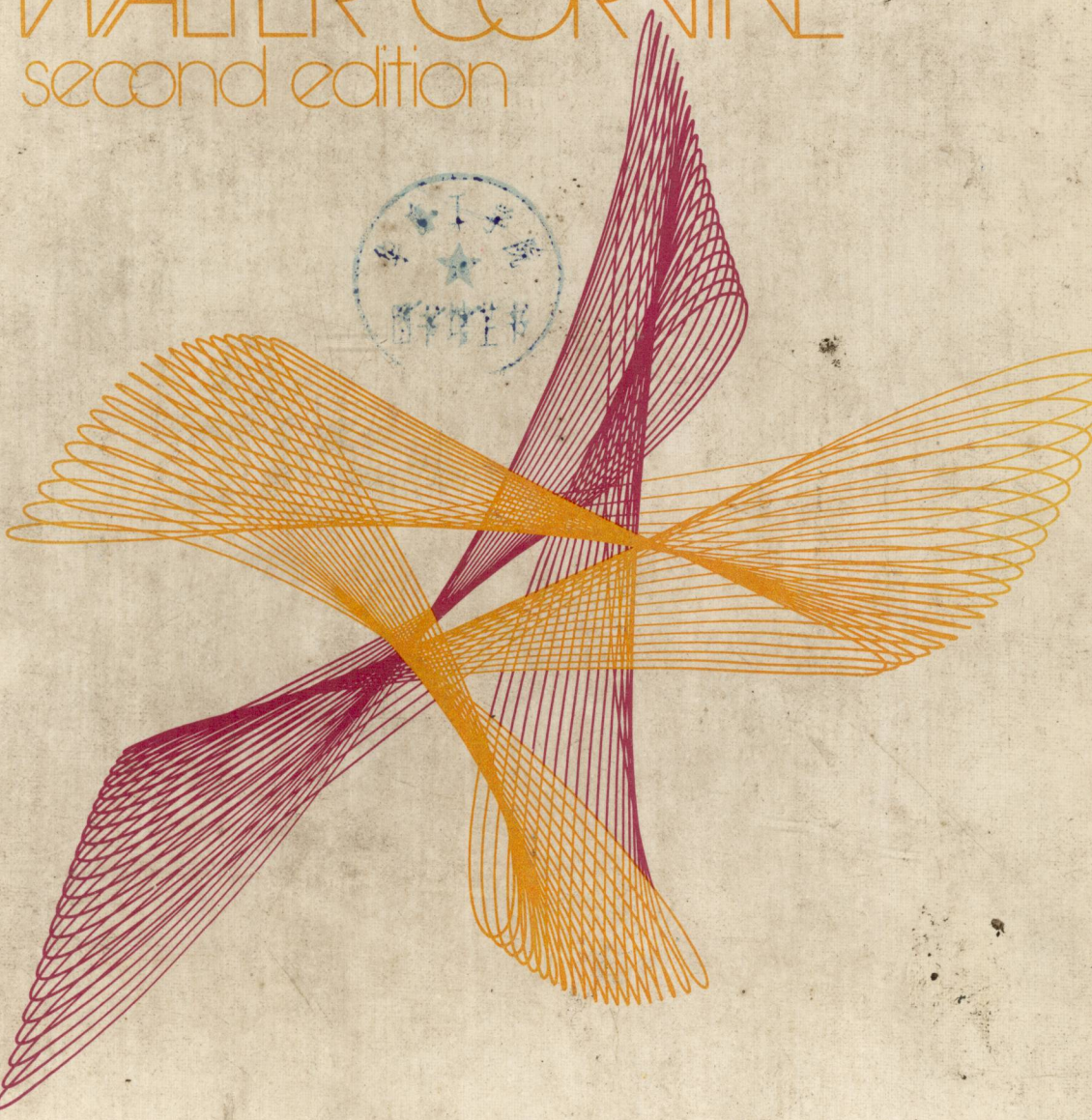
basic data processing

PETER ABRAMS &
WALTER CORNINE

second edition

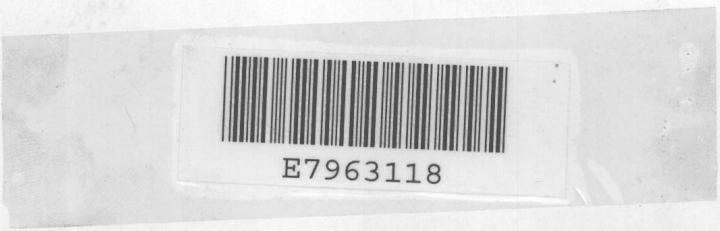


PETER ABRAMS
Northern Illinois University

WALTER CORVINE
Chicago State College

BASIC DATA PROCESSING

SECOND EDITION



RINEHART PRESS
San Francisco

Library of Congress Catalog Card Number: 76-140240
SBN: 03-084123-2

Printed in the United States of America

1 2 3 4 038 2 3 4 5 6 7 8 9

BASIC DATA PROCESSING

SECOND EDITION

PREFACE

Since the publication of the first edition of this text in 1966, many changes have taken place in the field of data processing—primarily in the hardware. The original purpose of the text was two-fold:

1. To present a survey of data processing and computer concepts as an area of general knowledge for the informed individual.
2. To present an introduction which might serve as the first step towards further study in specific fields of data processing.

