Integrating Knowledge Representation into a Multimedia Information System

Günther Görz
University of Erlangen-Nuremberg
FORWISS and IMMD VIII (Artificial Intelligence)
Am Weichselgarten 9, D-90425 Erlangen
Phone: (+ 49 9131) 85-9909; Fax: 85-9905
E-Mail: goerz@informatik.uni-erlangen.de

Abstract

The goal of many hypermedia systems, which are a synthesis of hypertext and multimedia systems, is to open complex information spaces through the use of associative networks with nodes containing texts, images, and acoustic signals. The central issue of this paper is to argue that the power of such information systems can be considerably enhanced by the integration of a knowledge representation component. We illustrate this thesis with an example hypermedia information system about the oldest existing globe of the earth whose creation was initiated by Martin Behaim in 1492. Displaying the pre-columbian image of the earth, the Behaim globe is a masterpiece of art of the European Renaissance. Today it is one of the most attractive objects in the collections of the Germanisches Nationalmuseum in Nuremberg. Our hypermedia information system consists of two parts, a graphical visualization component which can produce images of the globe from any desired view, and a hypermedia component which provides information about the globe, in particular different readings of the texts written on it, various comments on the texts and miniatures on the globe, images of contemporary globes and maps, a complete record of scientific publications about the object, and cultural and historical background information. Particular attentention is paid to the broad variety of relations between images and texts.

1 Introduction

The central thesis of this contribution is that

Hypermedia systems without knowledge representation are superficial.

The goal of many hypermedia systems, which are a synthesis of hypertext and multimedia systems, is to open complex information spaces through the use of associative networks with nodes containing texts, images, and acoustic signals. In those systems, usually little or no use is made of Artificial Intelligence (AI) techniques, in particular of knowledge representation. We argue that the power of hypermedia information systems can be considerably enhanced by the integration of a knowledge representation component. The representation formalisms available in most hypertext systems — if one can speak of such at all in a strict sense — are not expressively and inferentially powerful enough to provide a deep model of the

object domain. So users can access information only along paths which have been either explicitly or implicitly given by the system authors, a restriction which can be overcome by integrating AI knowledge representation and processing techniques.

We illustrate this thesis with an example hypermedia information system about the oldest existing globe of the earth whose creation was initiated by Martin Behaim in 1492. Displaying the pre-columbian image of the earth, the Behaim globe is a masterpiece of art of the European Renaissance. Today it is one of the most attractive objects in the collections of the Germanisches Nationalmuseum in Nuremberg.

To celebrate the 500th anniversary of the Behaim globe, research activities started in 1987 in order to prepare a big museum exhibition and the publication of a scientific monograph. First results were presented to the public with the two-volume exhibition catalog in 1992 [Willers 1992]. New insights about the production of the globe were gained through the application of innovative scientific methods like microchemical analysis, endoscopy, and computer tomography. Furthermore, a complete photographical record of the globe has been taken. Based on normalized segment images which were derived from these photographs, a simple computer animation of the globe has been developed for the exhibition.

The rich image and text material resulting from these research activities was the starting point for a new project between FORWISS, the Universities of Erlangen-Nuremberg and of Frankfurt/Oder, and the Germanisches Nationalmuseum, Nuremberg, aiming at the construction of a comprehensive hypermedia system to provide in-depth information on the Behaim globe.

2 Martin Behaim's Globe

In 1491, the merchant Martin Behaim, son of an ancient noble Nuremberg family, who lived in Portugal since 1484, visited the city of his fathers to settle a will case. During the two following years, he was one of the driving forces in the creation of a globe of the earth, which today is the earliest existing globe at all. The American continent is not shown on the globe, because Columbus' travels were not yet known at the time of its making. Behaim was born in 1459 and received his professional training as a textile merchant in the Netherlands. The reasons which led him to Portugal are unknown, but probably long distance trade with spices played an important role. Besides his business, he was also engaged in cosmography and navigation. In Portugal, he became a consultant of king João II. and participated in a voyage along the African coast in 1484. After that, he was knighted and married the daughter of an influental Portugese family.

