Natural Language Communication with Pictorial Information Systems

Edited by Leonard Bolc

Natural Language Communication with Pictorial Information Systems

Edited by Leonard Bolc

With 67 Figures

Springer-Verlag
Berlin Heidelberg New York Tokyo 1984

Leonard Bolc

Institute of Informatics, Warsaw University
PKiN, pok. 850, 00-901 Warszawa, Poland

ISBN 3-540-13478-6 Springer-Verlag Berlin Heidelberg New York Tokyo
ISBN 0-387-13478-6 Springer-Verlag New York Heidelberg Berlin Tokyo

Printing: Beltz Offsetdruck, Hemsbach/Bergstr.; Bookbinding: J. Schäffer OHG, Grünstadt
2145/3140-543210

Preface

This book contains the reports of selected projects involving natural language communication with pictorial information systems. More than just a record of research results, however, it presents concrete applications to the solution of a wide variety of problems. The authors are all prominent figures in the field whose authoritative contributions help ensure its continued expansion in both size and significance.

Y. C. Lee and K. S. Fu (Purdue University, USA) survey picture query languages which form an interface between the pictorial database system and the user and support information retrieval, data entry and manipulation, data analysis and output generation. They include explicit picture query languages that augment alphanumeric data query languages as well as languages and command sets which are implicitly embedded in a pictorial information system but perform similar functions. It is worth mentioning that some forms of query languages can be transformed from a given set of natural language sentences by using ATN (Augmented Transition Networks), which consequently allows for natural language communication with information system.

M. Hussmann and P. Schefe (Hamburg University, West Germany) describe a natural language dialoque system for scene analysis: SWYSS (Say What You See System). Recently, SWYSS has been constrained to include several modules that can serve as an interface to a scene analysis system, the output of which is a geometric scene description. In this system, a phrase-structure grammar is used, augmented by dependency rules to produce dependency structures for natural language inputs not constrained to complete and well-formed sentences.

M. Yokota, R. Taniguchi and E. Kawaguchi (Kyushu University, Japan), in their paper "Language-Picture Question-Answering Through Common Semantic Representation and Its Application to the World of Weather Report", present methodologies for constructing systems which can understand both natural languages and pictorial patterns. These methodologies involve preliminary analysis of linguistic and prictorial data performed by corresponding pictorial data to linguistic from the viewpoints of syntax and semantics. Their understanding system consists of six divisions, all of which are connected through the medium of meaning representation common to language and picture. They also describe the applications of the understanding systems to weather reporting in a system called ISOBAR (an Information understanding System of Basic weather Report). The most prominent feature of ISOBAR is that its performance is based on the semantic processing of the input.

O. Eriksson, E. Bengtsson, T. Jarkrans, B. Nordin and B. Stenkvist (Uppsala University, Sweden) introduce ILIAD, a comprehensive digital picture analysis system interacting with the user through a high level language. ILIAD runs on a general purpose computer,

but contains provisions for utilizing different kinds of special hardware. ILIAD's high level language makes it programmable at the monitor level. Thus, different picture processing algorithms can be programmed directly using a PASCAL-like command language. The system is equipped with a procedure facility: a self-contained piece of ILIAD code which enables each picture processing problem to be expressed in a structured, thus simplifying the use of the system. The full system can be implemented by an experienced user, althrough for novices it could be simpler to use the preprogrammed procedures. We present this publication with the conviction that the problem it includes will help disseminate the knowledge generated by research in the field of natural language communication with pictorial information systems, an area still too little known in the world today.

Warsaw, October 1983 Leonard Bolc

Table of Contents

Query Languages for Pictorial Database System

Y. C. Lee and K. S. Fu

School of Electrical Engineering, Purdue University, West Lafayette, IN 47907/USA

Preface

To store, retrieve, and manipulate a variety of pictorial data of huge volume, a pictorial database system is highly desirable. As the interface between the pictorial database system and its user, a picture query language is to support information retrieval, data entry and manipulation, data analysis, and output generation.

The main purpose of this monograph is to provide a survey of picture query languages. It includes languages that augment previous alphanumeric data query languages and have been explicitly labeled as picture query languages. It also covers languages or command sets which are implicitly embedded in a pictorial information system but perform similar functions.

