

## REPORT OF DISCUSSIONS<sup>1</sup>

G. A. Wilkins

Royal Greenwich Observatory, Herstmonceux Castle, U.K.

### SUMMARY

This report summarises the discussions on the topics of the five invited papers, and on other matters of general interest that were raised by the contributed papers. The principal headings are: standards for datafiles; the influence of acquisition and processing techniques; the critical evaluation of data; the designation of astronomical objects; the distribution of data; survey of facilities; the future role of data centres; and recommendations on IAU activities.

### 1. STANDARDS AND METHODS FOR DATA HANDLING

#### 1.1. Database management systems

In the first invited paper M.S. Davis (1) remarked that astronomers were amongst the first to use punched-card machines and computers for processing large data sets, such as the General Catalogue, but astronomers have not yet taken advantage of the facilities of data management systems that have been developed for other purposes. He made the point, which was later stressed again by Westerhout and others, that each datafile which is to be made available for processing by computer techniques should, like a good computer program, be self-documenting; that is the file should contain a full description of the data that it contains, and information about format, sources, precision, ranges of validity, peculiarities, formulae, etc., should be included. Such

---

<sup>1</sup>Papers are referred to by authors and numbers as shown in the table of contents.

information is now often, but not always, given in the prefaces and footnotes of printed catalogues, and is usually omitted when the catalogue is transcribed for computer processing; the new techniques of data management allow for the inclusion and retrieval of such information.

Davis, in reply to a question by Walter, said that the costs of file maintenance (deletions, additions and corrections) depend on the file structure; they can be enormous for a sequential file but modest for a list structure, and the need for file maintenance must be considered in the planning phase. When catalogues are transcribed from printed to computer-readable form, it is more efficient to omit computable quantities from the transcription; it would in any case be necessary to recompute them for checking purposes.

In a later paper [Harten and Spoelstra (11)], Spoelstra described the data formats used in connection with the Westerbork radio telescope, with special reference to the special format for "transport tapes", i.e. for magnetic tapes which are to be sent to other institutions. He considered that there should be standard formats and a restricted set of tape characteristics for such tapes; in particular, he suggested that the structure of the descriptive label at the beginning of a datafile should be the same for all branches of astronomy. He also suggested that the label should include the name of a person who could be contacted about the file, as well as the technical parameters of the instrument and the descriptions of the contents and format of the data.

## 1.2. Errors in datafiles

M.S. Davis (1) also drew attention to the nature of errors, to the techniques for their detection and to the vital importance of ensuring that datafiles are not corrupted or lost during use and maintenance (or by unauthorised alteration). R.J. Davis commented that errors occur during the initial collection, reduction and publication of data, and can only be detected later when they conflict with later, independently acquired data; bibliographic and personal problems can arise if such errors are corrected unilaterally. M.S. Davis considered that the author and the data centre that distributed the file should be informed, and other users should be notified through the publication of errata; problems need only arise if an author is negligent or obstinate. In a later paper Bidelman (17) pointed out that errors are sometimes not treated properly, and many authors prefer to allow their errors to go unnoticed. Dixon stressed that any errors found in a catalogue should always be reported to the author; it is unfair to him if this is not done.

In this connection Jaschek requested that details of errors in datafiles be sent to the Stellar Data Centre at Strasbourg and he offered to publish them in the Information Bulletin of the Centre. Duncombe suggested that the U.S. Naval Observatory would publish errata lists in its Circulars on the data available in machine-readable form. R.J. Davis suggested that each distribution centre should send details of errors to those to whom it had sent copies of the datafiles.

Collins asked whether experience had shown that optical-character-recognition machines had cut down the amount of proof-reading required when transcribing printed material for computer processing. M.S. Davis replied affirmatively, but pointed out that it will be a long time before it will be possible to read all documents automatically; in the meantime he recommended the use, where possible, of key-to-tape or key-to-disc systems, rather than the intermediate use of punched cards.

### 1.3. Other aspects

Many of the later papers and discussions were related to the need for standards for improving the quality of the data and for standards for facilitating the distribution of data. Section 3.2 on the presentation of data in the primary literature and section 5.1 on the publication of data on magnetic tape are of particular relevance.

Wilkins drew attention to the Task Group on the methodology for handling space- and time-dependent data that had been set up by CODATA on the recommendation of its Advisory Panel on the Geosciences, of which he is Chairman. The Chairman of the Task Group will be a geographer, Dr. R.F. Tomlinson, and it will contain scientists of many disciplines, including astronomy, who will be experts in the handling of data; it will seek to identify common problems and the possibilities for wider use of solutions developed in one discipline. Wilkins expressed the hope that astronomers would both contribute to, and benefit from, the work of this Task Group.

## 2. INFLUENCE OF ACQUISITION AND PROCESSING TECHNIQUES

### 2.1. The next 10 years

The principal theme of Westerhout's invited paper (9) was the necessity for astronomers to prepare for the great increase in the volume of data that will become available in, say, the next ten years because of the increasing sophistication of the instrumenta-

tion that is now coming into use for ground-based and space astronomy. On-line minicomputers will be used for the initial processing of the data, but even so the data for one night's work could fill one reel of magnetic tape. On the other hand, new computer devices will allow greater packing densities for the information and the new interactive techniques for data processing will lead to new uses for the data. He suggested that the IAU should set up a small panel to monitor the developments in computer hardware and techniques and to make recommendations about standard media and formats for data distribution. This suggestion was discussed again in the final session (section 8).

