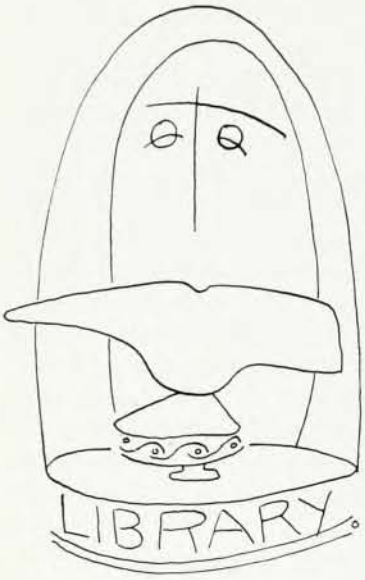


ABSTRACTS

Too little and too late is the reaction of most scientists to the short summaries intended to keep them informed of what is going on in their field now and to tag things nicely for investigators in the years to come. The American Institute of Physics and the American Physical Society are conducting a study of abstracting physics literature. The Director of this project reports on the points of view beginning to emerge in the first four months of the project's existence.

by ***Dwight E. Gray***

How important to American physicists are abstracts of the periodical literature in their field; how do they use abstracts; what do they think of the abstracts they now have; what kind of abstracts would they like to have? It is questions like these which the study now being conducted by the American Institute of Physics and the American Physical Society, under contract with the Office of Naval Research, hopes to answer. Although the work has not progressed far enough to warrant hard and fast conclusions on physicist opinion, it is likely that most of the major differences in viewpoint have been encountered by now.

How are Abstracts Used?

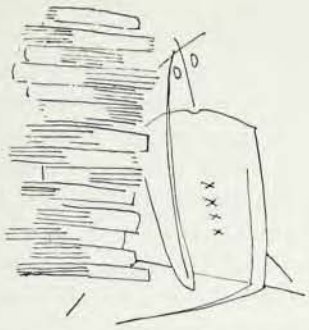
One of the surprising facts brought out at the Royal Society Information Conference, held in London last June, was that detailed, factual information about how scientific men actually employ abstracts, and for what purposes, was very limited. But there do seem to be two general functions which abstracts perform for physicists, and much of the searching for a base on which to build a plan for improved abstracting service must revolve about them. "Current awareness" is the name given to one of the functions of abstracts by a writer in Nature. This function is to aid the physicist in keeping up with what is going on, or to provide him with a definite research tool to give him either actual technical information or references to its source.

Another general function of abstracts is a "reference function," and concerns an abstract's usefulness one, two, five or more years after publication when it may serve in the preparation of bibliographies or in the study of all work done to date on a given subject or by a particular individual.

Some of the implications of these two functions are obvious. For example, prompt publication is considerably more important for current awareness than it is for reference, but extensive indexing is more important for reference than for current, general use. Other implications depend on what kinds of abstracts the physicist wants as well as why he wants them.

Kinds of Abstracts

Abstracts are usually thought of as falling into two general types, definitions of which were adopted at last year's Paris conference of UNESCO. An *indicative* (or descriptive) abstract is short, and is



written solely to enable the reader to decide whether he should read the original article. An *informative* abstract summarizes the article's major arguments and gives the principal data and conclusions which the abstractor considers make valuable contributions to knowledge or are likely to be of use to the specific class of readers for whom the abstract is prepared.

The separation between these types is not sharp and technical abstracts as a whole, instead of falling tidily into two well-defined categories, actually form a kind of continuous spectrum of varying informativeness. Also, by interpreting the term abstract a little more broadly one can include the idea of critical review which gives the reader, in addition to information about what the article contains, the abstractor's opinion of its importance and technical stature. A few of the existing abstracting services include this feature, usually combined with the informative type of abstract.

Most of the physicists queried have indicated a preference for indicative rather than informative abstracts on the basis that seldom if ever would they accept technical data and conclusions given in an abstract without themselves checking the original article. The ease with which original articles can be obtained varies greatly for different journals and usually this opinion is qualified in favor of some degree of informativeness in abstracts of articles which are hard to get.

The physicist who reads his abstract journal largely to help him keep up with what is going on in physics as a whole seems to prefer a kind of semi-informative product. Extremes of opinion encountered thus far have varied from the belief that simple titles and references would be adequate if the coverage of journals were sufficiently wide, to the desire for reasonably informative abstracts. But it is generally held that for any research application it is highly important to obtain the original article, if at all possible. The individual's opinion of the reviewer seems to determine his attitude towards the usefulness of a critical abstract.

