

Experimental Program for Online Access to ACS Primary Documents

S. M. COHEN,* C. A. SCHERMER, and L. R. GARSON

Chemical Abstracts Service, Columbus, Ohio 43210

Received August 14, 1980

Chemical Abstracts Service typesets 16 American Chemical Society primary journals using a computer-based composition system. In the production of these journals, an archived file consisting of the text and some graphic data in machine-readable form is generated as a byproduct. The availability of the archived file has prompted us to investigate ways in which it may be advantageously utilized in the future. This paper describes one aspect of this investigation involving the development and testing of an experimental online program which allows an operator to select textual subsets of primary documents. The preparation of articles for a synoptic journal is one possible future application of this type of program. Other possible future applications of the developed-text access and manipulation features are mentioned.

INTRODUCTION

Chemical Abstracts Service (CAS) operates a computer-based journal composition system for the American Chemical Society (ACS). A byproduct of the journal production process is an archival file of the text of which the journal articles are composed. In the case of one journal, the archived file contains the text of all issues since 1975. The availability of this accumulating file of machine-readable primary text has encouraged us to investigate ways in which this resource may be utilized in the future.

This paper reports upon one aspect of this investigation. An online program was developed and tested which allows the operator to access and select portions of text of the primary document. One possible future application of the text access and manipulation techniques developed in this program is the selection of segments of text for inclusion in article summaries or synopses. Such summaries could be used in a synoptic journal or for other purposes. While the experiment reported in this paper involved this specific application, a more general purpose in developing the program was to exploit the hierarchical structure of the archived file which seems well suited to providing online access to the text in a variety of ways. Other possible future applications of the online access and manipulation techniques tested in this study can be envisioned. Several of these are mentioned later in the paper. The following section provides some background on the ACS/CAS primary journal composition system.

BACKGROUND

In cooperation with the Books and Journals Division of the ACS, CAS began development of the journal composition system in 1972.¹ The January 1975 issue of *Inorganic Chemistry* was the first production of the system. Over the past 5 years the number of ACS journals produced by the system has gradually increased. With the addition of the *Journal of the American Chemical Society* in July 1980, a total of 16 journals comprising over 35 000 pages per year will be produced by the system on a regular basis. In the production of these journals, a computer-readable data base of manuscript text is generated. Both the journal issues and the archival file are later produced from this data base.

During the development period, the typesetting process and data base requirements were studied in detail. The aim of the development strategy was to use the computer for routine typesetting decision making while traditional interfaces between authors, editors, and the typesetting function were reserved. Computer programs were created to make typesetting style and page arrangement decisions.

CAS serves as the journal compositor, receiving author manuscripts and preparing typeset copy. The coordination of editorial review, technical editing, and interactions with authors is carried out by the ACS Books and Journals Division. The typesetting system is a man-machine system. Keyboarders, graphic artists, and proofreaders supply the input. The computer supplies composition and format decisions. All journal typesetting and page composition decisions are made under computer control in a noninteractive environment. In the page composition phase of the operation, the computer program arranges text and graphics on journal pages, according to a set of heuristics.

System Features. The publication system incorporates these features: (1) data base design, (2) content-directed data, (3) preservation of individual journal "personality", (4) menu of standard composition capabilities.

The use of data base design and content-directed rather than format-directed data elements was influenced by CAS's experience in secondary services publication. The concept of content-directed data allows data elements to be independent of the journal in which they appear. Within the data base, all titles, author names, text paragraphs, etc., are stored with the same data characteristics. Data element integrity has proven beneficial in normal system operations. The standardization of data input, edits, composition routines, and the multiple use of data elements are benefits derived from the content-directed data concept. This concept also appears to facilitate the future exploitation of the data base.

The need to preserve the individual journal's personality while maintaining an efficient and generalized system led to the development of individual "journal tables" containing specific composition and format details. These journal tables are computer-stored static programs which are referenced during the composition process. A separate journal table exists for each journal produced by the system. As journals were added to the system, many format routines developed for previous journals were "almost" usable. Rather than creating an entirely separate format package for each new journal, standard formatting decisions were separated from those which were specialized and journal specific. The standard format flow was controlled by a composition service routine, with specialized decisions being resolved through access to an individual journal table.

