

AN IBM CARD INDEXING SYSTEM FOR LITERATURE REFERENCES*

BY R. A. BONHAM, F. A. MOMANY, AND M. L. DRUELINGER

Chemistry Department, Indiana University, Bloomington, Indiana

A comprehensive set of abstracts of articles on molecular structure and certain related topics is being compiled on IBM cards at Indiana University. The details of the abstracting system are described below. A copy of the abstract card deck may be obtained from the authors for the price of postage and cards.

The proposed method makes use of IBM punched cards. These cards have 80 columns, each of which may be punched in 36 distinct ways. The information coded in these cards then may be retrieved easily by use of an IBM card sorter. If n columns are used to code a particular item then there exist at least 36^n different coding possibilities so that the amount of information which can be put on one card is quite large. As an added advantage a short abstract is typed on the back of each card before it is coded. It has been found that the punching of information onto a card does not usually lessen the legibility of the typed abstract.

The coding procedure outlined below for handling literature references in the field of molecular structure could be used in a modified form for other areas of interest. Columns one through three on each card are reserved for a three letter code representing the name of a journal. The code closely follows that adopted by Sutton¹ and a master sheet containing the journal codes is maintained for easy reference. Columns four through six contain the volume number, columns seven through eleven the page number and columns twelve through thirteen the year.

The authors names are coded according to their initials and in order of their appearance on the article. Columns fourteen through twenty-two are used for this purpose which allows three authors to be coded with first, middle, and last initials or first initial and the first two letters of the last name or some other suitable combination. Again a master code list is maintained for the author's initials.

Columns twenty-three through twenty-eight are used for coding the specific approach to molecular structure utilized by the article.

This space allows coding of up to two three letter codes for specific experimental and theoretical approaches to molecular structure. The code follows that utilized by Sutton as closely as is possible. Column twenty-nine uses an additional numerical code to subdivide the general area of molecular structure into broad sub-areas such as theory, structure, apparatus, and techniques.

The empirical formula of up to three chemical compounds may be coded in columns thirty through seventy-six. Eleven columns are used for the formula of each compound, coded as it is found in the "Handbook of Physics and Chemistry"² and four columns are used to designate isomeric features or pertinent ring systems of each compound. If an article did not deal with specific compounds this space could be utilized in a different manner. The last four columns (seventy-seven through eighty) may be used to label the cards numerically. This space could, however, also be used to store additional information, as it is not essential to keep track of the total number of cards.

In Fig. 1 a typical coded IBM card is shown from the punched side and also from the side with the typed abstract. When decoded the punched side would read the Journal of the American Chemical Society (JAS), Vol. 75, article begins on page 1542, year published was 1953, the authors names were J. W. Coutts, and R. L. Livingston (determined by use of the code sheet for authors using the punched initials JWC and RLL); the experimental method used was electron diffraction (E); the article deals with the experimental determination of molecular structure (1); and the compounds involved were tertiary-butyl chloride $[(CH_3)_3CCl]$, 2,2-dichloropropane $[CH_3CCl_2CH_3]$, and 1,1,1-trichloroethane $[CH_3CCl_3]$.

At the time of writing about four hundred articles had been abstracted and this total reference set occupies a space about three inches by four inches by seven inches. This system has the advantages that it makes use of readily available equipment and reference collections

*Contribution 992 from the Chemical Laboratories of Indiana University.