In the ensuing discussion, Nandy commented that the photographic plate is an economic storage device, and a plate can be measured quickly using new automatic measuring machines, which can reach the 24th magnitude. Westerhout agreed, but pointed out that photographic plates are not linear and that there are features in the universe which cannot be resolved by photographic plates; an array detector will be able to distinguish objects which have very large differences in intensity even when they are very close together. Baum added that there can be no doubt that array detectors will soon be providing much of our astronomical data and will therefore have a major impact on the operation of data centres; but not all array detectors are photon-counting devices with direct digital output. Some arrays, including charge-coupled devices, require analogue-to-digital conversion of the signal. Photon counting arrays, which directly provide a digital output with precise linearity, are severely limited in counting rate and are therefore suitable only for extremely faint sources; although improvements in photon counting arrays are foreseen, a major breakthrough will be required to fully overcome this limitation.

In his talk during the final session Baum (34) returned to the point that the conventional distinction between images and data is rapidly breaking down. Images now often exist in digital form, and this may even be the primary form (e.g. for planetary spacecraft and the Space Telescope). These digital images may be processed by computer to remove faults and apply calibrations; this initial processing is too costly and too specialised to repeat. This point of view was supported by Underhill, who remarked that a long learning process will be required, and by Westerhout who considered that it would be unrealistic to store raw, unprocessed data, but that it will be necessary to decide whether to hold both reduced data and the corrections for external factors that have been applied to the processed data to reduce them to standard form.

Mistrik questioned whether it would be more economical to represent the data by some appropriate mathematical expression (for example, by the coefficients of a polynomial), but Westerhout

considered that this would be limited by the presence of noise in the data. In his talk on data processing for space-based astronomy, Mistrik discussed in more detail the problems of where the data should be processed and gave an example of the rapid accumulation of data in a space project.

## 2.2. Other facilities

In presenting a joint paper with T.A. Nagy on retrieval techniques and graphics displays, Mead (24) gave an example of how overlays for use with Palomar sky-survey plates can be produced from a database formed by combining four catalogues and sorting the data according to plate area. The programs for this, and other tasks, have been written for an IBM 360/65 computer. Baum and Underhill also mentioned facilities for image processing and for interactive graphics that are available at the Jet Propulsion Laboratory and the Goddard Space Flight Center.

In describing the facilities offered by the Stellar Data Centre at Strasbourg, Ochsenbein [Ochsenbein, Egret and Bischoff (5)] drew attention to the way in which the observatories in France are linked to the IBM 360/65 computer at Meudon and hence have direct access to the data files of the Centre. Apart from this, there was, however, very little discussion of current applications in astronomy of interactive techniques.

Hauck later drew attention to the facility by which data may be printed on microfiche rather than on line-printer listings. Jaschek saw this as a way of making datafiles available to astronomers in institutions which did not have computing facilities, as well as a way of avoiding the high costs associated with printing by conventional means; one microfiche can contain up to 220 pages of text and costs less than 2 US dollars.

## 3. THE CRITICAL EVALUATION AND PRESENTATION OF DATA

### 3.1. Construction of catalogues

In an invited paper on the critical evaluation of data Underhill [Underhill, Mead and Nagy (15)] discussed the particular problems that arise in connection with catalogues that are available on magnetic tape. Many of these catalogues do not provide the information that is required if the user is to be able to judge the quality of the data. Even where the catalogue is derived from an earlier publication it is necessary that the information about methods of observation and reduction be included with the data, since users may not have access to the original publications,

especially if they work in a new institution that does not have a comprehensive library. She stressed that the error estimates for the data, and their basis, should be clearly given, and she gave specific examples of other relevant items that are necessary to a proper evaluation of the quality of astronomical data.

In the discussion Jaschek mentioned the situation in photometric photometry, where practically all the measurements before 1950 are lost because the authors did not specify precisely enough the transmission band used; the Lausanne group has, however, published the bands of most of the modern photometric systems. Gliese commented that details of instrumental techniques are necessary in the compilation of astrometric data to form a fundamental catalogue. The Astronomisches Rechen-Institut has asked that the observational catalogues to be used in the preparation of FK5 should be supplied on punched cards or magnetic tape, but it is not sufficient to be told that the catalogue was observed, say, "semi-absolutely"; rather the methods of determining azimuth, flexure and the magnitude equation should be stated explicitly.

In another session Dixon commented that it is still desirable to include computable quantities in the printed catalogues for the benefit of those who do not have computers at hand; those having access to large computers should not assume that everyone else is also well-equipped; the published material should be aimed at the widest audience.

The problems encountered in the formation of a homogeneous catalogue of photometric data were discussed by Nicolet [Nicolet and Hauck (16)]; in reply to a question by Nandy he stated that the observations were weighted according to their dispersions.

The following abstract, which was submitted by Mrs. K. Haramundanis (Smithsonian Astrophysical Observatory), was explicitly considered during the discussion on the role of data centres, but it is also relevant in this context.