The material that was presented originally emphasized concepts and principles. These have not changed significantly, substantiating our belief that there are basic concepts in the field of data processing which remain stable. The first edition was conceptual and not oriented to any specific hardware. This edition follows the same precepts.

The text can be used for the following courses or programs:

1. A broad introductory course (one year in length) in data processing at the secondary school level.
2. A general introductory course at the postsecondary school and senior college level.
3. A first course in any data processing sequence.
4. An introductory course at the graduate school level.
5. As supplementary material in those areas that may use data processing techniques as a tool (business, science, education, and so on).

There are four main units in the text, each providing an overview of one of the important subclassifications of the broad field of data processing.

Unit I, Unit Record Data Processing, has been abridged and revised, but still includes the early development in data processing which made use of unit record principles and unit record machines. Chapter 7 of this unit provides some applications of these principles, concepts, and techniques to case studies.

Unit II, Computer Data Processing, presents and discusses the development of the stored program computer system, the principles involved, and some of the general types of hardware associated with these devices.

Unit III, Computer Programming, introduces the techniques for the control of the computer through the vehicle of a hypothetical computer system. This "hypothetical" system, Idpac, although nonexistant, presents the typical problems incurred with existing hardware. Programming is presented in the steps necessary to solve computer-based problems. First, problem analyses and flowcharting techniques are discussed and typical mathematics and business problems are presented. These problems are first programmed (or coded) in "machine" or actual Idpac lanuage, then in an assembler language. The mathematics problem is then programmed in Fortran and the business problem in Cobol. PL/1 is used to program both problems. The final chapters of this unit deal with advanced programming techniques and systems analyses and design.

Unit IV, Overview, presents a summary of current applications of data processing together with a look into the future.

The Appendix includes sections on the new IBM System/3 and RPG, the specifications of the Idpac computer system, and the language specifications for Idpac machine and assembler language. Short sections contain the specifications of Fortran and Cobol. The final sections contain the glossary and a bibliography.

Time limitations may prohibit covering all the material in the text. At the discretion of the instructor, parts of Unit I can be omitted. However, Chapters 1 and 2 should be thoroughly covered. Some of Chapter 13 (Logic) can also be omitted without seriously disturbing the continuity of the material.

We wish to express appreciation for the technical assistance provided to us by the following manufacturers of data processing equipment and their representatives:

Burroughs Corporation
Control Data Corporation
Friden Division of Singer Company
General Electric Corporation
Honeywell, Incorporated
International Business Machines
National Cash Register Company
Radio Corporation of America
Univac Division of the Sperry Rand Corporation

Special thanks is due to IBM for permission to use their many fine photographs and diagrams.

The authors gratefully acknowledge the assistance and encouragement of their families, friends, and colleagues. We also thank the many teachers and students who used the first edition and suggested the additions and changes in this edition. The first edition was initially edited and typed by Mrs. Eleanore Rudolph. Since that time she has worked closely with us and has contributed substantially to the revised edition. Her assistance in rewriting portions of the text and final editing of this revision has made it better organized and more easily read.

Chicago, Illinois
December 1970

Peter Abrams
Walter Corvine

BASIC DATA PROCESSING

SECOND EDITION

INTRODUCTION

The Field of Data Processing

DATA PROCESSING

Data processing is the handling of data through the use of machines. While this definition is short, its meaning and implications reach into all areas of life. An explanation of the exact meaning of the various parts of this technique is essential.

The general term used to cover the broad area studied in this text is *data processing.* There are various subdivisions of the field, each of which has a unique name: unit record data processing, electronic data processing, and others. They are all, however, a part of the general area, and the general term *data processing* covers all divisions. As each division is introduced, it will be defined.

Data processing is an applied science; therefore, its value is based on what it can do, not on what it might be able to do. This value is measured in terms of the tasks it performs, the jobs it accomplishes, and the problems it solves. Because of this, data processing is of immediate concern to anyone who faces the problems to which its techniques are adaptable. It is presently a useful tool, by definition, and its usefulness continues to grow as more and more applications are devised.

