Technical-Abstracting Fundamentals. I. Introduction

By B. H. WEIL, I. ZAREMBER, and H. OWEN
Technical Information Division, Esso Research and Engineering Co., Linden, N. J.
Received February 6, 1963

Abstracts have long been used in the scientific and other intellectual communities. Today, millions of technical abstracts are being produced each year as reader aids, alerting tools, and retrieval media. They accompany original papers in primary journals, are published in reports, are collected into secondary journals, appear instead of or as keys to full texts, and are used as intermediates in information-retrieval systems. They are usually used or consulted to determine the need for reading the original documents, and they are sometimes used for current or background information in lieu of the complete documents. Basically, they are read to save time.

Abstracting is now "big business." The annual budgets of services such as *Chemical Abstracts* run into millions of dollars. So large is this business, indeed, that a National Federation of Science Abstracting and Indexing Services was organized in 1958 and is presently conducting a nationwide study to determine its proper role.

Everyone knows what an abstract is, of course. "Webster's New International Dictionary," Second Edition Unabridged, defines it as "that which comprises or concentrates in itself the essential quality of a larger thing. . . a summary or an epitome." Further study of this dictionary turns up such other intriguing synonyms as abridgement, brief, condensation, digest, review, and synopsis. Throughout many of the definitions of these words runs at least the constant thread of "fewer words, yet retaining the sense"; "condensation and omission of more or less of detail, but retaining the general sense and unity of the original"; and (as above) "essential quality of a larger thing." Some of the words branch off in the direction of "abbreviations," but of special interest in regard to "indicative abstracts" and "auto-abstracts" are two of the definitions of epitome: "a brief or curtailed statement of the *contents* of a topic or a work" and "a *part* which represents typically a large and intricate whole."

Types of Abstracts.—A number of distinct types of abstracts have come to be recognized on the basis of the information that they contain and who has prepared them:

1. The "informative," "informational," or "comprehensive" abstract, one that is still complete enough in its distillation "to communicate knowledge." This type of abstract contains the significant findings, arguments, and applications; states the scope; and usually at least indicates such other important aspects of the document as methods and equipment used. See Example 1.

Example 1. An Informative Abstract

ORGANIC COMPOUNDS CAN BE RAPIDLY TRITIUM-LABELED by incubation with tritium gas under a silent electrical discharge. For example, 500μl of benzene incorporated 0.67 mc tritium after 1 hr contact with a hydrogen-tritium mixture containing 40 mc tritium under a 20 kv/ma. discharge, 10^4 more than was incorporated without the silent discharge. Cobalt-60 gamma rays were much less effective, and were destructive.

R.M.Lemmon, B.M.Tolbert, W.Strohmeier, and I.M.Whittemore: "Ionizing Energy as an Aid in Exchange Tritium Labeling." (Science 129:1740-41 (1959)

2. The "indicative" or "descriptive" abstract, usually restricted to descriptive statements about the contents of the document. See Example 2.

Example 2. An Indicative Abstract

IONIZING ENERGY AS AN AID IN EXCHANGE TRITIUM LABELING. R.M.Lemmon, B.M.Tolbert, W. Strohmeier, and I.M.Whittemore (Science 129:1740-41 (1959)--The effect of exposure to silent electric discharge or to cobalt-60 γ-rays on incorporation of tritium in organic compounds, e.g., benzene, is reported.

3. The "annotation," in which a few words or a sentence are added to a title by way of further description, explanation, or even critical comment. Annotations are usually indicative; it is difficult to make them informative. See Example 3.

Example 3. An Annotation

IONIZING ENERGY AS AN AID IN EXCHANGE TRITIUM LABELING. R.M.Lemmon, B.M.Tolbert, W. Strohmeier, and I.M.Whittemore (Science 129:1740-41 (1959)--Effect of silent electric discharge and of Co-60.

4. The "title-only" abstract, where the title of a document is used without amplification to describe the document's contents. Titles usually state subjects, not findings, so are usually indicative rather than informative.
5. The "auto-abstract," a collection of actual sentences from the original document that have been selected as most typical of the contents by computer analysis of the frequency of use of significant words in the document and of the frequency with which these high-use words appear in the same sentences.
6. The "slanted" abstract, in which the information or or description reported is oriented to a specific "discipline" or to an industrial or governmental field or "mission." In such abstracts, emphasis may sometimes be placed on methods, equipment, or findings that were only incidental to the author's major purpose; the latter may even be omitted or minimized.

