sets of Termatrex cards are not distributed to the libraries, because it is totally impractical to attempt to keep several sets of cards up-to-date. The ability of any library to make a rapid response is provided by a priority telephone link between each library and the Corporate Technical Information Center. Several copies of the abstracts have been previously distributed and are on file in each library. Each abstract is assigned a serial number identical with the accession number of each document. The product of a search made with the optical coincidence cards is a list of the accession numbers of pertinent documents. This list is communicated by telephone to the library originating the inquiry. That library, in turn, hands the engineer a copy of each pertinent abstract, from which he selects the document he wishes to review in the original form. Problems of the acquisition of classified documents, based upon a need-to-know and security clearance, rest with the "needing" plant.

COST PER SEARCH

The average retrieval time for abstracts in response to an inquiry is 3 min., with a maximum time of 10 min., and the documents retrieved reflect better than 95% relevancy. Recently, 300 routine searches were made in a period of 1½ days; pertinent abstracts were retrieved from the file during the same period. The cost per search, including retrieval of the abstract, was approximately 11 cents. This cost does not include an allowance for the salary of the abstracters nor related overhead expenses.

The Use of Subordinate and Coordinate Indexes *vs.* the Scanning of Their Outputs*

FRED R. WHALEY
Information Systems Research Group, Battelle Memorial Institute, Washington, D. C.
Received July 6, 1964

INTRODUCTION

Users of information services, as well as the staff supplying such services, frequently scan lists of titles or abstracts or finger through cards in a card file. They may be studying lists or card files organized on a subject basis, such as a library card catalog, or lists or card files not so organized, such as those supplied by a current-awareness service. If we omit that part of their activity which is random browsing, we are left with activity that has something to do with an index. Part of such activity is actually using an index and part is studying the output of an index to decide which, if any, original documents should be procured.

It is submitted here that a better understanding of this over-all activity is possible if a distinction is made between index using and output scanning, so that each is seen in its relation to the other and to the interaction between them. This will result in system planning which is closer to system needs.

The following discussion is limited to nonnumeric systems or document retrieval, as opposed to data retrieval.

Definitions and Ground Rules.—Indexes operate in an environment consisting of:

(A) A population of items (for our purposes the file of documents or articles under consideration).

(B) A real or implied charter[1] indicating the purpose of the index and the point of view of the clientele to be served.

(C) An inquiry or interest profile, consistent with B, submitted by a client.

(D) A subpopulation of items, being a part of A, which contains information on the inquiry (C). This subpopulation may be zero even with a reasonable question.

When the index operates in the above environment, a new environmental element is created:

(E) A subpopulation of items, being a part of A, and being the output of the index when triggered by inquiry (C). Within the constraints imposed by cost, human error, linguistics, etc., E represents the index's best approximation to D.

A working definition of an index is now possible.

Within the environment as outlined above, an index is a device, which when triggered by a reasonable inquiry, points to a subpopulation of items (a small proportion of the total population of items in the collection) and indicates that most of the items in the collection dealing with this inquiry will be found in this subpopulation.

The subsequent step of examining this subpopulation for relevancy, whether done by the inquirer or by the staff of the information service or both, is held to be a post-index operation and not a part of index using.

Index Output, the Display Continuum.—Only in a hypothetical, ideal index is the output (E) for a given question identical with the actual answers (D). Real-life indexes are usually flexible enough so that inquiries can be negotiated into logical form so as either to minimize lost material with an increase in irrelevant material, as for a researcher with a "leave-no-stone-unturned" approach, or to minimize irrelevant material with some loss in wanted material, as for a practitioner wanting a good answer and quickly.

* Presented before the Division of Chemical Literature, 147th National Meeting of the American Chemical Society, Philadelphia, Pa., April 8, 1964.

The output (E) of an index comprises, as a minimum, a list of numbers identifying the articles in the subpopulation. Greater sophistication than this minimum is usually available in the output of the index, which is herein called the display, as exemplified in the following ascending order, starting with zero display:

(a) Identifying or accession numbers of items retrieved
(b) Same as a plus authors and titles
(c) Same as b plus key words used for indexing
(d) Same as b or c plus indicative abstract
(e) Same as b or c plus informative abstract or extract
(f) Same as d or e plus complete article

Examples of intermediate points are easily brought to mind, so that one can say that the display from an index falls on a point in a continuum from low to high sophistication. The user may have to proceed from the immediate output of the index to a secondary listing by accession number to get the maximum display of the system. This is true of most systems where document numbers are listed under terms, such as dual dictionaries and optical coincidence systems. In these cases a second listing in order by accession number gives a higher level display of the same entries obtained originally as mere accession numbers. We will consider the display sophistication as that of the system, rather than that of the early steps in the system.