The reasons for the creation of the globe are unknown, but there is little doubt that economic considerations played an important role, as can be seen from a lot of inscriptions about merchandise etc. on the globe. The globe was made by a group of craftsmen which is clearly documented by the final invoice to be paid by the city council of Nuremberg. This account gives clear impression of its complex production procedure. A ball of clay with a diameter of 51 cm was covered with laminated stripes of linen. Then the linen ball was cut in two pieces, the ball of clay was taken out and the two halves were put together again with a wooden ring. Then it was covered with parchment, and another layer of paper — eight segments and two polar caps — was glued on that. The map was drawn by a painter according to a printed map supplied by Behaim — this is the only mention of his in the final account. Possibly he had the plan for a series production of globes made

of printed segments so that the painted globe were a kind of prototype. Although Behaim was probably the main supplier of information, other Nuremberg scholars contributed to the final appearance of the globe. The series production was never realized, Behaim left Nuremberg in 1493, and the mentioned printed map has been lost.

The map image of the Behaim globe follows primarily the Ptolemaic tradition, but there are also influences from the medieval circular so called T-O maps of the world and from portulans which were used for navigation. Those contemporary maps which resemble the Behaim globe most directly are the maps of a German mapmaker working in Italy, Henricus Martellus Germanus. The Behaim globe is covered with a large number of small miniatures and carries a rich variety of texts, many of which deal with pecularities of foreign countries, their inhabitants, plants and animals, in particular with merchandise, with famous travels like that of Marco Polo, but also with religious issues, giving a rather comprehensive image of the world in Renaissance Europe. So the Behaim globe is one of the most important and valuable objects of art and cultural history from the beginning of modern times.

3 The Hypermedia Information System

The **"behaim 1492"** project was initiated with a feasibility study for hypermedia information system in 1993 and is now being continued as a basic research activity at our center. Intended users of this system are historians, cartographers, map collectors, and the interested public. The prototype system was conceived as an experimental and open development environment in the manner of an electronic encyclopedia, where images and texts are being organized as a hypermedia network, and presents to its users comprehensive graphical and textual information about the globe and its cultural and historical background.

Our hypermedia information system consists of two parts, a graphical visualisation component which can produce high resolution images of the globe from any desired view, and a hypertext component which provides in-depth information about the globe, in particular different readings of the texts written on it, various explanations and comments on the texts and miniatures on the globe, images of contemporary globes and maps, a complete record of the scientific publications about the object, and cultural and historical background information. Particular attention is paid to the variety of relations between images and texts. In order to enable complex queries the hypermedia structure is being linked to a conceptual model of the globe, implemented by means of a description logic based knowledge representation system. The graphics component runs on a *Silicon Graphics IRIS "Indigo"* and is connected to the hypertext component on a *Sun SPARC* workstation.

Our current work deals with the extension of the conceptual model which describes the globe phenomenologically and topologically, with porting our data base to an object-oriented data base system, and with the implementation of tools to generate a simplified version of the system for the World Wide Web.

4 The Graphics Component

The graphics component of the system displays the globe on the screen of a *IRIS "Indigo"* workstation under *OpenGL* [Schüller 1993]. The image data were generated from photographs covering the globe, preprocessed, and converted into

a format that can be processed by the workstation.

The original images of the globe are 34 photographs shot in polarized light with a *Linhof Technica* camera with a format of 6 by 9 cm. 32 of the pictures, showing the different segments of the globe, were taken with a constant offset of 30 degrees, providing a certain overlap. The remaining two pictures show the polar caps.

After 500 years the globe differs from the shape of a sphere at many points. For further image processing planar images of the globe are a prerequiste. Therefore, first a contour map of the globe was produced through the application of a triangulation procedure. On the basis of this map, an analog computer driven photogrammetric differential mapping process was used to produce orthophotographs of 24 segments of the globe and two polar caps. Planarity of the segments is achieved by projecting every segment on an elliptic cylinder. This normalization procedure, which is normally applied to process satellite images of the earth, was carried out at the Technical University of Vienna [Kraus 1992].

The resulting 26 orthophotographs of segments covering the globe were scanned with high resolution and 24 bit color depth corresponding to professional PhotoCD quality. These images had to be corrected for a proper display of the globe in several steps such as cutting, size correction and distortion. For performance reasons, the size of the image data was reduced from sixty down to four megabytes per segment. The reduction of resolution was necessary in order to be able to handle multiple pictures on the workstation. Otherwise, even an *IRIS "Indigo"* could process only one of the original images at the same time.

The images are being mapped on a sphere using the *texture mapping* function of OpenGL and displayed on the screen of the workstation. The view is oriented towards the middle of the globe, but currently no special illumination effects are taken into account. The user can choose an arbitrary view on the globe by specifying two angles and the distance to the globe surface. Furthermore, an angle can be chosen that rotates the viewing plane around the vector describing the viewing direction. Fig. 1 shows example views of the globe (Europe with the Atlantic Ocean and South Africa, respectively) as they appear on the graphics screen.