The most useful, reliable, trustworthy and effective compilations of astronomical data are prepared by specialists in their field, since critical evaluations are significantly more useful than lists. To assist the integration of critical evaluations, an international standard of identifications must be established and adhered to. Data should be disseminated in both computer-accessible and book form. The work and expense of preparing compilations of critically evaluated data could be effectively shared if each data centre were to be responsible for a particular sky area only.

### 3.2. Presentation of data in the primary literature

During the discussions on the papers by Davis and Underhill it became clear that there was general agreement that many papers containing original data do not contain sufficient information to permit readers to assess the quality of the data and results. Wilkins drew attention to the CODATA "Guide on the presentation of experimental data in the primary literature" (CODATA Bulletin No. 9, 1974) and suggested that it would be worthwhile preparing a corresponding guide for astronomical data obtained by observations. Lortet pointed out that it would be necessary to persuade editors to accept that such details should be given. Wilkins agreed and commented that the guide would be addressed to editors and referees as well as to authors. He expressed his disappointment that astronomers were not changing over to the International System (SI) of units more quickly, especially as the delay will cause unnecessary difficulties for those who are now being taught physics, chemistry, etc., in terms of SI units.

Van Altena drew attention to the multiplicity of units for angular measure by asking what is the SI unit for angle. Wilkins stated that it is the radian but the other units are recognised. McCarthy suggested that the change to SI units would be facilitated by the insertion of conversion tables in the next edition of Allen's *Astrophysical Quantities*; Wilkins said that he would like to convince Professor Allen that he should use SI units throughout the next edition. He also referred to the usage in the CODATA guide of the terms "imprecision" and "inaccuracy" for the measures of the internal and external errors of measurements and results; he considered that their logical appropriateness outweighed the disadvantage of their present unfamiliarity.

Westerhout agreed with the desirability of giving estimates of both "imprecision" and "inaccuracy" but went on to ask how many astronomers are represented on the CODATA panels that propose the new terminology; the long-established practices of astronomy should be taken into account by such international panels. Wilkins replied that, as far as he was aware, no astronomers were members of the panel that drafted the CODATA guide on experimental data, but that CODATA was encouraging the preparation of other guides that were suited to their fields of application. Underhill considered that astronomers were too few in number to carry such weight, and that it will continue to be necessary to be aware of internal errors, external errors, random errors, and systematic errors, whatever their "official" names may be.

Luyten expressed the opinion that the abbreviation pc for parsec is absurd, and that analogy with other abbreviations suggests that the abbreviation should be ps. (This would, however, mean picosecond in the nomenclature of SI.)

#### 4. THE DESIGNATION OF ASTRONOMICAL OBJECTS

##### 4.1. Current problems

Several of the papers presented at the Colloquium were concerned in whole or in part with the problems caused by the multiplicity and ambiguity of the designation systems for astronomical objects lying outside the solar system. There were also many other comments on these problems and several suggestions for their amelioration. As the following notes and selection of comments will make clear, there was no general agreement as to the designation systems that ought to be adopted.

Mermilliod's paper (2) on the principles of a coded numbering system for photometric data and its application to open clusters opened up the subject. In response to comments and questions he stated that the system did not introduce any new designations; it provides a translation of existing identifications for practical use in machine-readable files; in principle it allows for the inclusion of faint stars which are identifiable only by provisional numbers on charts; where an open cluster has two or more different numbering systems, one of them is adopted as a base and transformation tables are constructed for the others.

Ochsenbein [Ochsenbein, Egret and Bischoff (5)] described the catalogue of stellar identification that is available at the Stellar Data Centre at Strasbourg. This originally provided cross-references between the designations used in half-a-dozen or so catalogues but it had since been extended and now covered some 400 000 stars. The formation of this cross-index had revealed some errors in the original catalogues. Underhill [Underhill, Mead and Nagy (15)] and Mead [Mead and Nagy (24)] both referred to a comparable cross-index that had been started independently at the Goddard Space Flight Center; the two institutions are now cooperating on this work.

Collins (4) drew attention to the ambiguities in designations that he had found while forming a bibliography of astronomical catalogues published since 1951. He suggested that there should be control of new designations. He had given W175 as an example of a designation for several different objects, and Luyten added that it also referred to the proper motion star Wolf 175. In reply to a question, Collins stated that the computer file is not yet closed, but he hoped that the bibliography would be published by INSPEC before the end of 1976.

Spite (7) discussed the same problem and suggested that each author should be expected to give the full reference to each original catalogue that is not included in a list of designations that

would be recognised by the IAU. Jaschek commented that it is only necessary to convince the editors of the leading astronomical journals of the desirability of this proposal. If they refused to publish papers which did not provide correct identifications the problem would soon be solved. Mead considered that the editors should require authors to provide coordinates for all objects described in a paper.

Underhill made the suggestion that, as many of the early, small catalogues are in obscure publications that are not available in newer libraries, it is desirable that such catalogues be transcribed into machine-readable form. Spite thought that this could be done with a relatively small amount of extra work and that it would be better to give a reference to where such a catalogue is now available, or to a newer and better catalogue, rather than to an old publication that is now out of print.

