

groups which contain large numbers of compounds exhibit a regular pattern. This pattern may be used to predict the approximate number of compounds to be expected in any particular molecular formula group, an essential feature of the registry technique recently described by Lynch *et al.*<sup>1</sup>

#### ACKNOWLEDGMENT

We gratefully acknowledge a grant from the Office for Scientific and Technical Information, London. Part of this work was done by J. H. R. B. as a special study

in partial fulfillment of the Diploma in Librarianship of the University of Sheffield.

#### LITERATURE CITED

- (1) Lynch, M. F., J. Orton, and W. G. Town, *J. Chem. Soc., C*, **1969**, 1732-6.
- (2) Bernays, P. M., Statistical data on chemical compounds, AD-615-488, 1965.
- (3) Leighner, L. H., and D. P. Leiter, Jr. "A Statistical Analysis of the Structure Registry at Chemical Abstracts Service," Division of Chemical Literature, 154th Meeting, ACS, Chicago, Ill., September 1967.

## The GREMAS System, an Integral Part of the IDC System for Chemical Documentation

SIGRID RÖSSLER AND ARTHUR KOLB

Farbenfabriken Bayer AG, Leverkusen, IDC Internationale Dokumentationsgesellschaft für Chemie m.b.H., Frankfurt/Main, Germany

Received January 27, 1970

**The Genealogical Retrieval by Magnetic Tape Storage (GREMAS) system and the potential it offers for searches are described. The input and retrieval procedures of the system are explained as well as the integration of the GREMAS system into the IDC system<sup>1</sup>—i.e., machine generation of the GREMAS coding from topological input and of the superimposed bit code from the GREMAS coding.**

The GREMAS system serves to index low molecular organic compounds and compound classes. Essentially it is a fragment code which transcribes fragments of chemical structures into letter terms. These are then registered by a computer on magnetic tapes. Supplementary to the usual characteristics of the common fragment codes, the GREMAS system has additional features that enhance both the versatility and the hit rate in searching considerably. It was so designed as to give prime importance to a chemist's viewpoint; aspects of programming and machine processing became therefore subordinate. Consequently, the hierarchy of the system mirrors as closely as possible those principles of chemical classification that chemists use in their publications and inquiries. The rejection of too formalized tenets of classification forms one essential prerequisite for loss-free retrieval with minimal false drops.

The code is most selective in areas of chemistry comprising large numbers of compounds and having a high growth rate. But even a small percentage of false drops may become a problem with a file as large as the one which is formed by deep indexing of such a significant segment of the ever growing chemical literature. A still more specific system would then be desirable. By completely recording the topology of molecular structures—i.e., all atom-to-atom bonds—all the structural information of this compound is retained. But searching topological files takes so much computer time that it is economically unfeasible unless a preceding highly efficient and cheap

search reduces the number of compounds to a minimum. Therefore, the IDC system employs a combination of the GREMAS search and a topological search. The GREMAS search precedes the topological retrieval. To reduce cost, the GREMAS search comprises a first screen—the superimposed bit code<sup>2</sup>, a sort of "Abbreviated GREMAS"—which is computer-generated during input from each GREMAS term and is stored in front of the GREMAS file units on the magnetic tape.

The use of a polyhierarchical documentation system as sophisticated as the IDC system still remains expensive, but it is needed for selective retrieval. Although it is possible to look up specific compounds and parent structures in conventional card files and indexes, multidimensional processing is required for substructure searching. This can be handled adequately only by a computer when a large amount of literature has been stored. The GREMAS system allows selective substructure retrieval from a file containing currently about 900,000 compounds, including ones with alternative groups (Markush formulas). Moreover, reactions and types of reaction are searchable by themselves as well as in combination with inorganic reactants, catalysts, and nonstructural concepts.

