The Knowledge Engineering Review, Vol. 12:2, 1997, 219–223

Knowledge representation for interactive multimedia systems

G. A. VOUROS

Department of Mathematics, University of the Aegean, Karlovasi, Samos, Greece (email: georgev@aegean.gr)

1 Introduction

Intelligent multimedia systems (IMMS) assist users in choosing the information content, and to design the information presentation, that satisfies their needs and idiosyncrasies. Towards this target, IMMSs posses and utilize knowledge concerning the *means* and *function* (*purpose*) of a presentation. 'Means' denote communication techniques (information carriers and modalities), as well as information content and information characteristics. The 'function of a presentation' denotes *the communicative purpose of the presentation, methods for achieving that purpose, as well as a user's goals and tasks*.

Specifically, IMMSs must be able to:

- Assist users to navigate the information world, locate the appropriate information sources, retrieve, and assimilate and organize information efficiently and effectively. If the information world spreads around the globe and comprises huge amounts of incomplete, uncertain, heterogeneous and unstructured data, this task becomes more complicated.
- Communicate information in terms that users can readily 'consume'. To design a presentation, IMMS must reason about different communication means (e.g. how to choose between, and how to coordinate, different media and modalities), as well as about the function of the presentation.
- Adapt to changes' happening in the users' and the information world rapidly.

Interactivity plays a crucial role in addressing these requirements. Systems must enable users to communicate their goals, tasks, experience, abilities and preferences effectively. Users should be able to guide the system to select the appropriate information content, as well as the appropriate information presentation strategy and design, using techniques that seem most natural to them.

2 Interactivity in intelligent multimedia systems

Particular issues concerning interactivity in intelligent multimedia systems are the following (Roth & Hefley, 1993; Maybury, 1993; Grosz, 1994).

- *Specifying goals or tasks*. Users must use specific vocabulary and specific interface constructs for communicating their goals to the system. Users must communicate using their language, or any other way that is more natural and effective to them.
- *Selecting techniques and co-ordination methods for the interactive design of information presentation*. Users must be able to make design choices concerning either media selection or coordination between different media. Design choices should be made either in existing presentations or before a presentation.
- *Making rendering choices*. Users should be able to control rendering level choices in existing presentations, or before a presentation design.
- *Interactive modification of information*. Users should be able to change the values of database facts, update the existing information, and incorporate new sources of information into the system.
- *Interactive content selection and navigation*. Users must be able to navigate, assimilate and

organize information efficiently. Efficient and effective retrieval, disorientation and cognitive overhead are major issues concerning interactive content selection and navigation.

- *Dealing with contextual information*. This last issue requires the provision of valuable linkages to the user's world and to the current discourse to ensure that valuable and consumable presentations are generated. The users' world encompasses (a) the current information seeking task, (b) familiarity with the system and with the interface constructs it provides, and (c) the working environment.

3 KRIMS96

To be capable of dealing effectively with interactivity requirements, multimedia systems should posses and use knowledge about:

- The domain at hand.
- Tasks of users.
- Tasks that presentations are designed to support.
- Conditions for using and coordinating different media and modalities.
- Content of the stored information.
- Characteristics of the information.
- Context of interaction.

The main themes of the workshop 'Knowledge Representation for Interactive Multimedia Systems: Research and Experience (KRIMS96)' were the exploitation and extension of existing knowledge representation systems, architectures and organization principles for interactive multimedia systems, new paradigms for the representation of multimedia documents, and automatic segmentation.

KRIMS96 was held in conjunction with ECAI96 in Budapest, Hungary. It was organized by W. Klas, T. Sellis, M. Vazirgiannis and myself. Papers came from the United Kingdom, Germany, Austria, Greece, France, Italy, Belgium, the United States of America and Brazil. Sixteen researchers attended, discussed and criticized presentations in a friendly, cooperative and inquiring atmosphere.

This section organizes achievements and thoughts reported in the workshop under the four main themes of KRIMS96, which were also the titles of the workshop sessions.

Exploiting and extending existing technology

Utilization of description logics for the provision of conceptual, language and coding services, as well as the development of tools for effective communication with a system, are major issues discussed in the paper by Bechhofer and Goble (UK). The paper presents a preliminary exploratory work in the use of the GRAIL Description Logic and GALEN terminology server architecture for the classification and selection of art images based on their semantic annotations.

