BUSINESS DATA PROCESSING & PROGRAMMING

ROBERT H. GREGORY
&
RICHARD L. VAN HORN

BUSINESS DATA PROCESSING & PROGRAMMING

WADSWORTH PUBLISHING COMPANY, INC.
BELMONT, CALIFORNIA

L.C. CAT. CARD NO.: 63-14323

PRINTED IN THE UNITED STATES OF AMERICA

BUSINESS DATA PROCESSING & PROGRAMMING

PREFACE

Business Data Processing and Programming is a compact, introductory book on data processors and their programming for business use. It focuses on equipment, systems design, and programming. It covers data representation and organization into records and files, equipment components, systems analysis, systems design and implementation, editing, accuracy and auditing, file processing, and programming for business data processing.

Hardware components—control, arithmetic, storage, and input-output—are discussed to cover the basic features of electronic data-processing equipment. The myriad systems offered by manufacturers are not discussed as such, because yesterday's X038 is today's Y038.1 and tomorrow's Z038.14. All models are basically similar and new equipment may differ only in having a faster operating cycle, more storage, or faster tapes. While such changes are important to the user, they are not to a person just entering the field. Thus, rather than discuss the design features of the most-recently-announced equipment, the objective here is to deal with fundamentals worth knowing both today and tomorrow.

The treatment of systems analysis and design in Chapters 5 and

6 includes two new features: information flow analysis and decision tables. Unlike traditional systems-analysis work aimed at a single application, information flow analysis analyzes the flow of data concerning an *event* from its origin, through files, to output. This *event-chain* approach opens up the opportunity to use processors for analyzing the *actual flow* of data in an existing system and the *planned flow* in a proposed system. Thus it is both a systems-analysis *and* a systems-design technique at the data-flow level.

Decision tables are used to relate conditions and actions in a tabular form for the various rules that are applicable. Decision tables facilitate careful statement of conditions and actions, aid in developing correct and complete program logic, point up ambiguities and inconsistencies, and serve as a documentation device. Tables are a promising alternative to detailed flow charting in both the design and programming-coding phases of systems development.

This book treats programming in both problem-oriented and machine-oriented languages. COBOL-1, a simplified version of COBOL-61 (Common Business Oriented Language, 1961 edition), is covered first to bring out the fundamentals of processing business files. COBOL-1 is presented with the intricacies and options of COBOL-61 stripped away. It is more important, we think, to develop the concepts and drive home the fundamental points than it is to give encyclopedic but hasty sketches of the numerous features of programming. The coverage of COBOL-1, though short and disarmingly simple, introduces many features of problem-oriented languages. COBOL-1 is a real programming language, compatible with COBOL-61. Instructions for compiling and running COBOL-1 programs on one processor serve as a guide for compiling and running programs on other processors. COBOL-1 programs can be run on any processor for which the manufacturer has prepared a COBOL-61 compiler.

Programming in machine-oriented language, covered in Chapter 8, is discussed in terms of a fixed-word-length processor, WORDCOM, which is representative of this widely used class of machines. FIELDCOM, a "field" or variable-word-length machine, is also covered briefly to show the features of this class of machines. COBOL-1, WORDCOM, and FIELDCOM programming are covered at the introductory level to offer the reader some choice of where he wants to start. The sequence used here, and the one we recommend, is to study COBOL-1 first in order to learn programming fundamentals with the fewest embellishments. Then the reader

is ready, we think, to go to the machine-oriented languages of WORDCOM and FIELDCOM.

Business Data Processing and Programming is designed as a one-semester text for a basic course. Corresponding to each chapter are numerous questions and problems, carefully designed to emphasize the important points and to develop the student's facility in applying them. Ranging from simple to complex, these problems offer variety and potential enrichment for higher-level or more intensive courses.

This book is also oriented toward the person in business who is getting his first serious exposure to business data processing. Systems analysts, designers, and programmers who want to put data processing into proper perspective should find it most useful.

We wish to express our appreciation to several of our friends and associates who provided valuable help. Solomon Pollack made numerous detailed comments on the manuscript and suggestions on the use of decision tables. O. T. Gatto supplied information on AUTOSATE. Arthur Carlson, Richard W. Conway, James Gibbons, Elmer F. Judge, and Robert Schlosser read the manuscript and offered many useful observations. Edward Chappelear provided us with a variety of technical material, and Willard Harriss prepared illustrations. Mary Cole coordinated typing of the manuscript and handled a multitude of production details.

