

A Dialogue Agent for Navigation Support in Virtual Reality

Jeroen van Luin
University of Twente
Computer Science Department
PO Box 217, 7500 AE Enschede
The Netherlands
vanluin@cs.utwente.nl

Rieks op den Akker
University of Twente
Centre of Telematics and
Information Technology
PO Box 217, 7500 AE Enschede
The Netherlands
infrieks@cs.utwente.nl

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

ABSTRACT

We describe our work on designing a natural language accessible navigation agent for a virtual reality (VR) environment. The agent is part of an agent framework, which means that it can communicate with other agents. Its navigation task consists of guiding the visitors in the environment and to answer questions about this environment (a theatre building). Visitors are invited to explore this building, see what is there, ask questions and get advice from the navigation agent. A 2D map has been added to the environment so that visitors can make references to the locations and objects on this map, both in natural language and by clicking with the mouse, making it a multimodal system with cross-modality references.

Keywords

Navigation, virtual environments, natural language dialogue, multi-modality, agent technology

INTRODUCTION

We have designed and built a VR version of a theatre [3]. It has been built according to the drawings of the architects. In the theatre we have added agents. Karin is the main agent. She is standing behind an information desk and knows a lot about performances that take place and performers that will perform in the real theatre. Visitors can ask Karin questions in natural language. Karin knows how to access the performance and performers database and tries to extract and formulate answers. This virtual world has been made accessible to the audience. From the visitor's point of view the need of an other agent emerged. To whom do we address our questions about the environment itself? To whom do we address our questions about how to continue, where to find other visitors or where to find domain-related information?

At this moment we are following different approaches to solve this problem. The approaches are related and can be integrated since all of them are agent-oriented and are oriented towards a common framework of communicating

agents. In addition, we build this framework in such a way that different agents with different abilities can become part of it: a simple animated piano player, a baroque dancer that 'understands' the music she is dancing on, Karin who knows about theatre performances and our navigation agent who knows about the geography of the building.

The present paper is about navigation. How can we build navigation intelligence into an agent? What does navigation intelligence mean? How can we connect this intelligence to language and vision intelligence? Visitors of our environment are language users (or language animals). Visitors of our environment know (or interpret) what they see. There is a continuous interaction between verbal and nonverbal information use when interpreting a situation in our virtual environment. Modelling of this interaction, the modelling of the representation and interpretation of these sources and the generation of multimedia information from these sources are the main topics of our research.

AGENT AND MAP BASED NAVIGATION SUPPORT

To our virtual theatre we have added a map and an intelligent navigation agent. The reader is referred to [1,2] for observations on user's preferences for navigation support. Our work is in progress, meaning that the system is there, but no effort has been done to add 'graphic sugar' to the layout and the integration of the different windows that are used. When user experiments start this has to be worked on first. In Fig. 1 we display the current floormap and the agent window. In Fig. 2 a view on part of the

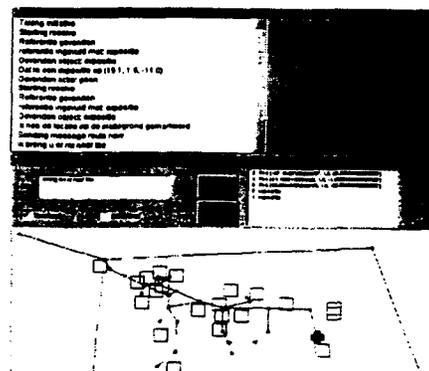


Figure 1: Floor map and agent window

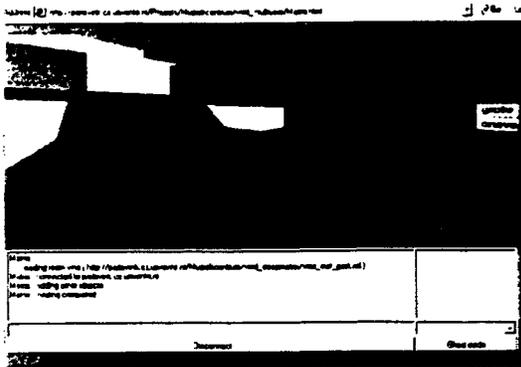


Figure 2: The visitor has been brought to the stairs

virtual world is presented. In this view the user is looking at the stairs going to the next floor of the building.