Cost—that most nearly universal of all boundary conditions—raises its unpretty head and keeps it high during consideration of all the other phases of abstracting. Other things being equal, the unit production cost obviously is lower for indicative than for informative abstracts.

Coverage

What and how many journals should be abstracted? What should be done about associated fields like mathematics and astronomy? How important should a journal be to make it eligible for coverage? How completely should borderline areas such as astrophysics, biophysics, chemical physics, engineering physics, and geophysics be abstracted? Many other questions arise all of which may be largely dependent upon available funds.

The most desirable criterion for coverage, from the physicist's standpoint (and unfortunately the most expensive), is to cover everything of physics interest, erring on the side of too much rather than too little. An abstracting service which follows this formula reasonably well has the very great advantage that its readers can assume that no abstract means no article.

Another possibility is complete coverage of a limited specific list of journals, with perhaps selected articles abstracted from other periodicals. In this case, for at least a known group of journals, the reader can make the assumption mentioned above. This kind of coverage is often employed by intra-company or laboratory services where the ob-

Dwight E. Gray, now on leave of absence from the Applied Physics Laboratory of The Johns Hopkins University, is in charge of the American Institute of Physics' current study of the problems of abstracting the periodical literature of physics. He is supervisor for technical reports at the Applied Physics Laboratory, and during the war was on the staff of the Harvard Underwater Sound Laboratory as a research associate and technical editor.

jective is to abstract the publications received in the organization's library for use by the staff.

A third basis for coverage is selective abstracting of all journals with no assurance that any one of

them is covered completely. How well such a procedure meets the needs of any particular physicist will depend on how many journals are considered by the abstracting agency and to what extent its editors happen to emphasize his field of interest. In general he cannot be certain that all important articles on any subject have been abstracted.

There is a general feeling that a major shortcoming of present physics abstracting is insufficient coverage, both of the number of journals and number of articles abstracted from the periodicals covered. What can be covered is closely tied up, of course, with budget. In most cases the cheapest service would be for indicative type abstracts and limited coverage. Indicative abstracts and wide coverage, or informative type abstracts and limited coverage, is more costly, and informative abstracts and wide coverage is the most expensive of all.

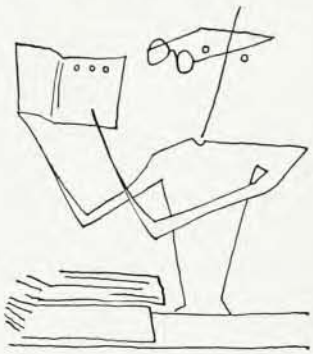
Who Should Abstract?

Members of an abstract journal's staff doing abstracting full or part time, working scientists who abstract in their own special fields on a part-time basis, and authors of the original articles are all possible abstractors. Existing abstract journals use all three of these authorship methods, and combinations of them, the most widely employed method being abstracting by working scientists on a part-time basis.

Maximum uniformity of style and reasonable promptness in publication is achieved with staff-written abstracts, but using professional abstractors, who get paid, is apt to make this method the most expensive of the three. Opinions differ widely on how much technical knowledge a professional abstractor needs to be able to write acceptable abstracts, but the need for such knowledge undoubtedly increases as one moves from the indicative toward the informative type of abstract. In the case of critical abstracts it is obvious that specialists must do the abstracting.

The working scientist abstracting in his own field is able to bring to bear valuable background experience. The advantage this gives him over the professional abstractor is probably partially offset by the fact that abstracting is a secondary consideration for him whereas it is the main job for the professional. The working-scientist abstractor is probably more objective than an article's author, since he is not subject to the parental myopia that sometimes affects one's treatment of one's own brain children.

Abstracting by working scientists introduces the greatest time lag between publication of the original article and appearance of its abstract. Reason-



able uniformity can be achieved by this method if desirable instructions are provided the abstractors.

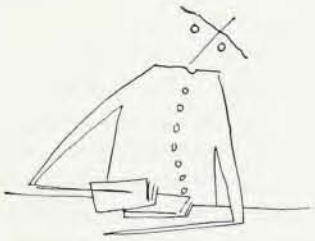
If all technical journals required that articles be accompanied by abstracts, and then these author-written paragraphs were reprinted in an abstract journal, the abstracts would appear most quickly and, presumably, at minimum expense. There is great difference of opinion, however, concerning the quality of abstracting that can be expected from authors. It is argued on the one hand that the author is by far the best judge of his article and of the significance of its contents and, therefore, is obviously the best man to write the abstract. On the other hand, it is claimed that the author's very closeness to the work described makes it impossible for him to view his article in proper perspective and write an adequate abstract. Also, the low cost of author-written abstracts may be largely cancelled by the additional editorial cost of putting them in some kind of standardized form; prior agreement regarding such form might be accomplished to a certain extent among some few of the

principal journals but would be very difficult to arrange with any large percentage of the periodicals that should be covered.

