offers convenient and powerful structure-based access to the prophetic and the generic substances presented in the chemical patent literature. Only one structure query need be constructed to search MARPAT or any of the other structure-searchable files on STN. MARPAT automatically matches the generic nodes and specific atoms in the query with the file structures and offers the user an option to control the level of specificity. All of the Markush structures available in MARPAT may be displayed online or printed offline with complete or selected portions of the CA File data for those answers. Significant enhancements are now under development in the areas of search response time, hit fragment highlighting, search precision, and query formulation. The file is updated biweekly and automatic current awareness searches are available to keep up with the current patent literature. Quality STN technical support is available to help searchers as needed. Searchers interested in comprehensive access to chemical patent information should consider performing MARPAT searches in the course of their normal structure searching efforts.

REFERENCES AND NOTES

(1) Barnard, J. M., Ed. *Computer Handling of Generic Chemical Structures*, Proceedings of a Conference organized by the Chemical Structure Association at the University of Sheffield, England, March 26-29, 1984; Gower: Aldershot, U.K., 1984.
(2) Barnard, J. M. Online Graphical Searching of Markush Structures in Patents. *Database* **1987**, *10*, 27-34.
(3) Fisanick, W. Storage and Retrieval of Generic Chemical Structure Representations. U.S. Patent 4,642,762, Feb 10, 1987.
(4) Fisanick, W. The Chemical Abstracts Service Generic Chemical (Markush) Structure Storage and Retrieval Capability. 1. Basic Concepts. *J. Chem. Inf. Comput. Sci.* **1990**, *30*, 145-154.

Searching for Simple Generic Structures†

ROBERT N. WILKE

Amoco Research Center, Amoco Corporation, P.O. Box 3011, Naperville, Illinois 60566

Received November 14, 1990

Markush structures representing millions of possible compounds have been indexed by using the Markush DARC system; the new MARPAT system offers similar capabilities. However, these files and their simpler versions (the CAS Registry system and the Generic DARC system) cannot easily provide the searcher with a searchable list of the many possible compounds represented by a simple generic structure. Examples of generic searches and how they are handled in the various structure files will be discussed. Comparisons between these files and the bibliographic files containing generic structure information will be shown, and some recommendations for the future will be examined.

INTRODUCTION

A lot of discussion has centered around the various methods of searching for chemical structures in patents and in the chemical literature. Papers and talks have been given on how the new substructure searching systems that employ connection tables are far superior to those that use codes to represent chemical fragments.[1-9] These substructure systems are shown to give very precise and mostly complete retrieval of complex chemical structures. With the recent introduction of MARPAT and the introduction of the Markush DARC system in 1989, we now have two substructure systems capable of retrieving the complex generic structures that are often found in patents. However, little or nothing has been said about how these search systems work in retrieving information about simple generic structures. One may wonder why anyone would want to search for simple generic chemical structures; but these searches are very common in companies that produce commodity chemicals and monomers, companies that are process oriented.

It is very difficult in connection table based systems to search for simple generic structures, which may contain only a single functional group. The problems that arise in searching these simple generic compounds are due to the large number of possible answers. For example, in searching for a new preparation of aliphatic diamines, any aliphatic diamine that is the product of a reaction could potentially have been prepared by this new method. So to obtain as complete recall as possible all aliphatic diamine containing products will have to be retrieved and then further qualified with the new reaction conditions and or starting materials.

PROCEDURES

In order to investigate how simple generic structures can be searched, the capabilities of six different structure searching systems were studied. These systems were the API Chemical Aspects system, the Chemical Abstracts Registry system, the MARPAT File, Derwent's CPI Chemical Fragmentation, the IFI Chemical Fragmentation, and the Markush DARC system. The Generic DARC system and the Beilstein File on STN were not covered in these examples, simply to cut down on the amount of data. These files consist of three fragmentation-based systems and three substructure-based systems. The first topic will discuss why they can still be searched in fragmentation-based databases. The second topic to be covered will deal with the problems associated with searching for these simple generic chemical substances in the substructure databases; the final topic will be on what the future may bring. To make things easier, databases where the structure is defined

† Presented at the 200th National Meeting of the American Chemical Society, Washington, DC, August 1990, Symposium on Markush Structure Files and Searching, CINF 32.

