*

KNOWLEDGE OF THE TERRITORY: A DATABASE WITH USER FRIENDLY INTERFACE

Roberto della Maggiore - Oreste Signore

CNUCE, Institute of CNR - via S. Maria 36 - 56126 Pisa (Italy)

ABSTRACT

The principle sources of information regarding national territories are gained from cartography, aerial photos, bibliographical description of the soil. Unfortunately such documents are managed with scarce uniformity.

The aim of this work is to provide a tool for knowledge of the land by collecting heterogeneous information in the same data base. Each piece of information is related to topological and geographical information of the territory concerned (covered area coordinates).

A database with the relevant items for cataloguing and retrieving information was implemented, and user interaction in the data entry phase is performed through applications written using ORACLE-FORMS. The database may be queried either using standard SQL language or FORMS written applications. In addition, the relevant items may be selected through political, administrative, or ad hoc defined areas which group together elementary territory units. However, one of the most important selection criteria are the coordinates of the area the user is interested in: therefore we implemented a user friendly graphic interface, which allows areas to be selected simply by moving the mouse over a topographic representation of the region the database refers to.

The system was developed under the supervision of the technical staff of CNUCE, using ORACLE in MS-DOS environment, following the directives of the "Centro Interregionale di coordinamento e documentazione per le informazioni territoriali".

INTRODUCTION

* Presented at the 9th European Oracle User Conference, 22-27 March 1992, Cannes, France.

Though knowledge of geographical areas is important, it is not easy to gain access to such knowledge. A basic tool is the cartography, but cartographic reproductions are not stored and catalogued in uniformly. Other methods are aerial photos, remote sensed imageries, and descriptive texts. In any case locating particular information entails knowing how it has been linked to the relevant geographical area, which may be laborious.

OUR APPROACH

We began on the premise that all information is related to a specific land area. This is true for aerial photos, thematic maps, as well as for bibliographic or descriptive information [dellaMaggiore88].

Traditionally every item is linked to the name of the localities (typically, in the case of the maps, to the most important towns or localities). This method is costly (an expert must perform the indexing) and has some drawbacks. The most important one is that finding an item may be impossible if the user is searching on the basis of a locality that the indexer decided should not be kept as an index term. This aspect may become relevant especially where the relative importance of two localities has changed, or the name of the town has changed or has variants in different languages, or a new town has grown up. The problem of “historical geography” is going to become very relevant, as recent political modifications have shown.

On the basis of these considerations, and bearing in mind that the user is normally thinking above a map with reference to the area it is pertinent to, we thought that it would be much more efficient to identify the covered area by using coordinates ([dellaMaggiore89], [Barducci90]). In this approach, users will simply identify the area they are interested in by specifying its coordinates.

THE STRUCTURE OF THE DATABASE

The database design process was conducted following standard methodological steps:

- analysis of user requirements;
- conceptual design;
- logical design;
- physical design.

Great emphasis was put on the conceptual schema, as this is recognized to be invariable in the mid-term, and so may be an important issue for future applications.

In this phase, following the Entity Relationship ([Chen76]) methodology, the entities and their relationships were identified. We used a CASE tool, IEW by Knowledgeware, which was very effective in sharing the information within the development team, and helped us to produce a complete and accurate documentation. However, the tool showed some drawbacks: the absence of the concept of the IS-A hierarchies, the lack of rich, user defined data types, and the impossibility to formally state integrity constraints. But these drawbacks, as far as the IS-A hierarchies are concerned, had to be resolved during the preliminary conceptual design phases, while the other problems were solved by specifying appropriate comments, and making reference to them in the subsequent design phases.

There are two main aspects of the resulting database: firstly, “information entities”, i.e. the entities that represent information such as maps (thematic or technical), bibliographic references, aerial photos, firms that for various reasons are related to the maps, homogeneous sets of maps.

Secondly, the “referential entities”, which act as a link between the administrative and geographical items. To be more precise, we defined punctual, linear and aerial elements, which are fitted by a set of rectangles, identified by the coordinates of their North-East and South-West vertices. We subdivided elements relating to the areas into “primary areas” and “extended regions”, the latter are defined in terms of the “primary areas” that they contain. In turn, there is a recursive association of the “extended regions” within themselves, as every “extended region” may be defined as a collection of other elements of the same type. This solution is a convenient shortcut because the user can easily specify particular areas, that may subsequently be used at the query formulation stage.