No strong consensus of opinion has been found among the physicists with whom abstract authorship has been discussed so far. A few feel very strongly that working scientists must play an important role in any satisfactory abstracting operation; others believe that the kind of indicative product they desire can be written entirely adequately by people with abstracting capabilities and experience, even though their actual technical knowledge of physics is quite limited. Several, admitting all of the alleged shortcomings and nonuniformities of authors' abstracts, have indicated their belief that a simple compilation of all such available abstracts, issued promptly, would be well worth while.

Indexing

Of the two general uses of abstracts cited earlier in this paper, indexing is most important for the reference function—"retrospective searching." Most abstract services issue some kind of periodic indexes to their subscribers, at least of the author-and-subject varieties which may be accumulated at five- or ten-year intervals. It seems generally accepted that an author index is desirable and, since there is probably only one or two ways in which such an index can be set up, it need receive no further attention here. The subject index, however, poses a real problem.

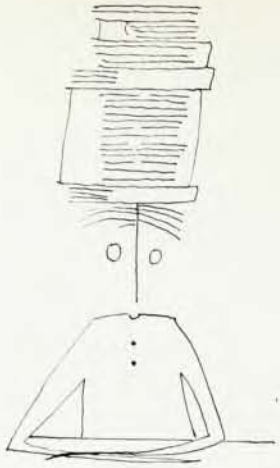


Nature, in reporting the Royal Society Information Conference, outlines four fundamental ways of attacking the problem of preparing a subject index. Two are pertinent for an index that is to be printed in journal or book format (the others concern symbol coding applicable to mechanical searching techniques). The first is "Indexing the names of subjects in alphabetical order. This is the usual system in books and at the end of each year's run of a journal; under certain conditions it is satisfactory also for card indexing. But it involves many more difficulties and moot points than are apparent at first sight. It has the effect of dispersing logically contiguous items according to the letters of the alphabet with which their names happen to begin, and as the names are not standardized so as to make them mutually exclusive, a searcher is liable to look in vain for something which the indexer has buried elsewhere under a synonym."

The second is "Classifying the subjects themselves under symbols which serve to pinpoint their positions in a logically constructed map of knowledge. Many such classifications are in use, some locally for special fields of subject matter and some generally like the Universal Decimal Classification, which is an internationally standardized extension of the Dewey system developed by librarians for shelving books. On this system, knowledge as a whole is divided into ten primary branches each denoted by a first decimal number, each of these into ten subdivisions denoted by a second decimal, each of the latter into ten sub-subdivisions denoted by a third decimal, and so on to whatever degree of particularization is required. As the sequence of subjects is ideographic, not alphabetic, an alphabetically arranged key (corresponding to the gazetteer in an atlas) is necessary for locating subjects when only their names are known."

It should be noted that the second of these approaches, although it emphasizes classification, does not preclude the issuance of a subject index in which the various heads and sub-heads in the classification system are arranged alphabetically; this is done, for example, in the case of the British journal, Physics Abstracts.

A physicist will find any given subject index adequate or inadequate depending on whether he is able to locate in it the references he wants under the subject headings he feels are logical. So it is not



surprising that most of the physicists consulted favor the first of the two approaches given above, a simple alphabetical arrangement of subjects, augmented by abundant cross references. They are not too concerned with the technicalities of whether or not the subject headings are mutually exclusive and in toto, they constitute a logical unified whole. Some of these physicists, however, have also expressed great enthusiasm concerning the possibilities of machine techniques for searching literature, where, as will be pointed out, a logically-constructed, generally-agreed-upon scheme of classification is a necessary preliminary.

It was mentioned earlier that the kind of abstract issued, indicative or informative, has important indexing implications. If, for example, the annual subject index is to be so complete that it is really an index of the original articles *and* if the indexing is to be done from the abstracts, then the abstracts will have to be highly informative in nature. If it should appear that indicative abstracts satisfy every requirement but indexing, then it would have to be decided whether the more expensive, informative kind of abstract could be justified purely as a mechanism for obtaining an exhaustive index.

Promptness

An ideal time for the appearance of an abstract would be at the same time as, or prior to, publication of the article. Practically, however, there are certain limitations which depend on other features of abstracting. With regard to the two general kinds of abstracts, for example, the very difference in their average lengths favors issuance of the indicative type more quickly than the informative. That this margin is made even greater by the authorship techniques usually used for the two kinds has already been pointed out.