Robert H. Gregory
Richard L. Van Horn

CONTENTS

BUSINESS DATA PROCESSING & PROGRAMMING

CHAPTER 1

INTRODUCTION TO BUSINESS DATA PROCESSING

Substantial changes in business activity during recent years have encouraged the creation of new data-processing systems. Conversely, changes in data-processing equipment and methods are suggesting new methods of doing business. Electronic data processing, operations research, and other new scientific and engineering developments are fast becoming a part of common business practice. And it is probable that data-processing and management-control methods will change more rapidly during the next decade than they have in recent years.

This progress will not make manual data processing and conventional control techniques completely obsolete, but it will make knowledge of automatic data processing invaluable for business managers, data-processing personnel, and systems and procedures analysts. All people who rely on facts for performing or managing business operations—people who need facts for answering questions and making reports—have a vital interest in the origination and processing of data.

WHY PROCESS DATA?

Data about business events are gathered, stored, processed, and reported for various reasons. Some of the more important reasons are to keep detailed *facts* about individual transactions; to produce routine *operating documents,* such as paychecks and purchase orders; to summarize masses of facts into useful, *informative reports,* such as reports on sales in given areas and on costs of production; and to make *analyses* of business problems, such as finding locations for factories and warehouses that will give minimum transportation costs.

Data Availability

Facts are the raw material of data processing. Initially, the collection of facts may be restricted to the minimum immediately required; but supplemental facts are often obtained at first in order to anticipate probable or possible demands for them in the future.

Often the solution of a problem has to be postponed until necessary facts are obtained. An efficient scheme for obtaining facts depends on balancing the *cost* of having them available—either too many or too few—against the *benefits* of having the necessary facts when they are wanted. The point is that it is usually impossible to obtain just the right set of facts about events "today." One must choose between trying to search out and reconstruct facts about yesterday, or waiting for tomorrow and gathering facts about events as they occur.

Document Production

The preparation of readable documents is an important phase of data processing. Documents are still a common means of communication among companies, and they are widely used within companies that have manual data-processing systems.

Where data collection is done by people and the output from processing goes to people, it is usually taken for granted that readable documents are required. This need is reduced, or even eliminated, when equipment is used for gathering and processing data. The output of machines at one stage—whether punched tape, cards, magnetic tape, or machine-readable printed characters—can be used as input at the next stage of processing, in either the same or a different company. The use of readable documents is sometimes restricted to small-volume operations where mechanization at the next stage is not feasible. Readable documents that are distributed

to a wide audience—bills to customers and reports to managers—will continue to be used for many years.

Management Information

An important reason for developing elaborate data-processing schemes is to supply the management organization with the critical facts needed to control operations. The word *data* might be used to cover *all* the facts obtained; the term *information* is useful for denoting the *particular* facts management wants to know.

Although information derives from raw data, information has certain qualities that can serve as a guide in processing data. A manager in a company is interested in getting facts about operations he is responsible for. He wants these facts to be accurate, timely, and related to problems he can solve by his decisions. Furthermore, he is more interested in learning about unpredictable than about predictable developments. He has no need to be told repeatedly something he already knows. In short, the information given him should be accurate, have an element of newness or novelty about it, focus on a selected area, and deal with the unexpected.

Some examples will help distinguish *data* from *information*. Newspapers are bought for information. Someone who feels that a particular issue has little news to offer, either because so little is happening or because too many other editions are published, will not buy it. To stimulate their sluggish circulation, newspapers often try to create an air of bright novelty around the humdrum, or revive interest in old mysteries; the Loch Ness monster, for instance, is said to be a summertime phenomenon that obligingly reappears for the benefit of Scottish newspapers. Likewise, reports to management about unexpected or undesirable events, such as the costs of production jobs exceeding their standard, select and emphasize items to increase the reader's attention. The costs of jobs not exceeding standard may be included in reports for completeness or omitted for brevity, according to the reader's immediate interests or his needs for later reference to the whole picture.

The difference between a mass of facts and a few critical facts is illustrated by the plight of a businessman who furnished 1350 pounds of records to a tax collector. He was brought to court by the collector, who wanted five additional books, weighing only 10 or 20 pounds, that he considered critical. Every day businessmen are given pounds of reports when they want only a few critical facts. Managers do not really care about source reports, documents, and data, however useful they may be elsewhere in the organization.

The nature of information can be described, but it is difficult to measure information itself as a quantity. Whether a new system will produce better information than the old system did is difficult to determine, and will be even after the new system is introduced. Information production and the related areas of management control, including automatic decision-making, are challenging and profitable areas of study.

HOW DATA ARE PROCESSED

The basic functions of processing data are well established. Managers and operating personnel in business, government, and other enterprises have long been accustomed to processing data to obtain facts about operations and information for their control.

