Domain Knowledge Organization for Encyclopaedia Design: An Object Oriented Approach

Abstract: In order to build an electronic encyclopaedia we propose a methodology based on the object oriented method which permits encyclopaedia domain analysis and decomposition, and objects, classes, operations, relations identification. We equally characterize a generic object called encyclopaedia basic document which, if instantiated, leads to encyclopaedia elementary scenarios. Grouping them together thanks to external links leads to the encyclopaedia final scenario.

1. Introduction

Human knowledge is geographically distributed in a natural manner. In fact, we find it in the expert brain, in books, papers, computer storage, ... Those various knowledge «containers» can be located in different towns, countries, in distinct buildings and offices, ...

To build an electronic encyclopaedia dealing with a particular or a general domain (art, history, sport,...) we need knowledge gathered from all over the world. This fact generates several problems such as knowledge availability which is strongly related to knowledge «containers» availability due to miscellaneous speaking languages, knowledge source separating distances, expert physical states,...

Most of those problems can be solved using networks, collaborative software, multimedia technologies, but first design methodologies must be reviewed to be adapted to the technological «explosion» touching computer science fields such as multimedia.

In this paper we propose a methodology based on the following approaches:

1. Object-oriented methods which permit domain analysis and decomposition, and objects, classes, operations, relations identification.
2. Elementary document composition methods which elaborate elementary scenarios.
3. Links method which permits the composition of the final scenario by linking elementary scenarios.

Steps 2 and 3 can be performed using either graph theory (Murota, 1987) or an hypermedia presentation model (Begoray, 1990).

According to the problem domain, the methodology can be performed by several teams which can interacted whether using local networks or the wide world web (WWW).

2. Encyclopaedia Life Cycle

To obtain an electronic encyclopaedia as a final product, some steps must be performed so as to correspond to an object oriented architecture (Booch, 1992). These steps called encyclopaedia life cycle are analysis, design, implementation and modification.

2.1 The Analysis

In this step, encyclopaedia users and developers meet for the first time. Discussions dealing with the problem domain can be led and will permit the definition and the use of a common vocabulary.

Moreover the problem and the behavior of the desired system description can be
determined. In this step, the more important objects and the relations between them are identified.

This step can be performed whether in planned «natural» meetings at a distance using communication technologies such as mail, news groups, GroupWare.

2.2 The Design

In this step the system details will be generated in order to implement the final system. A problem formal model will be obtained. The design process is characterized by the use of a set of prototypes, each of them concerned with an important aspect of the system architecture. As well as the first step, this one can be performed using current communication technologies.

The encyclopaedia is essentially composed of elementary objects such as words, image (static, animate), sound, … The binding of those elementary components leads to the instantiation of a generic object called an encyclopaedia basic document.

2.2.1 Basic Document, Basic scenario

A basic document is essentially defined by its components and its description. Its main components are the following three basic elements:

- Terminal component: text, image, audio, video
- Anchors. An anchor is considered as a basic document part which is composed of an organized set of terminal components.
- Links (Bloomfield, 1994) which can be internal or external. An internal link is associated with an anchor. An external one is associated with a basic document.

A more detailed model can be found in (Filho and al., 1995).

The organization of the terminal components constitutes the basic document description. This one can be performed using several tools such as (HTML, SGML, ODA, …) and programming languages (Visual Basic, Visual C++, …). The basic document description leads to what we call the basic scenario, which is defined as a mapping of the basic document onto a window. In that way, the external links can’t be temporarily activated and navigation is authorized only in a basic document, which means that the association to external links will be temporary undefined.

2.2.2 Encyclopaedia: A final Scenario

When writing all the basic documents we obtain an electronic book thanks to external links established between them. These documents must be indexed according to many research criteria: themes, articles, keywords, images, …

The final step needed to achieve the goal of having an operational encyclopaedia is to realize the user interface (Plaisant, 1989) which must include usual objects such as menus, dialog boxes, …

The search according to one or several criteria can lead to many basic documents. Choice and navigation through them will permit users to go deeply into their research by exploring new paths in encyclopedic knowledge.

3 Encyclopaedia Actors

The Encyclopaedia (from Greek egkuklopardia), is a book where we find a set of universal or specific knowledge related to one domain. It is considered as a complex system that is not easy to analyze, to design and to implement. The decomposition of the encyclopaedia domain allows a reduction in complexity. The complexity is usually hierarchical in which a complex system is composed of subsystems related to each other and
the subsystems have their own subsystems and so on, until we reach the lowest level of elementary components (Courtois, 1985). The choice of these elementary components is relatively arbitrary and widely dependant on system analyzers and designers. In our case, we suppose that an elementary component of the encyclopaedia corresponds to one specialist skill which has responsibility for the definition of a part of knowledge to include in the encyclopaedia.

Inside components, the links are generally stronger than those between the components themselves. This fact separates high frequency dynamics of elementary components (concerning interaction between elementary component specialists) from low frequency dynamics of components (concerning interaction between specialists of the whole elementary components) (Simon, 1982).

The specialists constitute the main actors of the encyclopaedia project. They have to work together using the World Wide Web for example. In addition to specialists, there are other actors who play an important role. Those actors are given in order:

- The group of persons who are at the origin of the idea.
- The piloting group who have responsibility for realization. The members of this group have the important task of defining the domain, its decomposition and its mapping on the specialist.
- The idea realization group which can be divided into two subgroups: design and implementation group.

All those groups must work together in a collaborative way.

4. The Encyclopaedia Implementation and Modification

The system implementation in the object oriented development life cycle combines the traditional aspects of code writing, test and integration. The development process has as a result, an incremental production of a set of prototypes which evolve to the final implementation.

In this step we introduce many changes such as new class addition, class implementation, interface and/or representation change, class structure reorganization.

5. Conclusion

We proposed a methodology based on the object oriented method which permits encyclopaedia design and implementation. Those tasks have to be performed in a distributed way using the world wide web tools. The proposed methodology must represent clearly the groups interaction, the documents mapping, and formalize the documents redaction. Those tasks are now being studied.

References
Bloomfield, H. (1994). Links in Hypertext: An Investigation into how they can provide on inter-node relationships. PhD, Queen Mary and Westfield College, University of London.

Advances in Knowledge Organization, Vol.6(1998)
Nodes. ACM Workshop on Effective Abstractions in Multimedia With Composite Nodes.