Access, Exploration and Visualization of Interest Communities: The VMC Case Study (in Progress)

Anton Nijholt
Centre of Telematics and Information Technology
University of Twente
PO Box 217, 7500 AE Enschede, the Netherlands
anijholt@cs.utwente.nl

ABSTRACT

This paper discusses a virtual world for representing information and natural interactions about performances in an existing theatre. Apart from mouse and keyboard input, interactions take place using speech and language. It is shown how this virtual environment can be considered as an interest community and it is shown what further research and development is required to obtain an environment where visitors can retrieve information about artists, authors and performances, can discuss performances with others and can be provided with information and contacts in accordance with their preferences.

Keywords: Virtual Reality, Speech and Language Interactions, Information Filtering, Agent Technology

1 INTRODUCTION

World Wide Web allows interactions and transactions through WebPages using speech and language, either by inanimate or live agents, image interpretation and generation, and, of course the more traditional ways of presenting explicitly pre-defined information by allowing users access to text, tables, figures, pictures, audio, animation and video. In a task- or domain-restricted way of interaction current technology allows the recognition and interpretation of rather natural speech and language in dialogues. However, rather than the current two-dimensional web-pages, the interesting parts of the Web will become three-dimensional, allowing the building of virtual worlds inhabited by interacting user and task agents and with which the user can interact using different types of modalities, including speech and language interpretation and generation. Agents can work on behalf of users, hence, human computer interaction will make use of 'indirect management', rather than interacting through direct manipulation of data by users.

In this paper we discuss a virtual world for representing information and natural interactions about performances in an existing theatre. Apart from mouse and keyboard input interactions take place using speech and language. It is shown how this virtual environment can be considered as an interest community and it is shown what further research and development is required to obtain an environment where visitors can retrieve information about artists, authors and performances, can discuss performances with others and can be provided with information and contacts in accordance with their preferences.

2 INFORMATION SPACES FOR INTEREST COMMUNITIES

Web-based digital cities have been around for some years. Like many computer games they have evolved from text environments to 2-dimensional graphical and 3D virtual environments with sounds, animation and video. Visitors, or maybe we should call them residents, of these cities visit libraries, museums, pubs, squares, etc., where they can get information, chat with others, etc. In these environments people get the feeling of being together, they are listening to each other and, in

general, they take responsibility for the environment and theirs and others behavior in such environments.

Today there are examples of virtual spaces that are visited and inhabited by people sharing common interests. With virtual spaces or environments we want to refer to computer accessible environments where users (visitors, passers-by) can enter 3D environments, browse (visual representations of) information and can communicate with objects or agents (maybe other visitors in the same environment). These spaces can for example, represent offices, shared workspaces, shops, class rooms, companies or cities. However, it is also possible to design virtual spaces that are devoted to certain themes and are tuned to users (visitors) interested in that theme or to users (visitors) that not necessarily share common (professional or educational) interests, but share some common conditions (driving a car, being in hospital for some period, have the same therapy, belonging to the same political party, etc.).

As an example we mention a virtual world developed at a cancer research institute in Seattle. This world enables people struggling with cancer to obtain information and interact with others facing similar challenges. Patients, families and friends can enter the three-dimensional world (a rendering of the actual outpatient lobby), get information at a reception desk, visit a virtual gift shop, etc. Each participant obtains an avatar representation. Participants can engage in public chat discussions or invitation-only meetings. A library can be visited, its resources can be used and participants can enter an auditorium to view presentations. Part of the project consists of developing tools to create other applications.

3 A VIRTUAL THEATRE COMMUNITY: A CASE STUDY IN PROGRESS

We present research on visualization and interaction in a realistic model of an existing theatre. This existing 'Muziekcentrum' offers its visitors information about performances by means of a yearly brochure. In addition, it is possible to get information at an information desk in the theatre (during office hours), to get information by phone (by talking to a human or by using IVR). The database of the theater holds the information that is available at the beginning of the 'theatre season'. Our aim is to make this information more accessible by using multi-modal accessible multi-media web pages. A more general aim is to do research in the area of web-based services, in particular interactions in virtual environments using speech and language.

Our virtual theatre has been built according to design drawings of the architects of the building. Part has been realized by converting AutoCAD drawings to VRML97. Video recordings have been used to add 'textures' to walls, floors, etc. Sensor nodes in the environment activate animations or start events (entering a dialogue mode, playing music, moving spotlights, etc.). Visitors can explore the environment of the building, hear the carillon of a nearby church, look at a neighboring pub and movie theatre, etc. They can enter the theatre and walk around, admire the paintings on the walls, enter the main performance hall, go to the balconies and, take a seat in order to get a view of the stage from that particular location. Information about today's performances is available on a screen that is automatically updated. Visitors may go to the information desk in the theatre, see previews and start a dialogue with an agent called 'Karen'. Karen has her looks against her. We are working on a more attractive – not necessarily human-like - appearance for Karen. In 1999 we aim at user evaluation studies that will concentrate on questions about, among others, the need of reasonable realistic representations of the theatre information and transaction service interactions that are offered. Another aim that will be explored is the embedding of this particular virtual environment in a virtual cultural arena where people can ask, retrieve and explore information about theatre and music performances in general. Agents, like Karen, will help the users with these tasks.

