

- 85[Z].—Food & Agriculture Organization of the United Nations, *Handbook on Data Processing Methods*, Columbia University Press, New York, 1959, vi + 111 p., 27 cm. Price \$1.00 (Part 1, Provisional Edition) Paperback.

This handbook is a well-written elementary discussion of the methods of processing statistical data from surveys or censuses. It reviews the various methods of data processing, with emphasis on hand and punch card methods. Only one page is devoted to stored program computers. A bibliography is included with the discussion of each topic. The general principles concerned with planning, organizing, administering, and operating data-processing services are adequately covered. Detailed instructions are given for the proper handling and preservation of records and for sorting records, using manual or punch card methods. The handbook is intended for statisticians or supervisors of data processing activities, and hence does not cover the special operating features and details of available machines. Although prepared primarily for the use of statisticians in countries unfamiliar with present data-processing tools and procedures, it could serve as a useful introduction to the subject for others.

A. EUGENE SMITH

Bureau of Ships  
Navy Department  
Washington 25, D. C.

- 86[Z].—FRITZ RUDOLF GÜNTSCH, *Einführung in die Programmierung Digitaler Rechenautomaten*, Walter de Gruyter & Co., Berlin, 1960, 144 p., 24 cm. Price DM 24.

This introduction to the programming of digital computers assumes very little previous knowledge of computational procedures. The various steps in the preparation of a computer program are explained with one particular computer, the Z22, in mind. A large number of specific problems are worked out in considerable detail by giving the algorithm, the flow chart, and the detailed coding. Much attention is paid to the discussion of loops, and subroutines are treated with considerable detail. Since the machine code of the Z22 contains only very few instructions—not even a multiplication instruction—nearly all practical coding has to be done in an external code. Therefore, all the illustrative coding in the book up to the second-to-the-last chapter is done in an external code.

The book is devoted exclusively to introducing the reader to programming. It, therefore, does not contain any chapters on numerical analysis or the details of machine logic.

F. THEILHEIMER

- 87[Z].—ALLEN KENT, Editor, *Information Retrieval and Machine Translation*, Part 1, Interscience Publishers, Inc., New York, 1960, xv + 686p., 23 cm. Price \$23.00.

This book contains twenty-one of the papers presented at the International Conference for Standards on a Common Language for Machine Searching and Translation, sponsored by Western Reserve University and the Rand Development Corporation, and held in Cleveland, Ohio, September 6–12, 1959. It is the first part of Volume III (in two parts) of the series entitled *Advances in Documenta-*

*tion and Library Science*. The papers, plus discussions and numerous references, afford a good indication of work now in progress on mechanizing the storage and retrieval of information, particularly as it relates to scientific documentation.

The first article (by Allen Kent) reviews the current state of the art in machine literature-searching and translation. It includes in tabular form a comprehensive survey of searching devices and analytical techniques under investigation. Additional tables provide a significant sampling of the recent literature bearing on machine translation. These are arranged according to institution and investigator, subject field, languages involved, and type of equipment used. It is interesting to note that either English or Russian is the target language in nearly every one of the more than one hundred different investigations cited, while the most frequent source languages are Russian, English, French, German, Chinese, and Japanese, in approximately that order of frequency.

The remainder of the book is devoted to topics connected with information retrieval. The papers are by workers from such fields as library science, patent law, engineering, chemistry, medicine, the behavioral sciences, linguistics, symbolic logic, and computer technology. The majority of them deal with problems in organizing, indexing, and coding information for mechanical processing.

The proclivities of each discipline are revealed by the varied approaches to the problems. The librarian is interested in enumeration of characteristics which serve to describe and classify things; the patent searcher is concerned primarily with structure and function, and he needs elaborate means for specifying interrelationships among assemblages of components; the chemist must be able to search in terms of properties, composition, and arrangement; the medical researcher struggles to elucidate information through techniques of statistical analysis and correlation of recorded observations; the linguist has to do with syntactic and semantic analysis of messages to discover the *intentional* as well as *intensional* implications which they convey; the logician manipulates information in symbolic form and tries to formulate a consistent set of rules for drawing valid inferences and selecting optimum strategies in a given frame of reference; and the systems designer seeks to provide a means for satisfying the multitudinous requirements of the potential users on an economic basis. Such a diversity of emphasis is characteristic of a field which is evolving rapidly.

Nevertheless, the problems of indexing and searching technical literature, patents, chemical structural formulas, medical records, etc., do have a common basis, even though they may differ in a number of nonessential features. There is general agreement that some kind of coordination of descriptors is necessary for efficient retrieval, but that something more sophisticated than simple conjunction of characteristics will be required to cope with the increasing size and complexity of subject matter files. More particularly, it has been found in encoding subject headings that the relations expressed by the product, sum, and difference of classical logic are not sufficient by themselves for achieving the desired degree of specificity. A reasonably complete representation also requires a means for denoting the syntactic relations between descriptors.

A number of schemes are discussed, particular importance being attached to the rules for analyzing subject matter and to the notation for expressing the results of the analysis. These include the use of modulants, interfixes, and role indicators to

specify the types of relationships between terms and the use of phrase boundaries, parentheses, or other special marks of punctuation to segregate groups of associated terms. Such devices greatly increase the discriminating power of an index by preventing false associations of terms.

However, there is an evident lack of agreement on the means for attaining these desirable ends. This stems from difficulties inherent in linguistic analysis. It is necessary to discover and to define the syntactic relationships which are significant in mechanical indexing and searching. These are sometimes elusive, although such categories as species, genus, part, whole, factor, process, product, and attribute are widely recognized. The problem of synonyms (or near synonyms) and recognition of specific-generic connections can best be handled through the preparation of technical dictionaries and thesauri. These will provide a semantic analysis of complex notions in terms of more elementary concepts.

Some concern is voiced that the potential gain in accuracy of representation may be offset by an increase in ambiguity of indexing as a result of the greater complexity of the task. No procedure can be considered satisfactory unless there is a reasonable prospect that two different workers indexing an item will arrive at substantially the same description. Further research on the development of information retrieval languages is needed.

The remaining articles describe the capabilities and limitations of several existing or proposed systems. They discuss the types of information services provided, the way in which the files are set up, and the types of mechanical devices used. Equipment mentioned includes Minicards, Magnacards, the WRU Searching Selector, the IBM 7090, and the new GE-225 computer system with special features for performing searches.

The book is recommended to those who are currently working in some area touching upon information retrieval and to any others with a special interest in this fast-growing field.

THOMAS S. WALTON

Applied Mathematics Laboratory  
David Taylor Model Basin  
Washington 7, D. C.