No Multimedia Without Representation

Bibliographic Reference:

Marc Davis, Daniel Russell, Catherine Baudin, and Smadar Kedar. "No Multimedia Without Representation." In: *Proceedings of Second ACM International Conference on Multimedia in San Francisco, California*, ACM Press, 181-182, 1994.

NO MULTIMEDIA WITHOUT REPRESENTATION

Organizer:
Marc Davis, MIT Media Laboratory and Interval Research

Panelists:
Catherine Baudin, NASA Ames Research Center
Smadar Kedar, Institute for the Learning Sciences
Daniel M. Russell, Apple Computer

Without representation multimedia will not happen. Computers are basically deaf, blind, and ignorant, and will remain so for quite some time. Only by creating representations for media content will we be able to construct large, robust multimedia systems that support content-based manipulation of video, audio, and images. Yet the research community that deals with representation (AI researchers) and the community that deals with media manipulation (multimedia researchers) have, for the most part, not come into contact. The purpose of this panel is to incite debate about the role of representation of multimedia content by bringing these two communities together. Our intention is to confront researchers at ACM Multimedia'94 with a variety of positions about the need for content representation in large-scale multimedia systems. We argue that work in knowledge representation is essential to the construction of robust multimedia systems to manage media content.

Marc Davis

Most contemporary work in multimedia systems suffers from the legacy of traditional approaches to media production and use. Video has largely been used in "single-use" applications (Hollywood movies) as opposed to "multiple reuse" applications (video archives and stock footage houses). Because of this historical bent, multimedia systems have not addressed the needs of building up representations which can enable media content to be reused by people other than those who originally created the content for purposes other than those for which the content was originally produced. The construction of large scale media archives, which can become common resources in a global digital network, demands that we solve problems of representational design for sharing and reusing media content. Unfortunately, approaches to content representation which rely solely on signal-based parsing are inadequate to the task of representing media so that it can be effectively browsed, searched, retrieved, and reused. What is needed is a hybrid approach of human and computational indexing of media content designed with sharable and reusable representation in mind. Without such content representation, multimedia will remain a toy application stuck on CD-ROMs and "titles." In order for multimedia to become a medium for everyday communication we must develop a language for the representation of multimedia content. This language requires that indexers be able to articulate a common substrate of domain independent media representation that can be shared so as to facilitate domain dependent reuse.

Catherine Baudin

The task of accessing the information in large repositories of multimedia documents and of building hypermedia networks relies on being able to describe WHAT is in these documents, as well as HOW this content is expressed in a given medium. Conceptual indexing languages focus on the important features of a domain (i.e., regularities in the set of questions posed by a class of users or key concepts of the domain) and on the RELATIONS between the indexing concepts. Such conceptual indexing schemes are difficult to extract automatically from the analysis of a document and need human intervention in their definition. The advantage however is that an Information Retrieval system can use its knowledge of the relations among indexing concepts to infer relations among information scattered in different places and different media and thus enhance its retrieval performance. "Good" indexing languages are best determined by studying the information seeking behavior of classes of users that are likely to query these documents. However, an information retrieval system should be able to dynamically modify and extend the scope of its representations to accommodate new classes of users and new applications. Information reuse occurs because a system provides facilities to extend the scope of the indexing schemes, rather than by providing a universal representation language that tries to accommodate every applications and all classes of users.

Smadar Kedar

The groundwork for the information superhighway has been laid, with a promise to deliver vast quantities of multimedia systems to America's homes. Yet we do not have the infrastructure to build and deliver large numbers of systems with vast media content. For one, even if we can build prototypes of such multimedia systems in our laboratories, we need the capability to enable those without computer programming knowledge to build such systems, allowing them to author those systems based on content instead of programming. Secondly, once those systems are built, their content will need to be modified and extended as the systems are reused for different applications and users. We need tools that will acquire and refine the way the multimedia information is stored. Thus, we need to strengthen the AI research on authoring tools for multimedia and on knowledge acquisition and refinement of indices of multimedia. Without an understanding of these areas, the promise of the information superhighway will remain just that.

Daniel M Russell

Just how far can we go without representation? It's clear that keyword indexes for multimedia resources have some uses; people use them now. But how much representation do we really need in order to make multimedia databases useful? The "strong AI" approach is to create deep semantic representations

of materials. Once such representations are created, then inference engines can make deductions to answer user queries. Rather than committing to a strongly representationalist position, I believe that a continuum of representational formalisms will prove to be most useful to multimedia users. Semiformal representations acknowledge that our representations of subject matter content are not complete and accurate. Instead of working to create deeply detailed representations of media content, we should be creating indexing systems that can function well with incomplete and partial knowledge.