Production Flow. Figure 1 illustrates the processing flow of primary journal documents. First, author manuscripts are entered on the primary journal data base. The data base itself consists of three physical files.

(1) Manuscript Text File: This file contains the text for all the journals. Each manuscript is individually identified

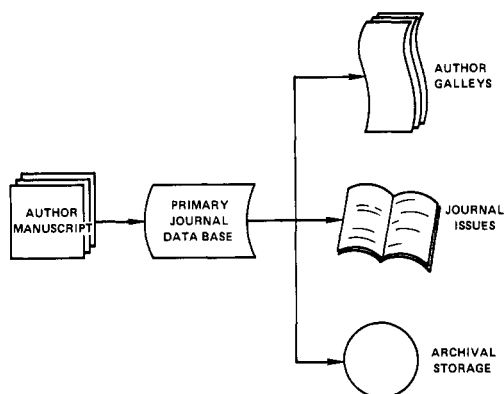


Figure 1. Processing flow.

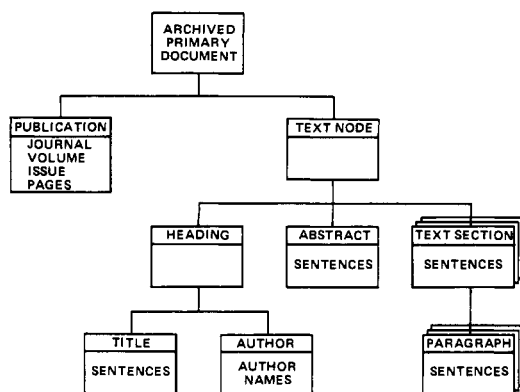


Figure 2. Hierarchical relationship of primary document components.

with its journal affiliation. Each manuscript follows the same structural hierarchy and contains the same basic set of identifiers.

(2) Manuscript Graphics File: This file contains all graphics that can be computer formatted with CAS's graphic system capabilities for primary journal data. Currently some equations are computer stored and formatted. Width and depth measurements are recorded for noncomputer stored graphics.

(3) Issue Ordering Information File: This file contains lists of manuscripts appearing in each journal issue.

Within its production cycle, the three different forms of output illustrated in Figure 1 are generated from the data base. The CAS journal production staff, rather than the system, determines when any of these output forms are to be generated for any given primary document. The output forms are (1) author galley (sent to authors and journal technical editors for approval and correction before final publication), (2) journal issues (the ordering conventions supplied in the Issue Ordering File are used in generating the issues), (3) archival storage (each archived manuscript contains all text, computer-stored graphics, and issue pagination). This paper reports upon experiments involving the text portions of the archived file.

Primary Journal Text Relationships. Figure 2 presents an abbreviated diagram illustrating the hierarchical structure of the various components of which an archived primary journal document is composed. For all data, except that in the publication component, the archived document contains the same relationships as the working document used in journal production. Only the most significant file components and data elements are illustrated here; the full text data base description is much more extensive.

Figure 2 contains the following file components (segments):

(1) Publication Segment: This is supplied from the issue-ordering information on the primary journal data base and

added to the text portion during archiving.

(2) Text Node: This is a collection of the segments identifying the manuscript text portion. The various types of text nodes are (a) heading segment—contains a subgrouping of bibliographic data, which is used to compose the manuscript heading for the document and table of contents entry; (b) abstract—identifies the author-supplied abstract, which is provided for most articles in ACS journals; (c) text section—groups paragraphs together under a specific subject heading, e.g., Introduction, Experimental Section, Discussion; (d) paragraph—identifies individual paragraphs within a text section.

The information within the title, author abstract, text-section heading, and paragraph segments is represented as individual sentences. There are generally multiple paragraphs within one text section and multiple text sections within a single document.

EXPERIMENTAL ONLINE PRIMARY TEXT ACCESS PROGRAM

The experimental online program is written in PL/1. It operates on one IBM 3277 terminal under TCAM (Telecommunication Access Method). The terminal is directly connected to an IBM 370/168 computer. The program utilizes a disk test file extracted from the ACS primary document archived file.