Data processing is concerned with machines. The machines, called *hardware,* are the essential elements that perform whatever is desired. This is not to slight man, but we must realistically recognize that more and more of our tasks are being done by nonhuman devices. Just as the automobile is an extension of man's ability to walk, data processing is an extension of man's ability to do other tasks. The fact that in this case the tasks are mental rather than physical should not be disturbing. Viewed as an extension of man's ability, data processing benefits both man and his civilization.

Data processing is concerned with the task of handling data, one of the most important problems that people and nations face today. The problem is of such concern that in the next section it will be discussed exclusively.

DATA HANDLING

An immediate question is, "What are data?" Data may be almost anything and everything. It is of concern to people both as individuals and as functional

members of society. It is constantly around us, and at times it is us. As defined in a dictionary, "Data are figures, words, or charts that refer to or describe some situation." Data may be the ages of everyone reading this book, the names of all the customers of a particular department store, or the first pictures of a distant planet. *It is always plural.* Any grouping of numbers, words, charts, or symbols with meaning for someone is a collection of data. Thus, data may be almost anything and everything.

Because it is a grouping of information, data handling has certain characteristics that need to be mentioned; namely, its descriptive factors and what can be done with it. Its descriptive factors are amount, processing, cost, personnel, error, and speed.

Data come in different amounts. The number of names in the phone book from Centerville, population 500, is quite different from the number of names in the phone book from Chicago. The census taken 1860 was quite different from the one taken in 1970. Both these examples testify to the fact that the amount of data varies greatly depending upon the specific grouping being considered.

Need determines what processing is to be done with any amount of data, and this processing differs greatly according to the circumstances. The name and address list of a company may be used either for sending out an advertising promotion or for billing customer accounts. In the first instance the amount of processing would be relatively small, while in the second instance processing might be quite complicated. A series of orbit readings used to calculate a hypothetical trajectory of a spaceship obviously requires more processing than the processing that determines the average age of the boys in a certain high school. Amount of data is one thing; processing of data is another.

In data handling, cost and personnel are related factors and therefore are considered together. Cost is the dollar amount necessary to process fully the data, and personnel is the number of people necessary to process fully the data. Depending upon the amount of data and processing, the cost and personnel needed vary as greatly as the factors of amount and processing.

Until now only the positive side of data handling has been considered; once the human element enters the picture other factors appear, the most pertinent of which are error and speed. Up to this point it has been assumed that there was data and that it was processed by people at a certain cost. This is an incomplete picture because requirements of data handling must be included. The first of these is accuracy. Expectations are that data should be handled without mistakes, but this is seldom the case. Any time data is manipulated error may occur and somewhat negate results. Error has two characteristics:

1. It may be insignificant (such as being a few cents off in adding up a sales purchase) or significant (such as going too fast around a sharp curve in a car).
2. It may occur seldom (such as once in every 500 sales) or often (such as almost every time someone speeds around the curve).

In either case the original goal desired is difficult to reach because of the influence of error. Some error might be reduced by the slower handling of data, but this is not always feasible or possible. Speed, then, is another factor to be considered. This is a variable that is not always under one's control, for

some data handling problems must be solved within a time limit. For example, when an astronaut is about to re-enter the Earth's atmosphere, results of the calculations that determine his safe landing must be available immediately, not two days later. Even if speed were not a factor, the number of people necessary to do and check certain data would be prohibitive. When one factor of data handling is altered the remaining factors are probably affected.

Data handling is not a simple problem, and the characteristics of data handling are many and interrelated. Each is a variable and changes in importance according to its individual characteristics as well as according to its relationship to the other variables. Therefore, amount, processing, cost, personnel, error, and speed make data handling a complex problem, but one that, nevertheless, must be recognized and solved in a framework that is feasible and applicable to particular goals.

THE NEED FOR DATA HANDLING

The student might wonder why this lengthy discussion of data handling is necessary. It is necessary because data handling is the problem of our time in business, government, and science. A solution to this problem supports our whole advanced economy and, in the broadest sense, our way of life. Let us survey some of the different types of problems common to our society. Imagine yourself the person whose job it is to solve them.

You are a member of the government, a supervisor in the Department of Defense. The inventory of this department consists of over 3 million items. You must supervise the processing of about 100 million transactions every year.

You are a member of the government, an administrative head in the Social Security Administration. The Administration's master file of individuals includes over 150 million names. You must supervise the processing of over 75 million transactions quarterly, about 9 million of which reach the Administration with some kind of error.

Again, you are a member of the government, the controller in the Post Office Department. You are responsible for the accurate processing of the 300 million money orders that are either issued or redeemed every year.

You are a member of the business community, working as an assistant cashier in an average-size bank. The number of accounts in the bank is approximately 80,000. Your job is to supervise the processing of the 90,000 check transactions that occur during an average day. If you were employed by a large bank, the daily transactions could number as high as 150 million.

