Chapter 1

The Information Environment

1.1 Automatic Information Processing

It has been claimed that we live in the information age, and our society is often called the information society. More information is produced and collected in our time than ever before: thousands of books, tens of thousands of journal articles, and innumerable informal studies and reports. Our capacity to absorb this information and use it in reaching intelligent decisions is stretched not only by the amount and variety of the available data, but also by the complex relationships among different types of information, and the resulting difficulties in interpreting the data.

Fortunately, although we are inundated by all sorts of information, improvements are being made in the ways in which information is stored and processed. In particular, modern information-processing equipment can organize and store large amounts of information and provide fast access to the stored records. Communications networks, used increasingly to reach the available information sources, also connect different information stores to large, often far-flung groups of users.

The use of modern computing equipment to process information has had a two-fold effect. On the one hand, it facilitates the generation, collection, and storage of more information, complicating the task of absorbing and using the available data. [1] On the other hand, modern equipment somewhat simplifies the problems of access to information by providing useful ways to search for and retrieve it.

This book deals with modern information processing, that is, the methods used to generate, analyze, store, retrieve, and handle information items using automatic equipment. Current capabilities in information processing are examined, and difficulties and conceptual problems in analyzing and understanding information are described. By distinguishing relatively routine tasks from more experimental, laboratory-type endeavors, and by considering future developments, the book also outlines the information-processing world of the future.

1.2 Types of Information

Information can take three forms: written texts, spoken utterances, and graphs and images. Text, the basic medium for formal communications between human beings, consists of notes, messages, letters, memoranda, books, newspapers, magazines, and so on. Speech is more informal than text and, unlike text, is also accessible to people who cannot read or write. Graphs and images may accompany written texts, but can also be used alone as illustrations, displays, movies, or paintings.

In dealing with these information types it is useful to consider two principal aspects of information processing. The first area is the *technical* problem of information representation and manipulation, including methods of introducing and storing information in computers, and of transferring the data and making them accessible to interested users. The second area relates to the *semantic* and *behavioral* aspects of information processing: the accuracy with which the stored information conveys intended meanings, and the effectiveness with which it affects users' conduct as intended.

From a technical point of view, stored information can be treated simply as collections of disconnected elements — for example, individual words in given texts, individual characters in particular words, or picture elements in graphs and pictures. For processing purposes, the information elements are not assumed to convey specific meanings or to be tied to particular contexts. Thus a text can be reproduced or copied without the text content ever being considered. In actual fact, however, the information elements do carry meaning, and are expected to generate specific responses by the information users. Ultimately the meaning of the information tends to be more important than the form of representation and the manner in which the data are manipulated.

Because meaning is often ill-defined and elusive, semantic and behavioral information-processing tasks are far more difficult to carry out than technical aspects of information processing. [2]

1.2.1 Text Processing

Processing problems are much simpler and better understood for text than for other types of information. For one thing, natural language text can be represented as one-dimensional character strings, whereas speech and pictures are inherently two-dimensional. Speech utterances are often represented by two-dimensional waveforms, known as spectrograms, which can be transformed into digital form for computer storage. Images and pictures are similarly two-dimensional; they can usually be decomposed into picture elements for processing purposes. A picture can then be represented in computer storage by the properties of the individual picture elements. Certain well-defined picture components — consisting for example of special lines and curves — can be described by mathematical equations.

The relatively simpler mode of representing text is reflected in substantially greater storage efficiencies. Table 1.1 gives the storage requirements for one page of information in terms of equivalent numbers of characters, or bytes of storage. It is assumed that a printed page contains about 500 words of text and that the average word consists of six characters, with each character encoded as one byte of eight binary digits (bits) — thus one printed page requires about 3000 bytes of storage. That page of information read aloud produces about four minutes or 240 seconds of speech. A typical conversion of spoken information to digital form takes place at a rate of 9600 bits of information per second of speech — equivalent to about 300,000 bytes of information. Spoken information thus requires a storage capacity two

Table 1.1 Storage requirements per page of information.

Type of Information	Amount of Information per Page	Assumptions	Total Information in Bytes per Page
Text	500 words per page	6 characters or bytes per word	3000 bytes per page
Speech	4 minutes (240 seconds) to read 500 words	9600 bit-per-second sampling rate of speech input	2.3M bits (about 300,000 bytes) per page
Pictures	500 bits per inch (250,000 bits per square inch)	93 square inches per page	23M bits (about 3,000,000 bytes) per page

orders of magnitude larger than required for the comparable written text. The situation is even less favorable for high-resolution graphic data, which typically can be represented by 250,000 picture points per square inch. Assuming that a typical page includes 93 square inches, 3 million bytes are needed to store a full page of pictorial information — three orders of magnitude larger than required for a comparable page of written text.