Extending description logics to provide a generic framework for modelling multimedia documents for effective interactive retrieval is the aim of the paper by Meghini and Straccia (Italy). Specifically, the authors present an extension of a description logic with the expression and handling of closure assertions. This is meta-knowledge about the completeness of knowledge concerning specific individuals or primitive concepts.

The paper by Rossi (Argentina) and Mere (Brazil) presents a temporal logic formalism in the context of the object-oriented design methodology, for the representation of dynamic and static aspects of the navigational structure in hypermedia applications. This work aims at building well-defined navigational structures, and mainly addresses issues concerning orientation.

Investigating new architectures and proposing new organizing principles

The role of description logics for building and exploring federated ontologies is discussed by Herzog and Petta (Austria). The authors propose a generic architecture for intelligent multimedia systems dealing with open-world domains. The conceptual foundation of the proposed architecture is the decomposition of the information space into architectonic and semantic spaces. Furthermore, the paper concerns case-based reasoning for supporting interactive content selection and navigation via a guided dialogue between user and system. The authors concur that description logics provide a framework for conceptualization-driven control for a system to satisfy the function of a presentation.

Conceptualization-driven interactive navigation is also investigated by Muller and Korczak (France). Their paper focuses on semantic relations, which are considered to be typed links with known properties. While conceptual navigation exploits the properties of links to access information nodes, contextual navigation facilitates accessing information according to a user's preferences (type, temporal and thematic characteristics of information).

Petrou, Martakos and Georgiadis (Greece) also use links as the primary structuring blocks of their models. The paper enhances links with inference capabilities using fuzzy set theory. Links can be deduced and utilized to built multi-layered semantic abstractions. Collections of links constitute 'views'. The paper shows how the representation can be exploited to support interactive exploration of a large amount of information, supporting a reader's varying degrees of knowledge about the contents and structure of hypermedia.

While the objective of all previous papers is to support human-computer interaction, the paper by Karakapilidis, Gordon, Papadias and Voss (Germany) in addition concerns collaboration among humans. The paper discusses ongoing research towards a generic architecture that provides services for interactive content selection, indexing of multimedia documents and accessing multimedia information in heterogeneous and distributed information systems. Moreover, the architecture aims to provide an issue-based conferencing and group decision support system that utilizes positions, constraints and arguments of different users.

New paradigms

The abovementioned works investigate how to conceptualize domain aspects, the structure of the 'content-bearing' artifacts, and contextual issues regarding a presentation. They also discuss how to exploit these conceptualizations to achieve the function of a presentation.

Minneman and Smoliar (USA) in their papers make the point that "the data structure nature of a document is actually secondary to the more fundamental nature of those processes which are required to the document to be 'produced' and 'consumed'." Therefore, the authors propose to search for knowledgebased tools to support the set of practices and strategies that give rise to a "content-bearing artifact" not to the artifact itself. Therefore, the authors emphasize representing knowledge concerning the function of a multimedia document rather than the document itself.

Also emphasizing the function of a presentation, Calvi (Belgium) argues that information designers must emphasize the temporal metaphor: users must experience information as a series of events unfolding in time. To achieve this in practice, the paper proposes the introduction of interface agents. Agents should suggest the next move within hyperspace and link information in a temporal-narrative way according to their point of view.

Automatic segmentation

Finally, the paper by House, Mani and Maybury (USA) addresses one of the major issues: enabling users to incorporate into a system new sources of information, and update existing information in an efficient and effective way. The paper reports on investigations to automatically segment, label

and summarize broadcast news video. The authors present results from a linguistic analysis of closed caption text, and compare a subject classification approach and one based on discourse cues.

4 Conclusions and future research agenda

Although the research community is well aware of the issues concerning interactivity in intelligent multimedia systems, few research efforts emphasize them, as far as their knowledge representation and reasoning requirements are concerned. This is especially true in cases where users interactively search large amounts of information. Interactive content selection and navigation is one of the major concerns of KRIMS96 papers. Few of these papers focus on this issue for open world domains.

To structure our discussion on specific issues concerning interactivity, we distinguish the five different layers within an intelligent multimedia system, as depicted in Figure 1. The first three layers, which are within KRIMS96 focus, "user profiles", "index layer" and "annotations", contain levels of abstraction concerning the information objects included in the "multimedia repository" and "information sources" layers. They also contain knowledge concerning the function of a presentation.