Going from original index output to another file giving more sophisticated display is not an index-use step, since it does not point to a subpopulation of items relative to the population of items just obtained from the index. This may be regarded as a step in the display perusal.

What Is Done with the Output of an Index?—What does the inquirer do with the output of an index in answer to an inquiry? He examines the display elements to determine which original items, if any, should be given detailed study and which can be safely ignored. A given level of display sophistication may comprise several consecutive steps, such as e shown above. Here we have beyond the accession number itself:

Step 1. Authors and titles
Step 2. All key words used in indexing
Step 3. Informative abstract

A person scanning a group of such display elements to decide which items to look up may decide to discard a given item from further consideration based on author and/or title, or he may be encouraged by what he sees at this display step to look at the key word display. (At least one key word was the point of entry to the index leading to this index-output, but the remaining key words are useful in making his decision.) The presence of some key words or the absence of others may lead him to discard an item at this step, or he may be encouraged to read the abstract. If he is still interested after making full use of the display at his disposal, he will now seek the original acticle, unless the informative abstract gives him all the information he needs. Obviously, steps 1 and 2 may be reversed, the easier or more applicable one being done first.

Battelle information centers[2] are examples of the above display sophistication level, if we use "clue words" in step 2, and "extract" in step 3.

It will be noted that the examination of display involves inspection of the complete subpopulation (E), unless the inquirer gives up short of this.** With respect to the subpopulation, it is not a selective or "look-up" search, because every item has to be examined; although it may be an organized search from the standpoint of looking first for authors, titles, or interesting key words before examining the rest of the display.

Again, let us emphasize that in using an index the point at which a population of items has to be serially examined for pertinency marks the end of the actual index or look-up steps and the beginning of a visual scan of display.

Subordinate Indexes and Coordinate Indexes.—To illustrate the points made above, let us refer to the *Chemical Abstracts* Index of 1961 (July–December) to look for items on *aluminum corrosion.* We find *aluminum* in the main alphabetical listing and in a subordinated alphabetical listing (indented and under *aluminum*) we find a group of adjacent and successive entries under *corrosion* as illustrated in Figure 1. All the articles under this subordinated

July – Dec. 1961 — Subject Index

corrosion (grain-boundary) of, in high-temp. water, oxide coating and H diffusion and overvoltage in relation to, 14231*c*.
corrosion inhibition of combustion gases by mixt. contg. powd., P 18546*g*.
corrosion of, 19716*i*.
in acids, temp. and. 18526*a*.
by alk. paste or soln., inhibition by 2-(2-carboxylphenyl)-azo-4-methylphenol, P 13284*b*.

Figure 1.

heading constitute the index output for this inquiry. The scanning of these displays would now proceed, involving almost one full column. If there had been no subordinated alphabetical listing, all of the display entries under *aluminum* would have to be examined (12 columns or 4 pages) because then the index would have been too shallow to have allowed the conjunction of *aluminum* and *corrosion*. Actually the index allows another conjunction by another subordinate indentation. For example, if the question is on *corrosion of aluminum by superheated steam*, under a further subordinated array is *superheated steam*.

A library card catalog, such as the one at the Battelle Library, offers the same opportunity for conjunction by subordination, in that under the principal term in the main alphabetical listing, *aluminum alloys*, is found a subordinated alphabetical listing, and conjunction can be implemented by a second alphabetical look-up in this ordered list for *corrosion*. The number of cards whose display is scanned as a result of this conjunction by subordination is much less than the number scanned if there were no subordinated alphabetical order.

Conjunction by subordination could be equally well implemented where the subject matter is organized as a classification scheme as illustrated in the Library of Congress Classification Schedule. Let us change our example to *corrosion of iron or steel*, since aluminum is

** If part of this display perusal is done by the staff of a central documentation center before given the predigested output to an inquirer, it is still a separate step from index-use, regardless of how the labor of display scanning is divided between central staff and inquirer.

not handled in the same detail. The subordinate headings are

T		Technology
A		Materials of engineering and construction
	459 (and on)	Metals
	464 (and on)	Iron and steel
	467	Corrosion

The complete designation would be TA467 for books on corrosion of iron or steel. Thus, if the schedule were used as an index, corrosion of iron or steel would be found by the subordinate conjunction of *iron and steel* with *corrosion*, after iron and steel were located in the semi-taxonomic relation of the schedule.