The capabilities of the experimental program are summarized as shown:

(1) **Document Access.** The operator can access documents on the test file sequentially and by directly requesting a specific document.

(2) **Access to Text within a Primary Document.** When the hierarchical structure of the archived file is exploited, the experimental online program provides access to the primary text in a variety of ways. The operator can request the program to display any of the following components of the primary text: (a) author abstract, (b) any paragraph within the text of the document, (c) any text section (e.g., Introduction, Experimental Section, Conclusions, etc.). While viewing the primary text in any of the three modes described, the operator can display the particular text segment a sentence at a time.

(3) **Selection, Display, and Reprinting of Text Subsets.** The operator can select subsets of the primary text for viewing on the terminal and/or reprinting.

The sections which follow provide a more detailed description of the text-accessing features of the experimental program from an operational standpoint. Several different types of screens are used. Each of these screen types is illustrated later in the paper. The flow of the different screen types in the online program is represented schematically in Figure 3.

Access to Specific Documents. Following logon, the first screen displayed by the program is a document selection screen. The document selection screen is illustrated in Figure 4 and is represented in Figure 3 by point 1. The asterisks which appear in the various screen diagrams were used to represent subscripts, superscripts, and other characters which could not be displayed on the terminal. This screen displays the title, author, and journal name for the documents on the test file. Each document is numbered so that the operator can select a particular article by keying the number corresponding to the desired document in the selection field.

The terminal used in this experiment has function keys which may be programmed to carry out specific operations in the application program. This capability was utilized in the online program for a number of high-frequency operations. The program has two options by which the operator may access specific articles on the test file. The operator may, by using a single function key, progress sequentially one document at

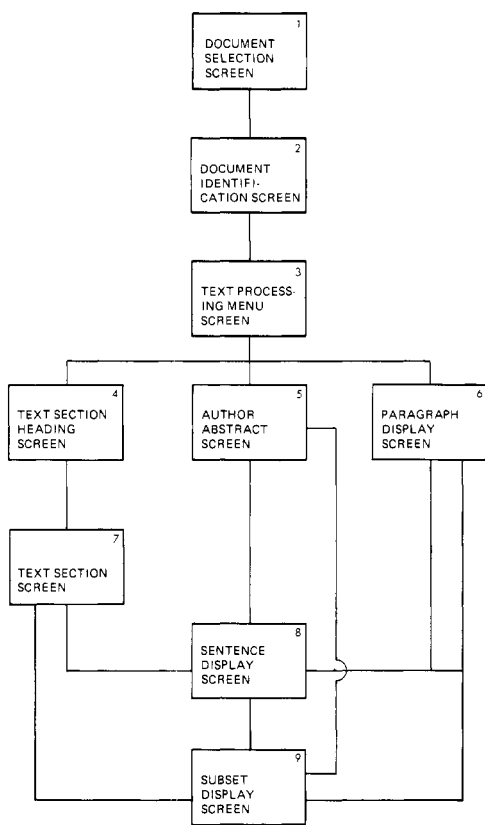


Figure 3. Flow of screens.

KEY SELECTION → 3 ACS PRIMARY DOCUMENT PROGRAM

DOC. NO. 1 of 20 THE JOURNAL OF PHYSICAL CHEMISTRY
TITLE: Dielectric Studies of Molecular Association
AUTHORS: G. Brink and L. Glasser

DOC. NO. 2 of 20 THE JOURNAL OF PHYSICAL CHEMISTRY
TITLE: Evidence for Near Zero Coordinate Silver(I)
AUTHORS: Y. Kim and K. Seff

DOC. NO. 3 of 20 THE JOURNAL OF PHYSICAL CHEMISTRY
TITLE: Reactions of Tritium Atoms with Fluoroform
AUTHORS: B. K. Min, C. *T. Yeh, and Y. *N. Tang

DOC. NO. 4 of 20 THE JOURNAL OF PHYSICAL CHEMISTRY
TITLE: Differential Heat of Chemisorption
AUTHORS: Nagao et al.

