System Design for Computer Applications

H. N. LADEN, PH.D.

Chief, New Systems Development
Chesapeake and Ohio Railway Company
also
Director and Chief Staff Consultant
National Computer Analysts, Inc.

T. R. GILDERSLEEVE

Manager of Programming
and Technical Services
for the UNIVAC III
Sperry Rand Corporation

John Wiley & Sons, Inc. NEW YORK · LONDON

CONTENTS

PART ONE

The Computer

1

INTRODUCTION

TERMS

TERMS OF REFERENCE—BUSINESS DATA PROCESSING

The general business organization can be conceived of as a system of channels connecting the organization's parts with each other and with the organization's environment—its customers; vendors and stockholders; the government; service organizations, such as Blue Cross, insurance companies, and banks; employee unions and numerous other facets of the external world. Money, material, and other resources flow through these channels to create the conditions necessary for the organization's survival in its environment. Business data processing is the part of the organization's operation that is concerned with the attempt to record, measure, and control the flow of money, materials, and other resources through the channels making up the organization's "circulation system".

SYSTEM DESIGN

Systems and procedures activity within the business organization is concerned with the critical analysis of the data processing operations and their redesign. The entire business may be made the subject of a system study in some far-reaching efforts. The philosophy for such an approach has come to be known as the Total System Concept. More commonly, but possibly less profitably, the business is segmented into subsystems on the basis of function, organization, region, product line, or other integral characteristics. The data processing operations are then studied for subsystems or for specific units or procedures within the subsystems.

Regardless of the specific devices that may be incorporated into a system in the course of redesign, the systems activity follows a well-defined course—a period of fact-finding and analysis followed by a period of creative synthesis of a new system, first in broad design and then in detailed design. Since our concern is primarily with computer-oriented systems, this course will now be run for such a system design project. Hopefully, this should clarify the terms of reference—system and procedures activity, computer system design, and programming.

SYSTEM ANALYSIS

The first phase of any system project is devoted to the analysis of the existing business data system. Facts are gathered about the old or existing operations within the system. Close attention must be paid to exceptions as well as to routine procedures. Almost invariably probing into function or purpose of the system leads to an early concentration within fact finding on the definition of system objective and on the collection and description of the existing output. This helps define the boundaries of the system.

In the determination of output requirements, one is concerned with the contents, format, timing of delivery, and frequency of updating of information supplied by the system. The data may be intermediate reports requiring later completion, finished reports, vouchers, checks, punched cards for further processing, incidental tabulations, phone calls, and so on. One needs to list all those to whom the output is distributed and needs to examine the use to which each puts his copy of the output. External distribution is especially noted, and legal or other uncontrollable restrictions imposed thereby are documented. Again, the broadness of definition of output is emphasized. It includes incidental summarizations which are fed backward rather than forward in the chain, such as work-in-progress reports, error listings, and so forth.

When the purposes and outputs of a system are established, one needs to establish how these are achieved in the system. What is the raw material for the data process? In answering this question, all inputs, their origins, and initiators are determined. Who outside the system also gets this same data? What is the phenomenon of the business being recorded and in what medium? What is the nature and format of the existing source documents? What in the input is fixed or predictable and what is variable and relatively unpredictable? What problems of accuracy (precision) and completeness (missing data) exist? What are the volumes of input data to be ingested? What legal or other restrictions exist? A comprehensive view of what constitutes input must be taken. For example, one must include various directories and reference files, changes thereto, standard rules and regulations presently used to make decisions in the handling of exceptional cases, telephoned reports, and other such data sources.

The analysis of facts gathered utilizes a variety of criteria and techniques. Output can be screened for relevance and purposefulness, redundancy, cost of a particular medium and format, and the convenience to the user of the output. In particular, many an output may be marked for disestablishment. Many intermediate outputs may well be disestablished in a redesigned system that produces a new output to more completely satisfy the ultimate purposes of the system. Similar considerations govern the analysis of input. Inevitably, there will occur frequently during this analysis ideas and notions for revision of the system under study. Hence, the creation of a new system is not truly distinct in time from the fact-finding and analysis phase.