#### THE GREMAS SYSTEM

The GREMAS system is a computerized storage and retrieval system for low molecular organic structures. It was developed by Farbwerke Hoechst AG<sup>3</sup> starting in

1957, and increasingly applied there during the following years. From 1962 on, Badische Anilin & Sodafabrik AG and Farbenfabriken Bayer AG participated in both its application and further development.

In 1967 it was given to IDC, and it is now an integrated part of the IDC system. The GREMAS system—like other systems which encode molecular structures—comprises a vocabulary (three-character terms) and additionally a syntax (Y- and Z-terms) to indicate how the individual structural entities of the compound had been linked.

Each carbon atom and the hetero atoms directly linked to it or substituents are described by a three-character term. If these terms are considered as building blocks, then the Y- and Z-terms are the corresponding blueprints. They specify how the individual fragment of the structural formula must be put together.

**The Three-Character Terms.** The smallest building blocks of a molecule are the atoms, primarily carbon atoms in organic chemistry. The obvious idea in designing the system was to use these building blocks as well as the various linkages a carbon atom can form. The first step is to differentiate between carbon atoms bonded exclusively to carbon and/or hydrogen and carbon atoms bonded to other atoms, so-called hetero atoms as well. In the first case, when only C and/or H are linked to the carbon atom to be described, it is defined as having zero degree of heteroorientation. In the second case, hetero atoms are linked to the C-atom to be described—four types of heterooriented C-atoms are defined: single heterooriented C-atoms, one bond between the C-atom and one hetero atom, double heterooriented A-atoms, two bonds between the C-atom and one or two hetero atoms, etc. Carbon atoms having the same degree of heteroorientation are further divided according to the type of hetero atom. The result is a classification into logical groups (aldehydes, ketones, carboxylic acid, etc.).

Table I shows the groups of single heterooriented

C-atoms. The respective C-atoms are specified by three-character terms the first letter of which is called Genus and is here A to H. The Genus thus characterizes a chemical compound class—e.g., alcohols and derivatives.

Analogously divided are double heterooriented C-atoms (ketones, aldehydes—Genera I to M), triple ones (carboxylic acid and derivatives—Genera N and O), quadruple ones (carbonic acid and derivatives—Genera P and Q) and zero heterooriented C-atoms—i.e., those not linked to a hetero atom (Genus R) (Table II).

The species, the second letter of the three-character term, distinguishes between the derivatives of the compound class described by the Genus—e.g., between primary, secondary, and tertiary amines (Table III).

The third letter of the three-character term, the sub-species, finally specifies whether the C-atom described—e.g., the amine group—is attached to or part of an aliphatic chain, an alicyclic, aromatic, or heterocyclic ring (Table IV).

Rings are not fragmented into atoms but are indexed as units with three-character terms of their own, the first letter of which is S, T, or U—S for all monocyclic rings, T and U for fused ones. Fused ring systems are either characterized by the individual rings, the terms of which are obtained by substituting S by T, or as an entity with U-terms. These are available for some frequently occurring fused ring systems to avoid reconstructing them again and again.

Terms of Genus V specify types and points of fusion; those of Genus W, the length as well as the relative position of substituents of aliphatic chains (e.g.,  $\beta$ -diketones) and points of substitution of heterocyclic rings.

Genus X describes the type and number of bonds between hetero atoms—e.g., nonmetal-nonmetal bonds and metal-nonmetal bonds—as well as charges and radical state of hetero atoms (Table V).

**The Syntactic Terms Y and Z** describe sections of mole-

Table I.

Heteroorientation	Genus	Examples
	X $\neq$ N,O,S,Hal	A-- CH <sub>3</sub> -Li,
	X = N	B-- CH <sub>3</sub> -CH <sub>2</sub> -NH <sub>2</sub>
		C-- CH <sub>3</sub> -NO <sub>2</sub>
		D--
	X = O	E--  CH <sub>3</sub> -CH <sub>2</sub> -O-CH <sub>2</sub> -CH <sub>3</sub>
	X = S	F--  CH <sub>3</sub> -CH <sub>2</sub> -S-CH <sub>2</sub> -CH <sub>3</sub>
		G--
X = Hal	H--	CH <sub>3</sub> -J,

Table II.