It is not claimed that pure taxonomic subordination, such as *iron and steel* being in the class, *metals*, is a case of subordinate conjunction. Classification schemes lead to subordinate conjunction only so far as the subordinated concept is not totally contained in the higher level concept. In the case of the subordination of *corrosion* to *iron and steel*, corrosion is not fully contained in (limited to) iron and steel, and, hence, the unit concepts are in conjunction. In the case of the subordination of *iron and steel* to *metals*, however, the complete taxonomic inclusion of the former in the latter precludes conjunction in the sense used in information retrieval.

In the above examples more or less complex concepts are analyzed by indexers into component concepts, so that the most important component is in the main organizational array, be it a listing or a card file. The next most important concept is in a subordinate (usually alphabetized) array that begins and ends under a particular heading in the primary array. Under a particular heading in this subordinated array may be another subordinated array. The user approaches the index with an inquiry which he also analyzes into component concepts. He looks in the primary array for what he judges to be the most likely term for the primary component concept, then under it in a subordinated array the next most important concept, and so on. When he reaches the maximum depth of the index, he starts the serial scanning of the display bracketed by his selective search. He may repeat these steps using other judgments as to the primary component concepts to use.

In recent years another means of attaining conjunction of unit concepts to build more complex concepts has grown in favor—the coordinate index. In these indexes the complex concepts are analyzed into component or unit concepts, which are stored on an equal basis. No guess work as to their order of being entered is needed because conjunction is by concept coordination.

Coordinate is defined[†] as "equal in rank or order, not subordinate," and hence is a good name for this type of index. "Equal in rank" means that either of two concepts in conjunction may be looked up first or both simultaneously and does not imply that the meanings of the terms are of equal importance as seen subjectively by an author or indexer. Refinements distinguishing between the relative weight of concepts, if deemed desirable, can be supplied by weighing factors, but have no bearing on the point that distinguishes a coordinate index; namely, the order of looking up concepts in conjunction is immaterial, or at most affects only the speed of retrieval, as opposed to a subordinate index, where order of looking up concepts may mean the difference between retrieval and no retrieval of the pertinent documents. This is because building subordinate indexes requires permuting of term orders (cross references) and such permuting is seldom as complete as might be desired.

We urge that coordinate indexes continue to be so called, in spite of efforts to find other names, because of the excellent way this name distinguishes them from their historical opposite, subordinate indexes. It is important, however, that the meaning of "coordinate index" is not narrowly interpreted as only one of the many subtypes (such as a terminal digit index) that rightfully belong in this class.

In grammar, a conjunction is "a connective having the special function of joining together sentences, clauses, phrases, or words and classifiable as *co-ordinating* or *subordinating*."[†] Conjunction in our sense is the intellectual joining of relatively simple concepts such as *aluminum* and *corrosion* to obtain more complex concepts such as the *corrosion of aluminum*. This can be done visually in a subordinate index by finding the term for the principal concept in an ordered array, and then in a subordinated order under it, finding the term for the other concept, or it can be done in a coordinate index by coordinating the terms for the concepts either visually or mechanically.

The two operations of index using and display scanning are easily distinguished in coordinate indexes, since the initial output is usually accession number only, and a second listing is consulted for display. For subordinate indexes, the inquirer uses the index to bracket its output, and then proceeds to scan the display, which is physically adjacent to and may include the terms used for look-up. The proximity of display elements to index elements may blind the user to the fact that he is performing separate operations.

A degree of subordination occurs even in coordinate indexes to the extent that "bound" terms are used. "Bound" terms were used by Taube[3] as preconjoined unit concepts, usually as a pair of simple terms such as "landing gear," or "electron microscope." Although these are often referred to as "precoordinated," they are actually preconjoined by subordination, since one term is chosen to precede the other, and the latter serves as an access point only when subordinated to the former.

Although coordinate indexes contain the above element of subordination, the two types of indexes are mutually exclusive if we define a coordinate index as one where coordination of terms is the prevailing and intended means for obtaining conjunction, even though some conjunction by subordination may also occur.

