Agent Assistance: From Problem Solving to Music Teaching

Anton Nijholt
University of Twente, CTIT
PO Box 217, 7500 AE Enschede
The Netherlands
anijholt@cs.utwente.nl

ABSTRACT
We report on our research on agents that act and behave in a web learning environment. This research is part of a general approach to agents acting and behaving in virtual environments where they are involved in providing information, performing transactions, demonstrating products and, more generally, assisting users or visitors of the web environment in doing what they want or have been asked to do. While initially we hardly provided our agents with 'teaching knowledge', we now are in the process of making such knowledge explicit, especially in models that take into account that assisting and teaching takes place in a visualized and information-rich environment. Our main (embodied) tutor-agent is called Jacob; it knows about the Towers of Hanoi, a well-known problem that is offered to CS students to learn about recursion. Other agents we are working on assist a visitor in navigating in a virtual world or help the visitor in getting information. We are now designing a music teacher - using knowledge of software engineering and how to design multi-modal interactions, from previous projects.

Categories and Subject Descriptors
K.3.1 [Computers and Education]: Computer Uses in Education. H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems - *Animations*. I.2.11 [Artificial Intelligence]: Distributed Artificial Intelligence - *Intelligent agents.*

General Terms
Design, Human Factors, Algorithms, Performance, Theory.

Keywords
Teaching, embodied agents, virtual reality, music teaching.

1. INTRODUCTION
In the past years we have developed several virtual environments for different domains. In these environments we have developed several agents. One of the embodied agents is Karin, who enters a natural language dialogue with a human agent about theatre performances in an environment where information is also visualized and with different ways for the user to explore the environment and information [2]; secondly, a navigation agent that knows about the geography of a theatre building and helps the user to explore this building [3]; thirdly, Jacob, an embodied agent who knows how to assist the user in a problem solving task by monitoring the user, correct the user and entering a dialogue about this particular, visualized, task [1]. To emphasize the role of visualization: in a spoken dialogue using the telephone it is hard to imagine a user saying 'Put the green disc on the right peg.' In a visualized domain this can be a natural utterance.

We survey the tasks of these agents, how they interact with the user, visitor or student and how and where we can use obtainments of this research to model learner-teacher interaction in a visualized world. We zoom in on Jacob in section 3 and give an examples of a dialogue between agent and user. One of the objectives of the Jacob project was to investigate the portability of such an agent to other domains. For that reason we are now working on two more projects. The first one is a navigation agent (section 4) that assists the user to find its way in a virtual environment. The agent has knowledge of its environment and allows multimodal interaction. In section 5 on future research we explain an even more recent project: the virtual music teacher.

2. AGENT INHABITED ENVIRONMENT
As mentioned in the introduction, we have built several virtual environments and have added agents to these environments that help, guide and teach the visitors and users. The main environment is the so-called virtual music theatre. This theatre environment is in fact our laboratory for integrating and experimenting with results from smaller projects and environments. In particular software engineering approaches, including agent-oriented design and usability issues are dealt with in this environment [4], [5]. In this environment the Karin agents is the embodiment of a natural language interface to a database containing information about theatre performances, performers and availability of tickets. Navigating in such an environment is a well known problem. A navigation agent is available (section 4), to help the user to find its way in this virtual theatre environment.

3. THE JACOB PROJECT
3.1 Introduction
The Jacob project investigates the application of virtual reality (VR) techniques and involves the design and construction of an animated agent in a 3-dimensional virtual environment. The agent is called Jacob and provides instruction and assistance for tasks that the user has to learn to perform in a virtual environment. The user interacts with Jacob by performing actions as well as by

using natural language. The use of a lifelike agent in an interactive learning environment has a strong positive impact on students, which has been shown by empirical studies. Such an agent can increase both the learning performance and the student's motivation. Eventually, versions of the Jacob agent will be integrated in the virtual theatre environment mentioned in section 2, but for different tasks than it has been designed for originally. One candidate task is navigation in the theatre environment, for which we already investigated problems associated with natural language navigation dialogue, multimodality and associating the interpretation of the visitor's utterances and actions with actions for a system agent that guides the user from its current position to the desired position (see section 4). In a recently started project we will also investigate whether we can make a translation from the Jacob agent to an embodied agent that acts as a music teacher (see section 5).

