Using Semantic Web Approach in Augmented Audio Reality System for Museum Visitors

Leila Kalantari, Jordan Willms, Marek Hatala
School of Interactive Arts and Technology
Simon Fraser University
Surrey, BC, Canada V3T 2W1
+1 (604) 268.7431
{lkalanta,jpwillms,mhatala}@sfu.ca

ABSTRACT
In this paper, we describe our work in progress on the reasoning module of *ec(h)o*, an augmented audio-reality interface for museum visitors utilizing spatialized soundscapes and a semantic web approach to information. We used ontologies to describe the semantics of sound objects and represent user model. A rule-based system for selecting sound object uses semantic description of objects, visitor's interaction history and heuristics for continuity of the dialogue between user and the system.

Categories and Subject Descriptors
C.3 [**SPECIAL-PURPOSE AND APPLICATION-BASED SYSTEMS**] *Real-time and embedded systems*

General Terms
Algorithms, Design, Experimentation, Human Factors

Keywords
Augmented-audio reality, ontologies, inference rules, user model

1. INTRODUCTION
ec(h)o is an augmented audio-reality interface utilizing spatialized soundscapes and a semantic web approach to information. The initial prototype is designed for a natural history and science museum. The platform is designed to create a unique museum experience that consists of a physical installation and an interactive virtual layer of three-dimensional soundscapes that are physically mapped to the museum displays. The source for the audio data is streaming digital sound objects. The digital sound objects originate in a network of object repositories that connect digital content from one museum with other museums' collections. The focus of this paper is the reasoning module of ec(h)o that is responsible for maintaining the user model, narration retrieval, and selecting background sound for narrations. The retrieval is based on continually updated user model and semantic descriptions of exhibition and sound objects. This paper presents basic components in the system and reports on our experience with implementation and early evaluation of the system.

2. VISITOR'S INTEREST
Sound objects including narrations are described using properties in ontologies that map objects to concepts, themes, topics, and environment. We created ontology for each of these types. A sound object can be also associated with Place, Event, and Time, which are already modeled in Conceptual Reference Model (CRM), a candidate standard ontology for describing cultural heritage information developed by the International Council of Museums (CIDOC).

Concepts are used to describe both information narrations and exhibits and artifacts. As concept ontology is specific for a particular museum collection we map concepts in each museum to Dewey Classification [1] to facilitate retrieval of objects from remote repositories.

We categorize museum visitors into 3 types: *greedy, selective,* and *busy.* A greedy visitor wants to know as much as possible, does not rush, and moves through the exhibit systematically. A selective visitor is only interested in certain concepts and exhibits. A busy visitor does not want to spend much time in front of an artifact, and only wants to get a general idea of the museum.

In addition to the type, the user model represents visitor interests in the form of concepts. Concepts the visitor is interested in are inferred based on the way the visitor interacts with the exhibition, i.e. which narrations visitor selects and how s/he moves through the exhibits. The following heuristics are implemented as rules:

1. When a greedy visitor enters an exhibit and moves slowly his/her interests will be matched to a primary concept of any narration that describes an artifact in that exhibit. For the greedy visitor we do not need to be picky about interests and we can assume that he is interested in almost any concept.

2. Interests of a selective visitor get overwritten only after s/he repeatedly chooses narrations with certain concepts.

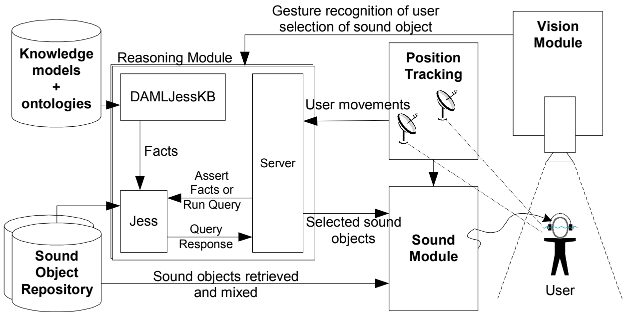
3. For each exhibit, we calculate a primary concept of an exhibit as the most frequent concept of narrations associated with a particular exhibit. The interests of a busy visitor can only be overwritten with those concepts at the time s/he enters an exhibit.

4. For any visitor, when s/he repeatedly refuses to listen to narrations with certain primary concepts, we infer his/her disinterest in those concepts.

To obtain the initial set of user interests we ask user to select those on the kiosk-type computer before entering the exhibition. From that point the user model continually updates the user's interests.