Figure 4. Document selection screen.

TITLE: Reactions of 4.5* and 6.0*eV Photochemically Produced Tritium Atoms with Fluoroform.

AUTHORS: B. K. Min
C. *T. Yeh
Y. *N. Tang

JOURNAL: THE JOURNAL OF PHYSICAL CHEMISTRY
VOLUME: 82
ISSUE: 9
PAGES: 971-975

- TO DISPLAY PROCESSING MENU ... HIT ENTER KEY
- TO ACCESS THE NEXT SEQUENTIAL MANUSCRIPT ... HIT PF3 KEY
- TO RETURN TO ORIGINAL (LOGOFF) SCREEN ... HIT PF1 KEY

Figure 5. Document identification screen.

a time through the documents on the disk file. Alternatively, the operator may select a particular document from a list of

KEY SELECTION → 2

1. DISPLAY TEXT SECTION HEADINGS
2. DISPLAY ABSTRACT
3. BROWSE THROUGH TEXT BY PARAGRAPHS

Reactions of Tritium Atoms with Fluoroform
- B. K. Min, C. *T. Yeh, and Y. *N. Tang

TO RETURN TO DOCUMENT ID SCREEN... HIT PF1 KEY

Figure 6. Text processing menu screen.

SCROLLING INCREMENT SCROLL UNIT

ARTICLE: Reactions of Tritium Atoms with Fluoroform

AUTHORS: B. K. Min, C. *T. Yeh, and Y. *N. Tang Page 1 of 1

TEXT SECTION: AUTHOR ABSTRACT

Tritium atoms with initial energies of 4.5 and 6.0 eV have been formed by the vacuum UV photolysis of TBr. Reactions between CF* and tritium atoms at these two energies, plus four lower ones, have been carried out. The threshold energies for the T* for *H and T* for *F substitution reactions in CHF* have been evaluated as 1.9 * 0.2 and 1.9 * 0.3 eV, respectively. The higher threshold and lower reactivity of T* for *F substitution of CHF* in comparison with those of CH*F indicate the absence of a possible inversion mechanism during the F substitution of CHF*.

Figure 7. Display of author abstract.

documents displayed on a screen and cause that document to be displayed directly.

Once a particular document has been selected, the program displays a document identification screen (point 2 of Figure 3). This screen, illustrated in Figure 5, provides the operator with more extensive bibliographic information for selected documents than is contained on the document selection screen. While the program is displaying the document identification screen for a particular document, the operator may, by using a single function key, proceed sequentially through the text file displaying the document identification screen for each document.

The operator can cause the program to transfer from the document identification screen back to the document selection screen by using a single function key. This same function key is programmed to transfer the display of any of the various screen types to the screen type immediately above it in the diagram in Figure 3.

Accessing Text within a Document. Having accessed a particular document, the operator can request a text processing menu to be displayed. Display of the text processing menu corresponds to point 3 in Figure 3. The text processing menu, illustrated in Figure 6, provides the operator with three choices for the display of primary text. The operator chooses by keying the appropriate digit in the selection field of the text processing menu. Choice one generates a display of the text section headings (e.g., Introduction, Experimental Section, Results, Conclusion, etc.) which are contained within the text of the particular primary document being accessed. Choice two generates a display of the author abstract, and choice three allows the operator to browse through the text paragraph by paragraph. The text processing menu also displays a condensed form of the title and the authors. This information serves as a reminder to the operator as to which primary document is presently being accessed.

Display of Author Abstract. If the operator chooses to display the author abstract, the program calculates the number of "pages" (screens) required to display the text of the abstract and then displays the first "page". If the abstract requires two pages, the screen appears with the notation "page 1 of 2"

| | | |
|--|----------------|-------------------|
| SCROLLING INCREMENT 1 | SCROLL UNIT: P | PARAGRAPH 7 of 27 |
| ARTICLE: Reactions of Tritium Atoms with Fluoroform | | |
| AUTHORS: B. K. Min, C. *T. Yeh, and Y. *N. Yang | | Page 1 of 1 |
| TEXT SECTION: Experimental Section | | |
| In a typical run, the discharge time through the mixture of T* and Br* was about 30 min, and the amount of TBr synthesized was around 10** mol. This transcribes into a pressure of about 10** Torr in the reaction cell which is about 10* times lower than the partial pressures of other components in the system | | |