You are a member of the business community and work as circulation director of a publication. Your monthly magazine circulation averages about 400,000 copies each issue. You must supervise the processing of the 20,000 new subscriptions, the 6000 renewals, the 4000 changes of address, the 500 cancellations, and the 3000 complaints that occur each month.

Again, you belong to the business community. You work as a manager of data in each of these areas: inventory control, file maintenance, payroll, accounts receivable, accounts payable, cash received, auditing, and taxes for local, state, and federal governments. The company employs 3000 people who work in two plants. Your gross sales of 18 million dollars represent 5100 items of production.

As a scientific investigator you are working for the space program. It is your job to solve the complex equations that control space flight. Not only must your work be completely accurate, but the thousands of calculations you do must be completed *seconds* after you receive the data.

These examples suggest only a few of the tremendous data problems that must be met in today's world. They obviously are not simple problems, but they cannot be ignored. The functional operation of civilization depends upon their solution.

THE SOLUTION

Never underestimate the mind of man, for if it can invent problems of the magnitude of those just presented, then it should also be able to devise their solutions.

The major requirement for the solutions is an aid, a helper, some mechanization. Man has continually invented and employed such devices whenever the need has been great enough. The aid needed for solving data handling problems must be extremely complex simply because the problems themselves are highly complex. Historically, the complexity of a problem has not prevented man from solving it. However, the aid for the data handling problem, in addition to giving the solution, must be accurate, extremely fast, and not too costly. Two major problems of data handling are the amount of data and the time allotted for processing. Therefore, any device to aid man must be fast. Error is inherent in large and complex data handling, but the aid must be accurate. Since we live in a society that does not have unlimited money, the aid must be economically practical. Such an aid has been developed to meet the needs of modern government, business, and science. This aid is a group of machines, the most eminent of which is the computer. These machines are fast, accurate, commercially available, and used by many segments of our society.

Data processors are found in government. Here they aid in the controlling of our national defense and in the processing of data for various departments and agencies. Data processors are found in business where big companies use big machines, little companies use little machines, and companies that cannot afford machines rent time on other companies' machines. Data processors are used by scientists for aiding in control of such things as the launching of rocket ships, the drilling for oil, the calculations of theoretical mathematics, and the study of human behavior. Data processors are found almost everywhere, and uses for them multiply each day. Our lives are more and more affected by them, but they still remain basically aids: machines that help us solve the problems that we must face. They are man's aids to help man solve the problems *he* has made.

SUMMARY

Data processing is the general term that covers a broad field of concepts and applications. Data are any collection of words or figures that have meaning. Processing is anything done to or with such data. Characteristics of such processing include amount, extent, cost, personnel, accuracy, and speed. In today's world the amount of data to be processed in government, business, and science has become so great that aids are necessary to assist man. One

category of such aids is machines that can process data. These are called data processing machines and this text is about their development, concepts, operation, and applications.

CONTENTS

UNIT III

Computer Programming

UNIT IV

Overview

APPENDIXES

UNIT I

UNIT RECORD DATA PROCESSING

CHAPTER 1

Unit Record Data Processing

BACKGROUND

The purpose of this chapter is to develop the background that eventually led to the emergence of current data processing. The "computer age" did not descend upon us suddenly, but rather it evolved from an unpretentious start over 150 years ago. Data processing is only a culmination of the attempt to relieve the monotony of repetitive tasks in a way that can produce identical results each time the task is done.

Joseph Marie Jacquard in about 1801 invented a punched card textile loom. This machine was capable of intricate weaving, and its operation was directed by punched cards. Jacquard and his machine were ridiculed when he first showed it to the public, and acceptance of the new device was slow. It was about 40 years later that a once-indignant public fully realized the potential of Jacquard's invention, and the punched card loom was put into operation. Jacquard received the Legion of Honor from the French government for this invention.

The weaving machine was the forerunner of all punched card controlled devices. It operated on the same basic principles used in modern machines, with the instructions being punched on strips of paper that were slipped into the machine. The loom was capable of sensing, or reading, each strip and weaving according to the pattern of holes that was in it. The finished product was a fully woven piece of material, and the instructions could be used again and again to guide the machine in producing other identical pieces.

It is interesting to note that Jacquard's rise to popularity could be attributed to the fact that the public eventually realized the commercial application of his invention. The punched card loom did much to raise the prosperity of the town in which he lived. When this fact was realized by the population, the new concept and its inventor were fully appreciated.

Logically, it seems that punched card looms would have been further developed and their use increased, but such was not the case. It was nearly 100 years later that the next major development in the use of punched cards took place. The United States Constitution provides that a census be taken every 10 years, and by 1880 census reporting was beginning to become a major problem. The 1880 census took about seven years to process fully, and someone in the Census Bureau became aware that the 1890 census might take well over 10 years to tabulate. Clearly, something had to be done.