This method of representing the areas gives users some freedom as they can access the information by using the name of a region or of a locality. This also frees the indexer from the burden of identifying the right names that have to be linked to the specific maps. The definition of these different areas, and the possibility of having a reference to them by name, acts as a conceptual representation of the data, and is the underlying software that will be responsible for mapping the information onto the appropriate set of rectangles, and therefore for coordinates.

As far as the implementation environment is concerned, we developed the “standard” management part, i.e. data entry, updating, query on the alphanumeric part, directly by using FORMS. Our decision was based on the well known advantages of IV generation languages, especially the speeding up of the development process.

Due to the peculiar way the applications are implemented in FORMS, we stopped using CASE at the database definition step, and made no use of the facilities for the definition of the Structure Charts and Action Diagrams. Another reason for doing so, was that basically the data entry part of the application as well as the user inquiry facilities were simply constituted by a main menu, followed by several minor menus, and interactions with the related tables.

THE GENERAL ARCHITECTURE

The structure of the data

The two main sets of documents constituting the data base, information and referential entities, share no predefined link. The fundamental characteristic of every document is the geographical reference to the territory, expressed in terms of the coordinates of North-East and South-West vertices of the rectangle containing the area which the document relates to. The dynamic link between documents of the two sets is the intersection of the rectangles, computed at query time.

In the implemented prototype, information entities are of two main kinds: maps and aerial photos. Maps may be thematic or technical (only a few attributes are different) and consist of sheets (at least one per map). Adjacent sheets form an entire map and for the sake of input simplicity sheets are grouped in sets called lots. Maps, lots and sheets are contained in tables and are linked to other tables containing utility information, such as names and addresses of corporations and firms in some way involved with maps.

Aerial photos may be single shots, stripes of photos, or sets of stripes; these documents too are related to the utility information.

Reference information consist of administrative entities whose vertices of the rectangle containing the land area pertaining to and the point of the administrative centres are known.

Data are organized on a regional basis: each region is responsible for the data of its territory. However, internal codes are assigned so that information of different regions may be collected (and queried) on the same data base. This is required in order to supply the central administration with a global view. It is also useful to gain knowledge of the situation near to the border of neighbouring regions.

Data management

The system has three main operating environments: under ORACLE FORMS two simple menus were implemented, one for data input and management and one for user queries. Queries may also be easily accessed from the third environment: the graphic interface.

Data may be easily inputted from the main administrative menu. FORMS triggers were largely used in order to support the user who must not necessarily be skilled in data base systems. Moreover, to speed up input operations, the user only needs to specify those data that are common to many documents once. In many cases it is possible to get dynamically a list of the possibilities of input on a window and to choose from them.

Data are maintained by selecting from the submenus to modify, delete or, in the case of maps, update documents. Here, too, it is possible to open a window with the list of the data in the data base. “Updating” maps is distinguished from “modifying” because it permits one to keep an historical sequence of the maps by providing a copy of the map under update. Two different copies of the document are thus made.

User queries

Two main kinds of query are possible through the query menu: geographical or hierarchical. The first is started by the name of the referential information of the area the user is interested in. It uses the intersection of its geographical coverage and those of the items of the data base with the same specified characteristics. This process is rather costly so we had to write an exit routine to perform it with an acceptable response time. Users set up their queries by choosing the name of the locality from those which the system knows; nevertheless they can define a new reference item at any time, with a new name, to search in an area which is not currently allowed.

The hierarchical query is of course faster but cannot give good results if the user is not well acquainted with the topology of the territory and, most of all, with the consistency of the data base. Documents can be retrieved by starting from the lists of the maps and going down to the single sheet.

The graphic interface

The data base is best exploited through the graphic interface which presents enhanced interactive features and is designed to perform geographical inquiries starting from the graphic visualization of the territory.

A topological map of the region is initially displayed on the screen, showing such information as administrative boundaries, rivers, main roads or whatever else is needed to recognize the territory. With the mouse the user can zoom out the area and the system shows the localization (and the names, if possible) of the administrative centres, provided there is sufficient detail level to prevent overlapping. The inquiry is started after enclosing the geographical area within a rectangle and after selecting the scale of representation of the documents to be found. The characteristic rectangles of the retrieved maps are overlaid on the topological map of the region and it is possible to skim through them and view the textual information attached to them.

