Chairman's Message

It has been said often that software is ten years behind hardware. This was a popular truism five years ago, and the span may have increased since then. Now it appears that the need to catch up has become acute enough recently to move us to do something about it.

One thing that must be done is to begin to handle information in a computer without bending the information so much that it is only coherent inside the computer. There are many ways to do this. All involve making a sharp and clear difference between the information outside and inside; all imply that present languages will be supplanted by ones that make it possible to manipulate information in a manner as free of the computer environment as is possible.

In order to do this, we will have to specify information much more clearly than we do now, both in terms of the relationships among parts and in the ways dynamic information structures can evolve. These are also the requirements of a D.D.L. for automatic data interchange or for data independence. Describing computer representations will be of little help if we are not sure exactly what is being represented.

The SICFIDET membership has demonstrated high competence in a wide variety of areas related to the use of data. It will be a great abuse of collective talent if we do not take some initiative in tackling the problem of removing the computer bias from computerized information.

Let us form some working groups in SICFIDET to attack specific description problems, compatibility problems, and computer independence problems. If these groups can be organized and clearly chartered, they can provide directions and goals to the industry to help eradicate the ten-year (or more) lag between hardware and software.

Some particular areas have come up in discussions over the past year among SICFIDET members. These areas involve defining information structures and storage structures, specifying requirements for data independence, development of a data descriptive language, categorization of the data-handling facilities of present software and hardware systems, and design of better ones.

If, with the help of the membership, we can identify a few precise areas which require attention, we will form research groups in those areas and begin working together.
Editor's Remarks

In the last issue I made a general statement of the purpose which I felt this publication should serve, specifically stating that this publication was to be an informal mechanism for the rapid dissemination of information and not intended to be in competition with publications of a more formal nature. In spite of the stated goal, both that issue and this second issue contain a number of papers which must be considered fairly well polished; and I am very pleased to have the opportunity to publish those papers and hope to have the opportunity to publish more of the same in the future.

My purpose now is to bewail the lack of less formal notes and communications which would help clarify issues and instigate discussions into aspects of the problems which arise in the description and transformation of files. One might even be so bold as to mention the need for a clearer identification of both the needs in that area and possible approaches to satisfying those needs as a starting point.

I now would like explicitly to encourage informal communication for possible publication and would be especially pleased to consider short, informal communications discussing the needs in this area, or describing possible approaches to general or specific problems. To initiate this activity and to stimulate further comment in the hope of clarifying the issue, I will be asking a few specific people to write their thoughts for publication. To stimulate such discussion further, I will discuss what I feel the role of SICFIDET and fdt to be and some areas that must be examined in pursuit of this role.

SICFIDET is chartered to work in the areas of file description and translation. There are many underlying motivations for describing and translating files, including the desire to transfer data from machine to machine, to access data prepared by other parties either remotely or on a common system, to permit common description of data and data structures at source language levels, to isolate the specification of logical data structures from the specification or need of physical representation on storage media, to isolate data specification from the specification of manipulations to be performed on the data, and to transfer programs which deal with data to other systems with different data structures and representations, and common data and data structure specifications across source languages. Perhaps the most critical of these goals, in the short run, is to support the access of data prepared or used by other parties. The second most critical goal is the need to deal with data at levels remote from the requirements of physical representation of the data for the purposes of human understanding of the data and data structures, and to permit description of data and processes on data to be more directly related to the needs of the problem rather than the needs of the environment within which the data will be stored and the problem solved.

Solutions to these problems are many and span a range which includes such diverse approaches as specification of common formats (to the level of identical bit patterns on storage media) to self-describing data in which the contents, organization and physical storage formats are specified with the data, in a manner usable by all the processors which will potentially come in contact with the data. It is clear to me that no single solution (such as a universal language for the description of data) is a viable solution in the short run (the next 5 to 10 years), and I also feel that such a solution is inappropriate in the long run.

In some cases is it reasonable to specify specific formats and stand by those — this is definition by standardization at its lowest level. In other
Editor's Remarks

Continued from page 8

In some cases it is reasonable to specify the appearance of data to a user and to provide mapping mechanisms (one per environment) between that appearance and actual storage techniques used. In some cases, more generality may be required because of the changing nature of the use of the data base or the equipment on which it will be used. In some cases, data bases may be maintained in a very general form but transformed prior to use to satisfy efficiency requirements.

I feel strongly that a collection of approaches must be identified, their properties examined, and the users (designers of data bases) given the information which will permit them to make appropriate choices, according to their identified needs and goals.

These comments identify some of my thoughts on the subject. I would like to open a discussion with people who feel differently, so that some of the issues can be clarified and useful areas of endeavor identified.

Harrison R. Morse, Editor
Journal of ACM SICFIDET