The graphics component exhanges data with the hypertext system and vice versa by means of Unix process communication facilities. At the present state, coordinate data are transferred between the two system components in both directions.

5 Design and Implementation of the Information System

Our data base consists of a complete high resolution scan of the globe's surface as well as of a complete record of the huge amount of text written on the globe in seven different historical readings. Furthermore, we collected scans of contemporary maps and a text corpus with commentaries, including all relevant scientific publications on the globe. To a large extent, these data have already been incorporated into a hypermedia system on the basis of *NoteCards*. NoteCards has been chosen because it is still one of the most powerful hypermedia systems, and for its extensibility through the available flexible programming interface, which is a salient feature for explorative and experimental system development. For the design of the presentation interface, the guidelines developed in [Hodges 1992] were our most inspiring source. Generality and adaptability to similar applications has been an important goal for the implementation of all software tools.

5.1 The Text Data Base

The Behaim globe contains a large number of inscriptions which we divided into three classes: inscriptions which are only one word long like names of cities, regions, etc., legends of the length of a phrase or a sentence, and long texts which comprise more than one sentence. Of these texts seven historical readings from the 16th to the 20th century are available which exhibit considerable differences. Some of these records and the secondary literature provide numerous comments on the texts and also on the miniatures painted on the globe, dealing with their readability and state of conservation, with the (five) different handwritings which can be recognized, and with the sources like Ptolemy, Marco Polo, Mandeville, etc. For the text versions and the comments directly referring to the texts it seems obvious to provide a regional grouping as the primary ordering criterion and to embed other references, e.g. to the sources, as typed links in the hypertext network. This led us to a layered structure for the text data base with the following levels: Readings in seven versions with the modern (1992) reading as the primary reference, commentaries, publications which refer directly to the globe or its parts, and publications which deal with the general cultural and historical background. All of these texts are now available in machine readable form, most of them for the first time. At present, the data base is being transferred from the actual file tree structure to the object-oriented data base system ObjectStore.

5.2 The Image Data Base

In addition to the preprocessed segment images in the visualization component, the 26 normalized high resolution segment images are stored in a separate data base which can be accessed directly by the hypermedia component. These images are given as uncompressed TIFF files to provide an unconstrained possibility of zooming down to the resolution of the film.

5.3 The Hypertext Component

Currently, the information system contains in the form of a hypertext

- an introduction into the object and its originator which comprises information about the technology of the production, the map image, the history and the reception of the globe;
- the texts written on the globe and corresponding explanations and comments;
- access to high resolution images of the globe's surface and of other contemporary maps, drawings, and printed images;
- a "library" containing all relevant scientific publications on the globe;
- "Guided Tours" through the information space, in particular by employing a conceptual model of the globe.

A typical screen is shown in fig. 2: In the top part of the screen there are two menus of frequently used operations. The selection of "Hypertext Start" pops up a table of contents. After clicking the "Guided Tours" button in it, the graph shown in the lower half of the screen is displayed which shows the class (i.e. concept) hierarchy together with selected instances, of which the island of Iceland has been

chosen. A node displayed as a window with a 19th century reproduction of the globe section containing Iceland and a link to another node containing a general description of Iceland's representation on the globe comes up. This node contains links to the reference readings of two texts about habits of the Icelanders written on the globe as well as to other readings, which are displayed on demand together with a translation into English and a commentary. At the same time, the graphics component generates a view on the globe centered around Iceland. Furthermore, one can either display the corresponding segment of the globe — through menu selection in the hypertext component — with the possibility to zoom into this image or to select an enlarged image of Iceland directly from the miniatures menu. A variety of text processing operations like full text search is provided within the NoteCards environment. To support navigation in the hypertext space and to prevent too many open windows on the screen, already visited nodes are shrunk to icons which are put on a stack on the left hand side of the screen. These nodes can be recovered by a simple mouse click. Clicking on the "Globus" button on the right menu opens a submenu with commands to turn the globe explicitly. The selection of one of these commands results in sending a message to the graphics component which generates a corresponding new view on the globe.

The system provides a variety of graphical operations which can be applied to all images, in particular to the globe segments, like zooming, color and contrast variation, and filtering. The application of image processing techniques makes it possible to exhibit details which can only partially or not at all be recognized on the original object. The system allows also to display MPEG-encoded animated images; e.g. a synthesized movie showing the rotating globe and an image sequence taken when the globe was moved through a computer tomograph have been included. The latter sequence of cuts exhibits clearly the different layers of the globe providing new insights about the techniques used for its production. Furthermore, a cartographical toolbox has been implemented which allows to produce different projections of a map from a standardized vector or bitmap representation. In the future, these tools will be augmented in order to display various superpositions of images like different maps, schemata of map types, travel routes, e.g. of Marco Polo or Columbus, etc.

