projects, products, and services planned. The major project now under way is a microthesaurus of plastics terms. The first phase is a controlled vocabulary project which will refine a list of terms to be used in the thesaurus. The second phase will involve term selection, classification, and the establishment of relationships between the terms. When the microthesaurus is produced, consideration will be given to computer storage with the ultimate goal of machine-aided indexing.

Now that PLASTEC's conversion to on-line is complete, products from the database are being planned. Starting with the January 1977 issue, PLASTEC now prepares a biweekly abstract bulletin with indexes. The indexes will be cumulated on a quarterly and annual basis. These products will be circulated locally at Picatinny Arsenal until PLASTEC is satisfied that the "bugs" have been worked out of the production system. Then, a commercial publication venture will be investigated, selling subscriptions to the general public. Another product will be off-line printouts of search results. These will be prepared at DDC and mailed to PLASTEC for forwarding to the customer. These products will be available in addition to the specialized reports published by PLASTEC's own subject specialists.

Services offered by PLASTEC will be expanding this year. PLASTEC now has on-line access to its own database as well as the technology and engineering databases offered by DDC. On-line access to the open literature will be possible through the Lockheed DIALOG system and the System Development Corporation ORBIT system. PLASTEC also has on-line access to a database it helped develop called COMPAT, which focuses on compatibility data between plastics, explosives, and propellants. By late 1977, preparation of another on-line service will begin. This database will be a materials property data bank which presents property data in tabular and graphic forms. The system will be interactive, so that engineers may specify performance requirements to obtain potential candidate materials for a given application.

PLASTEC also maintains complete and current microfilm files of military specifications and standards as well as the commercial and industry standards prepared by ASTM. Extensive collections of periodical and trade literature dating back to the 1950's are also maintained on microfilm.

Above all else there are six subject specialists who provide the really unique capability that makes PLASTEC more than just another information resource. This capability is their service in the evaluation of the literature and the generation of state-of-the-art reports. In addition they offer a wealth of experience in answering technical inquiries and in consultative advice.

REFERENCES AND NOTES

(1) A. M. Anzalone and G. Cohn, "Computer Techniques Applied to Compiling an Index of Ordnance Data", paper presented at the 142nd National Meeting of the American Chemical Society, Atlantic City, N.J., Sept 12, 1962.

(2) A. M. Anzalone, C. A. Brokars, and G. Cohn, "A Novel Index Tailored to Plastics Specialists", *Am. Doc.*, **15**, 191–195 (July 1964).

(3) J. L. Davis, G. Brincka, and D. W. Levi, "Computer Compatibility Data Retrieval Program", in the Conference on Compatibility of Propellants, Explosives and Pyrotechnics with Plastics and Additives, Picatinny Arsenal, Dover N.J., Dec 3–4, 1974, American Defense Preparedness Association, Washington, D.C., 1974.

(4) Defense Documentation Center, "IAC Input Procedures and Instructions", June 1976, unpublished draft.

(5) Defense Documentation Center, "Defense RDT & E On-Line System Terminal Operator's Manual", Appendix H, Remote Terminal Input Subsystem, DSAM 4185.13, June 1975.

(6) J. Nardone, "Computerized Material Property Data Information System", PLASTEC Note N31, Plastics Technical Evaluation Center, Picatinny Arsenal, Dover, N.J., June 1976, No. AD-A030675.

(7) R. S. Tompkins, "The PLASTEC Indexing System: A Consideration of Possible Adjustments Required by the DDC Program for Service Support to PLASTEC", Tracor Jitco 302-059-01, Tracor Jitco, Inc. Rockville, Md., Dec 15 1975, No. AD-A020031.

A Coded Data Bank for Chemical Instrumentation

REUBEN RUDMAN

Department of Chemistry, Adelphi University, Garden City, New York 11530

Received April 5, 1977

A coding system for describing low-temperature x-ray diffraction apparatus and techniques has been developed, along with a computer program which is used to prepare a coded and sorted bibliography. This system facilitates retrieval of various instrumental configurations and experimental techniques from the database. It can be extended to include other fields of chemical instrumentation.

INTRODUCTION

During the course of an experimental investigation, it is not unusual for a researcher to develop a new device or technique or to modify an old one. Very often these contributions do not warrant a separate paper and are included in the "experimental" section of a comprehensive paper which emphasizes the chemical aspects of the investigation. As a result, the apparatus and technique oriented portions of the paper are not abstracted and cannot be retrieved by later investigators. Consequently, one finds a great deal of duplication of effort in the area of instrument development.