In Chapter 1, an overview of picture query languages is followed by the discussion on the design considerations of a picture query language, including existing approaches, basic problems, and general requirements. In Chapter 2, related background is briefly reviewed in order to help the understanding of problems and techniques that have been addressed and proposed respectively. Following the general frame sketched in Chapter 1 and the background reviewed in Chapter 2, various picture query languages are discussed in groups.

Since a picture query language is either explicitly or implicitly built on top of a pictorial data management system, its design is influenced by the underlying system. Accordingly, a query language is discussed together with its supporting system but with a strong emphasis on the language aspects. In Chapter 3, several languages or systems that mainly support the management of image data are reviewed. They are EIDES, IQ, ADM, and IDMS. Approaches with either practical purposes or idealized goals are illustrated. Moving into Chapter 4, another group of four systems, namely MIDAS, GRAIN, PICCOLO, and an image database manipulation system, is described in order to introduce how image contents can be incorporated with images themselves. All four systems have distinctive features in system organization, data structure, and user interface. Discussed first in Chapter 5 is the QPE language which features the *query-by-pictorial-example* concept and emphasizes the importance of geometric computation. Introduced later in the same chapter is the GSYSR system, a relational database interface for graphics. Chapter 6 contains three geodata analysis systems including GADS, IBIS, and a cartographic database. Each of them shows new and useful design approaches.

The authors are fortunate to receive invaluable comments and encouragements from Dr. N. S. Chang, Dr. H. C. Lee, Dr. Y. F. Tsao, Dr. K. Y. Huang, Mr. W. C. Lin, Mr. M. A. Eshera, Mr. T. H. Chiu, and Mr. C. C. Liu. Mr. C. C. Liu and Mr. T. H. Chiu also helped redrawing some figures from original literatures. The authors owe a debt of gratitude for the support of the National Science Foundation. The text has been entirely prepared using the UNIX† operating system at Engineering Computer Network of Purdue University.

Y. C. Lee
K. S. Fu
Lafayatte, IN
August 1983

† UNIX is a Trademark of Bell Laboratories.

Chapter 1
Introduction

1.1. Pictorial Database system and Picture Query Language

Along with the advances in image processing and computer graphics during the past two decades, many problems have been addressed and approaches proposed for the storage, retrieval, and manipulation of pictorial data. A major recent attempt is to incorporate, more or less, existing database management system (DBMS) techniques[26,47,73] within a pictorial information system. A *pictorial database system* (PDBS) resulted from this attempt is thus a database management system which provides an integrated collection of pictorial data for easy access by a large number of users[20].

A *picture query language* is basically a computer language through which the user of a PDBS can pose queries related to pictures. In addition to supporting the fundamental tasks of information retrieval, a picture query language usually also provides facilities for the storage and manipulation of pictorial data. It may further include a

rich set of analysis and output generating functions, as the use of its underlying PDBS is possibly extended to support image processing, image understanding, and problem solving, etc..

Pictorial data items from various sources cover a wide range of varieties and can be encoded in various data structures. Each structure has its own advantages in representing certain types of data and in meeting requirements of some applications. Although pictorial data are also stored in alphanumerical format at the physical level, they are, in nature, quite different from alphanumerical data. Spatial information that is either implicitly or explicitly specified usually requires additional interpretations or a large amount of computation both in displaying the retrieved information and in processing the queries. A picture query language is thus characterized not only by its data structure but also by its pictorial manipulation capabilities.

As many PDBS's try to take advantages of the existing DBMS techniques, many picture query languages are designed as augmented versions of various conventional query languages[10]. These picture query languages therefore have certain features inherited from their original query languages whose nonprocedurality and descriptive power are heavily dependent on the underlying database models.

Although PDBS approaches in principle follow DBMS techniques, the result is rather disappointed when all the existing PDBS approaches are examined with respect to various developed DBMS disciplines such as data independence, data integrity, and concurrency control, etc.. This discrepancy is particularly significant in the aspect of picture query language. At one extreme, some languages are simply composed of a set of commands tailored to the existing system. At the other end, several languages strictly follow the design requirements of DBMS. It can be observed that system-tailored approaches usually aim at a short-term purpose and only stress practical considerations while rigorous approaches emphasize theoretic foundations yet appear to be

somewhat unrealistic.