Figure 8. Display of text by paragraph.

| | |
|--|-----------------------|
| KEY SELECTION → 3 | TEXT SECTION HEADINGS |
| 1. Introduction | |
| 2. Experimental Section | |
| 3. Results and Discussion | |
| Reactions of Tritium Atoms with Fluoroform B. K. Min, C. *T. Yeh, and Y. N. Tang | |
| SELECT NO. OF TEXT SECTION YOU WISH TO DISPLAY | |
| TO RETURN TO PROCESSING MENU ... HIT PF1 KEY | |

Figure 9. Text section headings screen.

in the upper right corner. The author abstract screen is illustrated in Figure 7. Using two function keys programmed for these specific purposes, the operator can page forward and back through the various pages of the abstract text.

Display of Text by Paragraphs. When the operator chooses to access the primary text by paragraph, the program displays the paragraph data on a screen as illustrated in Figure 8. In the upper right corner of the screen, the program shows the number of the particular paragraph being displayed as well as the total number of paragraphs in the document being accessed. The operator can scroll through the paragraphs one at a time (forward or backward) by using two function keys programmed for these purposes. The operator may scroll through the paragraphs one at a time (one is the default scrolling increment) or, by keying some other number "n" in the scrolling increment field, cause the program to display the paragraph "n" paragraphs ahead of or in back of the paragraph currently being viewed. The operator may also transfer directly to the first or the last paragraph of text by keying an appropriate command.

Display of Text Sections. Figure 9 illustrates the display of the text section headings contained within a particular document. The text section heading screen corresponds to point 4 of Figure 3. The operator may choose to display the text corresponding to any of the displayed text section headings by keying the appropriate digit on the text section heading screen. Figure 10 illustrates the display of a particular text section (the Results and Discussion section). As with the abstract text, the operator may page forward or backward through the various pages of the text section being displayed.

Display of Individual Sentences. At the option of the operator, an abstract, paragraph, or text section may be viewed a sentence at a time. The operator can display individual sentences by keying an appropriate command in the *Scroll Unit* field. Display of individual sentences is represented in Figure 3 by point 8. Figure 11 illustrates the screen used to display individual sentences. When displaying sentences, the operator can scroll backward or forward a sentence at a time by using an appropriate function key. Note in Figure 11 that the program indicates both the number of the sentence cur-

| | |
|---|-------------|
| SCROLLING INCREMENT | SCROLL UNIT |
| ARTICLE: Reactions of Tritium Atoms with Fluoroform | |
| AUTHORS: B. K. Min, C. *T. Yeh, and Y. *N. Tang | |
| Page 10 of 10 | |
| TEXT SECTION: Results and Discussion | |
| Relative Yields of Products from Photochemical Tritium Atom Reactions with CHF*. Fluoroform does not absorb light with a wavelength longer than 1200 **. The strong chemical bonds in this molecule render it as a good testing ground for the use of 4.5* and 6.0* eV tritium. In the present experiments, the photochemically produced tritium atoms might undergo either abstraction or reaction 1-4. The yields for all the expected products except TF have been measured. Since the experimental results in this work always indicate an HT yield which is much higher than those of the substitution products, it is necessary to ascertain the absence of any anomalous source of HT. (Some earlier work actually involved an anomalous HT yield. Those data have been discarded.) In the first place, the blank analyses of the reaction mixture before photolysis has been performed for every sample, and in most of the cases, the background correction for the T* impurity was less than 10% of the observed total HT radioactivity. Secondly, photolysis of TBr in the presence of CF* by the shorter wavelength radiations have been carried out, and the observed HT radioactivity was not more than the blank corrections (CTF* was also not observed.) Thirdly, several TBr-Br*-CH*F photolyses at 1849* | |

Figure 10. Display of a text section.