7. The "author abstract," an abstract required from the author of the document on the basis that the originator of the work knows the subject, purpose, findings, and implications. Author abstracts may be indicative or informative, depending on publisher requirements. In use, their readability and completeness have so far depended chiefly on the writing ability of the authors, as abstracters, to serve reader needs and not to "miss the forest because of the trees." Substantial help from reviewers and editors is mandatory but has been rare to date.
8. The "subject-expert-prepared" abstract, an informative or indicative abstract prepared by one who is a peer of the author of the document in discipline, education, specialization, and experience. Since subject expertness does not *per se* include abstract-writing ability, editing is necessary.
9. The "professional-abstracter-prepared" abstract, written by an individual who prepares abstracts as a professional career and who preferably is educated and experienced in some or all of the disciplines and fields in which he abstracts numerous documents. Ability to write concisely and to an audience is another prerequisite. Abstracts so prepared may be informative or indicative, depending on instructions and on the ability of the writers to comprehend the messages to be transmitted. Even here, editing is usually required.
10. The "rewritten" abstract. This includes abstracts of any type or origin that are reviewed and modified, if necessary, to improve accuracy and transmission of information.

In all of these definitions, the citation or availability of an adequate bibliographical description is implied.

Ideals and Difficulties.—At first glance, it would not seem difficult to select or synthesize a style of abstract suitable to a given purpose and to a given budget (time and/or money). Indeed, if the ideal could be approximated, most purposes could be served by providing an abstract with the original document that would:

1. Inform interested people about the forthcoming paper, if it is to be presented at a meeting.
2. Enable readers to determine whether they must read all or part of the full document to obtain needed or valuable information.
3. Be suitable for direct re-use in secondary journals and in retrieval systems.
4. At least serve as a basis for rewriting or editing for these latter and other uses.

Nevertheless, the number of types and purposes indicate some of the difficulties involved in preparing, publishing, and using abstracts, difficulties that seem numerous enough and important enough to warrant review and discussion in several papers designed to present the state of the art; to clarify phases of the subject by sorting out facts from opinions and wishful thinking; to derive standards and methods that reflect modern knowledge of communications, and to relate these to the special needs and problems of abstracts; and to suggest areas where objective research can yield important information.

Among the "difficulties" involved is lack of complete knowledge of readers' needs and desires; differences of purpose that make it difficult to kill several birds with one stone; the human tendency to make virtues out of necessities; and the very real financial problems of organizers of meetings, editors of primary journals, secondary journals, and reports, and operators of information-retrieval systems.

Aspects Warranting Closer Scrutiny.—First of all, it is desirable to look in some detail at abstract-writing principles and practices. Writing annotations and indicative abstracts is not particularly difficult. However, the writing of informative abstracts—preferable for most uses—involves knowledge and application of many factors. For example, the importance of writing findings-oriented topical sentences is little understood, despite the pressure of readers for key information as quickly as possible.

In regard to the use of abstracts in primary journals, attention needs to be paid to why the editors do or do not publish informative abstracts on the first page of each article or collect them in separate sections to aid readers in selecting items for intensive reading. Some journals give readers no quick-evaluation tools at all—not even a subtitle or annotation. Other editors use all types of "structuring" devices to amplify titles, but unfortunately these techniques are often designed to encourage readers to scan the entire document rather than to aid in selective reading. Recently, some engineering-journal editors have introduced another facet by combining the abstract for each article with information-retrieval descriptors, for use in the readers' information-retrieval systems. Additionally, because increased journal size means higher costs, editors of primary journals may be increasingly confronted with the desirability of publishing "abstracts" *instead of* full papers, with the latter then being available from central depositories. If abstracts are to play a more important role in primary journals, editors will have to do a better job of editing them.

Editors of secondary journals usually want better abstracts than are currently published with the original papers in primary journals, but some of the editors who use authors' abstracts make a virtue of necessity. Other editors strive for truly informative abstracts but differ on style and content, sometimes because they are also trying to meet the requirements of indexes.

Many designers of information-retrieval systems provide abstracts of presumably-pertinent documents to those whose questions they attempt to answer by manual or machine methods, to cut down further on the number of documents that clients must scan in their entirety. This is not effective if the abstracts fail to describe the documents stored in the system; this may be one of the reasons why other systems designers favor giving the original documents to the clients in all cases. However, skipping the use of abstracts as an intermediate step puts unusual stress on retrieval systems to avoid delivering a high percentage of non-pertinent documents, through which clients are understandably unwilling to wade.

Because of the complexities just described, and others, subsequent papers in this series will deal in some detail with abstract-writing principles and practices, and with aspects and problems involved in publishing abstracts in primary journals; in compiling secondary journals and internal bulletins of paper, patent, and report abstracts; in including abstracts in reports; and in preparing abstracts as an intermediate in information-retrieval sys-

tems. Other papers may consider costs, abstracter training, "auto-abstracting," etc.