Strictly speaking, the area of PDBS has been very interesting but not yet completely defined and developed. It is due to several reasons. First, different applications which impose different processing requirements have been involved in this area. The diversity of the attacked problems has resulted in different techniques that are successful in a limited sense. Second, there is still a great need of efficient and reliable techniques for the processing and manipulation of pictorial data. Without these techniques, none of the current PDBS approaches can be as disciplined or as feasible as conventional database management systems. And last, like in other research problems, a PDBS approach may look appealing but is not necessarily the best way to handle the storage, retrieval, and manipulation of pictorial data. But if the problem of PDBS can not be handled by other approaches, significant modifications to the existing DBMS techniques might still be needed.

Obviously, this is also why researches on PDBS as well as picture query languages remain extremely important as the number of applications which need to be supported increases. In the following section, principal PDBS supported applications are briefly discussed. Other applications are mostly based on these principal applications in the aspect of pictorial data management.

1.2. PDBS Supported Applications

Image Processing

Image processing has long been an important research area emphasizing methodologies that code, enhance, and restore images for human interpretation or machine perception. Processing techniques, usually in the form of algorithms, have to be examined and evaluated on a set of different images. Accordingly, a database needed in such an environment is required of convenient facilities for retrieving, storing, manipulating, and displaying images. A query language associated with it often focuses on the integration of image management commands and basic image processing routines.

In this kind of PDBS, Images or windowed subimages are the basic data items. Intricated relationships among data items, as in commercial alphanumeric database, are absent from this case. Therefore, the only aspect where the existing DBMS techniques can be directly applied is the management of image directory, or secondary registration information. Since only little consideration that is specific to picture has to be made for this aspect, PDBS approaches for image processing purpose tend to emphasize a convenient and useful command language which will practically relieve the user of tedious image file manipulation and provide the user sufficiently with various fundamental image processing tools. This is quite different from the focusing point of a conventional DBMS in which database content is retrieved with respect to different selecting conditions posed on a large number of entities and relationships.

Image Analysis and Understanding

One purpose of using image processing techniques is to support machine perception of images. Image analysis (and understanding) in this sense include all the processes that utilize the results of image processing and extract useful information from images. The extracted information in alphanumerical format can be used to support query-answering, decision-making, or further processing of images. Thus we see a

strong need of a database management system capable of organizing both nonalphanumerical images and alphanumerical image descriptions. Furthermore, the interplay between image data and their corresponding image descriptions necessitates a set of geometric operators to interpret spatial properties.

PDBS approaches for this purpose include research-oriented and application-oriented systems. Research-oriented systems are intended to support comparison and evaluation of different algorithms. Application-oriented systems are aimed at reducing repetitive information extraction tasks and organizing the extracted information for practical uses. Both kinds of systems have fairly complicated system hierarchy in order to achieve certain degree of data independence and to provide users with easy and concise interfaces.

Computer Graphics

While image analysis and understanding are a process that inputs images and outputs image descriptions, computer graphics implies an almost reverse process that generates images according to the descriptions. It plays an important role in computer-aided design and manufacturing. The design data which is interactively produced and visually approved by graphic systems will be used by simulation, analysis, or manufacturing programs. As a consequence, a supportive database for computer graphics is generally characterized by a systematic organization of data items, a convenient tool for data entry, and a powerful set of display commands. It also serves as a common data source for various applications.

Different from the previous two applications, nonalphanumerical image data items are not included in this sort of database. The database is distinguished from conventional DBMS's by geometric computational capabilities rather than different formats of data items.

Geodata Analysis

In a comprehensive sense, current geodata analysis can be considered as a typical application area which requires all the PDBS techniques for image processing, image analysis, and computer graphics. Geodata analysis has been traditionally using grid cell and polygon methods until when remotely sensed imagery was introduced. Satellite images and aerial photographs are important data for obtaining time-varying information about large areas. The combination of imagery data with existing maps and census data is therefore a crucial ingredient for an advanced geodata analysis system that is capable of extracting significant information from a vast amount of data in various formats.

A complete geodata analysis system appears to be much more complicated than those in the previous three kinds of applications, namely, image processing, image analysis, and computer graphics. It primarily deals with input data in matrix (image) format, polygonal (graph) structure, or tables. Format transformation and processing algorithms for each type of data are of particular importance in such a system.