| | | |
|--|----------------|--------------------|
| SCROLLING INCREMENT 1 | SCROLL UNIT: S | PARAGRAPH 12 of 27 |
| ARTICLE: Reactions of Tritium-Atom with Fluoroform | | |
| AUTHORS: B. K. Min, C. *T. Yeh, and Y. *N. Tang | | SENTENCE 3 of 12 |
| TEXT SECTION: Results and Discussion | | |
| The strong chemical bonds in this molecule render it as a good testing ground for the use of 4.5* and 6.0* eV tritium. | | |

Figure 11. Display of individual sentences.

rently being displayed and the total number of sentences in the particular text segment being accessed. The operator can transfer from sentence display mode back to the display of the total text segment currently being accessed by using a single function key. Such transfers are represented in Figure 3 by the paths 8 → 7, 8 → 5 and 8 → 6.

Selection of Text Subsets. The program allows the operator to select, display, and print textual subsets of the primary documents. Using a single function key, the operator can select for later display and/or printing any of the textual components or combination of components which the program can access. The operator can select abstracts, paragraphs, text sections, or individual sentences. The operator also has the capability of generating text subsets on the basis of any combination of components of the total primary document. Using the dual-intensity display capability of the terminal, the operator can highlight portions of any of the text components which can be accessed by the program (i.e., abstracts, paragraphs, or text sections) for selection and incorporation into text subsets. By positioning the cursor anywhere within the text of a particular sentence and striking a single function key, the operator can highlight the entire sentence. Once a particular sentence within a larger text segment has been highlighted, the operator can by repeated strikings of the highlight function key cause successive sentences to be highlighted. Highlighted portions of text can at any time be selected for inclusion in the text subset through the use of a select function key. At any time, the operator can display and/or print the text subset selected for a particular document. Figure 12 illustrates the display of an operator-selected text subset including an abstract and the introduction to a particular document.

TESTING OF THE PROGRAM

The experimental program in its present form is specifically structured to access primary text and to allow an editor to

| | |
|---|--|
| COMMAND: | TEXT SUBSET |
| ARTICLE: | Reactions of Tritium Atoms with Fluoroform |
| AUTHORS: | B. K. Min, C. T. Yeh, and Y. N. Tang |
| | Page 1 of 1 |
| ABSTRACT | |
| <p>Tritium atoms with initial energies of 4.5 and 6.0 eV have been formed by the vacuum UV photolysis of TBr. Reactions between CHF[*] and tritium atoms at these two energies, plus four lower ones, have been carried out. The threshold energies for the T[*] for "H and T" for "F substitution reactions in CHF[*]" have been evaluated as 1.9 ± 0.2 and 1.9 ± 0.3 eV, respectively. The higher threshold and lower reactivity of T[*] for "F substitution in CHF[*]" in comparison with those of CH[*]F indicate the absence of a possible inversion mechanism during the F substitution in CHF[*].</p> | |
| Introduction | |
| <p>The use of photochemical methods to form hydrogen atoms of certain unique energies and the study of their chemical reactions have been extremely successful during the past decade. Most often, the photodissociation of hydrogen halides, such as HBr and HI, is used to produce the high energy hydrogen atoms other molecules such as H[*]S, CH[*]SH, and H[*]O have also been used.</p> | |

Figure 12. Text subset screen.

select synoptic or summary subsets. However, the same online text access techniques used in the experimental program could be used in other applications (see Experimental Objectives). A major incentive for exploring the development of synoptic journals relates directly to the steadily increasing cost of producing primary journals with resulting increases in subscription prices. One can envision a future operation in which synoptic articles are distributed to subscribers with facilities for requesting the full text of articles of interest, available either in hard copy form or for online viewing at a terminal.

Experimental Objectives. The objectives of the experimentation described in this paper were to provide information on the following:

(1) How easy or difficult is it for technical editors accustomed to working only with hard copy to learn to access and manipulate primary document text online?

(2) What are the strengths and weaknesses of the experimental online program? How can it be improved? What additional features are needed? How useful are the program features listed?

- (a) Accessing and displaying of abstracts.
- (b) Accessing and displaying of paragraphs.
- (c) Accessing and displaying of text sections.
- (d) Accessing and displaying individual sentences.
- (e) Paging and scrolling through text.
- (f) Highlighting of blocks of text.
- (g) Synoptic subset selection techniques.
- (h) Display and printing of synoptic subsets.