	GENUS	EXAMPLES
$\text{II}$ $\text{C}$	I--	Ketones and derivatives
	K--	Thioketones and derivatives
	L--	Aldehydes and derivatives
	M--	Thioaldehydes and derivatives
$\text{III}$ $\text{C}$	N--	Carboxylic acids and derivatives
	$\emptyset$ --	Thiocarboxylic acids and derivatives
$\text{IV}$ $\text{C}$	P--	Carbonic acid and derivatives
	Q--	Thiocarbonic acid and derivatives
$\circ$ $\text{C}$	R--	Alkanes, Alkenes, Alkines

Table III.





EXAMPLES FOR SPECIES		
GENUS B		
BA-	prim. amines	e.g. $\text{CH}_3\text{-NH}_2$
BB-	sec. amines	e.g. $(\text{CH}_3)_2\text{-NH}$
BC-	tert. amines	e.g. $(\text{CH}_3)_3\text{-N}$
BD-	quatern. amines	e.g. $(\text{CH}_3)_4\text{-N}^+$
GENUS H		
HA-	fluorine cpds.	e.g. 
HB-	chlorine cpds.	e.g. 
HC-	bromine cpds.	e.g. 
HD-	iodine cpds.	e.g. 
GENUS R		
RA-	alkanes	e.g. $\text{CH}_3\text{-CH}_3$
RB-	alkenes	e.g. $\text{CH}_2\text{=CH}_2$
RC-	alkines	e.g. $\text{CH}\equiv\text{CH}$

Table IV.

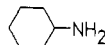
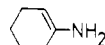
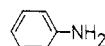
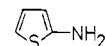
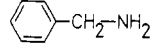
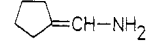





Examples for Subspecies of primary amines	
$\text{-CH}_2\text{-CH}_2\text{-NH}_2$	BAA
$\text{-CH=CH-NH}_2$	BAE
$\text{-C}\equiv\text{C-NH}_2$	BAM
	BAQ
	BAT
	BAR
	BAS
	BAD
	BAH

Table V.

Genera:	
A-- to R--	individual Carbon-atoms
S--	monocyclic rings, e.g.  SKC
T--	fused rings, e.g.  TCC + TKC
U--	often occurring ring-systems, e.g.
	 UDA  UBO  UBI
V--	points of fusion
W--	points of substitution, chain length
X--	heteroatom - heteroatom bond

cules. The GREMAS system differentiates between four types of sections

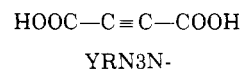
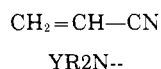
Aliphatic carbon chains (including branched ones)	YR----
Alicyclic rings	YS----
Aromatic rings	YT----
Heterocyclic rings	YU----

The terms for the respective sections are compiled by adding all of the Genus symbols of single to triple heterooriented C-atoms, attached rings, and points of fusion as often as they occur in this section.



Genera:	L.. I.. N..	L.. I.. N..
Section term:	YRLIN-	YRLIIN

Aliphatic double or triple bonds are specified by 2 or 3.



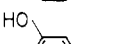
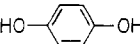
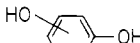
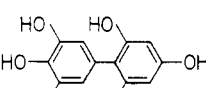


The relative position of two benzene substituents is encoded by an additional number in the YT-term (Table VI). The ZT-terms (Table VI) describe each C-atom of aromatic rings (monocyclic or fused) having at least three substituents. The notation of the ZT-term combines the Genus symbols including V (points of fusion) and O (for H as substituent) in consecutive order regardless of starting point and direction. The relative position of all substituents is thus exactly reproduced.

The Y- and Z-terms may get "tags." One type specifies how often a certain section is present in the molecule. For instance, the term YREE--.0812- assigned to an ethoxylated compound indicates that 8 to 12 moles of ethylene oxide have reacted.