The visitor can ask questions, give commands and provide information when prompted by the agent. This is done by typing natural language utterances and by moving the mouse pointer over the map to locations and objects the user is interested in. On the map the user can find the performance halls, the lounges and bars, selling points, information desks and other interesting locations and objects.¹ The current position of the visitor in the virtual environment is marked on the map. While moving in VR the visitor can check her position on this floor map. When using the mouse to point at a position on the map references can be made by user (in natural language) and system to the object or location pointed at.

Natural Language Access

As mentioned, the navigation agent can be accessed using natural language. We have annotated a small corpus of example user utterances that appear in navigation dialogues. On the one hand we have complete questions and commands. On the other hand we have also short phrases that are given by the user in reply to a clarifying question of the navigation agent. An example of a question is: "What is this?" while pointing at an object on the map, or "Is there an entrance for wheel chairs?". Examples of commands are "Bring me there." or "Bring me to the information desk." Examples of short phrases are "No, that one." or "Karin." From the annotated corpus a grammar was induced and a unification-type parser for Dutch can be used to parse these utterances into feature structures.

From Utterance to Action

Three agents communicate to fill in missing information in the feature structure (when the information given by the user in question, answer or command is not yet complete) and to determine the action that has to be undertaken (answering the question, prompting for clarification or

¹ Interestingly, one of these objects is a board in the virtual world displaying a map with chair positions in a performance hall. When the visitor clicks on a chair on this VR map she is teleported to this chair to get a view at the stage.

missing information, displaying a route on the map or guiding the user in VR to a certain position). This is done in co-operation by the navigation agent, the dialogue manager and the Cosmo Agent. The latter can 'talk' to the Cosmo-Browser using its EAI to retrieve the current position of the visitor. Not yet implemented is the possibility that not only the position but also what is in the eyesight of the visitor can be retrieved. This will allow references in the dialogue to objects that are visible (in VR) for the visitor.²

Navigation Dialogues

Three examples of navigation dialogues are given below:

Visitor: How do I get at the information desks?

Agent: Where did you want to go?

Visitor: The desk.

Visitor: Can you bring me to the information desk?

Agent: I have marked the position on the map.

Now I will bring you there.

Visitor: [Clicks on an object on the map] What is this?

Agent: That is an exposition.

Visitor: Where is it?

Agent: You can find it in the lounge.

Visitor: Let's go there.

Agent: I bring you there.

CONCLUSIONS AND FUTURE RESEARCH

The prototype navigation agent which we discussed here is certainly not our final solution in assisting visitors of our virtual environment. In the next phase of research we need to concentrate on the communication with other agents that are available in the VMC. How can we take care that a visitor's question reaches the appropriate agent? How can we model the history of interaction in such a way that different agents do not only know about their own role in this interaction but also about others? Unlike others, our environment allows the investigation of communication between active and passive agents that inform the visitor about the possibilities and the properties of an information-rich virtual environment.

REFERENCES

1. Darken, R.P. & Silbert, J.L. Wayfinding strategies and behaviors in virtual worlds. *Proc. CHI'96*, 142-149.
2. Höök, K. et al. Towards a framework for design and evaluation of navigation in electronic spaces. *Persona Deliverable* for the EC, 1998.
3. Nijholt, A. & Hulstijn, J. Multimodal interactions with agents in virtual worlds. In: *Future Directions for Intelligent Information Systems and Information Science*, Springer Physica-Verlag, 2000.

² And, indeed, it will introduce problems dealing with (not) being able to make clear distinctions between first and second person's perspectives on the environment.