0095-2338/91/1631-0036$02.50/0

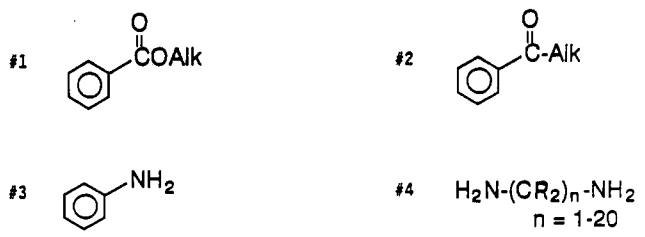

Figure 1. Study cases.

CASE #4 $H_2N-(CR_2)_n-NH_2$ $n = 1-20$

SS1 /IT MULTIAMINE LINK SATURATED CHAIN (9246)

SS2 /IT 1 AND METHANOL AND (VAPOR OR GAS) (19)

Figure 2. APIPAT strategy.

Table I. Results from Fragment Searches

	case			
database	1	2	3	4
API	7809/48	2698/8	10772/26	9246/19
IFI	12279/73	22630/50	56702/157	9403/36
WPI	7908/84	3668/81	10448/32	29110/97

by using connection tables will be referred to as substructure databases.

The study cases will be (1) A new method of preparing aromatic esters, (2) A new method of preparing acetophenone derivatives, (3) A new method of preparing aromatic amines, and (4) A new method of preparing aliphatic diamines. Aromatic refers to an aryl group only, not to aromatic heterocycles. These examples are shown in Figure 1. In these examples all carbon atoms could be substituted where possible. The new method in all of these cases refers to a reaction in the presence of methanol vapor. It is only used to show how process-related keywords will narrow the final search results.

RESULTS

Table I shows the results of doing these searches in the files that contain fragmentation coding. Each result has two numbers. The first number is the result from the fragmentation search. The second number is the answers after the process keywords are added. As one can see, these searches were very easy to do in the fragmentation-based systems and gave reasonable numbers of answers. These numbers show that a specific process will narrow the final search results. An examination of the answers from the structure searches showed that the fragmentation searches resulted in answers that usually contained the compounds in question. The search strategies for case 4 are given as examples for all the strategies used in the fragmentation files. These are shown in Figures 2–4.

When these same examples are run on the substructure-containing systems, very different results are obtained. Because of the complexity of these results, they will be discussed in more detail.

Case 1. A New Method of Preparing Aromatic Esters. In the CAS Registry system, this search resulted in 399 970 iterations, which is more than the 30 000 limit. It is still possible to do this search by ranging on the registry numbers. It would only require 13 range searches which would take about 4 h. If you got more than 10 000 answers, you would have to do ranging again to transfer the answers to the biblio file. The amount of time to finish the CA search was determined by doing one range of 30 000 iterations and determining the amount of time it took. To complete this search in the Registry system would have cost $1300 for the range

CASE #4

SS1 /CN 32757 [PRIMARY AMINE P-2+] (23415)

SS2 /CN 1 LINK NOT (30036 [CARBOCYCLIC] OR 34237 [HETEROCYCLIC] OR 34211 [FUSED]) (9403)

SS3 /CN 2 AND 50438 [METHANOL] (796)

SS4 /CN 65850 [VAPOR] OR 02413 [GASES] OR 07942 [VAPOR PHASE] (64456)

SS5 3 AND 4 (36)

Figure 3. IFI claims/c strategy.

CASE #4

SS1 /MO,M2,M3 H182 [2-ALIPH. AMINO] LINK (M380 [ATTACHED TO N; GENERAL PRE 1981] OR M383 [ATTACHED TO N]) (29110)

SS2 METHANOL OR ALKANOL# OR METHYL(W)ALCOHOL OR ALKYL(W)ALCOHOL# (34953)

SS3 1 AND 2 (806)

SS4 3 AND VAPOR# OR 3 AND GASES OR 3 AND GAS (97)

Figure 4. WPI strategy.

searches and about $180 for the connect time or $1480, not including print or display charges.

In MARPAT this search resulted in 5836 iterations which gave 1115 answers. On transferring the abstract numbers to the CA file and applying further keywords the answer set was reduced to 34 hits.

In the Markush DARC system this search resulted in 3936 answers at the RE level. An attempt to do an AA search was interrupted because there were more than 2000 answers. At this point the RE level answers were transfered to the WPIL File where they produced 2602 answers. After applying further keywords this set was reduced to 77 hits.