Other types state the presence of delocalized or complex

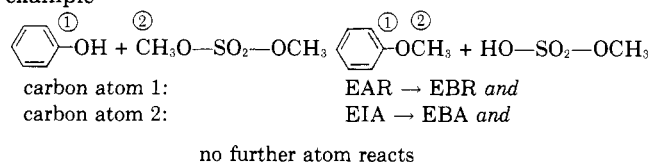
Table VI.

	„three character terms“	YT- or ZT-terms
	SAF, EAR	YTE---
	SAF, EAR	YTEE2-
	SAF, EAR	YTEE3-
	SAF, EAR	YTEE4-
	SAF, EAR	YTEE1-
	SAF, EAR	ZTEEE0T0 + ZTETE0E0

bonds and the alternatives of sections. Owing to the last mentioned option, even very generic Markush formulas of patent claims are represented very accurately.

#### ENCODING OF REACTIONS

Reactions are indexed about as precisely as structures with the GREMAS system. The reacting C-atoms of the starting and end product are characterized by their respective three-character terms as demonstrated in the following example



Clearly defined rules specify whether an atom reacts or not. Reactions indexed by this "detailed reaction code" are stored on a separate tape. The structural file units of the end product contain additional "short reaction terms" consisting of specific three-character terms for each reacting atom, the first character of which is 2 or 3, and the second and third one corresponds to the Genus and Species of the C-atom in question. In our example, the short reaction term is: 2EB.

#### NONSTRUCTURAL INFORMATION

Some of the nonstructural concepts are stored together with the structural terms on the master tape. These concepts are encoded as defined three-character terms and comprise: properties, pharmaceutical as well as agricultural applications, processing data, general chemical aspects, and spectra.

Others are stored alphanumerically on a supplementary tape (E-tape). These concepts are registered unstandardized as given by the analyst (free word selection). Work on establishing a comprehensive polyhierarchical thesaurus is in progress. After its completion, the computer will be used to combine synonyms and assign a standardized notation.

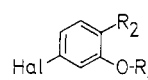
#### THE GREMAS SEARCH SYSTEM

A reliable criterion for selectivity in searching is how unambiguously a formula can be reconstructed from its encoding. Whenever a representation of a compound is complete and unique, a search for this compound will give only relevant answers and not false drops. To a high degree, this is the case with the GREMAS system. Normally, the number of false drops is insignificant and can be screened out by the inquirer without too much effort. Should the number of false drops at times be substantial, a subsequent topological search is possible if the input was done by this method.

The GREMAS search system can answer all the queries that can be handled by conventional files, such as card and punched card files. But mechanized retrieval is best applied for answering generic queries for compound classes or substructures. It is also quite popular to use mechanized retrieval of specific compounds for current awareness services. With each updating of the files the inquirer will automatically receive the literature pertaining to the compounds of his interest.

Its Potential Searches are:

Specific compounds, substructures, and compound classes—e.g., halosubstituted acid derivatives.



$R_1 = \text{H, alkyl, aryl, acyl}$   
 $R_2 = \text{CN, CONH}_2, \text{CO.hal, COOR}_3$   
 $R_3 = R_1$   
 Hal = halogen

#### Nonstructural concepts:

Compound is starting material of a reaction  
 Compound is product of a reaction  
 Compound is separated, purified, removed  
 Compound is a component of a specified mixture  
 Spectrum of the compound is given  
 Compound is a catalyst  
 Compound has (specified) pharmacological activity  
 Compound has (specified) pesticidal activity  
 Compound is a surfactant  
 Compound is used as a dye  
 Compound is used as a textile agent  
 Compound is used as a lubricant  
 Compound is used as a fuel additive  
 Compound is used in laquers, varnishes, paints, coating, etc.