Indexes for Serial Search and Indexes for Selective Search.—Another way that indexes can be split into two broad classes is according to whether the entire index (or a major portion of it) has to be serially searched in processing a question, or whether only that part of it involving the term or terms in the question needs to be selected for processing. The difference stems from the way the index elements are put together. A datum identifying a term being employed, or a document being examined,

† Webster's "New Collegiate Dictionary."

is useless in itself, but when a term (or its code) is tied to a document (or its code) the resulting index-couple is the smallest indivisible element of index. For example, if *aluminum* is an index term for *Document 5417*, an index-couple is a point in the index (card, tape film, printed page) that ties aluminum and Document 5417 together. How these index-couples are arranged relative to each other determines what kind of an index we have, if any.

These index-couples may be stored (1) in order by term (alphabetically or by code) or (2) not in order by term, but with all the index-couples for each document together. The first type allows selective search or look-up because of the ordered arrangement of the index-couples by term. It is often called an inverted index, possibly because the index-couples have to be removed from the environment in which they are born (the work sheet) and organized in an entirely different way.

The other type requires serial search. That is, every set of index-couples representing a document has to be examined for a term or for the conjunction of two or more terms. This index is not inverted, since the index-couples may enter the system's memory in exactly the same order as on the work sheet.

Two refinements available as tools to give greater depth to coordinate indexes are roles and links. These will not be discussed, since their only impact on the concepts under discussion is that an index-couple will now tie together a term-role with a document-link.

In addition to logical conjunction, two other operations may be used in processing inquiries[3]—logical alternation and logical negation. These three operations are frequently called by other names, such as logical product, logical sum, and logical difference, respectively. The complexity of the logical operations employed during retrieval has no impact on the concepts being developed in this article.

Some Known Indexes Typed by Both Above Procedures.— Since indexes can be split into two types by two independent parameters, this should allow all known indexes to fall somewhere on a simple 2 × 2 matrix. Examples are shown in Table I.

Entries in the quadrants in Table I are only illustrative and are by no means complete. The depth of an index, the nearness of subpopulation E to D, can vary from shallow to deep in any one of the three quadrants used above. Similarly, the quality of the index, how well it meets the aims of its charter, be it shallow or deep, can vary from poor to good in any one of the three quadrants used. It is not the purpose here to describe in detail the numerous subtypes of indexes within these quadrants, nor to compare critically their respective performances.

Entries in the upper-right quadrant involve manual-visual retrieval in every case although computers may be used in vocabulary control and print-out of indexes. The sponsors of these indexes may also operate central information services using a system in one of the lower quadrants.

Entries in the lower-left quadrant all involve mechanized retrieval varying in sophistication from using needles on edge-notch cards, through using selectors on internally punched cards, to using computers on magnetic tape.

Entries in the lower-right quadrant involve manual-visual retrieval as in terminal digit and optical coincidence systems, or mechanized retrieval varying in sophistication from card collating to random access computer systems such as the one used with the CIA Walnut system. Some indexes in this quadrant, after an initial selective search involving a term, may follow with a serial search on that term record to complete the logical operations required for the inquiry.[14] An example is the structural formula index used at Linde.[15] By selective search in an IBM card index, the deck of cards for a particular chemical fragment is found. Then by using a sorter on dedicated fields, that deck of cards may be serially searched for

Table I

Type as determined by how conjunction is obtained	Type as determined by initial search procedure: Serial	Type as determined by initial search procedure: Selective
Conjunction by subordination (or visual look-up with no conjunction possible)		Battelle Information Centers[2] Library Card Catalogs Engineering Index STAR, TAB Descriptor Index, Index to *Chemical Abstracts* KWIC, KWOC
Conjunction by coordination	ASM(WRU)[4] Marginal punched cards[5] Card selector systems (IBM 101 sorter)[6] Bureau of Ships Rapid Selector[7] Magnetic tape systems arranged under documents	Terminal digit systems[3,8] Optical coincidence systems[9] Card collating, Linde Co.[10] Random access computer systems[11] such as the index to the CIA Walnut system[12] Minicard System[13] Magnetic tape systems arranged under terms[a]

[a] These systems have lost pure selectivity, since the tape must be searched through the last of the terms involved in a particular question.

codes for other structural features. The selected cards show the codes for the chemicals meeting these structural requirements. Conjunction is still by coordination and can be obtained alternatively without serial search by collating the decks for fragment terms.[15, 16]

Note that where selective search is followed by serial search, the initial selective search eliminates most of the population of items in the file; hence, it is logical to class such indexes as selective search indexes, even though serial search devices may be used to complete the search. Where no device is used to complete the search, the index has gone as far as it can in pointing out subpopulations and the user visually examines the display for relevancy.