(3) How satisfactory are the selected synoptic subsets? How can their quality be improved?

(4) How important are the missing graphics data to the comprehension of the synoptic subsets?

(5) Does online selection and manipulation of primary text appear to be a workable method of preparing synoptic articles?

Experimental Approach. Three ACS primary journal editors were trained in the use of the online experimental program. While none of the editors participating in the experiment were experienced in using terminals and online programs, all three agreed that learning to use the program was quite straightforward. All three were able to manipulate and select text effectively in 3 h or less from the time they started learning the program operation. The training consisted of reading a manual, attending brief instructional sessions, and then practicing for several hours. Each editor spent about 4–6 h per week for several weeks using the program to generate synoptic summaries of primary articles from a test file of archived ACS primary articles. During the course of the experiment, test files of primary articles were generated to provide documents from each editor's particular area of expertise. Test files from the following three journals were made

available online: (1) *Journal of Medicinal Chemistry*, (2) *The Journal of Physical Chemistry*, (3) *Inorganic Chemistry*.

It was thought that these three journals were reasonably representative of the variety of primary manuscripts found in all 16 ACS journals. After a few hours of training, the participants started to use the online program to select synoptic subsets from the articles contained in their respective test files. Once the editors had completed the selection of the synoptic subsets, they requested a hard copy of the subsets via the online program. The listings were reviewed by the ACS primary journal editorial staff for evaluation of the quality of the selected subsets. In order to obtain the reaction of the editors to the experimental online program, each editor was asked to complete a questionnaire relating to the human interface and the various features of the program.

Modes of Access to Primary Text. The experimental program provided four access modes to the primary document text: (1) by abstracts, (2) by paragraph, (3) by text section, (4) by individual sentences.

The editors were in complete agreement on the usefulness of accessing the text by paragraphs and by displaying the abstracts. They also agreed on the relatively low utility of accessing the text via text section headings and individual sentences. When reviewing the text of a primary manuscript to select segments of the text for the synoptic subsets, the editors found it helpful in some cases to have a hard copy version of the article available for reference. Working in this manner, editors found accessing the text by paragraph to be the most natural and useful option of those available in the experimental program.

Text Highlighting Feature. The program feature which allowed the editors to highlight a sentence or group of sentences for the synoptic subset was found to be of high utility. The highlight feature allowed the editors to see the portion of text which was about to be selected within the context of the surrounding text before keying the final select command. The highlight feature as designed for the experimental program allowed the editors to highlight either an individual sentence or a continuous group of sentences with no interposed unhighlighted sentences. One editor indicated that the capability of highlighting discontinuous blocks of text would be useful.

Paging and Scrolling Capabilities. The editors were able to scroll from paragraph to paragraph and to advance or retreat through "multipage" screens (e.g., a paragraph which is too long to fit on one screen) without serious difficulty. The default scrolling increment was one paragraph. However, the program does allow the editor to specify any scrolling increment desired. This enables the editor to transfer directly to a specific paragraph.

Programmed Edits and Other Safeguards. The experimental program contains a variety of automatic edits which display error messages when an inappropriate action is requested by the operator. The program requires the operator to correct the erroneous situation before any further program action is executed. The editors considered these programmed edits to be quite vital to the efficient use of the program.

In addition to the programmed edits, the experimental program was structured so that many of the major program actions required a two-step action consisting of an original request followed by a confirmation. The intention of these safeguards was to minimize the accidental execution of a major program action by the unintentional striking of a function key. Two examples of program actions to which this type of safeguard was applied are (1) selection of a text segment for inclusion in the synoptic subset, (2) print request for a synoptic subset.

Table I. Quality of Synoptic Subsets

| journal | percent of subsets in category | | | |
|--|--------------------------------|------|------|------|
| | excellent | good | fair | poor |
| <i>Journal of Medicinal Chemistry</i> | 10 | 70 | 10 | 10 |
| <i>The Journal of Physical Chemistry</i> | 11 | 56 | 22 | 11 |
| <i>Inorganic Chemistry</i> | 0 | 86 | 14 | 0 |
| overall | 8 | 69 | 15 | 8 |

The editors considered this type of safeguard to be of moderate to vital importance to the efficient operation of the program.