#### Reactions

The options A to C may be combined in any way whereby a high selectivity is attained. Due to the principles of the GREMAS system only those documents will be retrieved in which the asked for compound or a more specific one is described. A publication quoting a more generic compound, which comprises the asked for one, will not be recalled. Should an inquirer want to get these publications too, the search has to be set up as two searches, a specific and a generic one. A query for ethylbenzene (specific query), for instance, will not be answered by alkylbenzenes. Their retrieval would necessitate searching generically for alkylbenzenes, whereby the "alkyl" may be limited to "lower alkyl" to reduce the number of answers. Besides these single searches for preparation, use, etc., of one chemical structure there is another important type of search, called "coordinate search," in which several single searches are intersected, thereby requiring the presence of 2...16 compounds or compound classes in one document.

A search for the preparation of terephthalic acid from *p*-xylene would require an intersection of

*p*-xylene as reactant, and  
terephthalic acid as final product,  
i.e., both compounds have to be  
described in the same document (e.g., patent)

An intersection of compounds naturally presupposes the registration of starting materials as well as intermediates as is being done at IDC. The coordinate search for reactants and products is a kind of reaction search. The same results are obtained by correlating a search of the master tape for "terephthalic acid as product" with the supplementary tape (E-tape) for "oxidation of a CH<sub>3</sub>-group on an aromatic ring to a COOH-group." Why this correlation? Because the reaction encoding of this example would specify only the conversion of one CH<sub>3</sub>-group on an aromatic ring to one COOH-group, but say nothing about the aromatic ring which could be benzene, naphthalene, or a higher fused system. The search of just the reaction file would therefore recall all of the documents reporting this oxidation. Only by correlating this search of the supplementary tape with that of the master tape, which adds the restrictions of the ring system and further substituents, false drops will be excluded.

#### INQUIRY CODING

Inquiries are translated into system language—namely, three-character terms and/or Y- and/or Z-terms. The inquiry program provides for the following logical operations: union, intersection, negation, and combinations thereof. Although Markush formulas as such—i.e., without dissolving them into specific compounds—are being registered, it is possible to apply negations since the terms specifying alternative groups are stored in a file unit part which does not answer negations. There is no risk that a compound asked for will not be recalled just because it happens to be combined as a Markush formula with a negated one.

#### SEARCHING WITH THREE-CHARACTER TERMS

The faceted hierarchy of these terms makes it possible to apply different aspects to generic searching. Primary, aliphatic amines are, for instance, not just retrieved by asking for "BAA" (aliphatic, primary amines) but also by asking for BA\* (primary amines), for B\*A (aliph. amines, primary, secondary, and tertiary ones as well as quaternary ammonium compounds) and for B\*\* A-B A-E. The notation B\*\* A-B A-E recalls all terms, the first character of which is B, the second one is A or B, and the third one is A, B, C, D, or E (primary or secondary alicyclic, including those where a ring—alicyclic, aromatic, or heterocyclic—may be attached to the carbon bearing the amino group) (see also Table IV).

#### SEARCHING WITH Y-TERMS

Y-terms cannot only be requested or negated as entities, but the presence or absence of genera within one Y-term may be specified, even the exact number or minimum number of the Genus symbols. It is, for instance, possible to require the presence of "at least three" or "three and only three" oxy-groups on one aromatic ring and simultaneously negate amines (on the same ring). Unions of Y-terms may be formed as of three-character terms and

the presence of tags may be required. For example, a search for ethoxylated compounds with at least 12 (—CH<sub>2</sub>—CH<sub>2</sub>—O—) units will retrieve only those file units in which the ethoxy group is present at least 12 times or in a range comprising 12, such as 8 to 16.

#### SEARCHING WITH ZT-TERMS

The exact positions on an aromatic ring having at least three substituents or points of fusion are searchable with the help of ZT-terms. Starting point and direction of the consecutive specification of all Genera may be different in the ZT-term representing an asked-for compound than in the filed one, the computer will establish equivalence anyhow.