Given the simpler representation and much greater storage efficiency available for text, not surprisingly it is much easier to manipulate text than speech or graphics. Indeed a wide range of text operations that depend principally on the recognition of individual text words, and individual characters within text words, are now used routinely in many types of applications. [3,4] These operations include automatic editing and formatting methods for texts; detection and sometimes correction of spelling errors; compression of text into reduced representations requiring much less than 3000 bytes of storage per page; encryption of text into a form that hides the normal appearance of the text words; comparison of words occurring in texts with stored dictionary information to determine such word properties as the syntactic functions or translated forms of the words; and retrieval of texts containing certain words or combinations of words.

1.2.2 Speech Processing

The text manipulations just mentioned depend on the presence or absence of words or word patterns in texts, and possibly on statistical properties of texts such as the frequencies with which particular text words or patterns of words occur. It appears simple in principle to extend these methods to speech and graphics information. This is unfortunately not true: There is no single, standard form of representation for all speech utterances and picture components, and these components are much harder to isolate and recognize than individual characters and words of text.

In speech processing, it is necessary to deal with many different forms of speech input, which depend on the type of speaker and the environment in which the speech sample is produced. The same spoken utterance can be represented by different wave patterns, and many otherwise clear utterances are difficult to capture under noisy conditions. In addition, distinct words such as "seen" and "seem" exhibit similar speech forms. [5,6] Since speech is normally produced without detectable pauses between sounds or between the words of an utterance, identifying separate units of spoken information is also a major problem.

It is also necessary to distinguish between speech *synthesis* and speech *recognition*. Synthesis, the production of speech output obtained from various forms of stored information, is used in such appli-

cations as talking instrument panels and question-answering systems — for example, automatic stock-quotation systems, which furnish the values of particular stocks over telephone lines on demand. Speech synthesis is also used in reading machines for visually impaired people, which automatically convert written texts into spoken output.

Conversely, speech-recognition systems use spoken utterances or speech waveforms as input, producing written versions of the input — or at least analyzed forms specifying the structure and possibly the meaning of the input utterances. Speech recognition can be used in telephone transaction systems in which goods are ordered by telephone, but input keyboards are not available to the user. In other applications, machinery is controlled by speech when the operators' hands are not free to use an input keyboard.

Speech synthesis is obviously much simpler than speech recognition: In a synthesis system only one type of voice with known characteristics must be produced, and the desired properties of the voice output can be accurately specified. In speech-recognition applications, however, the quality of the voice input may not be under the control of the designers. In such a case the system must be adjusted to many different input forms, some difficult to handle and analyze. Applications of speech synthesis are therefore much further advanced than those of speech recognition.

Speech synthesis is now possible for particular voice types, especially in environments where the input is limited to particular types of voices and particular classes of vocabulary. But synthesis applications are not always straightforward: In addition to producing individual speech sounds, the "glue" that relates the sound units must also be supplied, as must other speech characteristics that give ordinary speech its human quality — stress, pitch, and intonation for example. The disadvantage of speech recognition systems is that they are available only for small vocabularies (up to a few hundred words). Moreover, the preferable input consists of discrete, as opposed to continuous, speech in which pauses are forcibly introduced between adjacent words of the utterance. Also, the recognition equipment must be trained in advance to the particular qualities of the input voices.

1.2.3 Graphics Processing

The problems of speech processing also apply to graphic information. It is usually difficult to decompose pictures and images into discrete, recognizable elements, and hence manipulations that depend on identifying individual characters or words cannot be easily used in graphic-information environments. As in speech handling, image synthesis must be distinguished from image recognition, and synthesis is again much easier than recognition. Many applications require only simple images to be produced: For example, bar charts showing varia-

tions over time in the price or quantity of goods, or curves outlining statistical data. In such cases, it is relatively easy to process the corresponding graphic elements and substructures, and to transform the pictures by processes like translation, rotation, reflection, duplication, or superposition of certain elements. [7,8] Some applications allow even relatively complex pictures such as shapes of automobile bodies or airplanes to be generated from component parts; image-synthesis programs can then be used for the computer-aided design of these products.

Image recognition presents a much more complicated situation. As in the speech environment, it is difficult to standardize picture input. Thus many different kinds of images, ostensibly representing the same input, must be dealt with. For example, in recognizing handwritten alphabetic characters, as many distinct forms of input may have to be processed as the number of subjects submitting handwritten samples. Furthermore, as in the speech illustration, the individual picture elements appear in context with specific background data. The context may substantially affect the ease with which the elements can be recognized.

Picture-generation systems are usable when there are only a small number of different picture elements and when individual components are well defined. Picture recognition is usually limited to special situations and controlled environments. The technical processing environment for text, speech, and pictures is summarized in Table 1.2.

1.2.4 Semantic and Behavioral Processing

In many applications the detection of words and features within document texts will not solve the problem at hand. A deeper analysis of document content, one dealing with the meaning and the intent of the items, may then be required. Thus in information search and retrieval, it is not always sufficient to identify items containing particular text words. Instead the attempt must be to find records that actually cover a particular topic.