The use of a hypermedia network as the basis for an information system provides a convenient way to represent relations between text and images (and other media like acoustic signals as well). First of all, texts and images can both be content elements of hypermedia nodes — as it is the case with the Iceland example —, whereas their relations to each other can be expressed by typed links. Because we use further visualization tools separate from the NoteCards environment, like the graphics component described above, display and image processing tools from *xv*'s library, and an MPEG-player, our system provides special nodes which spawn a corresponding process being executed in parallel.

5.4 The Knowledge Representation Component

Even for the task of structuring the text and image data base whose elements contain a lot of mutual references to each other, it is useful to provide a systematic phenomenological model of the domain in the first place. In order to provide a general solution we decided to combine AI knowledge representation techniques with the hypermedia system. In this way, the power of a hypermedia system can be augmented considerably by adding a possibility to process complex queries which reach far beyond the associations stored in its links.

Therefore we specified a knowledge base on the basis of a subsumption (superconcept-subconcept) and an aggregation (part-of) hierarchy which provides a conceptual model of the globe. First of all, this model is a phenomenological one because it deals primarily with the visual appearance of the globe (cf. [Pfannenstein 1994]).

As a first step, both hierarchies were implemented directly in the hypermedia system as subnetworks consisting of simple hypertext nodes and typed links. So they can be easily accessed following appropriate links from other parts of the network; a rather convenient way is through a "Guided Tour" as illustrated in fig. 2 for the subsumption hierarchy. The nodes of the subsumption hierarchy represent concepts, the edges an inheritance relation, and the terminal nodes point to corresponding instance objects, given as text or image units. The same holds in analogy for the aggregation hierarchy. Following the edges simulates an inheritance mechanism in its most simple form, but inferences, e.g. those which take transitivity into consideration, cannot be executed directly in an automatic way. Inheritance of features — attributes or relations which are associated with concepts — is far beyond any possibility which can be achieved by means of navigation and network extension.

This experience indicated clearly the need for a more powerful knowledge representation framework. We came to the conclusion that a description logic, which is probably the best understood family of AI knowledge representation formalisms, would satisfy our needs. We chose the description logic systems YAK [Franconi 1993] as the modelling tool, a knowledge representation language of the KL-ONE family, and modelled the globe domain according to the method proposed in [Brachman 1991]. This knowledge engineering method consists in an enumeration of object types, a distinction of concepts and roles (attributes), the development of a concept hierarchy, the determination of individual objects and composite objects and their parts, and the specification of type and number restrictions for role values. So, for example an object type would be "geographical object", a concept would be "island", a role "geographical name" with the restriction that it has to be a string, and an individual object (instance) is "Iceland". The subsumption hierarchy as implemented in YAK is shown in fig. 3 which was generated by the knowledge base browser. This browser is interactive, i.e. the nodes of the network are active elements which allow to generate various graphical representations of the knowledge base elements — concepts and instances — via selection of commands from submenus which are presented on demand.

Knowledge bases of description logic systems consist of two parts: The conceptual or terminological part ("T-Box") consisting of the subsumption and aggregation hierarchies, and the assertional part ("A-Box") containing the instances and the relations ("axioms") which hold among them. The biggest effort which had to be spent in the construction of the knowledge base was "programming" the descriptions of hundreds of instances, although we concentrated on only one globe segment showing Europe and North Africa. This segment is representative in the sense that it contains instances of all object types which can be found on the globe. All attribute values had to be recorded manually with the exception of coordinate data which where gathered automatically through the graphics component by pointing to the corresponding pictorial instances on the segment image. We expect that the extension of the A-Box to the remaining segments will become easier by applying a cut-and-paste process to already existing instances.

A decisive step in the knowledge engineering process was the development of a catalog of typical questions users might address to an information system of this kind. Basically, one can distinguish between two classes of questions: Whereas

simple questions, typically regarding single objects, can be directly answered by navigating along links in the hypermedia network, complex questions, e.g. for sets of objects or properties, require more complex inferential capabilities. Characteristic examples for simple questions are

- on images: *What is a (displayed) geographical object? Where is a certain non-geographical object?*
- on texts: *What is a certain reading of an inscription? What is the source for a certain inscription?*

Examples for complex questions are questions for

- properties of objects: *What are the colors of the tents shown in Africa?*
- objects and sets of objects: *What does the Portugese flag look like? Where are the Nuremberg coats of arms? To which class belongs a certain object? Which are the members of the Azore islands?*
- object classes: *Which coats of arms are there? How many flags are there? What are the signs on the zodiac? Which animals are shown in Africa? In which region can a certain object be found?*
- composite properties of objects and object classes: *Where are named town silhouettes?*
- explanations: *What is the zodiac? Why is a certain object at a certain place?*

In these questions, the word "certain" is a placeholder for the designation of a particular instance, e.g. "elephant" would be an actual instance of "certain object".