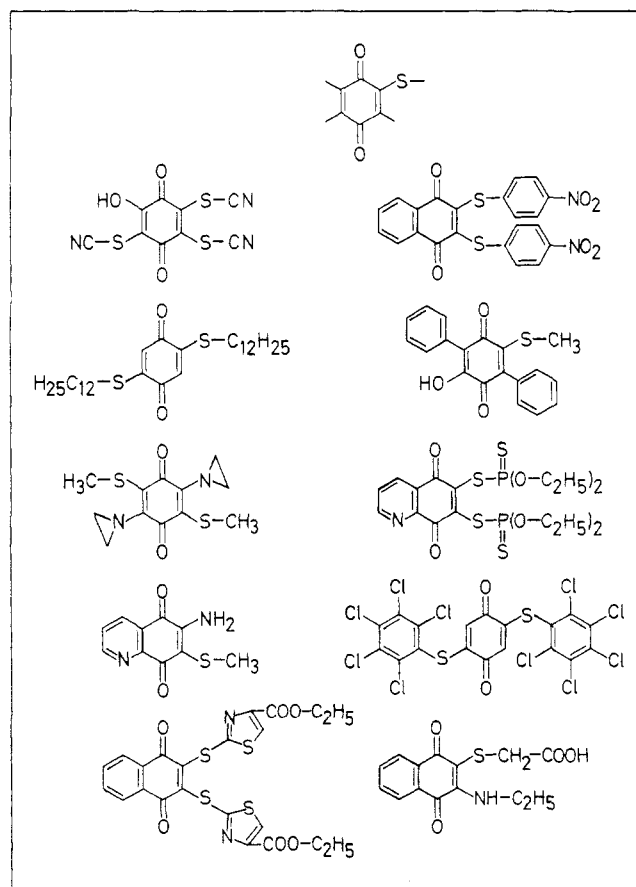
Here again we have the option to ask for a union of Genera at one point of substitution, negate a Genus at another point, and leave everything else open.

Since it is impossible to present in this paper all the combinations the GREMAS search system offers, the following three examples may help to demonstrate its versatility.

The first example (Table VII) is an inquiry for *p*-quinones substituted with at least one mercapto functional group without further restrictions as to other substituents and the ring system comprising the quinone. The search coding would read:

F\*T = mercaptans or derivatives thereof, (F..) at a C-atom  
which is part of an alicyclic ring and is double-

Table VII.



bonded to another carbon (. . T); \* indicates that the second character is not specified and may be any letter.

IIV = ketone group of *p*-quinones.

YSFII # = an alicyclic ring (YS) is substituted by at least one mercapto group (F) and by two and only two ketone groups (II #). Without the extra character, #, structures with more than two ketone groups would be recalled as well (which, of course, is not possible for our example); it provides thus an additional specification.

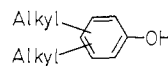
The answers to this search, all of which are shown in Table VII, are pertinent. This had to be the case because the asked for substructure can unambiguously be reconstructed from its GREMAS coding.

In the second example (Table VIII), we have an inquiry for semicarbazones of aldehydes or ketones in which the terminal N-atom is at least once substituted with alkyl or aryl. The encoded search comprises two independent unions of terms. In translating an inquiry into the GREMAS search coding, all intersected and negated—e.g., constant—terms are written down first in alphabetical order followed by unions of terms—e.g., alternative ones—which are being introduced by the symbol, /

PFA = semicarbazide and (intersect)  
/CFA = alkyl or (union)  
CFR = aryl and  
/IE\* = semicarbazones of ketones or  
LD\* = semicarbazones of aldehydes