Reference information may be accessed as well by asking the system to visualize the rectangles of administrative centres and then using them as a guideline to set up the inquiry.

The use of different colours is essential to distinguish the semantic meaning of rectangles.

The technical environment

The prototype system was implemented on an IBM compatible Personal Computer 286 running DOS 3.0, equipped with an extended memory of 2 Megabytes, a hard disk of 40 Megabytes, colour monitor, Enhanced Graphic Adapter (EGA) card and a three keys mouse. ORACLE version was 5.1A and 5.1B. Of course the hardware configuration should be revised if installing version 6.

The graphic interface was developed using Microsoft C and ORACLE's Pro*C.

CONCLUSION AND FUTURE WORK

We have implemented a user friendly and effective application for the management of the territorial information. We succeeded in having a simple, user independent way of coding the relevant items relating to maps and the other data. In addition, the graphic interface, which permits the user to interactively select a particular area, gives the entire system a friendliness and a flexibility which exceeds the more usual way of interaction, based on the identification of the most relevant place names in each map. Even more flexibility is possible by defining special interest areas in terms of other areas.

One of the major issues of the implementation environment was defining the constraints, either at the conceptual design level, or at the implementation level. At the conceptual level, constraints were defined as annotation, or comments, on the description of the relevant entities. At the implementation level, the absence of a way for the definition of the constraints directly at the schema level, and the FORMS trigger mechanism forced us to define the same constraint more than once, with the obvious risk of possible inconsistencies. In conclusion, even if the usage of FORMS led to a simpler and faster approach in the implementation of the application, the need to implement a rather sophisticated enquiry system showed that the traditional approach may sometimes lead to a more effective and efficient code. We found that the use of C language exit routines speeds up drastically the enquiry process when we use geographical intersection formula to retrieve information on a specific area.

At present we are working mainly on the upgrading of the graphic interface. The main topics are:

- visualizing images associated with a document or a group of documents;
- increasing the interactive possibility of the system, giving users the opportunity to "take note" of some documents they have found;
- making overlaid graphic visualization of retrieved documents possible.

Our aim is to set up a flexible and user friendly data base system that incorporates an image data bank suitable for the knowledge of the territory (in this case we suppose that such images as aerial photos, remote sensed imageries, map samples, and so on are available). Nevertheless a system with these characteristics could be used to manage all kinds of information and visual data related to a territory.

Another possible area is the addition of some "hypertext" characteristics, which will make it possible to move from one document to another, following some referential or organizational links (for example a link between one "extended region" to another, or one "primary area" to the different "extended regions" it belongs to).

An interesting technical matter currently under study is the connection of a DOS station running the graphic interface to a mainframe running ORACLE. This could be achieved through ORACLE's networking facilities; however an even better method is to use X protocol to develop the graphic interface, thus abandoning the constraints imposed by the DOS system. However this seems to be too troublesome to implement, owing to the apparent incompatibility between ORACLE and X environment.

ACKNOWLEDGEMENTS

A warm thanks is due to the “Centro Interregionale di coordinamento e documentazione per le informazioni territoriali” who proposed the work and helped us in the refinement of the specifications of the application.

Thanks also to Alberto Barducci, who wrote the code.

BIBLIOGRAPHY

[Barducci90] Barducci A.: Una base di dati per la gestione di un repertorio cartografico, Ph. D. Dissertation, University of Pisa, Academic year 1988-89

[Chen76] Chen P.P.: *The Entity-Relationship Model: Toward a Unified View of Data*, ACM Transaction on Database Systems, Vol. 1, Nº. 1, (1976), pp. 9-36

[dellaMaggiore88] della Maggiore R., Mogorovich P.: *Repertorio Cartografico Interregionale - Ipotesi di realizzazione,* Internal note C88-19, (CNR - CNUCE, 1988)

[dellaMaggiore89] della Maggiore R., Signore O.: *Una base di dati per la conoscenza del territorio*, Documenti del territorio, Nº.15-16, pp. 62-66, (Centro Interregionale, 1989)