Needed Program Capabilities. The online program allowed the editors to select sentences from the primary text for inclusion in the synoptic subset but not to manipulate the text of the subset. The editors considered the following additional capabilities with respect to the online generation of synoptic subsets to be of high importance: (1) deletion of whole sentences and parts of sentences, (2) rearrangement of the order of sentences, (3) insertion of editor-generated connective text.

Some other program capabilities suggested by the editors as necessary or useful are (1) option to scroll by half-page increments, (2) capability of holding a partially completed synopsis from one terminal session to another, (3) a split screen capability allowing the operator to view two different segments of text at the same time, (4) option to remove text section headings from the selected synopses.

Quality of Synoptic Subsets. A total of 26 synoptic subsets were prepared by the editors during the course of the experimentation with the online program. ACS editorial staff evaluated the quality of each of these synopses. Each subset was placed into one of four categories. The results of the evaluations are shown in Table I.

As can be seen in Table I, 84% of the subsets were judged to be in the good and fair categories while the excellent and poor categories each accounted for 8% of the selected subsets. A major factor affecting the quality of about half of the synopses generated was the lack of graphical data. Table II shows the relative importance of the graphical data to the quality of the selected subsets as judged by the ACS editorial staff.

Missing graphics data were considered to be of high to moderate importance to the comprehension of the synoptic article in 46% of the subsets and to be unimportant in 54% of the subsets. The second major factor affecting the quality of the subsets was the need for the inclusion of editor-generated connective text as well as other text manipulation capabilities relative to the selected subsets.

Workability of Online Selection of Synopses in a Production Environment. The three editors who participated in the ex-

Table II. Importance of Missing Graphics Data

| journal | percent of subsets in category | | |
|--|--------------------------------|--------------------------------|------------------|
| | very im- portant | mod- erately im- portant | unim- portant |
| <i>Journal of Medicinal Chemistry</i> | 10 | 40 | 50 |
| <i>The Journal of Physical Chemistry</i> | 22 | 22 | 56 |
| <i>Inorganic Chemistry</i> | 14 | 29 | 57 |
| overall | 15 | 31 | 54 |

periment were asked, on the basis of their recent experience with the program, to assess the workability of the online method of synopsis selection. In general, they considered the method to be workable providing the following conditions were satisfied:

(1) Extension of the online program to incorporate text manipulation capabilities relative to the synoptic subset as stated earlier in this paper.

(2) The inclusion of graphics in the synopses. Development work is under way which should result in portions of the primary graphics information being stored in the archival file. Presently, much of the equation information is stored in machine-readable form.

(3) Provision of instructional sessions for the editors regarding the content of synopses.

Possible Future Applications. Although no decisions have yet been made regarding the use of the online text access techniques discussed in this paper, a number of possible future online applications can be envisioned. These include (1) review of articles retrieved during an online search,² (2) browsing through "electronic journals", (3) selection of portions of primary articles for inclusion in synoptic journals or other summaries, (4) correction of primary manuscript during the editing and proofing steps of journal production.

REFERENCES AND NOTES

- (1) Schermer, C. "The Primary Journal System: A Case Study", *GCCA J.* 1978, 2, 19-24.
- (2) In early 1980, the ACS implemented an experimental online text search file based upon the full text of ACS primary documents. The file was implemented under an arrangement with Bibliographic Retrieval Services, Inc. A group of approximately 12 organizations was selected to conduct test searches against this file. Our purpose in this experimental effort is to gain a preliminary indication of the utility of online searching and/or display of a file consisting of the complete text of primary documents. The test file was extracted from the archived primary text file generated by the journal composition system. It consists of articles appearing in the 1976-1978 issues of the *Journal of Medicinal Chemistry*. While some of the features of the experimental search system are common to most online text searching implementations, it has other features specifically designed to facilitate searching the relatively large amount of text associated with most primary documents.