Case 2. A New Method of Preparing Acetophenones. In the CAS Registry system, this search resulted in 712 064 iterations. To range search this many iterations would require 24 searches, which would take about 6 h. Again, if you had more than 10 000 answers, you would have to range search them again to transfer them over to the CA File.

In MARPAT, this search resulted in 9280 answers to iterate. This required two range searches, which resulted in 2961 answers. After transferring to the CA File and applying the further keywords, this was reduced to 34 answers.

In the Markush DARC system, the RE search gave 1628 answers, which resulted in 1203 answers after the AA search. Transferring these compound numbers to the WPIL File resulted in 23 hits after applying the keywords.

Case 3. A New Method of Preparing Aromatic Amines. In the CAS Registry system, this search resulted in 1 000 000 answers to iterate. To do range searching on an answer set this large would take about 12 h!

In MARPAT, the search resulted in 8867 answers to iterate. This gave 2819 answers, which were transfered to the CA File. After further keywords, this was reduced to 43 hits.

In the Markush DARC system, the RE search gave 4434 answers. The AA search had more than 2000 answers and was aborted. Transferring the RE set to the WPIL File gave 4239 answers. This was reduced after applying keywords to 8 hits.

Case 4. Preparation of Aliphatic Diamines. In the CAS Registry system, the search resulted in 954 425 answers to iterate. To do range searches on this many iterations would take about 11 h!

In MARPAT, there were 13 406 answers to iterate. The iterations "timed out" twice, so it was necessary to do three range searches. This resulted in 775 answers and took 45 min. After transferring to the CA File and applying the keywords, this was reduced to 20 hits.

```
GRA  C3
NOD  1 N, 3 N
BON  1-2 2-3 SE
NOD  2 G1, REP G1=(1-20)C
HCO  1 E2, 3 E2
END

1              3
 N-----G1-----N
E2             E2

REP G1=(1-20)C
```

Figure 5. CAS Registry/MARPAT search.

Table II. Summary of Substructure Searches[a]

database	case 1	case 2	case 3	case 4
Registry	failed	failed	failed	failed
MARPAT	1115/34	2961/34 2 ranges	2819/43	775/20 3 ranges
Markush DARC	3936/77 RE level	1203/23 AA level	4434/8 RE level	failed

[a] Answers from structure search/answers in bibliographic file.

In the Markush DARC system, this search could not be run online at all. A RE search got the answer "TOO LARGE QUERY - PRECISE IT". A SB search was interrupted because it had more than 10000 answers.

The results of these test cases are shown in Table II.

As you can see from this list of examples, it was more difficult to do these searches in the online substructure files than in the files where the structures were fragmented. None of the examples could easily be run in the CA Registry system.[10] The Markush DARC system also had more difficulties. One search completely failed, and it was necessary in two cases to transfer the results from the RE search. The RE results are a very coarse answer set compared to the AA results. Figures 5 and 6 give as examples the structures that were coded for case 4 for the CA Registry and MARPAR searches and the Markush DARC search. In the Markush DARC search, two structures were tried since the CHK superatom does not include specific alkyl groups.

Both CAS and Questel offer batch searching for very large queries. However in the case of the CA Registry system, batch limits were still not high enough to handle them. In both the Registry system and MARPAT Files this limit is 100000 iterations. For Questel there is no limit to the size of a batch search, but they have a limit of three batch searches/user/day. Both of these producers promise next day delivery on batch searches. This is very reasonable. However some clients and searchers are very impatient and will not want to wait even this long.