The basic operations in processing data are (1) to originate data, (2) to manipulate, according to some plan, the new data and files prepared in an earlier cycle, and (3) to report results.

Origination of Data

The origination of data in a form suitable for processing includes three necessary stages—collection, conversion, and verification.

Data Collection. Data collection captures facts when they are available; they may be processed later, when needed. For example, the time an employee starts or stops work on a particular job may be recorded in *writing* by a timekeeper, *stamped* in numerals by a time clock, or *punched* into a card. A storekeeper identifies and counts any material received in the stockroom in order to write a receiving report. Requisitions for material, on the other hand, specify desired quantities and serve as the original records of issuance. Employees of utility companies record customer meter readings by marking cards that are later run through equipment that "senses" the marks and punches them.

Data collection often starts with the *manual operation* of keyboards that punch cards or paper tape, or that record data on magnetic tape. Several devices recently developed for business use are capable of automatically collecting data in *machine-processable* form, though one class of these devices yields data in a machine-processable form that people cannot read. Examples of devices producing machine-readable media are point-of-sale recorders, transaction recorders, and time clocks that punch tape or cards. Another class of automatic data-collection devices produces nu-

merals and letters on paper or cards in a form readable by both machines and people. Character-reading machines "read" the characters and convert them to a form suitable for automatic processing.

Other important techniques for data collection are pre-preparation of constant data, by-product preparation as a part of other operations, and duplication from card, plastic, or metal plates. A simple time clock records the basic facts for a transaction; a more complex clock might record all the facts—worker, time, job involved, and units produced—and thereby deserve the name "automatic transaction recorder."

New input data may be only a small part of the total data handled at each cycle. The inventory file, for example, is used repeatedly to introduce necessary changes for receipts, issues, and price changes, and to introduce and delete items. Data already in files are much easier to handle than new data, because master cards or tapes may be selectively duplicated for use as input data. The date, batch number or transaction serial number, and fixed data can be supplied automatically during the data-collection operation.

Data Conversion. Data collected on the medium—paper, cards, or tape—that is most efficient at the first stage must often be converted to a different medium for further use. Some companies, for example, use audio tape recorders to record inventory counts. Until automatic voice-to-digital converters become practicable, conversion by people is required for further processing. People are able to work with oral, handwritten, or typed data, but equipment usually requires that data be recorded in a carefully prescribed form. Data from oral recordings or handwritten documents may be manually punched into cards or tape or written on magnetic tape for input to automatic processors. In many cases, punched cards or tape are prepared manually and the data are then converted automatically to magnetic tape for input.

Machines are used now to read typed, printed, and handprinted characters and to convert them to a form suitable for machine processing. Enough research has been done to indicate the feasibility of automatically converting the spoken word into a form suitable for further processing.

The form in which data are originally captured, and the manner in which data are converted, depend on several things: the volume

of data involved, the number of points of origination, the permissible financial investment at each point of origination, transmission requirements, conversion costs, and the most efficient form for use in subsequent stages. Large volumes of data originating at a few points warrant big investments in elaborate collection and conversion facilities.

Data Verification. Data are verified to obtain a desired degree of accuracy. The standard "a desired degree" does not imply perfection. Some inaccuracies—a slight misspelling of a name, for example—may be trivial, but other inaccuracies, such as crediting the wrong customer for a collection on account, can cause trouble unless they are corrected soon after they occur and before much processing is done. Verification includes checking data to determine whether they are in the approved format, convey the correct meaning to the reader, and will lead to the appropriate action.

The simplest type of *format* verification is to ensure that each data field contains the correct kinds of characters: numeric, alphabetic, or both. An alphabetic character in a numerical statement of amount—say, "$1,2A3.46" on a check—violates the rules of format, and it defies further processing because an alphabetic quantity cannot be handled arithmetically. Such an error in format must be corrected before processing.

A broader problem in format verification concerns the completeness of data input. Each business event requires these elements of data to describe it: (1) identification of parties to the transaction (one being the firm itself), (2) description by name and number of what is involved, (3) quantity involved, (4) the time of the transaction, and (5) the unit price or dollar amount, if money is involved. A receiving report, for example, must have the first four elements in order to be complete; it can be easily examined for correct format without referring to anything else.

More difficult problems of data verification arise from failures to originate documents for *all transactions* or to receive all of them for processing. Mechanical recorders, document numbering, and careful control over valuable items are used to assure complete reporting and correct handling.

The *meaning* of some data in terms of reasonableness serves to verify it. It is possible for an employee to work 80 hours in a week, but it is unreasonable to assume unquestioningly that he did, even with a two-shift operation. Detection of a possible error need not stop the checking process. After the questionable data are flagged