The two classes of questions define two different types of access to information, *access by navigation* and *access by descriptive conceptual queries*, which require different processing methods. In order to process queries of the second type, an explicit (logical) domain model is required, upon which conclusions can be drawn by an inference mechanism. The implementation of a conceptual model with a system like YAK provides just these services: Besides typical inferences like complete inheritance and consistency tests complex queries of the types shown above can be posed in a logical interface language and are executed by a specialized prover. As yet, the results of inferences drawn are given as expressions of the logical interface language.

In order to *integrate* the conceptual model with the hypermedia network, we follow an approach which associates generators for corresponding hypertext nodes with special concept roles. So, the hypertext subnetworks described above representing the subsumption and aggregation hierarchies without any inferential capability need not be generated manually in parallel to their representation in YAK — which would cause a maintenance problem. Instead, they can be generated automatically from the YAK knowledge base. In this way, both parts, the hypermedia network and the conceptual model, are kept separated, but compatible with a clearly defined interface, hence preserving their particular capabilites. The automatically generated representations of the hierarchies as subnetworks in the hypermedia system show clearly the advantage of the synthesis of both components.

The next modelling step will be to augment the phenomenologically oriented model with a model of geometrical and topographical relations in the same formalism. We do not expect problems in the extension of the subsumption and aggregation hierarchies with a representation of topological concepts. More difficult is the

question of geometric inferences. For the sake of coherence — but possibly at a performance price — we are working on a propositional representation of topological relations. Whether we would need to include special inference mechanisms as it is the case in Geographical Information Systems, is still an open question.

Last, but not least, to facilitate complex queries a menu based query interface to the representation system has still to be created. With respect to the output of queries, we are planning to provide a more easily readable representation than expressions of the query language and an interface to the graphics component in order to highlight respective objects on images of the globe (e.g. outlining all Portugese flags as a result to a question like *"Where are "Portugese flags?"*).

Acknowledgements. The author owes special thanks to Jürgen Pfannenstein, Thomas Schüller, Hans-Peter Seidel, and Christoph Taegert-Kilger for contributions to the implementation and critical remarks.

References

[Brachman 1991] Brachman, R.J. et al.: *Living with CLASSIC: When and How to Use an KL-ONE-Like Language.* In: Sowa, J, (Ed.): *Principles of Semantic Networks: Explorations in the Representation of Knowledge.* San Mateo: Morgan Kaufmann, 1991, 401–456

[Franconi 1993] Franconi, E.: *The YAK Manual.* IRST Technical Report, Trento, 1993

[Goerz 1994] Görz, Günther et al.: *Machbarkeitsstudie für ein Multimedia-Informationssystem über den Behaim Globus (1492).* Universität Erlangen-Nürnberg, FORWISS, Report FR-1994-002, Erlangen, Feb. 1994

[Hodges 1992] Hodges, M.E., Sasnett, R.M. (Eds.): *Multimedia Computing — Case Studies from MIT Project Athena.* Reading, Mass.: Addison-Wesley, 1992

[Kraus 1992] Kraus, K.: *Photogrammetrie und Fernerkundung angewandt auf den Behaim-Globus.* In: [Willers 1992], 301–308, 1992

[Pfannenstein 1994] Pfannenstein, J.: *Wissensmodellierung am Beispiel des Behaim-Globus.* Universität Erlangen-Nürnberg, IMMD VIII, Studienarbeit, Erlangen, Juli 1994

[Schüller 1993] Schüller, Th.: *Visualisierung eines digitalen rotierenden Behaim-Globus.* Universität Erlangen-Nürnberg, IMMD IX, Studienarbeit, Erlangen, September 1993

[Willers 1992] Willers, J. (Ed.): *Focus Behaim-Globus. Ausstellungskatalog.* 2 vols., Germanisches Nationalmuseum, Nürnberg, 1992

Figure 1: Synthesized views of the globe (Europe/Atlantic, South Africa)

Figure 2: System screen with guided tour and information on Iceland

Figure 3: The subsumption hierarchy implemented in YAK