3.2 Project Objectives

The research focus of the Jacob project is the use of VR techniques and the design and implementation of VR based systems. Important questions addressed in this project are:

- How can different interaction modalities like natural language, gestures, gaze, and manipulation of objects be integrated in a task oriented VR system?
- How can agent technology be used in VR systems?
- Can traditional HCI and software engineering technology be applied for designing and building a VR system or does it require different technology?

To support the research activities, a prototype system is being developed that has to meet a number of requirements. First, the interaction between the user and Jacob should be multimodal. An important issue is how natural language dialogue and nonverbal actions are to be integrated and what knowledge of the virtual environment is needed for this purpose. Second, the Jacob agent should behave in an intelligent way, helping the user proactively and learning from the interaction. Third, visualisation of Jacob plays an important role, including animation of the body, generation of facial expressions, and synchronisation of lip movement and speech. Fourth, both the user and Jacob should be able to manipulate objects in the virtual environment, e.g. by using a dataglove. As the domain for the project, we have selected instruction of tasks. This concerns tasks that consist of manipulating objects, e.g. moving an object, pressing a button, or pulling a lever. More in particular we have chosen the task of solving the Towers of Hanoi probem in interaction with a learner.

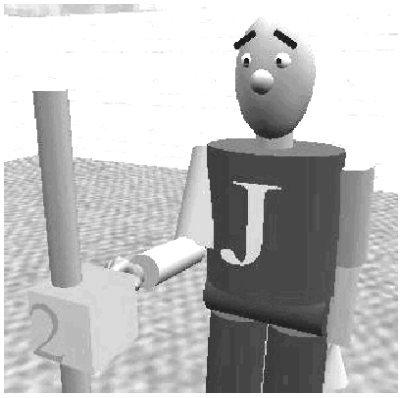

Figure 1: Jacob Showing a Move

User: "What should I do?"
Jacob: "The green block has to be moved from the left peg to the right peg. Should I demonstrate it?"
User: "Yes please"
Jacob: "I will demonstrate the next step. The green block has to be moved from the left peg to the right peg."
Jacob moves the green block from the left to the right peg; User moves the red from the left to the middle peg.
Jacob: "That is correct. Keep up the good work!"
The user moves the green block from the right peg to the middle peg (on top of the red block).
User: "Am I doing right?"
Jacob: "Yes, you are doing fine!"
The user moves the blue block from the left peg to the middle peg (on top of the green block).
Jacob: "The blue block is not placed correctly: it is not allowed to place a larger block on top of a smaller one."
Jacob moves the blue block back to the left peg.
User: "That was not right?"
Jacob: "No"
User: "Why not?"
Jacob: "Because the blue block has to be moved from the left peg to the right peg."

Figure 2: Jacob Instructing the User

3.3 Approach

We have used existing knowledge, theories, and frameworks from different areas like intelligent tutoring systems, computer graphics, and multi agent technology. Software engineering plays a prominent role in the Jacob project. We apply object oriented techniques, design patterns, and software architecture knowledge. For the current version of the prototype, we have applied the following technology: the basic virtual environment and the Jacob agent have been defined using VRML. The 'intelligent' part of the system has been written using Java. The Java part is linked to the VRML world through the External Authoring Interface. The system runs in a web browser with a VRML plug-in. In this way, the system is highly portable and can be executed on any PC or workstation. In Fig. 1 we have a screenshot showing Jacob while he is moving a block.

3.4 Natural Language Processing

As Jacob and the student should be able to have a dialogue about the task that is instructed, the Jacob system also contains a *natural language dialogue manager*. Seen as a black box, Jacob's dialogue manager processes and interprets utterances from the student and produces utterances and actions based on the state of the world and the dialogue. In an utterance, the student can refer to physical objects like a block, abstract objects like the task, agents, actions like moving a block, and conversational and didactical actions performed by Jacob. An example of a dialogue fragment is presented in Fig. 2.