There were other suggestions for short-term remedies. For example, Luyten urged that if a star is contained in the Bonn Durchmusterung (B.D.) or the Cordoba Durchmusterung (Co.D.) no other designation should be used; these catalogues contain about one million stars and are now almost universally used. He later expressed his concern about the coining of new designations for stars that have been in the literature for a long time; for example, new designations are often given when a star is observed by, say, a spectroscopist even when it has previously been designated as a white dwarf; some such stars now have five or six different names. Bidelman suggested that there should be a recognised list giving the order of preference for the use of designations when a star was listed in more than one catalogue; in particular he considered that common names should be used in preference to catalogue numbers. Others felt that it would be difficult to get general agreement to such a list since the usefulness of a catalogue depends on whether it contains the information required for the application.

#### 4.2. Use of coordinates in identifiers

There seemed to be general agreement to the view that for the identification of data held in computer files a system based on the use of the coordinates of the object would be appropriate, but there were differing views as to the desirability and practicability of having a single system for all astronomical objects outside the solar system and as to the particular form that the identifier should take. The following notes give a selection of these views.

In his paper, Eichhorn (3) suggested that the identifiers should be formed from galactic rather than equatorial coordinates.

M.S. Davis questioned whether it might be better to use an inertial frame defined with respect of extra-galactic objects rather than a galactic frame which is defined with respect to the equatorial system. Westerhout commented that any fixed frame could be used and that the necessary coordinate conversions can be carried out at tape-reading speeds. Dixon pointed out that galactic coordinates are not particularly significant for extra-galactic objects. Van Herk suggested that space astrometry will provide an alternative system. Bidelman considered that it is more important that everyone should use the same equinox. Heintz drew attention to the imminent changes in the equinox and precession coefficients of the fundamental system. In his replies, Eichhorn emphasised that whatever the choice of system it will always be necessary to make transformations between the moving equatorial system in which we observe and the chosen system, and this transformation will be uncertain for coordinates of high precision.

Dixon later returned to the question of the choice of the standard epoch for use in catalogues. He argues that the proposal to be considered at Grenoble for the adoption of 2000.0 as the new standard epoch would involve much unnecessary work in changing existing catalogues and that therefore the epoch of 1950.0 should be retained indefinitely. Wilkins pointed out that the proposal was specifically related to the standard epoch of the new fundamental catalogue FK5 and of new planetary theories. Bidelman commented that the adoption of 2000.0 for the fundamental catalogue did not preclude the continued use of 1950.0 for the epoch of coordinates intended primarily for identification purposes; it will be necessary to continue to provide the appropriate precession coefficients for transformation between 1950.0 and current epochs. Westerhout supported this view and, on this basis, Dixon did not press his point further.

In connection with the problem of identifiers for non-stellar objects whose positions cannot be specified precisely, Wilkins drew attention to a technique used in a geographical data project in which the identifiers were formed by combining the degrees of latitude and longitude, then the minutes, and finally, the seconds; with this technique, the identifier need be no longer than the known precision demanded. Underhill commented that we need a technique that can be used efficiently in the computer, and the view was expressed by Mead that we should not perpetuate the use of sexagesimal measure in datafiles for computer use. It may be noticed that Eichhorn suggested publishing direction cosines, which also have the advantage that they avoid the singularity at the pole of a spherical coordinate system. McCarthy and others suggested the magnitude should be included in each identifier since this would help to guard against misidentifications.

There were divided views about the identification of stars in photometric sequences, clusters, and other dense fields. At present stars are often identified only by a number on a particular chart. Bidelman suggested that it ought to be practicable to determine coordinates for such stars with sufficient precision for identification purposes, using, for example, Palomar sky-survey charts; a central service could do this. He considered that the coordinates should be determined so that the data can be integrated into a general astronomical datafile. Hoffleit considered that charts will continue to be useful and that such coordinates would not be sufficient. Westerhout agreed on the need for charts and considered that coordinates should be given to the highest possible accuracy; they are important because they can be put in machine-readable form and avoid the necessity for publishing finding charts, which can be produced by computer when required. Argue pointed out that some finding charts do not resemble very closely what is seen at the telescope eyepiece; no difficulty over identification arises if the telescope has a setting accuracy of about 1"; when a telescope does not set to this accuracy it may be possible to set on an easily identifiable bright object, and to offset from this. He claimed that positions can be easily read from survey plates to an accuracy of  $\pm 3''$  using simply an eyepiece and to 0".5 using a simple X,Y measuring machine. Baum commented that the coordinates of stars belonging to a local magnitude sequence will not provide the complete information needed for a user to construct his own finding chart, because only a fraction of the stars in a field are selected for photometric measurements; if the author provides no finding chart, he will have to provide coordinates for many more stars than he measures photometrically.

## 5. THE DISTRIBUTION OF DATA

### 5.1. The publication of data

In his invited paper Hauck (23) considered the basic problem of how astronomical data should be published in view of the increasing rate of acquisition. He considered that there will be a continuing need for printed catalogues but that the basic support for them should be on magnetic tape. He described a scheme whereby tapes may be refereed through the journal system - the referee examines the description of the datafile that is published in the journal while a cooperating data centre verifies that a copy of the tape satisfies its specification and makes further copies available on request. The scheme allows for an author to get credit for his work, provides a reference and an adequate source of information about the data, and should filter out datafiles that are based on poor observing and reduction techniques. Such a scheme is offered by Astronomy and Astrophysics Supplements

in cooperation with the Stellar Data Centre at Strasbourg. Other centres will be required if the scheme is successful.

