Storage of gas chromatographic data using key-punched IBM cards

During the past few years, the use of gas chromatographic techniques has grown at a very rapid rate. As a result, more and more data are being published, but little work has been published on compiling the data in a form convenient for use. SPENCER AND JOHNSON[1] have reported a method of storing data on punched cards; however, their method is limited to hydrocarbons. This paper is a brief description of a system using IBM cards for the storage of gas chromatographic data for a wide variety of compounds.

Two decks of IBM cards are used. One deck consists of name–formula cards (NFC) and the other of gas chromatographic data cards (CDC). The NFC contains the compound name in standard punched card alphabet, molecular formula, and a cross-reference serial number. The original NFC were reproduced from the NFC being used by infrared adsorption spectroscopy groups. Cross-reference serial numbers were assigned to these cards after cards for inorganic materials, polymers, salts, complex compounds, and unstable materials had been removed from the deck. The CDC contains key-punched codes for the liquid phase, solid support or active solid, reference material, relative retentions, column temperature, type of compound (such as ester, aldehyde, ketone, acid, alcohol, amino acid, or steroid), bibliography reference including journal, volume, page, year, and main subject of article, and cross-reference serial number of the compound.

The purpose of the NFC is to identify the compound for which the chromatographic data are listed. A given compound will, therefore, have the same cross-reference serial number, regardless of the number of times data have been coded for this compound. Also, if a given compound can be named in different ways, each name will have the same cross-reference serial number. The NFC are arranged in order of molecular formula; thus, the serial numbers are also in order of molecular formula.

Once the decks of cards have been produced, they can be reproduced, sorted, and rearranged rapidly by means of conventional punched card handling equipment. Thus, the information stored in the system can be rapidly retrieved. The cards can also be used to produce printed lists of all or any part of the information stored, and in almost any order desired. The following printed lists have been useful here:

1. NFC in order of molecular formula.
2. NFC in alphabetical order.
3. CDC arranged as follows:
   a. Major sort; liquid phase.
   b. Intermediate sort; reference material, solid support or active solid, and column temperature.
   c. Minor sort; relative retentions.
4. Compound name or names from merged NFC and CDC in order of molecular formulas. Under the name are all the data for that particular compound, as in 3.
5. A bibliography on gas chromatography in order of the subject of the article.

Lists 1 and 2 are useful for coding new data. Lists 3–5 are used for rapid retrieval of data.

The system can be applied to practical problems by using either the cards and machine sorting or the printed lists. So far, the authors have used the printed lists for:

1. Selection of the liquid phase for separation of a mixture of two or more substances.

2. Tentative identification of unknown materials by comparison of the measured relative retention of the unknown with pertinent relative retentions in the lists. Two or more columns are used for a more positive identification[2].

3. Easy access to literature in which work of interest is described.

4. Comparison and correlation of data from different sources.

Several spaces on the CDC are unused in the present system. It is expected that some of these will later be assigned to data not now being stored in the system. The ratio of liquid phase to solid support and the specific retention volume, for example, are not coded at present; however, since data are being reported more precisely than in the past, it is planned to code this information.

We are grateful to H. W. PATTON and M. V. OTIS for helpful suggestions on setting up this system for storage and retrieval of data; to A. D. ALLEN, C. S. MCCONNELL, Mrs. DORIS W. MOREHEAD, W. E. ROBERTSON, J. N. ROPER, JR., J. L. SANKS, A. L. STONE and M. R. WHITLEY for assistance in preparing the original NFC; to M. H. LEIMKUHLER and A. G. FISH for co-ordinating the work done with the IBM machine card handling equipment; and to Mrs. FRANCES P. HUFFAKER for coding data from the literature.

Research Laboratories, Tennessee Eastman Company,
Division of Eastman Kodak Company,
Kingsport, Tenn. (U.S.A.)

J. S. LEWIS
G. T. MCCLOUD
W. SCHIRMER, JR.

[1] C. F. SPENCER AND J. F. JOHNSON, *Anal. Chem.*, 30 (1958) 893.
[2] J. S. LEWIS, H. W. PATTON AND W. I. KAYE, *Anal. Chem.*, 28 (1956) 1370.

Received February 2nd, 1961

Fractionation of sulphosalicylic acid filtrates on diethylaminoethylcellulose

In the course of a study on serum mucoproteins one of the aims of the authors was to investigate the high molecular substances that are not precipitable by sulphosalicylic acid, after the removal of the precipitant.

A method, formerly[1] developed for the separation of urine mucoproteins was used. It consists in the chromatographic separation of the substances on a column of