In an attempt to alleviate this situation, several specialized bibliographies have been prepared in the past. However, the use of such bibliographies is often hampered by limited or incomplete descriptions of the apparatus and by poor, if any, indexing. The use of a system which encodes the pertinent descriptions of each device would (a) aid in identifying the major features of the particular device, and (b) permit the use of data handling techniques in locating and indexing various instruments and techniques.

Although computer stored databases have been used extensively for the storage and retrieval of chemical and structural information, e.g., in x-ray crystallography,[1] very little effort appears to have gone into the development of similar systems for the easy retrieval of instrument design and laboratory techniques. However, during the course of preparing a monograph on the apparatus and techniques used in low-temperature x-ray diffraction (LTXRD),[2] a coding system and associated data-handling computer program were developed. This system, which can be adapted for use with other types of apparatus, is described in this article.

CODING SYSTEM

Six major categories can be used to describe the cooling apparatus used in LTXRD studies[3] (Table I). Each low-temperature device can be assigned a six-digit code number,

Table I. Code Numbers Used in Classifying Low-Temperature X-Ray Diffraction Apparatus and Techniques

(A) Apparatus Code Numbers

First digit: Type of sample that can be studied
1. Any type
2. Single crystal
3. Powder
4. Metal
5. Protein
6. Neutron diffraction

Second digit: Type of cooling used
1. Cold-gas stream
2. Conduction (cryogenic fluid as coolant)
3. Conduction (thermoelectric cooling)
4. Conduction (mechanical refrigeration)
5. Joule-Thomson expansion
6. Immersion of sample
7. Immersion of camera
8. Use of cold room

Third digit: Method of frost prevention
1. Dry gas stream
2. Dry chamber
3. Evacuated chamber
4. Not given
5. None

Fourth digit: Minimum temperature attainable (Kelvin)
1. Less than 20
2. 20–78
3. 78–200
4. 200–260
5. Greater than 260
6. Not available

Fifth digit: Type of x-ray instrument mentioned

0. Any type
1. Debye-Scherrer camera (includes back-reflection)
2. Flat-cassette and Laue cameras
3. Guinier camera
4. Oscillation-rotation camera
5. Weissenberg goniometer
6. Precession camera
7. Diffractometer
8. Small-angle
9. Topographic studies

Sixth digit: Special characteristics

0. None
1. Horizontal
2. Vertical
3. Back-reflection
4. High-temperature also
5. High-pressure also
6. Weissenberg goniometer accessories
7. Cold-working at low temperatures

(B) Technique Code Numbers
1. Sample preparation if solid at room temperature
2. Sample preparation if liquid at room temperature
3. Sample preparation if gas at room temperature
4. Crystal growth if liquid or gas at room temperature
5. Techniques for randomly aligning powder sample
6. Reactive or radioactive sample
7. Applications

11. Purpose of low-temperature methods
12. Choice of temperature
13. Choice of cooling method
14. Temperature calibration
15. Characteristics of x-ray film at low temperatures
16. Alignment of crystal and/or apparatus
17. Special programming of automatic diffractometer
18. Correction for absorption due to sample holder
19. Refinement of disordered models
20. Review article
21. General low-temperature techniques

following the scheme outlined in Table IA, which characterizes the major features of the apparatus and becomes part of the database. In addition, a separate set of numbers is used to indicate any discussion of specific techniques employed in LTXRD studies (Table IB).