Batten expressed his concern about the role of the referee; the compiler of a catalogue may have been in touch with potential users and be in the best position to judge the best format in which to present the data; it would be unfortunate if a referee with strong opinions could force a compiler to adopt a presentation that was later found to be less suitable. Underhill had similar opinions and suggested that, since preparing and publishing a catalogue is very expensive, the refereeing should be done on the basis of trial data. Hauck agreed that it is necessary to define precisely the tasks of a referee (see note at end of section 5.4), but considered that his task is a very useful one and that he should examine the realisation as well as the intention. Argue favoured the scheme because of its implications for students; at present the submission of data is not regarded as having the same weight as the publication of a research paper, but the status of data compilation would be enhanced if it were known that the data had been submitted to a proper system of refereeing.

Collins reported a survey which had shown that the half-life for the citation of stellar data is the same (5 years) as that of the research papers themselves; he concluded that there is an urgent need to make the magnetic tapes available in a very short period. It is possible that the unexpected result of the survey is due to the incorporation of the older data in later catalogues or it indicates that astronomers do not, in general, search the older literature for data.

## 5.2 Sources of information

Several of the papers were concerned with services and publications intended to provide information about where data is to be found. Collins (4) described work on a bibliography of recent catalogues, including quite short lists of data. Spite [Ochsenbein and Spite (27)] described a computer-accessible bibliographic file on stellar data that has been formed through collaboration between the Paris and Strasbourg observatories. Walter commented that the consistent usage of the abbreviations for astronomical journals as adopted in Astronomy and Astrophysics Abstracts would simplify the interpretation of the stellar bibliographic file, and he enquired why a different notation is employed. Spite replied that the aim is to save punching time and computer storage, and so to save money; in his opinion such a limited number of abbreviations cannot produce serious ambiguities, and it is easy to build and publish a short table of the correspondences between the adopted and official abbreviations. In reply to questions by Mistrik and Straižys he stated that the contents of a file can be output to

microfiche using a standard package, and that the file includes all variable stars that have a designation or coordinate; there is no limitation on magnitude.

When presenting his survey of facilities Jaschek (33) drew attention to the existence of many examples of observations that had been unnecessarily duplicated; Bidelman and others confirmed this view that the data that are available are not used to the full extent that is possible. There are many indexes to data that are at present only available on hand-written cards; for example, Bidelman has an index on stellar spectra and Wood (21) described such a catalogue for eclipsing variables. There was general agreement that these indexes, in particular, and there may be others, should be transcribed to magnetic tape so that the information in them can be made more widely available. Parsons [Parsons and Wray (14)] reported that work had started on the punching of Bidelman's datafile.

Dixon (25) described the progress on the preparation of a master list of non-stellar objects. The list will give both the basic data and a bibliographic reference to the original source for each object. No attempt has been made to evaluate the data, and a single object may be given several times if the published positions are different. Jaschek suggested that it would be useful to collect later bibliographic references for all objects, but Dixon felt that there should be some consensus that this would be necessary. Bidelman said that he felt that this would be useful, and he had started to compile such a file; he asked how Dixon handled NGC or IC objects that are actually parts of other galaxies. Dixon replied that the NGC, IC and all galaxy catalogues are included in the master list; by looking at the positions and angular diameters of each object it is easily determined if one object lies within another; in many cases such objects are specifically identified in the remarks column. In reply to a question by Westerhout, he also indicated that there are plans to try to combine all entries that refer to the same object, but this is a major project in which many people can take part; the master list is the first step.

Lortet (26) described a project for documenting and cross-identifying non-stellar galactic objects, such as dark clouds and reflection nebulae. She drew attention to the need for a set of keywords for classification of such documents. Wilkins suggested that the revised version of the Universal Decimal Classification for astronomy (just prepared by a Working Group of IAU Commission 5) would provide a satisfactory alternative to the use of keywords; it includes numerical codes to indicate, for example, the type of object and the wavelength range of the observations.

Wood drew attention to the desirability of determining what the Commissions of the IAU are already doing in the matter of data dissemination, and gave the example of Commission 41 which distributes, twice a year, lists of references and research on individual objects. Some information on these activities has been recently published in the Information Bulletin of the Strasbourg Centre.

### 5.3. Requirements for new documentation services

In the discussion on Jaschek's survey of data centres (see section 6.1) several participants drew attention to the need for additional documentation services for special types of object. Bidelman noted that a new fairly extensive list of shell stars, together with a bibliography, will be included in the proceedings of IAU Symposium No. 70 on Be and shell stars: he continued by remarking that there are many other types of stellar object that merit continued cataloguing - pulsars, X-ray stars, infra-red stars, etc., to name but a few examples; further, galaxies, quasars and extra-galactic radio sources deserve extensive documentation; this does not appear to be being done at the present time, except for the brightest objects of these classes.

McCarthy commented on the catalogues for emission-line stars; the researches sponsored by the IAU at Prague had proved fruitful and, with the publication of the Henize Catalogue of Southern Be stars, this preliminary phase of the data collection is finished; a single working catalogue integrating the best features of all existing catalogues plus the new discovery lists can now be assembled at a data centre.