At the present time, most of the examples I studied ran in MARPAT and the Markush DARC system, although in the latter I had to transfer the rough RE set, in most cases, to the biblio file. Even though the MARPAT File was able to handle these searches, we need to remember that it only covers Markush structures, and it is necessary to also search the Registry system for a complete search. Both of these files will become larger, and as they do it will become increasingly difficult to search them for simple generic searches. The CA Registry system now contains more than 10000000 compounds. At this time MARPAT covers Markush structures from 1988 and only contains about 30000 structure records.[11] The Markush DARC system dates from 1987 and only covers 85000 structures. This shows that there are a lot less Markush structures to search in the Markush databases compared to the single structures in the Registry system.

```
GR  1-2-3, 4
AT  N 1,4, G1 3
FS  2 2, 1 4
GM  1
GR  1,2,3
AT  H 1, C 2, CHK 3
FS  3 2, 2 3

N-----C-----G1    N (FS1)
      (FS2)

G1 = H, C (FS3), CHK (FS2)

RESULTS:  MORE THAN 10000 ANSWERS IN SB
          TOO LARGE QUERY - PRECISE IT IN RE


GR  1-2-3
AT  N 1,3
AT  G1 2
GM  1
GR  1,2, 3-4, 5-6-7, 8-9-10-11, ETC UP TO C13
AT  CHK 1
FS  2
FS  0 1
AP
No. ATOMS? 1,2,3,5,8,12,17,23,31,39,48,58,69,81
ATT. = ATOM 3 OF GROUP NUMBER 0
No. ATOMS? 1,2,4,7,11,16,22,30,38,47,57,68,80,93

RESULTS:  MORE THAN 10000 ANSWERS IN SB
          TOO LARGE QUERY - PRECISE IT IN RE
```

Figure 6. Markush DARC strategy.

CONCLUSIONS

There are two reasons why some of these examples would not work in the substructure files. The searches exceeded the system limits for iterations or answers or they exceeded the processing time for a query. In the CAS Registry system and MARPAT this is 30000 iterations and 30000 answers online and 100000 iterations and 100000 answers in batch. The processing time limit is 75 min for the Registry system but 15 min for MARPAT.[12] In some of the examples it was necessary to range in MARPAT because the search "timed out".

In the Markush DARC system the limits are 10000 answers in RE or SB and 2000 answers in AA. Questel has since raised the limit to 4000 answers in an AA search. There is no limit to either in batch mode. The Markush DARC system has two different time limits. There is a limit on the time to do an AA search on a single compound and a total time limit for each AA search. These numbers vary depending on the CPU load. One thing that really limits using the Markush DARC system for simple generic searches is that there is no provision for ranging compound numbers. At least on the Registry system and on MARPAT you can search ranges to obtain your answers if you want.

Compared to these substructure files, the files that use a fragmentation system were easily searchable. Why is this? Simply because all of the structural information is contained in the biblio file. This allows you to create answer sets of structure information and your additional terms as large as the online vendor will allow. In Dialog, this is 5000000 total postings. In Orbit, this limit varies depending on the database and the type of search terms, but it is typically about 1500000

postings per session. In STN, there is a limit of 1 000 000 answers/online session with a limit to the number of intermediate postings of 8 000 000. Questel also has a limit of 1 000 000 postings.

Even though searches for simple generic structures can be easily run in the fragmentation-based files, this does not solve all of the problems. The structure information for the Chemical Abstracts File still needs to be searched in a substructure file if you want to cover all possibilities. These problems will increase in the near future to include the WPI File, since Derwent has announced that they will eventually stop fragmentation indexing the chemical sections of their databases in the near future. This will confine fragmentation indexing to the backfile and force users to use the Markush DARC system for structure searching in the current file.

The problems that result from trying to run substructure searches on simple generic structures has left us with only two options when we need to run a search in a file that indexes its structural information using connection tables. An example would be a search for the preparation of a new monomer or for a new process to prepare polyamides. These options are to run the substructure search with all the ranging it requires (if possible) or to use keywords. Substructure searching with extensive ranging is very expensive and time consuming, so one is forced to rely on keywords in the biblio files for sructure searching. Fortunately, we can still do this search in the WPI, IFI, and API Files using fragmentation codes or chemical aspects. And when the fragmentation codes are no longer applied in the WPI Files one will also have to use keyword searching there. This will leave only the IFI and API databases as files where you can still do complete simple generic structure searches. However, since the API File covers only petroleum and petrochemical related patents and the IFI database only U.S. patents, our subject and country coveraged would be seriously effected.

The problem with keyword searching is that you never feel confident that you have a complete search. Chemical information that is contained in abstracts, titles, or index fields is limited. You have to rely on the indexer and the abstracter to decide what is the important chemical information to include. They cannot include all the chemical information that is in a patent as keywords. If they tried to do this, the database producers might as well include the full text of a patent with all the problems that this would bring.

FUTURE POSSIBILITIES

There are a number of possible things that can be done to keep this bleak scenario from happening. One, Derwent does not drop the fragmentation coding. This is not likely to happen because they cannot continue forever to double code; and this does not solve the problem with the CAS files.