SYSTEM SYNTHESIS

Synthesis of a new system begins with a statement of new or validated old objectives and purposes. This tends to provide at least broad direction and constraint to output. With this preliminary vision of output and its general content, one re-examines input possibilities. Often one is confronted with numerous outputs stemming from a multiplicity of overlapping source data. A major portion of the system synthesis is the study of inputs used in a variety of related sections and the provision for coordination of their collection. This can be the source of improvement in completeness of the input. It makes each input record more generally available to all potential outputs and users. It is also the source of major economies in processing. Elimination of partial or total overlap in reporting, pooling of related information fed in from a variety of sources, and mechanization of data recording as close to a source as possible, so that further extraction and transcription can be handled automatically, are among the principles used. If part or all of a body of source data is really fixed, then procedures are established to avoid reporting these data anew as if they are unknown. If the source data are predictable, *exception reporting* or other techniques exploiting statistical bias may be used. In these, only changes from prediction are reported *in extenso*. Unchanged source data are assumed present or reported in abbreviated form. Numerous document design tricks and machine devices exist for supplementing fixed and predictable data with unpredicted and variable data at substantially lower cost than that of full reporting of source data.

Input redesign includes considerations of the control of quality, timing, field procedures for the recording of source data, standardization of format, and so forth. The subsequent problems and costs of encoding, converting from one medium to another, and the multiple uses of the document are considered in the redesign. For example, it may include introduction of prepunched and preprinted forms for reporting fixed or predictable information. Recording, communications, and data conversion procedures for delivering input must be economically provided consistent with volume, quality, and timing factors.

Of course, the same concerns exist in the delivery of output to users. In the core of the new system, one provides for coordinating the production of all those outputs which relate to or are dependent upon the same reservoir of basic input information. Intermediate products are eliminated and outputs are tailored to the specific needs of the ultimate user. This causes the core of the new system to provide for bridging gaps between sections engaged in sequential activity rather than independent activities. The medium and format of ultimate output may be revised to make it more economical to produce by a particular mechanism. New methods and procedures may be designed for the production and distribution of output. Extension, discard, and redesign of previous output is almost a certain consequence of intensive system study in an area if one takes advantage of the opportunity to make numerous incidental methods improvements, fundamental changes in accounting and management control systems, desirable organizational changes, or similar revisions.

A major step in the synthesis of the new system is the derivation of an overall, logical block diagram representing the gross steps for going from inputs to outputs. These steps are themselves complexes of operations such as file maintenance, computing, editing, and so forth. The steps must be feasible in the light of the machinery and techniques envisioned for incorporation into the system. The designation of the steps is often defended by analogy or experience. Actually, attempts at this diagram begin almost with the initial analysis, since it represents one way in which the system analyst expresses, from time to time, his success at correlating the facts he has gathered. When the block diagram is finished it represents a gross logical description of the application.

With this logical structure in mind, the input and output requirements can be re-examined for compatibility, not only with the machinery and techniques to be incorporated into the system, but also with the organizational problems peculiar to the company. Decentralized or centralized input; source document creation—supervised or not; communications needs; central reports distribution or regional autonomy and review—all these problems and many more of an equipment, organizational, or procedural nature need to be resolved.

Wherever reference is made here to the design of new procedures, there is implied an important task of training personnel of the revised organization in the new procedures, preliminary to their test and installation, as well as the task of conducting the test and installation.

COMPUTER SYSTEM DESIGN

Up to this point, the description of systems activity applies quite generally without regard to machinery and techniques to be used within the system. It is only when the designer of the system is confronted with the problem of optimum employment of the specific machinery that the activity becomes specialized.

Computer system design begins with the consideration of a computer as an alternative solution to the problems presented by the operational requirements of a proposed system. Prototype system designs may be prepared using a variety of approaches—hardware and procedural. The probable costs of each of these should be estimated and due weight given to the differences in degree of fulfillment of requirements. In any event, once a computer solution is selected for installation, computer system design may proceed further.

In the design of the prototype system that incorporates a computer, computer orientation is necessary for meaningful results. It is extremely important, for instance, that the synthesis of the new system not be only the speeding up of the procedures of the old system. The new system must reflect a reorganization of the entire operation so that pieces are related to each other in much more direct, although more complex, logical patterns. That is, we must exploit not only the speed, but also the tremendous versatility and capacity of the computer to effect real economy. Thus, the synthesis of the new system is likely to be satisfactory only if the system and procedures personnel have been properly oriented to computer thinking. Even the analysis of the old system proceeds more satisfactorily with people who have the computer constantly in mind. Without this, the system analyst tends to get lost in detailed flow diagrams of the internal procedures in an area. These internal procedures reflect the limitations of precomputer tools and are largely irrelevant to the computer system design.

Within the framework of a system design which includes a computer, the system design activity has not yet been completely described. The next phase begins with the logical block diagram of complex operations. This must be broken down into acceptable steps, called runs, involving as input specific reels of tape, bundles of machine interpretable documents, decks of punched cards, and keyboard insertions, and producing as output still other specific reels of tape, printed copy, punched cards with or without imprinting, and so forth.