Why is the upper-left quadrant vacant? Let us recall that an index in this quadrant would not be designed for any conjunction by coordination and would have no selective search capabilities. Index entries would be organized by document records and a user would be required to examine visually every entry from beginning to end. An example would be a bundle of work sheets emanating from the analytical phase of index building, and to my knowledge no one considers a bundle of work sheets to be an index. The index output for any inquiry would be the entire set of records, since no device points to a subpopulation. No index worthy of the name and worthy of the charter that set it in existence would have an output for a reasonable inquiry that is coterminous with the entire file. Such an array is more aptly called a "table of contents," or a "shelf list," or an "accessions list," and may comprise a rather sophisticated display, but it is a display of the whole, and not the display of an index output.

Our definition of an index includes the presence of a *device* indicating a subpopulation. Such a device may be physical (needle, computer, etc.) or intellectual (an ordered array allowing selective search), as long as it obviates the need for the user to scan visually a list of items or item numbers from beginning to end. Any definition of index that includes in its scope the latter deviceless undertaking would have difficulty in the real world. It would predict indexes in the upper-left quadrant, where there are none. By including serial scanning as index using, all juxtaposed items or item numbers would become indexes. In contrast, our definition conforms to the observed absence of indexes in the upper-left quadrant, and allows a sharp, objective, demarcation line between an index and some other grouping.

Index Use *vs.* Output Scanning for Various Indexes.— KWIC (key work in context) and KWOC (key word out of context) indexes are examples with single-entry terms; namely, those terms in the well-known alphabetical array. At the present state of the art, conjunction is not possible, in that there is as yet no device which can operate on the population of items under a given term entry and point to a subpopulation involving another term, as can be done with an index like that of *Chemical Abstracts*. The involvement of a second term (in KWIC or KWOC) along with the one in the ordered array can be found only by a serial search of all entries under the latter term. Only by perusing the display of all 69 entries under *aluminum* can one be sure of finding all possible entries on aluminum corrosion. Only three are found, or about 4% relevancy. Inability to obtain conjunction of terms, as a conventional subject heading index or a coordinate index can do, tends to limit KWIC and KWOC indexes to collections with a relatively low population of items, such as current awareness bulletins.

The "Descriptor Index to Technical Abstracts Bulletin for 1962" contains large portions for which there is no subordinate array under an index entry, and thus affords an excellent example of a rather involved scanning of index output. (This format was discontinued by TAB in 1963.) The issue of January to June 1962, includes 35 entries under *Aluminum* and 61 entries under *Aluminum Alloys*, each containing 20 to 25 terms listed consecutively in 8 to 10 lines. A search for corrosion under *Aluminum* or *Aluminum Alloys* is a display search and not a look-up or index-use search. We have described two types of conjunction, by subordination and by coordination. If someone insists that serially searching a display such as in the above example for "corrosion" among the key words is another type of conjunction—we have a name for it, "brute-force" conjunction. This will be dramatically illustrated if one attempts this in the example given. Indexes are orderly devices, and such a "brute-force" step does not belong in the concept of index-using, but only in display perusing, where "from beginning to end" scrutiny is the order of the day.

Since many inverted indexes with selective search have the elementary index-couples on cards, there is room on the cards for more display than the minimal document identification. In these, and other cases, a much higher level of display is available, including key words, abstracts, or extracts. The fact that the record on these cards looks like (and may be) material copied directly off a work sheet, or duplicated from a descriptive catalog card, should not blind one to the fact that the index is still inverted, since the index-couples are arranged for selective search. What is on the cards, beyond the all-important index-couples that establish their places in the ordered file, is only display for use by the inquirer on the index output after the index has done its work.

Significance in Systems Design.—This post-retrieval necessity of serially examining the output display is common to the output of all indexes, whether of the serial or selective type, and whether of the subordinate or coordinate type, and whether the index is large, small, deep, shallow, good, or bad.

This distinction between the activities of using an index and examining a display of its output is important, because it allows better understanding of the unit work processes that make up a documentation system.