1.3. Design of a Picture Query Language

In this section, issues related to the design of a picture query language are outlined. Three different design methodologies characterizing the existing approaches are first discussed. Basic problems associated with the language design are then examined with respect to these approaches. These discussions thus suggest a list of general requirements which can be used as a guideline in studying advantages and disadvantages of an existing language or in developing a new one.

Existing Approaches

It is worth noting that some languages described in this monograph are not necessarily initiated as database approaches, as discussed in the following. However, these languages have dealt with the same problems that distinguish a picture query language from a conventional query language, namely the manipulation and analysis of pictorial data, and some of the techniques employed can be of significant help in the design of a picture query language.

1). *System-tailored Approach* - In order to facilitate certain data processing tasks, it is usually a straightforward and economic way to develop a set of frequently used commands and/or subroutines by following the specification of the supporting operating system. Without exceptions, such an approach, which creates very few difficulties to a noncausal user with appropriate programming and system knowledge, is adopted by many pictorial data processing systems. As indicated in the previous section, a large number of I/O routines and standardized primitive operations are needed to support the processing of pictorial data of which the size and complexity are extremely large. By this approach, the commands and/or subroutines are developed whenever needed in an ad hoc fashion that only short-term considerations are taken. Without a unified communication between existing commands and the liberty for adopting new software and hardware facilities, such an approach usually fails to establish flexible and portable language packages.

2). *Structured Top-down Design* - The structured top-down design emphasizes a disciplined software management technique that systematically organizes the hierarchy of commands in order to localize the device dependence and maintain the flexibility to include new facilities. This approach starts with the overall consideration of the system requirement, which may rely on the experience accrued from the system-tailored approach. After examining all the tasks that will be supported, the design basically

includes:

i). extracting common primitive operations from all tasks;

ii). defining a simple, unified format to be followed by every command; and

iii). constructing the command hierarchy.

In general, the resulting system, concerning the user, is still file-oriented even though the format of each working unit, file, is well specified. In rigorous database approaches, files, if exist, are under the control of the management system at physical level and are completely transparent to the user. The following approach which originates from conventional query languages is somehow different from the preceding two.

3). *Augmentation of Conventional Query Languages* - Image descriptions or other alphanumerical yet spatially related data are regular data items and can be organized and manipulated in conventional database management system. Naturally, a good approach to designing a picture query language is to enhance an adequate existing query language so as to include the additional facilities needed for picture query processing. These required facilities range from fundamental geometric computation capabilities to sophisticated image analysis algorithms, depending on the applications supported. Ideally, this approach tries to design a high level, nonprocedural query language which can handle different formats of pictorial data uniformly. For example, the data items should not be limited to only alphanumerical data in specifying the retrieving condition of a query.

Basic Problems

Although problems of various degrees of difficulty have been associated with different applications, most of them actually result from the fact that pictorial data items are often not a perfectly ready source from which answers to picture queries can be directly retrieved. This implies the need for a very rich set of operators aiming at bridging the gap between the raw pictorial data and the query answers. In fact, a similar need also happens, but at a much less crucial level, in conventional query languages in which aggregate, arithmetic, and output generating functions are required to enhance the manipulation capabilities or the appearance of query answers.

A common comparison operator *equal to*, which is used to specify certain selecting conditions, can be examined here in order to realize this need. Suppose that A and B are two images, and the notation = stands for the operator *equal to*. A simple expression, $A = B$, will involve a time-consuming, point-by-point matching problem and first of all, the use of this expression will be meaningful only when A and B are identical images. Instead of this, an equivalence in the sense of feature-matching seems to be more realistic and feasible. For example, $A =(w.r.t.\ feature\ x)\ B$ can be evaluated by a certain algorithm if provided. Such an algorithm may be much more complicated than a point-by-point matching procedure but is more realistic, considering all the possible translational, rotational, and brightness differences when an identical scene is pictured repeatedly. If more distortion is allowed, the equivalence may be defined on a similarity measure. A number of PDBS approaches chose to precompute several desired features so as to reduce the computation load of query processing. This is part of the reason why image descriptions are the actual data items being manipulated in several systems.

It appears to be that all problems are raised in an effort to achieving that images can be treated uniformly as other data items while their contents can also be accessed