3. NARRATION RETRIEVAL
When user enters an interaction zone of a museum artifact, three narrations are selected and a visitor is presented with three prefaces, short audio clues that correspond to selected narrations. Three initial narrations are inferred based on user's interests and

WWW 2004, May 17–22, 2004, New York, New York, USA.
ACM 1-58113-912-8/04/0005.



concepts represented by the exhibit. After user selects and listens to one of the narrations a new narration is inferred and substitutes the previously selected one. Other two narrations that were not selected are repeated in a next offering. The narration is repeatedly offered at most three times before it is replaced.

The mechanism for inferring narrations considers several aspects to maintain a meaningful semantic continuity of in the visitor's experience. A main driving factor is the visitor's interests. Other factors that contribute to this continuity come from the selected narrations properties: primary concept, secondary concept, theme, topic, and in some cases an artifact/exhibit the narration refers to and how specific or general the narration is. Inference rules are activated for those properties where user interests, properties of the artifact in front of the user, and properties of previously selected narrations match. A narration with highest matching score is selected. The selection is checked with respect to the redundancy of the information in already presented narrations. This is achieved with a help of a partial ordering of the narration objects when such a relation exists.

4. BACKGROUND SOUNDS

Background sounds are integral and ever present part of echo experience. Background sounds are selected based on the properties of the selected narration such as event, place, environment type, primary or secondary concept. When a narration is associated with a specific event (such as World War II) then the reasoning module tries to select a background sound object with this specific event. If an event property is not present in the narration then the reasoning module looks at an environment type associated with the narration and tries to select a background sound corresponding to the environment type.

5. IMPLEMENTATION DETAILS

ec(h)o architecture is shown in Figure 1. A fundamental duty of the reasoning engine is to select narration objects in reaction to external events created by the other ec(h)o modules. Four varieties of communication exist: asynchronously receiving tracking information from the Tracking module, asynchronously receiving Narration/Sound object selection from the Vision module (user selects a narration by manipulating an object which is monitored by the camera), the insertion of initial visitor concepts of interest, and responding to requests for selected narration objects from the Audio module. An application-level protocol was devised for inter-module communication, and a server was implemented on top of the ec(h)o reasoning module. The server is responsible for parsing incoming requests and inserting them as facts or executing queries in the reasoning engine.

The ec(h)o reasoning engine is implemented using rules in the Jess Expert System Shell, a Java tool for embedding an expert system technology to Java-based software systems [2]. We have modeled ontologies using DAML+OIL ontology language and used DAMLJessKB [3] to convert ontologies into Jess facts so they can be accessed by the rules.

6. ISSUES AND CORRECTIVE ACTION

The first prototype for the reasoning module was problematic and we faced a plethora of issues. First and foremost was the fact that the query for selecting narrations took an immense amount of time to complete and generated a non-negligible delay. We have resolved that issue by analyzing the rules and carefully designing rules with an understanding how the RETE network inside Jess is built and updated.

In the first round of evaluations with test users we have identified several issues in the dialogue flow, such as repetitive offering of the same narrations and offering narrations with redundant information. We have addressed those by introducing partial ordering into the model.

On technical side we have encountered several problems with system performance when JESS and DAMLJessKB took a large amount of time to initially load due to the vast amount of facts being generated from the ontologies and from the rules. The solution was two partite. First, it was noticed that the RETE algorithm was very sporadic with its memory usage during the initial loading phase. With a special memory monitoring tool, we noticed that the algorithm would often spike up to a maximum of 420MB of RAM usage and then return to normal memory use. This was easily overcome by increasing the Java Virtual Machine initial heap size and maximum heap size to a value higher than the memory spike. Secondly, we noticed that our module had a greater efficiency when we loaded up the inference rules first, in contrast to loading the ontologies into Jess first. These two simple modifications resulted in a vast speedup of over 700% and into an acceptable range for our application

The last evaluation was conducted from the demonstration of the system in Nature Museum in Ottawa. We collected data (both log files, and questionnaires) from the 5-minute interaction of 15 subjects. More extensive testing was done with 7 subjects. All subjects were briefly trained how to use the system (learning phase), than they had an opportunity to ask questions. Then they used the system on their own for a period of 10-20 minutes. Given the complex interdisciplinary nature of ec(h)o, the project that has to be evaluated from different perspectives. We are still processing evaluation data. However, an anecdotal evidence from the participant interviews indicated high satisfaction with the system.

REFERENCES

[1] Dewey Decimal Classification, http://www.oclc.org/dewey/

[2] Friedman-Hill, E. (2003). Jess In Action. Manning Publications Co.

[3] Kopena, J., and Regli, W. C., DAMLJessKB: A Tool for Reasoning with the Semantic Web. In: Proceedings of ISWC2003. Sanibel Island FL, 2003, Springer, 628-643