Two, increase the iteration and RE limits, the number of answers limit, and the number of answers that can be transfered to the CA File from the Registry system and to the WPI File from the Markush DARC system. This would have to be in the area of a 1 000 000 iterations or RE answers allowed and 100 000 answers total with this number transferrable. This may be possible, but with the present state of computers, none of us would want to wait for such a search to finish.

Three, add the bibliographic information to the structure file and make it searchable. It has to be searchable or the only thing you would gain would be to not have to transfer records. The same system limits on iterations or RE's or answers would still be there. And if you increased the limits you would still have the same problems with time. If the bibliographic material is searchable and the set from such a search can be used as a subset for a structure search, the system limits would not need to be increased as much or maybe not at all. CAS claims they will make the biblio information searchable in MARPAT. Let us hope that they do this before the file becomes too large. This still will not help us search the Registry system unless they add the structure information to the bibliographic file.

Four, allow information from the bibliographic files to be transfered to the structure files and used for subset searching. This would work quite well as long as the limits on transferring data were large enough. In the CA File, you would have to transfer RN numbers, and there are a lot of those per record. Questel has announced that they are working on transferring compound numbers from the WPI Files to the Markush DARC system.

Five, find a way to include generic structure fragmentation information in the bibliographic file. This could be a very simple system like API's Chemical Aspect system. This would be double coding, but there ought to be a way to generate simple fragments from the information in the structure files and automatically apply this to the bibliographic files. Some database producers have said that this would result in vastly overcoded records. As we have seen, this would not be a problem with broad generic searches, and one could use the substructure system for the narrow searches. This would be especially important for the CA File with its large backfile of structure information in that it might be possible to generate generic structure information for the backfile from the substructure records.

Six, another entirely different approach would be to increase the coverage of the reaction-based files (CASReact, CRDS, and others) to the point that they cover the same amount of material as the bibliographic files. These files are excellent for doing the type of searches that have been described, since they are functional group based systems. However, presently, their coverage of chemistry is so incomplete that they cannot be used for chemical searches except to complement the other chemistry files.

SUMMARY

The ability to index and search complex structures has been addressed first by the CAS Registry system and the Generic DARC system, and then enhanced by the addition of Markush structure searching capability with the Markush DARC system and MARPAT. However, in designing these systems the database producers did not appear to consider the need for simple generic structure searching. With the Registry system and MARPAT, we simply did not get this capability, nor did we have it before. With the Markush DARC system, the capability will be taken away. This will leave out a whole class of chemical searching and will deny a large group of searchers the benefits that their subscriptions have provided others.

We can hope that all database producers will look at their systems and institute some method for simple generic searching if one does not exist; or if one exists they will make sure that they retain it if they plan to upgrade their structure searching capability.

ACKNOWLEDGMENT

I thank Robert Buntrock of Amoco for his help in setting up the CAS substructure searches and Michael O'Hara of Questel for his help in constructing the Markush DARC queries.

REFERENCES AND NOTES

(1) See other papers from this symposium published in this issue.
(2) Schoch-Grubler, U. (Sub)structure Searches in Databases Containing Generic Chemical Structure Representations. *Online Rev.* **1990**, *14*, 97–108.
(3) Meurling, A. CAS ONLINE and DARC: A Comparison. *Database* **1990**, *13* (1), 54–63.

(4) Buntrock, R. E. Chemical Database Evolution. *Database* **1987**, *10* (5), 143-144.
(5) Barnard, J. M. Online Graphical Searching of Markush Structures in Patents. *Database* **1987**, *10* (3), 27-34.
(6) Stobaugh, R. E. Chemical Substructure Searching. *J. Chem. Inf. Comput. Sci.* **1985**, *25*, 271-275.
(7) Simmons, E. S. Generic Structures and Chemical Patents. *Chem. Inf. Bull.* **1985**, *37* (1), 27.
(8) Lavagnino, M. R. Computer Access to Chemical Information—The Information Specialist Perspective. *Chem. Inf. Bull.* **1985**, *37* (1), 27.
(9) Kaback, S. M. Online Patent Searching: The Realities. *Online* **1983**, *7* (4), 22-31.
(10) It may be possible in some of these examples to reduce the number of structures to be iterated by the use of selected screens. However the determination of the correct screens to use is not straightforward even to those who should know (i.e., the STN Help Desk) and definitely beyond the end user.
(11) In September 1990, MARPAT contained 66 000 Markush structure records that are found in 22 000 patent citations.
(12) This has now been increased to 30 min for MARPAT.