To determine document or record content, a broad background is necessary. Obviously the subject area under consideration must be mastered. Thus when requesting records dealing with information retrieval, the user must be aware that information retrieval is related to library science and to information organization and classification. Further, because information searches leading to retrieval are often carried out automatically, information retrieval is also related to aspects of computer processing. In addition to knowing specific subject areas, an interpretation of document content also requires commonsense knowledge and the wide background normally available to peo-

ple in the world-at-large. Unfortunately, methods are not yet available for identifying the information actually needed in a particular case and for introducing apparently unlimited stores of knowledge into computer memory.

Attempts have been made to design structured *knowledge bases* for particular areas of discourse by specifying the main entities of interest in each area and then presenting specific relationships between entities — synonym relations, whole-part relations, cause-effect relations, and so on. Also, rules have been developed to relate record contents and document texts to the entities included in the knowledge base. However, little success has been achieved in actually analyzing the content of unrestricted texts by computer acceptably: Document content cannot usually be captured by means of limited and circumscribed amounts of prestored knowledge.

Table 1.2 Problem characterization for various information types.

	Technical Problem: Transmission or Processing	**Semantic and Behavorial Problem: Conveying Meaning and Affecting Conduct**
Text	generally under control for areas such as editing, formatting, compression, encryption, message handling, dictionary preparation, storing, retrieving	tractable for severely constrained subject areas or for special-purpose documents in particular environments
Speech	spoken utterances take on many forms *synthesis*: under control for limited vocabularies and limited voice types *recognition*: restricted to limited vocabularies and discrete speech, with training sequences	may be tractable in severely constrained subject areas; more difficult than for text
Pictures	*generation*: easy when needed components are prestored; understood somewhat for graphs *transformation*: difficulties arise in transforming two-dimensional representations to three-dimensional forms *recognition*: difficult in general; easier for limited predefined objects	more difficult than for text because of wider variety of possible relationships among components

For practical purposes, then, content-based information processing is possible only in special circumstances: [9]

1. When the environment is severely limited and can be represented adequately by a few entities and their relationships, or
2. When the documents fulfill special functions that automatically place the texts in particular contexts from which most of the usual ambiguities in interpretation are absent.

Content analyses have thus been carried out for certain controlled *microworlds*, but there is no indication that the methods applicable in these special circumstances can be extended to more realistic situations.

Content analysis of speech samples and pictures is even more elusive than the analysis of written texts because the basic entities and the relationships among them are more difficult to determine. Once again, content analysis is much simplified when the origin and context of the incoming samples are specified in advance. A baby's cry, inherently unintelligible, is often easily interpreted by the mother, who knows the child and the environment in which he or she cries. The same is true of drawings produced by four-year olds: They may appear to be jumbles of lines to uninitiated outsiders, but are accessible to informed observers.

Because written texts are basic to human information processing and can be manipulated easily by computer, text processing is emphasized in the rest of this book. There are three main sections: word processing, including text editing, document preparation, text compression and encryption, and file accessing (Chapters 4–7); information storage and retrieval, including document organization and retrieval modeling (Chapters 8–10); and language processing, including the syntactic and semantic aspects of language manipulation, as well as text transformations such as abstracting and text translation (Chapters 11–12.) In Chapter 13, the processing of multimedia items consisting of text, speech, and images is again examined, together with the design of electronic communication systems, such as mail and messaging systems, electronic conferencing, and automatic library and publication systems. The material of Chapters 4–13 is introduced by an examination of the existing computer hardware (Chapter 2) and modern office environments (Chapter 3).

References

1. I. de Sola Pool, Tracking the Flow of Information, *Science*, 221:4611, 12 August 1983, 609–613.

2. C.E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, University of Illinois Press, Urbana, IL, 1949.

3. A.B. Tucker, *Text Processing: Algorithms, Languages and Applications*, Academic Press, New York, 1979.

4. G. Sampath, *An Introduction to Text Processing*, River Valley Publishing Co., Jeffersontown, KY, 1985.

5. H.L. Andrews, Speech Processing, *Computer*, 17:10, October 1984, 315–324.

6. D.H. Klatt, Review of the ARPA Speech Understanding Project, *Journal of the Acoustical Society of America*, 62:6, December 1977, 1345–1366.

7. J.D. Foley and A. van Dam, *Fundamentals of Interactive Computer Graphics*, Addison-Wesley Publishing Co., Reading, MA, 1982.

8. W.M. Newman and R.F. Sproull, *Principles of Interactive Computer Graphics*, McGraw-Hill Book Co., New York, 1979.

9. P.J. Hayes and J.G. Carbonell, *A Tutorial on Techniques and Applications for Natural Language Processing*, Carnegie-Mellon University, Report CS 83-158, Pittsburgh, PA, October 1983.