

FUNDAMENTALS OF DATABASE SYSTEMS

Ramez Elmasri

Department of Computer Science
University of Houston

Shamkant B. Navathe

Database Systems Research and Development Center
Department of Computer and Information Sciences
University of Florida



The Benjamin/Cummings Publishing Company, Inc.
Redwood City, California • Fort Collins, Colorado
Menlo Park, California • Reading, Massachusetts • New York
Don Mills, Ontario • Wokingham, U.K. • Amsterdam • Bonn
Sydney • Singapore • Tokyo • Madrid • San Juan

Sponsoring Editor: Alan R. Apt
Associate Editor: Mark McCormick
Production Editor: Mary B. Shields
Text and Cover design: Hal Lockwood
Copy Editor: Mary Prescott
Composition: Graphic Typesetting Service/Coordinator: Sharon Squires
Typesetter: Rebecca Herren

Cover Art: "Sandstone". Original hand pulled limited edition serigraph by Tetsuro Sawada. Image size, 35" x 22-3/4". Edition size, 60.
With thanks to Galerie de Metropolitan, La Jolla, CA. Exclusive Sawada publisher and distributor: Buschelen/Mowatt Fine Arts Ltd. 111-1445 West Georgia Street, Vancouver, Canada V6G 2T3.

The basic text of this book was designed using the Modular Design System, as developed by Wendy Earl and Design Office Bruce Kortebein.

Copyright © 1989 by Ramez Elmasri and Shamkant B. Navathe

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher. Printed in the United States of America. Published simultaneously in Canada.

Library of Congress Cataloging-in-Publication Data

Elmasri, Ramez.
Fundamentals of database systems.

Bibliography: p. 751
Includes indexes.

1. Data base management. 2. Data base design. I. Navathe, Sham. II. Title.
QA76.9.D3E57 1989 005.74 88-35001

ISBN 0-8053-0145-3

BCDEFGHIJ-MU-8932109

The Benjamin/Cummings Publishing Company, Inc.
390 Bridge Parkway
Redwood City, California 94065

CHAPTER 1

Databases and Database Users

1.1 Introduction

Databases and database technology are having a major impact on the growing use of computers. It is fair to say that databases will play a critical role in almost all areas where computers are used, including business, engineering, medicine, law, education, and library science, to name a few. The word "database" is in such common use that we must begin by defining what a database is. Our initial definition is quite general.

A **database** is a collection of related data.* By **data**, we mean known facts that can be recorded and that have implicit meaning. For example, consider the names, telephone numbers, and addresses of all the people you know. You may have recorded this data in an indexed address book, or you may have stored it on a diskette using a personal computer and software such as DBASE III or Lotus 1-2-3. This is a collection of related data with an implicit meaning and hence is a database.

The above definition of database is quite general; for example, we may consider the collection of words that make up this page of text to be related data and hence a database. However, the common use of the term database is usually more restricted. A database has the following implicit properties:

- A database is a logically coherent collection of data with some inherent meaning. A random assortment of data cannot be referred to as a database.

*We will use the word *data* in both singular and plural, which is common in database literature. Context will determine whether it is singular or plural. In standard English, *data* is used only as the plural; *datum* is used as the singular.

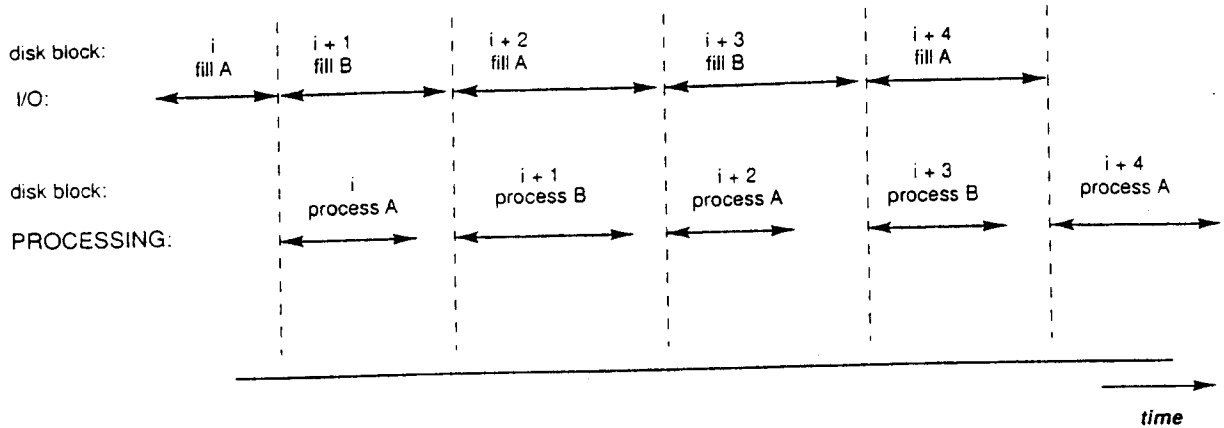


Figure 4.4 Use of two buffers A and B for reading from disk

4.4.1 Record Types

Data is usually stored in the form of **records**. Each record consists of a collection of related data **values** or **items**, where each value is formed of one or more bytes and corresponds to a particular field of the record. Records usually describe entities, their attributes, and their relationships. For example, an **EMPLOYEE** record represents an employee entity, and each field value in the record specifies some attribute or relationship of that employee, such as **NAME**, **BIRTHDATE**, **SALARY**, or **SUPERVISOR**. A collection of field names and their corresponding data types constitutes a **record type** or **record format definition**. A **data type**, associated with each field, specifies the type of values a field can take.

The data type of a field is usually one of the standard data types used in programming. These include numeric (integer, long integer, or real number), string of characters (fixed length or varying), Boolean (having 0 and 1 or **TRUE** and **FALSE** values only), and sometimes specially coded **date** data types. The number of bytes required for each data type is fixed for a given computer system. An integer may require 4 bytes, a long integer 8 bytes, a real number 4 bytes, a Boolean 1 byte, a date 4 bytes (to code the date into an integer), and a fixed-length string of *k* characters *k* bytes. Variable-length strings may require as many bytes as there are characters in each field value. For example, an **EMPLOYEE** record type may be defined—using **PASCAL** notation—as follows:

RECORD TYPE NAME	FIELD NAMES	DATA TYPES
type EMPLOYEE = record	NAME	: packed array [1..30] of character;
	SSN	: packed array [1..9] of character;
	SALARY	: integer;
	JOBCODE	: integer;
	DEPARTMENT	: packed array [1..10] of character
		end;

4.4.2 Files, Fixed-Length Records, and Variable-Length Records

A file is a *sequence* of records. In many cases, all records in a file are of the same record type. If every record in the file has exactly the same size (in bytes), the file is said to be