We are aware, of course, of the many studies on abstracting, and will cite their pertinent findings in our papers. Among these studies are those of ASTIA on a standard abstract for Department of Defense technical reports[1]; the NFSAIS's "Draft Proposal on Abstracting Standards"[2]; UNESCO's "Guide for the Preparation and Publication of Synopses"[3]; the System Development Corporation's study of "Criteria for Acceptable Abstracts: A Survey of Abstracters' Instructions"[4]; studies of the Rome Air Development Center, Thompson Ramo Woolridge, and the American Institute of Research[5]; and the abstract-descriptor-combination programs of the A.I.Ch.E. and the Engineers Joint Council.[6,7] Interestingly enough, at least two of these studies[4,5] are designed to establish criteria that may eventually lead to better computer-prepared abstracts than the initial type of "auto-abstract."

Two recent contributions to the general literature on abstracting are Stevens' pamphlet entitled "Writing Better Titles and Abstracts"[8] and Arnold's paper on "The Writing of Abstracts."[9]

Validity of Abstracts.—In closing this introductory paper, it is important to consider the validity of using abstracts. Most members of the scientific community have blithely assumed that the *only* dangers involved in condensation of information into abstracts are those of omission of vital information, lack of clarity, and inaccuracy of statement. These are bad enough possibilities, of course, but is is also possible that the compaction of information inherent in writing an abstract may lessen information transmission even if the facts are still presented accurately and clearly. If true, this is of concern to all users of abstracts, particularly to those who read many abstracts in sequence and to those who read them in lieu of originals.

The evidence so far available is conflicting. Freedman[10] found that an equivalent amount of information is transferred irrespective of whether an individual is reading the original document, a half-length digest containing the salient features, or an abstract one-sixth the length of the original, although he reports that reading speed decreases with compaction. Similar findings have been reported in a more recent study by the Operations Research Group at Case Institute of Technology,[11] who found that no significant loss of information resulted from any of three degrees of article compaction (to two-thirds of original length, to one-third, and finally to an abstract).

On the other hand, Rath, Resnick, and Savage of IBM[12,13] found that condensation to an abstract reduces information transfer by 50–75%, and that the use of titles reduces it almost to zero. Similar alarms have been sounded by Bloom and Goldstein,[14] with emphasis on the possible skipping of information by too-rapid reading of dense material, and by Wason,[15] who found a loss in both the quantity and accuracy of the information retained from abstracts.

As pointed out by Henley,[16] the experiments at Case and IBM involved significant differences in control of such variables as subjects, documents, information-transfer measure, condensation method, and time. The resultant conflicts in findings and interpretation will therefore require further experiments in which these variables are better controlled.

In this context, we must be careful to distinguish between actual information transfer from the document read, be it abstract or original, and the determination by the reader that a given document is relevant to his interests. Titles are reported to be as good as abstracts for the latter purpose.[17,18]

Whatever the case, it is important that "the reader (of abstracts) must understand, by warning if necessary, that he is reading a condensed version and must slow down his normal reading speed to match the density of the information."[10] The possibility of inherent danger should also alert writers of abstracts to the necessity for clear and complete presentation of facts.

In all this attention to abstracts, we must also be careful to remember that abstracts, even well written abstracts, are not basically an easy, enjoyable type of reading in large collections. As J. Hilary Kelley pointed out at the American Documentation Institute's 1962 Annual Convention, the high concentration of information in abstracts often causes them to be instinctively shunned by readers, who need or will tolerate some redundancy in what they read.

Nevetheless, despite these inherent dangers and problems, and despite the competition from other alerting media and reader aids, we are convinced that abstracts have a basic validity that ensures their continued use—the saving of precious time to their users while also imparting substantive information. There are indications that those who must write, process, and read abstracts are becoming more alert to their possibilities and less afraid or unwilling to measure up to the work involved. The many studies in progress are a sign that the communication and documentation professions intend to supply the support and guidance needed. This series on abstracting fundamentals is designed to encourage these trends.

BIBLIOGRAPHY

(1) Armed Services Technical Information Agency, "Abstracting Instructions for Technical Reports" (Draft), Arlington, Va., September 14, 1962. 8pp.

(2) National Federation of Science Abstracting and Indexing Services, "Draft Proposal on Abstracting Standard," Committee of NFSAIS Abstract Editorial Working Group, Kenneth L. Coe, Committee Chairman, The Chemical Abstracts Service, Columbus, Ohio, Revised March 29, 1962, 2pp.

(3) United Nations Educational, Scientific, and Cultural Organization, "Guide for the Preparation of Synopses," Paris, 1957, 2pp.

(4) H. Borko, and S. Chatman, "Criteria for Acceptable Abstracts: A Survey of Abstracters' Instructions," TM 759, System Development Corporation, Santa Monica, Calif., November, 1962, 32pp.