In a run, one generally tries to include the maximum amount of processing of each item of data one can accomplish within the existing limitation of input and output facilities. The number of these holds down the variety of input items and output product. The memory size and speed of the computer may limit the complexity of the run, as will the availability of output facilities. In some computers, minor print-outs can be produced on the Supervisory Control or Console Printer to make available an additional product. All the instructions may sometimes be dumped into memory to clear external facilities for additional output or input use.

More often, however, one cannot make the run more complex without bringing in some supplementary input from still other input media; there are not, however, enough input handling facilities available. Exploiting the full capability of the computer has as yet not been completely reduced to a science. There is room for substantial virtuosity!

When the block diagram has been refined into a run structure, the individual input and output streams are assigned names and numbers and assigned to specific external facilities. Estimates of number of items to be handled in each input and output stream and on each facility are made. The item itself is designed; an *item layout* is the counterpart of a card format in punched card technology. If files are to be stored internally, as on drums or discs, speed and capacity requirements are checked for conformity to equipment limitations. The number of reels of tape in each data variety which uses tape can then be estimated. This should be completed for all tapes associated with a run. Also, the number of tape handlers should be prescribed for each tape set (tape-swap or no, and so on). When this assignment of facilities and identification with corresponding input and output streams has been completed, one can turn to the run itself.

Each run is assigned a name and number. A detailed verbal description of the processing to be accomplished in the run is prepared and filed in a Run Book with the item layouts for both input and output items as well as for perpetually stored files such as exist in random access computer applications.

PROGRAMMING

Beyond this, the narrower technical functions of programming begin. With a rough conception of the way in which the available memory will be used, and sometimes with a specific memory-allocation chart developed early in the run programming, a detailed flow chart is developed. The distinction in a flow chart is that it bridges the gap between a verbal run description and the actual step-by-step instruction coding, by grouping standard complexes of computer code steps. It is essential that such a flow chart show all branchings of the manipulative streams, all logical decision points, and alternative choices.

With a flow chart in hand, the programmer proceeds to write precise sequences of instructions in a rigidly prescribed *source language* vocabulary on code sheets. The lines of coding are supplemented with reference data and informative annotations. These instructions are prepared in such a way that they can be entered readily into the computer. All the material related to the run—code sheets, print-outs of instruction tapes properly annotated, item layouts, memory allocations, operator instructions, flow charts, and so on, are filed in the Run Book. Of course, if the processing involves external input or output conversion, there is another variety of activity in some computers. The programmer must prepare card layouts for direct card input/output or for the tape-to-card conversion, plugboard wiring diagrams for the converter, and operator instructions; for card-to-tape conversion, one needs a plugboard wiring diagram, operator instructions, and test cards. Similar material must be provided for other auxiliary equipment such as the High Speed Printer, document readers, and perforated tape readers and punches.

In order to make tests of our coded programs, we are obliged to have sample data and even to handle actual data for a period of time. This implies that certain nonprogramming tasks are underway or completed and that sufficient progress has been made in the installation of new procedures and in the training of the line clerical organization so that such data can be made available. Also, the computer center must be staffed and the staff familiarized with the new instructions provided by the programmer.

The programmer must know what constitutes an adequate sample of data with which to test the coded program. Certainly, such a sample must provide for the occurrence of all the exceptions whose treatment the programmer proposes to mechanize.

The practice is to *debug* individual runs using synthetic test data in small volume. Beyond this, sequences of related runs are integrated into subsystems using samples of data. Finally, the entire system can be tested with sample data. The next step is to try the computer system in competition with the existing one. This is called parallel operation. The computer is obliged to handle typical volumes of actual data and the results of computer and existing system are compared. Hopefully, few errors crop up and last-minute patching is then effected. Also, timing trials presumably indicate the computer can meet its commitments. The computer system is then ready to assume the responsibility for production.

Meanwhile, of course, preparation elsewhere in the business for a rapid reshaping of the existing data processing system in the organization, and possibly even of the entire organization, has been in process. This is rough going, because initially there must be expansion to provide such services as test data and checking to the computer group; expansion to release personnel from normal production activity in order that the training and installation of new procedures be planned for and effected; and expansion to provide for planning and effecting the inevitable post-computer reorganization and cost reductions. These points are emphasized because it is only too easy to obscure the fact that computerization calls for intense, organization-wide involvement and the closest team work. Credit for its success seldom attaches to any one or any few individuals.