McCarthy also enquired about the present situation regarding lists and catalogues of stellar associations and open star clusters. Jaschek replied that the activity has effectively ceased. There is a great need for continuing the work on the Prague catalogue, but no persons seem to be willing to undertake this. Lynga agreed that the Prague catalogue of open clusters had served a useful purpose and that a similar bibliographic catalogue should be kept up; he suggested that making a tape of the references given in the IAU Reports on Astronomy by the President of Commission 37 would make a good start.

Westerhout, in reply to a request by Jaschek for information about documentation work in radio astronomy, stated that he knew of only Dixon's list of radio sources, Terzian's list of pulsars, and a private abstracting service for the members of the CSIRO Division of Radiophysics in Sydney; staff members scan all journals weekly for articles on radio astronomy and related subjects, write small abstracts with critical comments, and put them on tape.

Dixon stated that he receives a copy of the abstracts by airmail every two weeks or so, and that he believed the service was open to others. Jaschek commented that this is a nice example of a service which exists, but which is known only to the specialists.

Wilkins wondered whether the three principal abstracting services could assist in the collection of material for such specialised files. In some sciences the presence of data in an article is indicated by the attachment of a 'flag' or a 'tag' to the abstract. The extra effort required to identify data could, perhaps, be found by increasing the cooperation between the three services.

R.J. Davis suggested that the abstracting services, and institutional libraries, should receive information about catalogues that are available on magnetic tape in order that astronomers may be directed to these data sources by means of the common techniques for searching the literature. This would occur automatically if details of all such catalogues were given in the primary journals as suggested by Hauck.

#### 5.4. The "ownership" of data

During the discussion on Underhill's invited paper, Jaschek drew attention to the personal problems that arise as a result of the availability of new techniques for copying, updating, and merging catalogues. The authorship of a printed catalogue is normally clear and well recognized, even though the authors' names may not be given in citations. There are, however, examples where datafiles have been copied and then updated, or expanded, and no recognition of the original author has been given when the new datafiles have been used. Underhill commented that this is one more reason why every catalogue, whether printed or only available on tape, should be accompanied by documentation that gives the sources of the data; proper credit must be given to those who prepare catalogues; machine-readable catalogues are a very necessary form of publishing; all rights envisaged in copyright law that are commensurate with the requirement of open transfer of scientific information must be preserved.

It may be remarked here that one of the tasks of the referee of a machine-readable catalogue (see section 5.1) should be to verify that the author has given proper recognition of the sources of the data. Further, Dixon suggested that when a catalogue is duplicated by a centre for use at other institutions, the original author should be sent a list of those to whom his catalogue has been sent; this will enable him to know how widely his work is being used and to whom he should send announcements of updates and corrections.

## 6. SURVEY OF FACILITIES

### 6.1 General survey of data centres

In the fifth invited paper Jaschek (33) gave a summary of the replies that he had received in response to a questionnaire that he had distributed earlier in the year. The survey revealed an extremely wide variety of functions and policies. It showed that work on non-stellar objects is less well organised than that on stars, and this in turn is less well organised than that on solar system objects and phenomena. It also showed that astronomers are very slow to adopt new ideas and they do not utilise fully the services available from the data centres. It is the intention that the results of the survey will be published by CODATA, as part of the production of a directory of data sources for science and technology. The discussions on documentation services that followed this paper are reported in section 5.3.

Jaschek had previously drawn attention to the setting up at Potsdam of a centre containing a copy of the datafiles available at Strasbourg; this is intended to facilitate the use of these data by the countries of Eastern Europe; and other similar centres could be established elsewhere. Duncombe suggested that the data sets at USNO and RGO could also be made available to such centres; he also drew attention to the increasing efforts to coordinate the work of the centres, and drew attention to the service provided by the Bureau des Longitudes, in Paris, which distributes cards giving information about the availability of astronomical ephemerides and catalogues.

### 6.2 Individual services and projects

During the Colloquium, many of the participants described briefly data services and projects with which they are concerned. Some matters of general interest that arose during these presentations and the ensuing discussions are mentioned in other sections of this report. The following notes are therefore intended only to record a few points of specialist interest; further details are given in the papers themselves.

Terashita (30) described the activities of astronomical data-systems groups in Japan; its first task was to find out what datafiles were already available on tape in Japan; he felt there is a need for a set of standards for coding data, such as spectral types.

Abalachine read a paper by Polojentsev (31) on the computing facilities at the Pulkovo Observatory; the suggestion was made

that astronomers should be provided with algorithms to compute ephemeris data. Wilkins commented that the Nautical Almanac Offices of the US Naval Observatory and the Royal Greenwich Observatory were planning to publish such algorithms but he considered that there would still be a requirement for the publication of the ephemerides.

Worley presented the paper by Fiala and Seidelmann (32) on the data services now provided by the US Naval Observatory; a new Federal service for supplying copies of tapes might be recommended.

Gliese (6) described plans for the third edition of his catalogue of nearby stars; he expected it to contain about 1700 stars with parallaxes greater than  $0''.045$ . Walter asked whether it would be possible to label the stars that are emitting at radio wavelengths; Gliese replied that this is under consideration. Luyten stated that he had just completed a new catalogue of stars with annual proper motions greater than  $0''.5$ ; it contains 3600 entries, of which about 2500 are in neither the B.D. nor the Co.D.; he hoped that it would be possible to prepare and publish identification charts for all of these stars.