Based on the meaning of the utterance and based on actions performed by the student, an action to be performed or an utterance to be produced is selected. The dialogue manager uses and updates a dialogue state and an instruction state (e.g. how many times Jacob has explained something). The focus list is also up-

dated. Jacob's utterances are produced using predefined sentences and templates in which appropriate information is put. It can be observed that the instruction involves dialogue and that actions and utterances are mixed. The dialogue is actually part of the instruction. Both the instruction model and the dialogue manager react on actions and utterances and can produce actions and utterances themselves. This implies an *integration of dialogue manager and instruction model*. We are developing a formalism that uniformly represents instructional rules and methods and the process of selecting actions and utterances to be produced. The dialogue system keeps also track of Jacob's mood, represented as a 4D vector, which is mapped onto a mood descriptor, e.g. "happy". The information is sent to the Jacob applet to produce an appropriate change in the facial expression.

4. A NAVIGATION AGENT IN VR

4.1 Introduction

As mentioned in previous sections, a first version of a navigation agent has been designed and implemented in our virtual thetare environment. The current agent is not yet embodied. This prototype version has been made to get aware of problems concerning navigation dialogues before introducing a Jacob-like architecture for a navigation agent in the virtual theatre. 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 a 2D map associated with the virtual world the user can find the performance halls, the lounges and bars, information desks and other interesting locations and objects. 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 and system to the object or location pointed at.

Three agents communicate to fill in missing information (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 missing information, displaying a route on the map or guiding the user in VR to a certain postion). This is done in co-operation by the navigation agent, the dialogue manager and the Cosmo Agent. The latter can 'talk' to the CosmoBrowser using its EAI to retrieve the current position of the visitor.

4.2 Helping to Navigate

Two examples of navigation dialogues are given below:

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.

The dialogue manager keeps track of part of the history of the interaction and places that have been visited. This makes it possible to make the user aware of what he has done before, when appropriate. Some of the additions to the agent deal with giving him the initiative in the interaction [5]. In a next phase we will start with introducing a Jacob-like architecture for this particular agent in the virtual theatre environment.

5. TOWARDS A VR MUSIC TEACHER

We took the decision to start working on a second, again modest, project in which we can further experiment with a similar agent, that is, an agent that assists the user in an information-rich visualized environment. What generalizations are possible from the previous tasks and domains in order to obtain an agent that teaches playing a particular musical instrument? Apart from knowing how to teach according to different strategies, this agent is expected to interact with the user about and using music. This will be done by listening and interpreting musical input from the learner (input from synthesizer keyboard and other interactions that can be converted into midi-sequenes) and simple speech commands. Presently we are investigating how we can use a virtual 3D piano, where the virtual animated teacher and maybe also the student can jump to play the keys. Remember the famous piano dance of Tom Hanks and Robert Loggia in the movie *Big*. Although many differences can be mentioned, we rather like to explore the similarities between this task and the task our Jacob agent performs. We need an animated agent, we need some kind of interaction (not only verbal interaction, but also by animation, playing and comparing a melody), etc.

6. REFERENCES

[1] Evers, M. & A. Nijholt. Jacob - an animated instruction agent for virtual reality. In: *Advances in Multimodal Interfaces - ICMI 2000*, LNCS 1948, T. Tan et al. (Eds.), Springer, Berlin, 526-533.

[2] Nijholt, A. & Hulstijn, J. Multimodal Interactions with Agents in Virtual Worlds. *Future Directions for Intelligent Systems and Information Science*. N. Kasabov (ed.), Physica-Verlag, 2000, 148-173.

[3] Luin, J. van, Nijholt, A. & Akker, R. op den. Natural Language Navigation Support in Virtual Reality. ICAV3D, May-June 2001, Mykonos, Greece, to appear.

[4] Schooten, B.W. van. A specification technique for building interface agents in a web environment. In: *Learning to Behave: Interacting Agents*. A. Nijholt et al. (eds.), University of Twente, 2000, 83-98.

[5] Zwiers, J. et al. Design issues for navigation and assistance agents in virtual environments. In: *Learning to Behave: Interacting Agents.* A. Nijholt et al. (eds.), University of Twente, 2000, 119-132.