COMPUTER PERSONNEL

Thus, many classes of people and types of skills are involved in a major project involving the study and reshaping of the data processing operations of a business system; the study may possibly involve a revision of the fundamental aspects of management's approach to the problems of the business. Various activities have been described in a general system and procedures framework as well as unique activity that is specifically oriented toward application of a computer within data processing operations. The work of programmers and computer-oriented system designers has been described. Of course, there are machine operators both for the computer, its peripheral equipment, and for other supporting equipment in the office machinery field.

The emphasis here is on delineating the tasks of both programmer and computer system designer. The programmer is the man who prepares the instructions telling the computer what operations to perform and in what sequence to perform them. The system designer is the man who studies the company data processing operation proposed to be handled by the computer, determines what has to be done to implement the operation, decides how it is to be done on the computer, and directs the programmer in the programming of the computer to effect the implementation of the operation.

OBJECTIVES AND SCOPE

Briefly stated, this book is a text for computer system designers. To understand the problems faced by the computer system designer and to appreciate the tools available for his use, one must be conversant with programming. This programming knowledge, on the reader's part, is assumed here. If the reader does not have experience in the programming field, it is recommended that he initiate his study of this book with the perusal of a programming text such as *Programming Business Computers* by Daniel D. McCracken, Harold Weiss, and Tsai-Hwa Lee, John Wiley and Sons, New York, 1959.

More specifically, the objective of this book is to instruct in the principles of computer system design and their application. It is assumed that general system activity has resulted in an agreed on definition of the scope of the system. Within this definition, our objective is to produce efficient computer systems. It is assumed that the reader not only has some familiarity with programming, but also has some background in general system and procedures activity without particular orientation toward any type of hardware. In the bibliography of this book several texts about systems work are included for those who feel inadequately conversant with the terms of reference reviewed. However, many of the matters discussed in this book are as completely pertinent viewed in the context of general systems activity as they are within computer system design. Further, the increasing incorporation of computer devices within new designs for business data processing systems makes the principles enunciated here of relevance virtually to every systems and procedures staff.

Similarly, this is not a programmer's handbook; however, it does concentrate on the principles underlying computer system design. These principles and their application result in designs which must be implemented by the programmer. One can expect many programmers to have a curiosity about the rationale for the design and to seek an understanding of the basis for the constraints and directions provided in the run descriptions furnished to them by the system designer. In some systems organizations, design activity may not be clearly delineated from programming. In short, there is much in this book which should be of distinct interest to the programmer although it is primarily addressed to the computer system designer.

There are computers and there are other computers, even when one is restricted to digital computers as is the case here. Considering the fundamentally different characteristics among some of the varieties, it is difficult to evoke principles fully applicable to all types. A less awkward solution has been adopted here. It is to place emphasis on (magnetic) tape-handling, batch-processing computers. This emphasis is not an exclusion of other types of systems, since the material presented often applies to several types and the text generally states and identifies each one. Also, an alternative version or corresponding principle for the remaining types of computers is frequently provided. If this is not done invariably, it is sometimes because no corresponding situation exists for the types other than the one covered.

The reader should have no illusion that this work is encyclopedic enough to cover all the principles applicable to each computer. Certainly other books will be forthcoming specializing in each variety of computer. There is, however, much in this volume applicable to punched card systems, random-access serial processors, real-time on-line systems, so-called satellite computers, computers with off-line peripherals, computers with on-line peripherals, word-addressable and character-addressable computers, computers with serial or parallel arithmetic, conventional and concurrent-processing (multiprogramming) computers, and so forth. This is the era of modularity, the custom-tailored hardware configuration, and the time-shared multiplexing of applications. The type labels listed here are no longer so neatly differentiating as they once were. For example, the multiprogrammed tape computer, in which the main high-speed memory (generally magnetic core storage) has been supplemented by intermediate speed memories (drums or discs) of longer access times but much higher capacity, becomes a real-time on-line system for random access processing as well as a batch processor. Why, it even becomes an automatic message switching and relay center in the hub of a multistation intercommunication system, as well as a computer in the more traditional sense! Thus, there is no need to be excessively apologetic about not exhausting the principles of computer system design peculiarly dependent upon the equipment characteristics of each variety of computer installation. It is a comfort that many of the principles cited are so firmly based on fundamentals in information theory and on the essential nature of data processing that they are indeed pertinent directly or with minor modification to virtually every variety of computer mentioned.

STRUCTURE OF THE TEXT

This book is divided into three parts. Part one deals with system design considerations which result from the fact that a computer is being used as a data processor.

Part two is concerned with documents: the source documents in which the data to be processed by the computer are gathered, and the reports which are generated by the computer. Although much of the information in this part is general to the subject of form design, its emphasis is on those form-design considerations that are taken into account because the documents represent part of a computer system.