Walter (12) gave the preliminary results of a search for observational data from radio interferometry that would be suitable for astrometric purposes.

Nandy (13) compared systems of stellar classification from low-dispersion ultraviolet spectra as a preliminary to the preparation of a catalogue based on observations made in a satellite experiment.

Parsons [Parsons and Wray (14)] discussed the preparation of a catalogue of ultraviolet spectroscopic and flux data for early-type stars; it will be based on merging new satellite data with existing catalogues, and will be accompanied by a bibliographic file.

Abalakin brought greetings from B.V. Kukarkin, who was unfortunately not well, and then read the paper [Kukarkin, Kholopov and Kireeva (18)] describing the work in the USSR on variable stars, including the tentative plans for the fourth edition of the General Catalogue on Variable Stars. Bidelman congratulated the Soviet astronomers for their extremely fine work on the documentation of variable stars and looked forward to the new catalogue; he considered that the new tables of alternative designations would be very useful. In reply to questions Abalakin said that he expected that the catalogue would be available on magnetic tape and that the introduction would be in both Russian and English.

Argue (19) described the way in which he and Miller had prepared a catalogue of photometric sequences; the aim was to put the "consumer" in contact with the data and the observer, and no attempt was made to provide critically evaluated data; the observers could probably provide better current data than could be found, less quickly, in the literature. The ensuing discussion on finding charts is reported at the end of section 4.2.

Batten (20) then discussed the preparation of a catalogue of spectroscopic binaries; in contrast, the aim here is to select the best set of elements to represent the behaviour of each system. The notes on the individual systems represent an important part of the catalogue and justified the production in book form. In reply to Mead, he said that the new catalogue will be available on magnetic tape, but he did not consider that it would be practicable to include the notes. Westerhout argued that the notes should go with the data on the tape, and could easily be included in a separate alphanumeric subfile. Jaschek suggests that the notes should mention any photoelectric observations made when looking for eclipses; he stated that the group at Toulouse will publish a bibliography of all orbits for each star.

Wood (21) described the history and use of the manual card catalogue on eclipsing variables now maintained at the University of Florida. He hopes to produce a new edition of the finding list for observers and to transcribe the file to punched-cards - this led to the suggestion that it is better to avoid the intermediate use of punched cards, and to edit the file via a computer terminal.

Hoffleit [Hoffleit and Jaschek (22)] asked for suggestions as to what changes should be made in the next edition; she said that the catalogue aims to cater for users who do not have access to computers. Slettebak asked that rotational velocities be considered. McCarthy expressed the hope that the temptation to reduce the size of the type would be resisted since the catalogue is used during the making of observations; Hoffleit replied: "Older astronomers of my age group are well aware of this problem!" R.J. Davis suggested that photometric data should be given in the *ubvy*  $\beta$  system. Walter asked that both Córdoba and Cape Durchmusterung numbers should be given when both have been allocated.

Worley (28) described the arrangements for the maintenance and use of the master copy of the catalogues for visual double stars at the US Naval Observatory; there is an Index Catalogue and an Observation Catalogue; old as well as new observations are being punched to make the catalogue as complete as possible. Gliese asked whether the catalogue included star pairs of common proper motion even if their angular separations are large. Worley stated that there is now no arbitrary separation limit, and previously omitted wide pairs are being recovered where possible.

Van Altena [van Altena, Hoffleit and Smith (29)] discussed the systematic differences between trigonometric parallaxes obtained at different observatories and described plans for the construction at Yale of a new general catalogue of trigonometric parallaxes. Gliese drew attention to Strand's view that parallaxes based on the use of astrometric reflectors are more likely to be free from systematic error and that the true parallax system would be about 0".002 greater than the Allegheny system. Van Altena commented that it will be very difficult to determine the zero point of the system of parallaxes until the causes of the systematic difference are identified. Modern parallax observations may define a reasonably good, uniform system of relative parallaxes, since there appear to be little, if any, differences between the parallaxes of, for example, the USNO and Yerkes.

## 7. THE FUTURE ROLE OF DATA CENTRES

Throughout the colloquium there were brief discussions on the future role of data centres and so the discussion during the final session took on the character of a short extension and review of the earlier discussions.

Baum (34) used the example of the planetary-photograph centres as the basis for his talk, which included the following points: data centres operate best if staff at the centre are engaged on research that uses the data; the conventional distinction between images and data is rapidly breaking down; we have to distinguish between raw data, processed data, and reduced data; and the initial processing is too costly and too specialised to be repeated (see also section 2.1). In addition we have standard data that are obtained by combining together data from different sources that have been subject to a critical evaluation of their random and systematic errors.

