Exploiting Semantic Networks to Generate Hypermedia

Yannis Tzitzikas

Department of Computer Science,University of Crete
and
Institute of Computer Science,
Foundation for Research and Technology - Hellas
email : *tzitzik@ics.forth.gr*

Abstract

Hypermedia is a hot research topic nowadays with a wide application domain. In this paper we study the problem of generating hypermedia presentations from semantic network-based repositories. After discussing the application domain of this approach, we focus on issues concerning the mapping of semantic networks to hypermedia structures. In addition, we present briefly our experiences from implementing a prototype system which employs the above approach. Finally, we present some possible extensions and related open issues.

1 Introduction

Hypermedia constitute a very effective and friendly way to present information which is employed nowadays by many popular systems (www servers are the most indicative) and is subject of intensive research by many research areas including hyperbases, open hypermedia systems, hypertext protocols and operating systems. Research is being done concerning many issues of hypermedia development such as design, architecture, representation, expressiveness, flexibility, production automation, extensibility, integrity, consistency and other quality issues.

Hypermedia models are characterized by big differentiation according to structuring and implementation. For example some of them support link types whose multitude can be small (KMS, HDM) or large (gIBIS, TextNet). In addition, links can be *extensional* (explicitly declared) or *intentional* (they are computed each time an attempt is made to traverse them). There are also many different approaches concerning their implementation (eg: embedded links vs link servers).

On the other hand *semantic networks* and *semantic data models* [2, 9] are a means to represent knowledge expressively in a more natural way to human, in comparison to other data models like the relational and the hierarchical. They do not encode the knowledge as a raw kind of data, but they structure them using meaningful relations (semantically interpretable) which allow further logical deductions. Semantic networks are suitable to represent collections of structured heterogeneous and highly interrelated information objects, therefore they are suitable for applications like scientific catalogs and engineering applications (CAD/CASE/CAM). Although semantic networks have been studied for over twenty-five years, they are not yet widely adopted since some crucial related issues are still open (eg: implementation efficiency, query languages).

2 Hypermedia Structures and Semantic Networks

We are studying the problem of generating hypermedia presentations for data represented in a semantic network. This is a quite reasonable approach since both hypermedia structures and semantic networks are knowledge representation paradigms and they are in their core graphs. Moreover, semantic networks offer a well-established, well-formalized and rich means to represent knowledge. On the other hand, the hypermedia models that have been proposed lack of semantic meaning preventing their usage in applications where advanced retrieval methods are required (eg: semantic queries). Furthermore, data semantics are very crucial because they could be exploited by the hypermedia applications in order to support content-based hyperlinking [14], incremental knowledge acquisition [11], and to provide different kinds of presentations of the same data for different users, thus, reusing the existing data while respecting the human personality and the cultural diversity.

There are more than one ways to employ a semantic network in hypermedia. According to Rada [13], the semantic net can be used in a hypertext structure in an *independent* way where the semantic net is independent of the text blocks, thus can be seen/traversed separately (hypertext nodes and links are related to nodes of the semantic network). This is a two-level model of hypertext [10]. Alternatively, the semantic net can be used in an *embedded way* where text block are the destination of semantic network links (in this approach the semantic network cannot be traversed separately).

3 Problem Statement

We focus on the development of hypermedia presentations of data already represented in a semantic network which has been constructed in order to serve the purposes of advanced and general purpose applications. That is, the modeling is being determined by the domain of discourse and not by the requirements of one specialized application. In case of unstructured data (plain text [1], audio, video, image) the semantic net includes their semantic description (if any), plus logical pointers to their physical location.

To generate hypermedia presentation for these data we need: (a) to define an *abstraction view* of the semantic network which determine the part of the net that is going to be presented in the hypermedia application, plus the way that hyperlinks are generated by the links of the net, (b) to *enrich* the objects and links of the semantic net with names suitable for different presentations (eg: multilingual), plus with data specifying the presentation layout. We will refer to the above metadata (which define the abstraction/enrichment of the semantic net) with the term *presentation specification data* (hereafter PSD).