For example, a scientist notices that he or his colleagues emerge from using a conventional (subordinate) index with really pertinent references, whereas when they use a coordinate index, or are given the output from one, the pertinent references are diluted with a lot of extraneous material. This looks at first sight as though the coordinate index has less resolving power (is less deep) than the conventional index. However, they are comparing the wrong things. They should either compare the results after both the indexing step and the output scanning step has been completed in each index, or they should compare the results after the indexing step only and before the output display is examined in each index. Better system planning is possible by the latter comparison, because the two work processes—index operation and output scanning—can be worked on separately. The frequent confounding of these operations by the user is because the index-couples, which are at the heart of index using, frequently carry the display also, as with card catalogs or subject heading indexes such as *Chemical Abstracts*. Thus, the user subconsciously proceeds from using the index-couples to using the display, to which the former have guided him. This intertwining of the operation in use should not blind one to the fact that the operations can be separately designed in system building.

To give a further example, a good display in a small system with a very shallow index can lull people into a false sense of security. The shallow index gives the user a large proportion of irrelevant material, but the system is small and he has a good display. If the act of perusing the display is erroneously considered part of the act of using the index, the following error could easily be encountered. A user poll might indicate high satisfaction with the index, with practically no false drop. But the users eliminated the false drop during the display scanning, thinking this was part of using the index. On the basis of the low false-drop report, management might expand into a large mechanized system, with the indexing depth frozen at the point of the user survey. At large volume a good display ceases to be a substitute for a deep index. This is not to argue against high-level display; it simply warns against confusing display with index.

Confusion of another sort has resulted when an index entry leads to a pertinent article, but the abstract serving as a display is misleading causing the inquirer to eliminate the item. This loss of a pertinent reference may be erroneously attributed to a weakness in the index, but is actually a weakness or misuse of the display.

In summary, the two steps of index using and display scanning are often confused, particularly in subordinate indexes. It is submitted that distinguishing between these two work processes results in better system design.

REFERENCES

(1) J. C. Costello, Jr., *J. Chem. Doc.*, **4**, 12 (1964).
(2) R. J. Runck, *ibid.*, **2**, 129 (1962).
(3) M. Taube, *et. al.*, "Studies in Coordinate Indexing," Vol. 1, Documentation Inc., 1953.
(4) J. W. Perry, and A. Kent, "Tools for Machine Literature Searching," Interscience Publishers, Inc., New York, N. Y., 1958.
(5) C. W. Brenner, and C. N. Mooers, in "Punched Cards," 2nd Ed., R. S. Casey, *et al.*, Ed., Reinhold Publishing Corp., New York, N. Y., 1958.
(6) C. K. Schultz, in "Punched Cards," ref. 5.
(7) J. P. McMurray, *Am. Doc.*, **13**, 66 (1962).
(8) J. C. Costello, Jr., *ibid.*, **12**, 20 (1961).
(9) W. E. Batten, in "Punched Cards," 1st Ed., R. S. Casey and J. W. Perry, Ed., Reinhold Publishing Corp., New York, N. Y., 1951, pp. 169–181.
(10) F. R. Whaley, in "Information Systems in Documentation," Interscience Publishers, Inc., New York, N. Y., 1957.
(11) J. J. Nolan, *Am. Doc.*, **10**, 27 (1959).
(12) P. D. Bradshaw, *ibid.*, **13**, 270 (1962).
(13) J. W. Kuipers, A. W. Tyler, and W. L. Myers, *ibid.*, **8**, 246 (1957).
(14) F. R. Whaley, *ibid.*, **12**, 101 (1961).
(15) C. M. Lauer, and F. R. Whaley, *J. Chem. Doc.*, **3**, 150 (1963).
(16) F. R. Whaley, *Special Libraries*, **53**, 65 (1962).

The Generation of a Unique Machine Description for Chemical Structures—A Technique Developed at Chemical Abstracts Service

H. L. MORGAN
Chemical Abstracts Service, The Ohio State University, Columbus, Ohio 43210
Received January 15, 1965

I. INTRODUCTION

As part of the development of a computer-based chemical information system at CAS, it has been necessary to devise techniques for the registration of drawings of chemical structures. A major purpose of the CAS registration process is to determine whether a particular structure has already been stored in the system. The ability to make this determination makes it possible to utilize a computer to assign to every chemical structure a unique identifying label. This identifying label, referred to as a registry number, is the thread that ties together all information associated with a particular compound throughout the developing CAS computer system. It is because of this association, made possible by the registration process, that CAS will be able to provide multiple-file correlative searches with assurance that all information on file for a particular compound has been located.