The symbol, /, in front of a concept signifies a union of all concepts up to the next /. A query, as in the following example, restricted to dialkylphenols requires the negation of other substituents to avoid retrieval of a multitude of compounds which would have additional characteristics to the ones asked for (benzene, one hydroxy,

and two alkyl groups). The symbol for negation is <. The search for

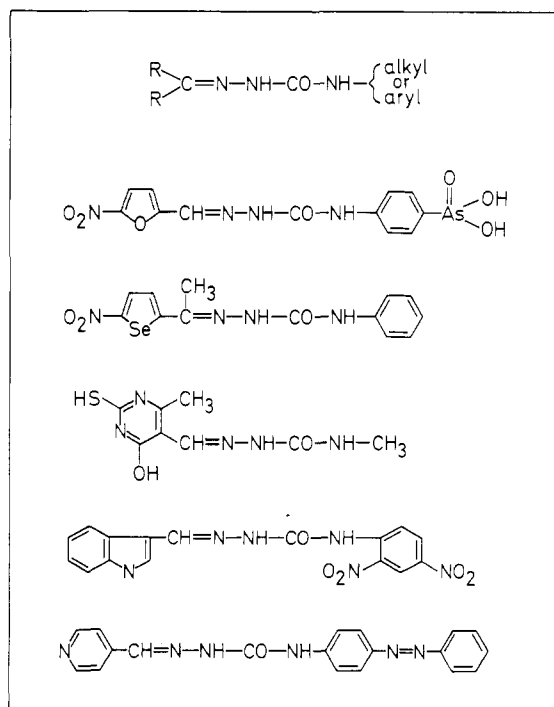


is then encoded as:

A<D = Genera A to D are negated, i.e., a three-character term, the first letter of which is A, B, C, or D must not be present  
E<< = all three-character terms of Genus E are negated except the one the second letter of which is A and the third is R—namely, EAR, the descriptor for phenols—which must be present (intersect).  
F<P = Genera F to P are negated  
RA\* = alkyl on an aromatic ring. The C-atom attached to the ring may be linked to 0 to 3 H and/or C-atoms.  
K-T  
S<< = all three-character terms of Genus S (monocyclic rings) are negated except SAF which is required (combination of a negation and an intersection of SAF).  
F-F  
T<V = Genera T-V are negated and  
YTE#RR# = term of an aromatic section (YT) comprising Genus E once and only once (E#) and Genus R twice and only twice (RR#)

It was not considered necessary to repeat the negation of Genera within the YT-term; that had already been achieved with the negation of the corresponding three-character terms. The three examples should suffice to demonstrate the efficiency of the GREMAS system for structure and especially substructure searching.

Table VIII.



## REACTION SEARCHES

The GREMAS system offers several options for retrieval of reactions. The condensation of heterocyclic aldehydes with aliphatic nitro compounds, for instance, can be searched by the following methods:

By combining the terms for the structural search of the product III with the respective short reaction terms. This search will give quite selective answers as shown in Table IX.

The notation for this search is:

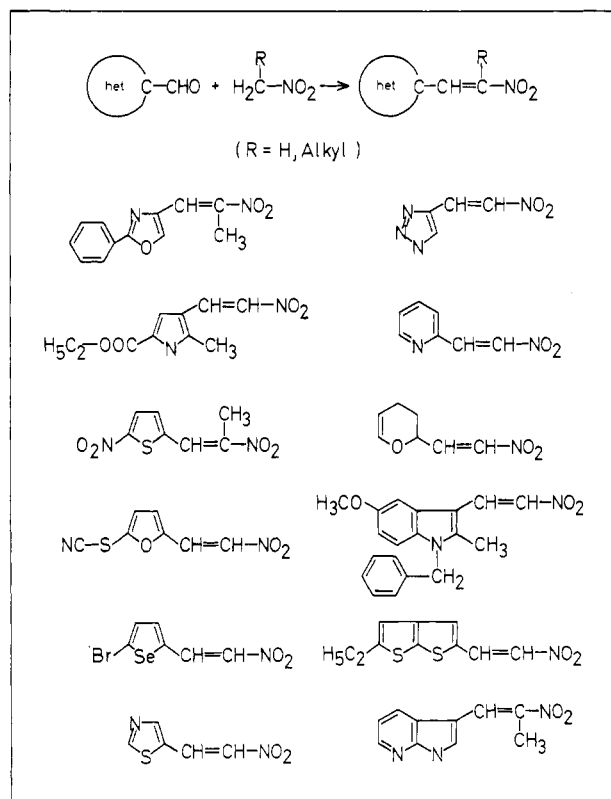
CMF = NO<sub>2</sub> on an olefinic C-atom,  
RBQ = olefinic C-atom with one H as substituent of a heterocyclic ring and  
YRCU2 = aliphatic section substituted with NO<sub>2</sub>, a heterocyclic ring, and containing a C=C double bond.