3.1 Motivation

Developing hypermedia with this approach is advantageous because the same (rich and expressive) knowledge base can be used in order to produce different presentations facing different needs which may be exploited by different more simple or more focused applications. Thus, this method promises data extensibility and reusability and development efficiency [6]. In addition, mappings to different hypermedia data models are feasible.

One possible application is the development of different on-line hypermedia presentations which are addressed to different kinds of visitors of a museum [2], exploiting a knowledge base of cultural goods which serves the complex needs of museum curators and scientists.

An other ambitious and irritating application of this approach could be the classification of the universal data (at first those resided in www-servers) in a complex detailed semantic network from which automatically constructible and customizable hypermedia views according to the user's profile and interests would be feasible (obviously this imposes the confronting of many design and technical problems).

3.2 Analysis

The mapping of a semantic net to a hypermedia structures should support: multiple presentations/mappings, PSD integrity, consistency between the domain data and the PSD, separation of the domain data from the PSD, simplicity and flexibility, extensibility and reusability, and utilization efficiency. We can distinguish the possible mapping approaches according to three general aspects: *definition*, *declaration* and *representation*.
Definition. The mapping may be defined in a per object basis (fine grain) or in a per set basis (gross grain). The former permits flexible declaration of presentations, while the latter offers uniformity and result in declaration brevity.
Declaration. The declaration of the mapping may be determinded either by the designer, or it may be fixed. The former offers control and greater tailorability, hence, presentation quality, but requires more effort by the designer. The latter approach may take into account path shapes [11], or in order to provide tailorability it may take into account weights [8] attached to links of the semantic net.
Representation. The PSD could be represented either inside or outside the semantic net. The "inside" approach promises consistency and utilization efficiency but requires rich structuring mechanisms from the data model in order to keep the information model simple and clear.

4 Our Approach

4.1 Representation Framework

Our information representation framework is the Semantic Index System (hereafter SIS) [5, 4]. SIS uses SIS-Telos as the information representation framework. Telos [12] is an object-oriented knowledge representation language that supports a number of structuring mechanisms as well as an assertional and temporal reasoning sublanguage. In SIS we confine ourselves to a version of the structural part of the Telos language, the SIS-Telos.

Objects in SIS-Telos are named and organized along three dimensions: attribution, classification and generalization [1, 2, 9, 12]. A distinctive feature of Telos , and consequently of the SIS data model, is the uniform treatment of individuals and attributes. This allows attributes to be organized in classification and

[1] It could be either an autonomous document or a document part.
[2] eg: children, ordinary visitors, interested users.

generalization hierarchies and to have attributes of their own, which provides great expressive power and flexibility.

Multiple classification is allowed, supporting the separate representation of multiple modeling aspects. An open-ended classification hierarchy is possible. Classes within a given instantiation level are also organized in terms of generalization (or isA) relationships. These can be multiple and give rise to hierarchies that are directed acyclic graphs. They induce strict inheritance of attributes, in the sense that inherited attributes cannot be overridden but only restricted by the definition of the subclass.

The SIS has been developed by the Information System and Software Technology Group of the Institute of Computer Science (FORTH [3]).

4.2 The DOMENICUS Prototype System

DOMENICUS is a prototype hypermedia presentation engine which maps at run-time a cultural information base, which is based on the CLIO [3] model (represented in SIS), to a set of core hypermedia functionalities appearing almost in every application domain including subject catalogs, alphabetic lists, complex queries, guided tours and hypermedia/multimedia cards. In order to offer user-familiar card contents and not impersonal record structures, a limited form of natural language generation is supported. The current application of DOMENICUS is a prototype electronic presentation of the painting exhibition "From El-Greco to Cezzane" held in the National Gallery of Athens in 1992.

The mapping of the knowledge base to hypermedia structures was first hardwired in the code of DOMENICUS, but now we are currently working on developing a *Presentation Model* (to be stored in the knowledge base itself), which will be interpreted by the kernel of DOMENICUS in order to specify that mapping.