Part three deals with the broader aspects of the system designer's responsibility: how he conducts a study of a company operation, and how he implements the computer system he has designed.

ACKNOWLEDGMENT

In 1954–1955, The Chesapeake and Ohio Railway Company accelerated its pioneering efforts in business applications of electronic computers. There was then manifest a complete lack of formal training courses for personnel newly oriented to systems and procedures activity in general, and a lack of training with computer orientation in particular. Computer manufacturers provided limited customer training in programming techniques only.

To fill the training gap, C & O personnel began preparation of a textbook on systems and procedures and another on principles of computer system design. The former was never completed, but the draft was used to influence the Finance Division of the American Management Association to establish its now well-known courses in Systems and Procedures.

Repeated efforts were made over the next several years through various groups of computer users to interest manufacturers in the design of training courses transcending the coding aspects of computer activity. Although these efforts were not successful, the wide circulation of proposed course material was beneficial. Within Sperry Rand, experienced training officials were aware that the scope of the new training proposed would have a greater influence on the efficiency of computer applications than would coding dexterity. Some of the proposed material found its way into UNIVAC training and especially into technical literature directed to the experienced computer user.

With the onset of "second-generation" and even "third-generation" computers, there is manifest elaborate prefabricated programming libraries, English language programming, and a host of automatic programming schemes. The dependence on individual coding ingenuity as a method of achieving computer application efficiency has become reduced even more. Given the scope of a defined system, computer application efficiency has more and more come to be achieved through broader design aspects rather than through coding tricks.

It has become increasingly more common in organizing data processing activity to distinguish between system analysis, design and coding, and programming. With this demarcation, a clarification of training needs is provided. Training requirements for programmer responsibilities in such an organization have been met readily by computer manufacturers, schools, self-study, and so forth. Suitable programmer manuals are available. Courses in business systems are widely offered by numerous universities. However, none of these courses is specifically oriented toward efficient exploitation of a computer within the design of a business system. Most design know-how continues to be handed down as technical gossip or to be rediscovered.

This book draws together training material developed over the years by Dr. Laden and Mr. Gildersleeve. It incorporates the fruits of numerous intellectual exchanges among various members of Dr. Laden's staff, between them and other computer users, and among employees of Sperry Rand UNIVAC. Indeed, many of the principles cited and illustrative examples provided in this book stem from these exchanges. The present documentation stems from a system design course given by Sperry Rand UNIVAC for The Chesapeake and Ohio Railway Company during 1960 and 1961. Much of the material has been circulated in a series of internal memoranda within C & O and Sperry Rand. The authors express their gratitude to Sperry Rand for free access to the information in these publications. Significant results have also been selected from the literature which already exists on the subject of system design. In such an agglomeration, the ability to reference and acknowledge specific contributions is often lost. It is hoped that none is slighted thereby. Any shortcomings in this book should be attributed to the authors whose gratitude is extended to both the UNIVAC Division of the Sperry Rand Corporation and to The Chesapeake and Ohio Railway Company for providing access to information in company publications and files and for indulgence during the period of final preparation of this book.

2

INTRODUCTION TO COMPUTER SYSTEM DESIGN

COMPUTERS IN SYSTEM DESIGN

Systems activity has been defined as a period of fact finding and analysis followed by a period of creative synthesis of a new system, first in broad design and then in detail. Fact finding is documentation. Analysis consists of a critical evaluation of the existing business data system, its objectives, outputs, and inputs. System synthesis consist broadly of a restatement of objectives, a definition of output in terms of these objectives, and a reorganization of the input data gathering and processing required to produce this output. There is literature covering these subjects. References are included in the bibliography at the end of the book.

As indicated in the previous chapter, when computing hardware is to be used in the implementation of the system, there comes a point in system synthesis at which consideration must be given to the computer, its characteristics, and the opportunities these characteristics offer for increased data processing efficiency. It is in this area of interaction between data processing goals to be achieved and the computer hardware to be used in reaching these goals that the knowledge referred to here as computer system design is applicable.

Computer system design is almost totally unformalized and undocumented. Knowledge in this area consists of principles, maxims and rules of thumb existing in overlapping but not congruent sets in the minds of people whose work is computer system design. These rules generally take the form of lessons distilled from the experiences of some project in which each person has engaged. They lack any statement in general form. A purpose of this book is to document many of these rules so that the novice computer system designer no longer will be dependent upon having them transmitted to him informally and unpredictably by word of mouth. Another purpose is to bring some of these principles, case histories, and examples together and group them into categories.

Computer system design begins with the block diagram of the data