Wilkins drew attention to the abstracts submitted by himself (36) and by Mrs Haramundanis (see end of section 3.1), and then put forward a view of how data centre operations might be organised for astronomy. He suggested that there should be several principal data-dissemination centres which would hold copies of the main datafiles and which would be prepared to provide copying and search services, and possibly facilities for visiting astronomers who wish to use the data interactively. They would cooperate with each other so as to share the tasks of merging catalogues, forming cross-identification indexes, etc. Each of these principal centres would serve a particular area and would, where necessary, make arrangements to obtain specialist data from one of the other centres. These centres should be in institutions in which astronomers are using the data for research purposes and providing new

data in their own fields. In addition there should be data-analysis centres which would be concerned with the compilation of catalogues of critically evaluated data in specialist fields; they would not be expected to have extensive computing facilities, and so they would normally pass their data to a data-dissemination centre, and would not deal directly with requests for copies of the files; they would, however, answer enquiries of a technical character about the data, and perhaps supply new data that had not yet been incorporated in the main file. Such an analysis centre might even be located with a dissemination centre, or have formal links with one, but the main aim would be to build on the informal arrangements which have worked so well in astronomy in the past, and which have permitted small institutions to make very valuable contributions to the astronomical database.

## 8. RECOMMENDATIONS ON IAU ACTIVITIES

During the final session there was an informal discussion about the action that should be taken to implement some of the suggestions for future activities that had been made during the course of the colloquium. On the Chairman's suggestion it was agreed that no attempt should be made to draft and pass formal recommendations; rather, any such recommendations should be considered by IAU Commission 5 (Documentation) at Grenoble, during the session concerned with the Working Group on Numerical Data. Bidelman wondered whether the name of the Group is sufficiently general to encompass the types of astronomical data that are not strictly numerical, such as descriptions of spectroscopic peculiarities, and he suggested the name 'Working Group on Astronomical Data'. Baum also considered that the name should avoid the inference that the group is interested only in conventional tabulations of reduced data. Eichhorn suggested that the Group should become a separate Commission but Wilkins drew attention to the considerable overlap with the other interests of Commission 5 and to the fact that the decision to associate the Group with Commission had been made only in 1973. Wood remarked that it would be easier to persuade the Executive Committee to accept a change of name than to establish a new commission.

There was general agreement to the Chairman's view that the Group should prepare a guide to the presentation of astronomical data in the primary literature.

The meeting then discussed Westerhout's suggestion that there should be a group to keep up to date with hardware developments. M.S. Davis spoke in support and suggested that the group could also produce broad guidelines as to standard formats, choice of data management systems, etc. Duncombe drew attention to the financial constraints that limit the equipment available to data

centres, and said that when equipment is upgraded efforts should be made to maintain compatibility so as to assure continued ease of exchange between data centres. Westerhout thought that the Group should report frequently and Jaschek offered to publish the reports biannually in the Information Bulletin of the Strasbourg Centre. The Chairman asked Westerhout and Davis to formulate recommendations to only these lines.

The discussion of the action required to overcome the problems caused by the proliferation of designations for astronomical objects led also to some discussion of technical matters that have been reported in section 4. Jaschek considered that the best approach is to convince the editors of leading journals that adequate identification should be given in every paper. Dixon drew attention to the danger that a resolution on nomenclature may be misunderstood; at the last IAU General Assembly, Commission 28 adopted a resolution that all future catalogues of galaxies, UV objects, etc should be designated by Parkes-type numbers, eg 2024-18; unfortunately some authors misunderstood this and created new names for objects that already had other designations; such action leads to confusion and provides no guide to readers as to where to find the original catalogue. Bidelman favoured the use of ordinary designations for stars brighter than the magnitude limit of the B.D. catalogue, since these would lead most easily to further documentation. Eichhorn and Westerhout both took the view that every object should have a precise, unambiguous designation. It was generally agreed that the problem needed detailed study by a small group, and that a new system could not be introduced until the General Assembly in 1979. The Chairman asked Bidelman and Eichhorn to prepare a recommendation on this matter.

There was no dissent to Parsons' view that the bibliographic files on stellar spectra and eclipsing binaries maintained by Bidelman and Wood should be made more widely available in computer-accessible form, nor to Lynga's suggestion that the Prague catalogue on stellar clusters should be similarly treated. There were, however, divergent views on the suggestion by R.J. Davis that an attempt should be made to compile on tape a comprehensive catalogue of photographic-plate collections. Underhill felt that the publication of such information should be left to the institution concerned. McCarthy commented that not all of the plates would be worth examining. On the other hand, Westerhout considered that a lot of valuable telescope time is wasted in duplicating observations. Baum said that the listing of plate collections and the dissemination of a merged list would be an exact analogue of the 118 000-item listing of planetary photographs already compiled and widely distributed by the Planetary Research Centre at the Lowell Observatory; this planetary "catalogue" has substantially increased the utilisation of the plate collections and has not

created any problems; he therefore did not agree with the reluctance to encourage observatories to make plate listings freely available and to merge them in due course. The matter was left for further discussion within the Working Group.

## 9. CONCLUSIONS

Delhaye had to return unexpectedly to Paris and so was not able to present his Concluding Remarks on the Colloquium. Instead, Garstang (38) reviewed some of the ways in which data could be made more readily available and intelligible to those who will wish to find and use the data. Heintz (37) preceded his vote of thanks to the organizers by discussing the role of the abstracting services and of libraries in the dissemination of data.

The Chairman expressed his thanks to Jaschek and the other staff at Strasbourg who had contributed to the organization of the Colloquium; he thanked the participants for their contributions in the presentation of papers and in the discussions; and he looked forward to the continuation of the discussions on the improvement of astronomical data services at the Grenoble meetings.