a

```
     RUDMAN,R.          SEE LIPPMAN AND RUDMAN (1976)

     RUDMAN,R.          SEE SILVER AND RUDMAN (1971)

     RUDMAN,R.          SEE SILVER AND RUDMAN (1972)

408  RUDMAN,R. AND GODEL,J.  (1969)
     J. APPL. CRYST.  2, 109-112
     AN AUTOMATIC LOW-TEMPERATURE APPARATUS FOR
     SINGLE-CRYSTAL DIFFRACTOMETRY
              111300

409  RUDMAN,R. AND POST,B.  (1968)
     MOL. CRYST.  5, 95-110
     POLYMORPHISM OF THE CRYSTALLINE METHYLCHLOROMETHANE
                                2    4                 COMPOUNDS

410  RUHEMANN,B.  (1935)
     PHYSIK Z. SOWJETUNION  7, 572-582
     NEW CAMERA FOR LOW-TEMPERATURE X-RAY DIFFRACTION WORK
              323110

     RUHEMANN,B.        SEE RUHEMANN AND RUHEMANN (1937)

411  RUHEMANN,B. AND SIMON,F.  (1931)
     Z. PHYSIK. CHEM.  B15, 389-413
     THE CRYSTAL STRUCTURE OF KR, XE, HI, AND HBR
              323210         3

412  RUHEMANN,M.  (1932)
     Z. PHYSIK  76, 368-385
     X-RAY INVESTIGATION OF SOLID NITROGEN AND HYDROGEN
              323110         3

413  RUHEMANN,M. AND RUHEMANN,B.  (1937)
     LOW TEMPERATURE PHYSICS
     CAMBRIDGE UNIVERSITY PRESS,  PAGES 103-113
                             20

414  RUHL,W.  (1954)
     Z. PHYSIK  138, 121-135
     X-RAY INVESTIGATION OF CONDENSED THIN FILMS
     AT LOW TEMPERATURES
              323111

     RUSAKOVA-LUKOVSKAYA,N.Y.   SEE BAGARYATSKII ET AL.
                                                     (1951)
     SABINE,T.M.        SEE COPPENS ET AL. (1967)

     SAINI,H.           SEE WEIGLE AND SAINI (1936)

415  SAKURAI,T. AND SUZUKI,T.  (1959)
     NORELCO REPORTER  6, 122,128
     A LOW TEMPERATURE ATTACHMENT FOR THE NORELCO X-RAY
              312372                          DIFFRACTOMETER
```

b

```
111300  -  522
           488
           467
           437
           408
           407
           372
           307
           305
           297
           243
           240
           231
           221
           212
           180
           177
           162
           145
           133
           114
            86
            69
111306  -  401
           353
           144
111343  -  265
           103
111400  -  197
112300  -  384
           327
           158
             4
             2
112306  -  117
112340  -  264
112342  -  505
           392
112356  -    6
112371  -  479
112390  -   65
             9
113600  -   12
114300  -  474
           296
           213
```

Figure 1. Sample output of program. (a) Bibliographic listing: note the program-assigned reference identifier numbers and the code numbers printed as the last line of each entry. (b) Sorting on the basis of the six-digit number: reference number 408 is listed under code number 111300.

For example, the apparatus code number 111312 following a reference means that "the LTXRD device described in that article can be used with any type of sample, cools the sample with a cold gas stream, prevents ice from condensing on the sample by means of a concentric warm outer gas stream, can reach a minimum temperature between 78 and 200 K (i.e., the liquid nitrogen range), and was designed for use with a Debye–Scherrer camera with a vertical sample axis". Similarly, the technique code number 4 indicates that the cited paper contains a discussion of "crystal growth at low temperatures if the sample is a liquid or gas at room temperature". These code numbers were included as part of the bibliographical reference list which was stored on a disk file.

COMPUTER PROGRAM

The literature citations are entered into the database in the following format:

Last Name, Initials (Year of Publication)
Journal Title, Volume Number, Pages (inclusive)
Title of Article
Code Numbers

As each citation is encountered, the program assigns it a sequential identifier number and prints this number along with the complete reference in the above format. Each citation can use as many records as are necessary (e.g., for multiple authors or long titles), with the record preceding the code-number record having a "1" punched in column 2. This "look-ahead" feature warns the program to read the following record (code numbers) in a special format. The apparatus code number is read twice, once as a single six-digit number and once as six individual code numbers. The various code numbers for each reference are stored in several arrays along with the identifier number.

After all the references have been read, numbered, and printed in alphabetical order (Figure 1a), the program sorts the various arrays according to increasing value of the code numbers. The following sorted listings are printed: (a) six-digit apparatus code number, (b) tables (6) in which sorting is done for each individual code number, and (c) technique code numbers. In each table, the code numbers are printed followed by all the identifier numbers of those references in which that type of apparatus or technique is described (Figure 1b). As new references are added to the database (in alphabetical order), the program-assigned sequential identifier numbers will be changed. Thus one must always use the bibliography and numerical tables from the same database.