A Comparison of Three Online Markush Databases†

KATHLEEN A. CLOUTIER

Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, Indiana 46285

Received November 15, 1990

A comparison of the content of three Markush structure databases is presented. The scope of WPIM, Pharmsearch, and MARPAT is examined by looking at the patent documents chosen for indexing for a specific publication date. The Markush structures found in the databases for five patents are also analyzed. The three organizations producing these databases are found to have different policies and approaches in both the choice of the patent documents covered and the scope of the Markush structures indexed for a given document.

The chemical information community has recently been presented with three new databases of generic (Markush) substances from the patent literature. These Markush structure files are being created by three organizations. Derwent Publications Ltd. is building its World Patents Index—Markush (WPIM) File, the French Patent and Trademark Office (INPI) is developing Pharmsearch, and Chemical Abstracts Service is producing MARPAT. WPIM and Pharmsearch are loaded on the Markush DARC software system. MARPAT is available on STN and runs under its own, registry-compatible, software. This paper looks at and compares the content of the three databases of Markush structures. First, the coverage in terms of subject matter, patent countries, and time periods will be briefly explored. The actual coverage of patent documents[1] from a specific date will then be examined. After a review of the indexing policies of the three organizations, the actual indexing of five documents covered in all of the databases will be analyzed.

COVERAGE OF THE DATABASES

Subject Areas. INPI's Pharmsearch is the most focused product, specializing in pharmaceutical patents. Derwent's WPIM covers Pharmaceutical/Veterinary, Agricultural, and General Chemical patent documents. These areas are known to subscribers as classification sections B, C, and E, respectively. The MARPAT File from Chemical Abstracts contains Markush structures of organic or organometallic molecules for patents indexed in *Chemical Abstracts*. The documentation states that "all chemical subject areas are included".[2] Both Derwent and Chemical Abstracts have announced that they are not yet able to index some classes of structures. Currently, WPIM is not handling polymers, oligosaccharides, and polypeptides. MARPAT is not currently indexing alloys, metal oxides, inorganic salts, intermetallics, and polymers. *Chemical Abstracts* Section 74 (Radiation Chemistry, Photochemistry, and Photographic and Other Reprographic Processes) is not completely covered at this time.

† Presented at the 200th National Meeting of the American Chemical Society, Washington, DC, Aug 29, 1990.

Countries. Pharmsearch covers patent documents from the European Patent Office (EP and PCT), France, and the United States. Documents that are equivalent and identical with one already indexed are simply posted as equivalents, but equivalents that are slightly different are reindexed.[3] WPIM indexes Markush structures from the "Basic" patents in Derwent's major countries. MARPAT uses the patents selected for *Chemical Abstracts* except those from the U.S.S.R.

Time Periods. Eventually, Pharmsearch will cover French patents from 1961, and United States and European documents from 1978. The database has been kept current within 6 weeks for European patents and is being extended backwards in time. Data from 1986 is being added now. It has approximately 1500 records from 1960 to 1961 as well. WPIM goes back to Derwent week 8701 which has patents from the latter part of 1986. MARPAT has indexed patent documents dated 1988 forward.

TREATMENT OF PATENT DOCUMENTS DATED MAY 16, 1989

The remainder of this paper examines the patent documents actually covered in the three databases for a given time period. A specific date was chosen as the most manageable period. Derwent has announced to their subscribers that as of indexing week 8920, all pharmaceutical and agricultural patents (sections B and C), excluding the polymers, oligosaccharides, and polypeptides, have been covered in WPIM.[4] The data used for comparison is one from Derwent week 8923 on which U.S. patents appeared. The next section of this paper will examine the coverage of patent documents published on May 16, 1989.

Both Derwent and Chemical Abstracts index only one of the first members of a patent family that they receive. The scope of this analysis must be restricted to documents which are treated as basic in these databases.

Derwent. The WPIL File has 3957 families with at least one patent dated 5/16/89. Of these, 621 are classified in the sections included in Markush DARC (B, C, and E). Three hundred thirty-eight have code M904, indicating that they have a companion structure in the WPIM File. Of these, 124 are the basic (indexed) patent document.

0095-2338/91/1631-0040$02.50/0