If, as is usually the case, the abstracting is done after the technical journal is published, the irreducible lag is the time required to make the abstract assignment, to do the abstracting, to edit the abstract, and to print and distribute the abstract journal. One could increase the efficiency of the operation and, perhaps, obtain articles in preprint, page proof, or other advance form. Thorough investigation of the possibilities of the latter approach was one of the official recommendations of the Royal Society Information Conference.

The importance to any given physicist of prompt publication of abstracts seems to vary considerably with the use he makes of them. Those who employ them principally to keep abreast of what is going on, think it is very important to improve this phase of the problem. Physicists who employ abstracts largely for reference purposes appear to be very much less concerned with how promptly they receive them.

Microfilm, Cards, and Machines

Less fundamental to the abstracting problem as a whole than those considered above, and not as closely linked to possible first-approximation improvements in physics abstracting, are the possibilities of making original articles available, preparing abstracts on cards, or various mechanical search techniques.

One service which some abstracting agencies offer their readers is the privilege of purchasing, essentially at cost, copies of articles abstracted, usually in microfilm or photostat form. This can be valuable in the case of indicative abstracts of articles in relatively obscure journals since otherwise the abstract may serve only to remind one of something he wants and cannot have. A feature of this kind also tends to give the abstract service more nearly the same value for the subscriber located in

a small college or laboratory without extensive library facilities that it has for the man near a large university or public library.

Most abstracting services distribute their product to the consumer as a journal in which the abstracts are grouped according to some scheme of subject divisions. A few issue cards on which are printed the usual library-card information and short abstract or annotation. This system permits the easy maintenance of a permanent, up-to-date card index of one card per article *if* the subscriber files the

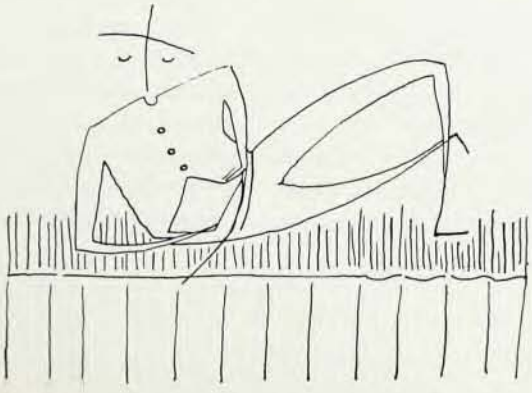
cards promptly; but the possible length of abstract is limited and, unless filing is kept current, cards are likely to be mislaid. A journal is probably somewhat the more convenient of the two for the subscriber interested primarily in keeping up with developments in his own or other fields; it has obvious reference disadvantages, however, particularly during the time between publication of periodic indexes. One or two services are combining the good points of both systems by issuing a journal in which the abstracts are printed in typical library-card format on one side of the page only. By cutting and pasting, a subscriber can set up and maintain his own card reference file.

The potentialities of punched cards and machine sorting to facilitate search for technical papers is today the subject of much speculation. In physics, for example, one might visualize a system in which some agency would prepare for every article or book of physics interest a punched card which would code the author's name, several subject headings, and other pertinent items. The center of the card could carry an abstract of the article, or perhaps the whole article, if microprinting were used. These punched cards might be supplied to all agencies having facilities for machine sorting, and similar cards, adapted for one of the rapid manual sorting methods, might be made available to individual physicists.

But push-button library research is hardly around the corner. One must first get the information on the cards. Too many ignore this less glamorous phase of the job and begin their thinking at a point following card punching. It has been pointed out by Dr. J. W. Perry, chairman of the American Chemical Society's Punched Card Committee, that much of the discussion on this question appears to assume that punched cards and sorting machines can think, whereas, of course, one can get from the cards only what someone has first put on them. Here lie the very difficult large problems associated with developing a standardized and generally accepted classification and coding system.

Whether machine techniques can ever satisfactorily replace the more conventional methods of library search is doubtful. It seems more likely that, at the most, the two will supplement each other. Machine sorting would be highly desirable where all articles on a given subject or by a given author are to be culled from the total accumulation of material. On the other hand, when only one or two items are desired, conventional library search is undoubtedly preferable to running hundreds of thousands or millions of cards through a sorting device.

In any event, it is necessary to learn what physicists think about abstracting before one can go about improving it intelligently. We have intended to highlight the principal phases of the problem, outline their wheels-within-wheels relationships, and state the major points of view encountered early in the Institute's study of the question. Comment is invited.



Sketches by Paul Bond