Books

Managing Gigabytes: Compressing and Indexing Documents and Images, Ian H. Witten, Alistair Moffat, and Timothy C. Bell, Van Nostrand Reinhold, New York, 1994. 429 pp. (ISBN 0-442-01863-0).

One seldom encounters a book that can be recommended without hesitation, but this work belongs easily within such a category.

The three authors are all major contributors in the field of textual storage, compression, and retrieval systems, and their collaboration has resulted in an excellent and extremely well-structured book, capable of satisfying the needs of a wide variety of readers.

At one level, the first and last chapters provide an easily readable overview and historical perspective of the current information problem in its broadest sense, its past development, present magnitude, and likely future growth. Some comparatively recent advances in computer and telecommunications technologies have had a twofold effect. They have made access to specific information much easier, but have at the same time stimulated even more dramatic growth in the total volume of information available. The central problem, then, becomes not one of simple storage and access to the data, but the determination of what is useful or relevant to the purpose at hand, and where it is located.

The authors argue convincingly that indexing methods based on the use of full text retrieval, in conjunction with automatically generated concordances used as part of an overall compression scheme, offer one promising approach to the problem of storing and indexing such large amounts of information.

Having outlined their basic premise, the other chapters then discuss in depth the various components of such a retrieval system. There are chapters on text compression itself, on indexing, on querying, and on index construction. It is expected that large document databases in general may contain more than coded text, so subsequent chapters cover image compression, textual images (such as scanned original documents), mixed text and images, and implementation issues for such systems.

The chapters are extremely well structured, beginning with an excellent first section to overview the content. This is followed by several subsections, to break out the topic and add more detailed discussion. Within a subsection, there is a further division achieved, by marking some of the text with a light grey bar in the margin. This material is intended for use by the more technical or more theoretically inclined reader, and can be skipped if necessary.

The end result is a book that can be used at several different levels. As an overview of a significant problem facing our society, it is excellent, yet it also describes in very readable detail how one can implement a system capable of solving a large part of that problem, giving well-reasoned arguments that justify the various individual elements of the system. Finally, the book can be used as a comprehensive reference, since it goes into sufficient depth for this purpose in almost every topic covered.

Each chapter ends with a section on further reading, which is in most cases a very useful literature source, full details of the referenced materials be-

ing gathered together in a special section at the end of the book.

An appendix discusses the mg system, an experimental, public domain, full-text retrieval system that runs under the UNIX** operating system and demonstrates some of the key ideas described in the book. It is written in ANSII C and can be obtained via anonymous ftp by following the directions given.

Overall, this book is a pleasure to read, covers some very fundamental issues in a manner well suited for the nonspecialist, yet has sufficient detail to serve as a very convenient and compact reference for those working with information retrieval, compression, or related topics.

David J. Craft
IBM Development Laboratory
Austin
Texas

**Trademark or registered trademark of X/Open Co. Ltd.

Asynchronous Transfer Mode: Solution for Broadband ISDN, Second Edition, Martin de Prycker, Ellis Horwood Limited, Hertfordshire, UK, 1993. 331 pp. (ISBN 0-13-178542-7).

This is an excellent presentation of the structure and design goals of Asynchronous Transfer Mode (ATM). Intended for individuals already familiar with telecommunications and networking, it provides a solid foundation for further reading. It is the background perspective provided by the author that makes this book interesting.

The first section is primarily concerned with various transfer modes and discusses the basic characteristics of ATM. The author provides an excellent description of ATM, including the issues underlying the basic structure. He utilizes clear examples to explain the influence of certain design parameters on end-to-end system performance. Error control, delay, and efficiency in ATM are described in an understandable manner and provide an appreciation of the problems encountered with designing a solution for integrating circuit and packet switching. The subject of variable versus fixed-length packets is treated in a similar fashion. These introductory concepts all lead up to not only a good explanation of the "what" of the basic ATM model, but the "why" that many authors neglect.

The next section covers the basic principles of the CCITT (International Consultative Committee on Telegraph and Telephone; now International Telecommunications Union, ITU-T) ATM standard. It concisely describes the basic functions of the multiple layers that comprise ATM. The author starts with the fundamental lower physical layer and takes the reader systematically through the top adaptation layer. Each layer description includes an explanation of the necessary capabilities, while continually referencing the overall ATM structure. Details are not isolated but are combined with general concepts to provide a better perspective of the process. A slight flaw with this section is that there have been recent extensions to the ATM specification that are not addressed. (For the newer extensions, the reader should reference the *ATM User-Network Interface Specification* published by the ATM Forum. Version 3.0 was published in 1993 and Version 3.1, approved in 1994, should be available some time this year.)

Martin de Prycker's treatment of ATM switching architectures follows a nice format. He provides a discussion of basic switching requirements and alternatives, followed by a comparison of multiple existing ATM switching systems, their capabilities, and reasons for specific advantages and disadvantages. The switching systems discussed include Knockout, Roxanne, Coprin, Athena, Starlite, Moonshine, and St. Louis. Nonengineers will find this an especially revealing section if they have not had previous exposure to hardware switching.

The book concludes with general discussions on topics such as variable bit rate video coding, statistical and service multiplexing, ATM LANs, metropolitan area networks, and ATM traffic control. The sections on variable bit rate and traffic control, statistical multiplexing, and ATM LANs are valuable additions and provide additional insight into the overall ATM solution. The discussion of metropolitan area networks provides a background of alternative approaches and how they potentially complement the overall ATM solution.

This book provides a clear overall description of ATM, its goals, capabilities, and areas needing further investigation. A complete and current explanation of all the technical facets and changes is really not possible in any one book. The interested technical person must constantly monitor the activities of the standards groups and major publications to stay current. This book provides a good starting place.

James M. Cavin
Telelead Corporation
Raleigh
North Carolina

Note—The books reviewed are those the Editor thinks might be of interest to our readers. The reviews express the opinions of the reviewers.