DOMENICUS showed that our approach is feasible and beneficial since it offers development efficiency, presentation quality, extensibility and easy, flexible adaptation.

5 Open Issues and Future Work

Open issues include the enhancement of the Presentation model in terms of expressiveness and effectiveness. We are also working on some other features of the SIS which could be exploited by the hypermedia applications: context-based updating [7] and naming [15] and integration with external tools.

In the framework of the AQUARELLE project [4] SIS (and CLIO) is going to be used for the storage and management of multimedia folders and will be integrated with SGML editors in order to create, store and make accessible documents (folders) with referential integrity to formal knowledge entities and other document parts.

References

[1] G. Attardi and M. Simi. "Completeness and Consistency of OMEGA, A Logic for Knowledge Representation". In *Proceedings of International Joint Conference on Artificial Intelligence*, Vancouver, 1991.

[2] M. Brodie, J. Mylopoulos, and J. Schmidt, editors. *"On Conceptual Modeling: Perspectives from Artificial Intelligence, Databases and Programming Languages"*. Springer-Verlag, 1984.

[3] Panos Constantopoulos. "Cultural Documentation: The CLIO System". Technical Report 115, Institute of Computer Science Foundation for Research and Technology Hellas, January 1994.

[4] Panos Constantopoulos and Martin Doerr. "Component Classification in the Software Information Base". in O.Nierstrasz and D.Tsichritzis, eds.,Object-Oriented Software Composition, Prentice-Hall,1995.

[5] Panos Constantopoulos and Martin Doerr. "The Semantic Index System : A brief presentation". Institute of Computer Science Foundation for Research and Technology Hellas, May 1994. (http://www.ics.forth.gr/proj/isst/Systems/sis/).

[6] Panos Constantopoulos, Manos Theodorakis, and Yannis Tzitzikas. "Developing Hypermedia Over an Information Repository ". to appear in the 2nd WorkShop on Open Hypermedia Systems at Hypertext'96, Washington, DC, USA,March 16-20, 1996.

[7] Panos Constantopoulos and Yannis Tzitzikas. "Context-Driven Information Base Update". to appear in CAiSE*96, Heraklion, Crete, Greece (20-24 May 1996).

[3] FOundation for Research and Technology - Hellas.

[4] AQUARELLE is a project of the EC Telematics Programme aiming to the sharing of cultural heritage.

[8] Michael J. Hu and Peter Kirstein. “An Intellignet Hypertext System”. In James Mayfield, editor, *Proceedings of the Workshop on Intelligent Hypertext*, 1993.

[9] Richard Hull and Roger King. “Semantic Database Modeling”. *ACM Computing Surveys*, 19(3):202–260, September 1987.

[10] James Mayfield. “Two-Level Models of Hypertext”. in R. D. Sammel, eds.,Advances in Software Engineering and Knowledge Engineering, World Publishing, 1995.

[11] James Mayfield and Charles Nicholas. “SNITCH: Augmenting Hypertext Documents with a Semantic Net”. *International Journal of Intelligent and Cooperative Information Systems*, 2(3):335–351, 1993.

[12] John Mylopoulos, Alex Borgida, Matthias Jarke, and Manolis Koubarakis. “Telos : Representing Knowledge about Information Systems”. *ACM Transactions on Information Systems*, 8(4), October 1990.

[13] Roy Rada. “Hypertext Writing and Document Reuse: The Role of a Semantic Net”. *Electronic Publishing-Origination, Dissemination and Design*, 3(3):125–140, August 1990.

[14] Li-Cheng Tai. “Architecture Support for Content-based Open Hypermedia”. to appear in the 2nd WorkShop on Open Hypermedia Systems at Hypertext’96,Washington, DC, USA,March 1996.

[15] Manos Theodorakis. “Context-based Naming in Semantic Networks”. to appear in the 3rd Doctoral Consortium on Advanced Information Systems Engineering, Heraklion, Crete, Greece, May 20-21, 1996.