(5) A. DeLucia, "Index-Abstract Evaluation and Design," Paper presented before the American Documentation institute, Hollywood, Fla., December 13, 1962 (Rome Air Development Center, Griffiss AFB, N. Y.).

(6) B. E. Holm, "Information Services of the American Institute of Chemical Engineers," Paper presented before the American Association for the Advancement of Science,

Philadelphia, Pa., December 27, 1962 (Engineers Joint Council, United Engineering Center, New York, N. Y.).
(7) E. Wall, "An Action Plan for the Engineering Profession," Paper presented before the American Association for the Advancement of Science, Philadelphia, Pa., December 27, 1962 (Engineers Joint Council, United Engineering Center, New York, N. Y.).
(8) J. D. Stevens, "Writing Better Titles and Abstracts," Bulletin 254, Washington State Institute of Technology, Pullman,Washington, 1962, 11pp.
(9) C. K. Arnold, "The Writing of Abstracts," *IRE Transactions of Engineering Writing and Speech EWS-4*, 80–82 (December, 1961).
(10) W. A. Freedman, "A Study in Communication," *STWE Review*, **5**, No. 3, 18–20 (1958).
(11) Operations Research Group, "An Operations Research Study of the Dissemination and Use of Recorded Scientific Information. Part III. The Effect of Condensation on Comprehension of Journal Articles," Case Institute of Technology, Cleveland, Ohio, December, 1960; NSF Grant No. G-8434.
(12) G. J. Rath, A. Resnick, and T. R. Savage, "Comparison of Four Types of Lexical Indicators of Content," *American Documentation*, **12**, 126–130 (1961).
(13) A. Resnick, and T. R. Savage, "A Re-evaluation of Machine-Generated Abstracts," *Human Factors*, **2**, No. 3, 141–146 (1960).
(14) J. H. Bloom, and S. A. Goldstein, "Five Experiments in Technical Writing," *Product Engineering*, **31**, No. 18, 55 (1960).
(15) Staff, "What Good is a Summary?" *The Technical Writer* (STW-ATWE), p. 5 (October, 1957). An account of a meeting paper by P. C. Wason of University College, London.
(16) C. B. Hensley, private communication to I. Zarember, August 27, 1962.
(17) A. Resnick, "Relative Effectiveness of Document Titles and Abstracts for Determining Relevance of Documents," *Science*, **134**, 1004–1006 (1961).
(18) T. R. Savage, and A. Resnick, "On the Consistency of Human Judgements of Relevance," Paper presented before the American Documentation Institute, Hollywood, Fla., December 13, 1962 (International Business Machines Corporation, Yorktown Heights, N. Y.).

Selective and Consecutively Arranging Triple-Notch Code

By KENZO HIRAYAMA
Research Laboratory, Fuji Photo Film Co., Ltd., Minamiashigara, Odawara, Japan
Received November 29, 1961

1. FOREWORD

Codes for edge-punched cards can be classified into two kinds, the direct or single-notch code and a combination or multiple-notch code.

The combination code is a method of representing a concept by notching two or more holes. One such method, the *selective code*, is a form of code in which the number to be notched in one field is always made constant, and the cards sorted do not make the errors of the first and second kinds, *i.e.*, the required card drops out and the unrequired card does not.

The number of items I that can be represented by notching the number of holes N in one field with H number of holes, will be

$$I = {}_H\mathrm{C}_N = H!/N!(H-N)!$$

When H is constant, the number of holes N for the most effective, selective code is obtained by $N = H/2$ (where H is an even number), or $N = (H-1)/2$ (where H is an odd number) and the maximum I can be calculated.

In the present article, a method of representing the most effective selective code on two-dimensional paper, and the method of arranging the cards notched by this code in a consecutive numerical order will be examined.

2. STANDARD SORTING FOR CONSECUTIVE ARRANGEMENT

Figure 1 illustrates one of the simplest two-notch consecutive codes.

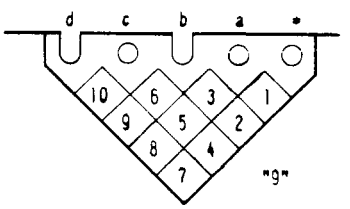

Fig. 1.—Double-notch code.

The way to put the cards notched by this code into consecutive order is to sort the cards at the hole "a"[1]; cards that drop out are replaced on the back of the cards in hand; all the cards are re-sorted at the hole "b", without changing the order of the cards; cards that drop out are replaced on the back of the cards in hand; and similarly with the holes "c" and "d". This would be apparent from Fig. 2.

(1) The hole at the extreme right (marked with * in figures) is not sorted for consecutive arrangement, and the process is started from the next hole, going to the left.