In addition to the structural terms are the following short reaction terms

2CM = a NO<sub>2</sub>-substituted C-atom has reacted  
2RB = an olefinic C-atom has been formed  
3ZL = a C—C bond has been formed

By a coordinate search, in which the simultaneous presence of several specific compounds or substructures in one document (e.g., a patent) is required. In our example, the compounds could be intersected as follows:

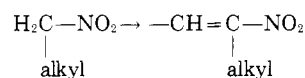
Table IX.



I and II and III or  
 I and III or  
 I and II or  
 II and III

The third option (I and II) is preferred whenever no information is given or can be deduced from the product. Additionally, each compound or substructure in a coordinate search may be further specified—e.g., compounds I and II are reactants, compound III is the product of a reaction.

By searching the detailed reaction description on the E-tape for:



whereby a  $-\text{C}=\text{C}-$  bond is formed.

## LITERATURE INPUT AND OUTPUT

Since 1959, abstracts of journal and patent literature have been indexed—e.g., ones appearing in "Fortschrittsbericht" (published by Farbenfabriken Bayer AG), in "Patentbericht" (published by Farbwerke Hoechst AG until 1962), and in "Patentschnellbericht."

The latter is published weekly by Badische Anilin- u. Sodafabrik AG, Farbenfabriken Bayer AG, and Farbwerke Hoechst AG and contains abstracts made by the Patent Documentation Group (PDG). With the formation of IDC in February 1967, IDC started indexing on a more complete base. Besides the above mentioned abstracts, IDC encoded the abstracts of "Schnellreferatedienst" of Chemisches Zentralblatt (SRD) with the help of complete indexing records, especially prepared for IDC by the abstractor of SRD and containing more information than the published abstracts. This deep analysis and indexing are aimed at for patents too. Starting in 1970, "Fortschrittsbericht" and SRD are continued by a joint publication of Farbenfabriken Bayer AG and Gesellschaft Deutscher Chemiker entitled "Chemischer Informationsdienst," which is also encoded.

IDC is currently encoding about 35,000 abstracts of journals and patents a year.

The magnetic tape file contains 900,000 file units. One file unit represents a specific compound or a Markush formula.

Initial search answers are computer printouts of accession numbers of relevant abstracts sorted according to year and source. The abstracts, which are on 16 mm film, are copied and forwarded to the inquirer. To improve economics several searches should be processed simultaneously since the increase in machine time is insignificant by rising numbers of searches. The search programs are so designed that up to 125 searches can be processed simultaneously. On an IBM/360-40 computer, it takes about 60 minutes for an average of 20 to 25 questions to search the entire file, sorting and printing time included.

On an IBM/360-65 (512K) OS/MVT, and an average of 14 questions processed simultaneously, job times of 84.3 seconds and CPU times of 18.8 seconds per question are required.

## LITERATURE CITED

- (1) Meyer, E., "The IDC System for Chemical Documentation," *J. CHEM. Doc.* 9, 109 (1969).
- (2) Meyer, E., in "Mechanized Information Storage, Retrieval and Dissemination," Proceedings of the EID/IFIP Joint Conference, Rome, June 1967, pp. 280-8, North-Holland Publishing Co., Amsterdam, 1968.
- (3) Fugmann, R., W. Braun, and W. Vaupel, *Nachr. Dok.* 14, 179-90 (1963).