We must note that Figure 1 does not imply a specific implementation, nor the use of specific representation and reasoning techniques. For instance, it is possible to have a federated implementation using agents for migrating data to different schemes, maintaining ontologies, translating and processing queries across multiple ontologies, and communicating with the user.

The "user profiles" layer contains abstractions of information and supports different views according to users' idiosyncrasies (experience, culture, language, etc.), information seeking goals, tasks and methods for achieving these goals. To support interactivity at this layer, information must be structured in the "appropriate" level of detail. This will enable users to deal with a manageable vocabulary. On the other hand, systems must assist users without constraining their information seeking tasks, and should not fail to provide assistance for unprecedented usage patterns. To achieve these goals, and to support changes happening in the users' world, systems must effectively enable "switching" between idiosyncratic views.

Furthermore, learning about the user is one of the critical issues at this layer. This will enable systems to adapt the communicative purpose of any presentation to the user's task, without having to be told explicitly what to do.

The "index" layer provides conceptualizations of the data available by means of domain-specific

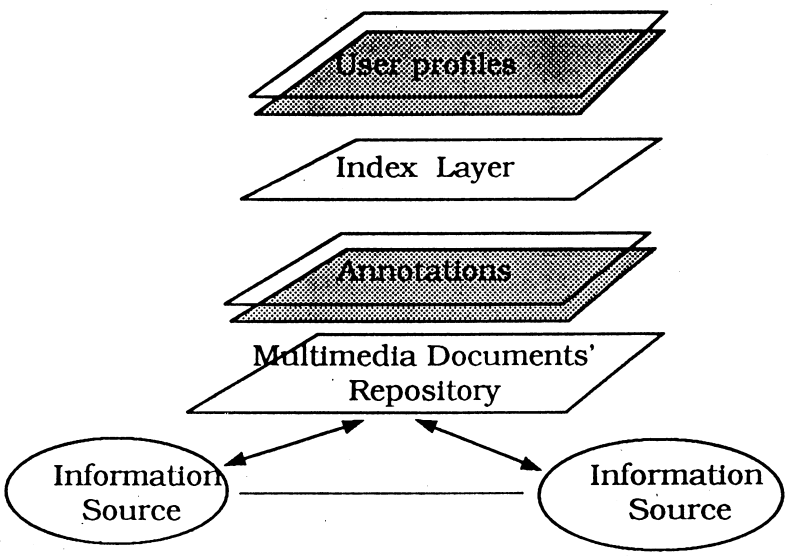

Figure 1 A "layered" framework for discussion

ontologies. Building generic ontologies even for well circumscribed domains as well as, federated ontologies and the appropriate translation services between the internally used representation formalism and a declarative form are critical issues for future research. Furthermore, to support the required functions in the ''user profiles'' layer, systems must handle their content in a flexible way: lack of a once and for all commitment for structuring of the data is critical in intelligent interactive systems.

Moreover, Minneman and Smoliar's paper poses an important question: should we emphasize the strategies and processes for ''producing'' and ''consuming'' multimedia information? Then, how would information in this layer be encoded, and what will be the relation between this and the ''user profiles'' layer?

Concerning the ''annotations'' layer, we must determine the kinds of annotations needed to provide efficient and effective retrieval. Techniques for automatic production of annotations, as well as services to represent multimodal questions, are major research issues concerning this layer.

Moreover, the following question arises: should this layer be combined with the ''multimedia repository'' layer or not? If it has been structured according to some well-defined principles, then it can be utilized directly by a reasoning mechanism and should appear separately. In that case, what is its relation with the ''index'' layer?

Interactivity in intelligent multimedia systems is a challenging and promising area for research in reasoning and representation. While many techniques investigated in KRIMS96 papers, such as description logics, case-based reasoning, reasoning about time and fuzzy logic, as well as other techniques such as nonmonotonic reasoning and planning, are promising in this direction, we need to integrate them into architectures that will be customizable in different domains and applications.

References

Grosz, B, (ed), 1994. ''A report to ARPA on twenty-first century intelligent systems'', *AI Magazine*, **15** (3) 10–20.

Maybury, M, (ed), 1993. *Intelligent Multimedia Interfaces*, AAAI Press/MIT Press.

Roth, SF and Hefley, WE, 1993. ''Intelligent multimedia presentation systems: research and principles'', In: *Intelligent Multimedia Interfaces*, M Maybury (ed), AAAI Press/MIT Press, 13–58.