These tables allow one to locate quickly all references in which a particular type of instrument, device component, or technique is described. It is also possible to "design" an instrument by constructing a suitable six-digit number. The appropriate table can then be checked to see if a similar device has been described previously in the literature.

The program also contains an option to print the coauthors as cross-references. These names are contained on separate records (identified by a "3" in column 2) and are not assigned an identifier number by the program. Thus, even though they will be printed in the reference list, they will not appear as separate entries in the sorted tables.

Recently, a program has been written to reformat the database so that one complete citation is written per record. It is now possible to obtain selected bibliographies, in which the complete literature references are printed, for all articles dealing with any specific category. For example, one can list all articles in the database that were published in 1976, or all those describing a particular piece of apparatus, or some subset of components, or LTXRD technique.

ACKNOWLEDGMENT

This project was supported in part by Grant No. CHE 75-13935 from the National Science Foundation. The assistance of Frank Ely of the Adelphi University Computer Center and of Stuart Olson is gratefully acknowledged.

REFERENCES AND NOTES

(1) R. J. Feldmann, G. W. A. Milne, and S. R. Heller, "Crystallographic Data Retrieval and Display", *Trans. Am. Crystallogr. Assoc.*, **12**, 75–83 (1976).

(2) R. Rudman, "Low-Temperature X-Ray Diffraction: Apparatus and Techniques", Plenum Press, New York, N.Y., 1976.

(3) R. Rudman, "The Proper Description of Low-Temperature X-Ray Diffraction Apparatus", *J. Appl. Cryst.*, **10**, 209–210 (1977).

CASSI and the Compression of Journal Names in an Information Retrieval System

J. HEILIK and R. H. BURTON*

Canada Institute for Scientific and Technical Information, Ottawa, Canada K1A OS2

Received June 21, 1977

A data compression technique for journal names is described. This technique is enhanced by use of the Chemical Abstracts Service Source Index (CASSI). Implications of CASSI's use in an information retrieval system are then discussed.

The Canada Institute for Scientific and Technical Information (CISTI) began offering an on-line information retrieval system (CAN/OLE) on a national scale in April 1974. Experience with the system quickly indicated several potential bottlenecks which could restrict the usefulness of the system and/or increase its costs. Of critical economic importance was the one of limited availability of disk storage space for data.

CISTI buys computing services from the National Research Council (Canada) Computation Centre. With file sizes soaring into the millions of records, the disk space at the Computation Centre was quickly filled, and additional disk drives, dedicated to the sole use of CAN/OLE, had to be acquired. These devices are expensive, so priority was placed on keeping the requirement for them as low as possible. One way is to find methods of compressing data. Several techniques were found suitable and are used currently; the subject of this paper is the one dealing with journal names.

A journal may have several articles per issue, several issues per volume, and several volumes referred to in OLE. Each article gives rise to an OLE record; each record contains a journal name as part of its bibliographic citation. Consequently, all records coming from a given journal will have identical data in the journal name field. Early in the life of CAN/OLE, it became evident that although the files contained many records, the number of journals was relatively small. Hence, this duplication of data became fairly significant.

Based on the above observation, a simple compression technique was developed. All journal names were removed from the base files, and replaced with a five-character CODEN (the standard ASTM CODEN stripped of its check digit). A central file of journal names was created in CODEN order. Whenever a record is displayed at a terminal, a name is automatically fetched from the journal file and displayed with that record. To the person at the terminal, it appears that all records have journal names. The link between records and the journal name file is, of course, the CODEN.

Removing journal names reduced storage requirements by 13%. In May 1977, there were approximately 4.5 million records in the on-line files; this modest 13% is thus equivalent to 585 000 records for which no additional space is needed. Evidently, when very large numbers are involved, even the smallest gains are significant.

A common problem faced by a customer of CAN/OLE, or most other bibliographic information retrieval services, is that having found a desirable set of references, how does he find the corresponding documents? Creating a central journal name file provided CISTI staff with an opportunity of alleviating this problem somewhat. Being relatively small and easily controlled, the name file was amenable to experimentation.

In mid-1975, CISTI approached Chemical Abstracts Service (CAS) to suggest an experiment with the Chemical Abstracts Service Source Index (CASSI). This Index was of interest because it contains not only unabbreviated journal names, but a record of which libraries subscribe to the various journals. Further, it contains all journal names found in two of the CAN/OLE data bases (*Chemical Abstracts Condensates* and