4 AGENTS AND INTERACTIONS

4.1 A NAVIGATIONAL AGENT

Clearly, the WWW-based virtual environment we are developing allows navigation input through keyboard and mouse. Such input allows

the user to move and to rotate, to jump from one location to an other, to interact with objects and to trigger them. In addition, a navigation agent has been developed that allows the user to explore the environment and to interact with objects in this environment by means of speech commands. Obviously, we do not want completely separated modalities. It should be left to the user to choose between the interacting means or to use both, sequentially or simultaneously. A smooth integration of the pointing devices and speech in a virtual environment requires means to resolve deictic references that occur in the interaction, and the navigation agent should be able to reason (in a modest way) about the geometry of the world in which it moves. We slowly extend and improve the interaction and navigation intelligence of our present navigation agent. At this moment we are exploring the possibility of speech recognition for several clients on a central server and the advantages of making the navigation agent visible for the user. One of our conclusions is that current web technology hardly allows smooth integration of speech recognition and browsing a virtual world.

4.2 AN INFORMATION AND TRANSACTION AGENT

As mentioned before, a second agent called Karen allows a natural language dialogue with the system about performances, artists, dates, prices, etc. Karen wants to sell tickets. She is fed from a database that contains information about performances in some of our local theatres. Developing intelligence for Karen, in this particular environment, is a main aim of our project. Presently the input to Karen is keyboard-driven natural language and the output in our for the general audience WWW accessible virtual world is screen and menu based ([3]). In a prototype system we allow Karen to use a combination of speech synthesis and information presentation on the screen ([4]). Based on the user utterance, the context and the database, the system has to decide on a response action, consisting of database manipulation and dialogue acts. In our experimental system Karen's spoken dialogue contribution is presented by visual speech, that is, a 'talking face' on the screen, embedded in the virtual world, mouths the questions and part of the responses. If necessary, information is given in a window on the screen and the user can click items to get more information. The virtual face that has been designed allows animation of lip and face movements and animation of some simple face expressions ([1]). A Dutch text-to-speech synthesis system has been used to give Karen a voice ([2]). For speech-image synchronization 3D images of visemes are called when corresponding phonemes are spoken. Since our application is web-based, it requires the solution of technical problems dealing with sending and compressing sound files, commands and synchronizing sounds and animations.

4.3 FUTURE AGENTS: CONVERSATION, RETRIEVAL, FILTERING

As may have become clear from the previous sections, our approach to designing a virtual environment for an interest community is bottom-up. At this moment the system has two agents with different tasks and with no interactions between them. Moreover, the agents do not employ a model of a user or of user groups. In general, when we talk about interface agents we mean software agents with a user model, that is, a user model programmed in the agent by the user, provided as a knowledge base by a knowledge engineer or obtained and maintained by a learning procedure from the user and customized according to his preferences and habits and to the history of interaction with the system. In this way we have agents that make personalized suggestions (e.g. about articles, performances, etc.) based on social filtering (look at others who seem to have similar preferences) or content filtering (detect patterns, e.g. keywords) of the items that turn out to be of interest to the user. These agents can be passive that wait until they are put into action or they sense changes, take initiative and perform actions, e.g. to inform the user without being asked about new information.

Our first concern in the near future will be the introduction of a conversational agent (which has some general knowledge about well known artists and some well known performances). In this way we have obtained three kinds of dialogues (information & transaction dialogues, command-like dialogues and conversational dialogues). A slight sharing

of knowledge (in particular, preferences of the user) between agents will become possible. In a next step this knowledge of preferences should be exploited, not only in the interactions with the user, but also in designing an agent that retrieves information that matches with the users profile. Visualization of the domains accessible to users may help to guide the interpretation of questions and requests for retrieval (we give a different interpretation to a question about artists when we are in an opera building then when we are in a music hall).

REFERENCES

[1] M. van den Berk. Visuele Spraaksynthese: Een tot de verbeelding sprekend gezicht. M. Sc. thesis, October 1998.

[2] J. Hulstijn & A. van Hessen. Utterance Generation for Transaction Dialogues. In: Proceedings 5th *International Conf. Spoken Language Processing (ICSLP)*, Sydney, Australia, 1998, to appear.

[3] D. Lie, J. Hulstijn, A. Nijholt, R. op den Akker. A Transformational Approach to NL Understanding in Dialogue Systems. Proceedings *NLP and Industrial Applications*, Moncton, New Brunswick, August 1998, 163-168.

[4] A. Nijholt, A. van Hessen & J. Hulstijn. Speech and Language Interaction in a (Virtual) Cultural Theatre. Proceedings *NLP and Industrial Applications*, Moncton, New Brunswick, August 1998, 176-182.

[5] A. Nijholt, J. Hulstijn & A. van Hessen. Interacties in Virtuele Web-gebaseerde Omgevingen. Proceedings Conference *Informatiewetenschap '98*, (in Dutch) Antwerpen, December 1998, to appear.