Virtual Malls for Web Commerce: Observations and Case Study

Anton Nijholt

Centre for Telematics and Information Technology (CTIT)
University of Twente, PO Box 217
7500 AE Enschede, the Netherlands
anijholt@cs.utwente.nl

Abstract In the real world there seem to be essential differences between shops, tourist offices, city halls, theatres and museums. However, when we try to design virtual environments allowing services, transactions, explorations and communication, are these environments really that different? Doesn't the 'virtual' in virtual environments allow us to introduce possibilities for visitors that extend those that are available in each of the mentioned physical environments and that can be become common for each of them? In this paper we survey our research on these issues and our attempt to integrate this research in a virtual environment devoted to the presentation and exploration of theatre information. We try to make clear that the topics we have to deal with do not differ from those that deal with access to commercial environments, museum and educational environments and other service and transaction environments.

1 Introduction

Many distinctions can be made between shopping, going to a theatre performance, visiting an art exhibition in a gallery, going to a museum, asking for information at a tourist office, buying a house and visiting a town hall to arrange a new passport or to register a newly-born child. These distinctions are useful and really make a difference. Nevertheless, we can also distinguish behavior and interactions of the people involved in these activities that is common, rather than distinct. When the activity is not completely routine, people need help and advice. "What else is there to see?", "Do you have similar and cheaper products.", "Is there a Van Gogh painting from the same period?", "Why isn't this allowed?", "Do I have to pay for this?", etc.

In all these utterances the user is making references to information that has not been mentioned or shown explicitly. Generally, this is one of the main obstacles to natural language understanding. However, unlike telephone conversations or menu-based computer-human interaction, in a virtual reality environment we can make an attempt to resolve such references by looking at the 3D context we have designed ourselves. The possibility to match user utterances with both a language model and a visualized context model makes it interesting to explore the embedding of dialogue systems in virtual environments.

In this paper we argue that in virtual reality environments we have the tools both to emphasize and to blur all kinds of distinctions in the natural world. Here we want to

explore similarities in using and visiting virtual environments, rather than differences. However, before doing that, it should be clear that different domains have different audiences; such audiences may require different interaction strategies and such audiences may require different types of help in order to get the information they want, to perform a transaction, to successfully conclude a negotiation or to be able to communicate with other users interested in the same environment. Clearly, accidental meetings may lead to interesting and useful contacts between people and when we design information and transaction systems in virtual worlds, we certainly should take care that such contacts not only remain possible, but also are natural, perhaps more natural than in reality. Nevertheless, we should be aware of the fact that for the majority of people buying a pack of milk from a supermarket, is an act that emotionally is very different from the act of discussing which performance, which art exhibition or which other kinds of cultural events will be visited this evening.

2   What Interactions Are Desired by Visitors?

We take the point of view that despite differences, there are so many similarities for which we need research results that, without neglecting the special needs for the domain of museums, the domain of theatres, the domain of commerce, the domain of (commercial and governmental) information services, etc., we first need to explore similarities rather than differences. Similarities we want to exploit are:

- Visitors need to be able to explore an environment; this requires both visualization of an environment and the available information and the possibility to start dialogues about information, transactions and environment;
- Visitors need to be able to change an environment; changes may differ from changing the number of available tickets for a performance, changing the number of books displayed in a bookshop, changing the color of the furniture in a showroom, becoming a member of the noticeable audience (for a performance, an exhibition or a particular painting) to being able to influence the composition of an exhibition, the lightning during a performance, the behavior of actors during a performance, etc.
- Visitors need to be able to express themselves in their own, natural ways; the environment should be able to detect and interpret the different ways a visitor tries to communicate with the environment. That is, the environment has to understand multimodal user input (keyboard, mouse, language, speech, gestures, touch, eye tracking, facial expressions, body movements, etc,).
- Visitors need help. This may be provided by traditional means (e.g., context-sensitive pop-up windows), but it is more interesting to look at domain agents that know about the environment, that know about the information that is available, and that know about the preferences of a particular visitor. It should be defined how such personal assistants (or butlers) can act, communicate with others, serve their master and decide which private information others can access.
- Visitors should be able to communicate with each other. That is, not only communication with domain agents, but communication with who ever shows interest in the same environment, the information it offers, but much more importantly,

information that is not explicitly available in the environment, but that is available in the shared knowledge of domain agents, visitors and world wide web.

In what follows we don't make a distinction between customers, visitors, users, etc. We present a case study in which we discuss an environment in which we can see and anticipate goal-directed behavior and browsing behavior of people using WWW. These different kinds of behavior follow from the fact that user goals can be more or less definite. In some situations shopping involves stress, in other situations it might be entertainment. The product that is looked for is also important. This shows in the differences between buying the usual pack of milk, the daily newspaper, a pair of shoes, a long awaited book or the ingredients for tonight's dinner. In the latter case, during shopping and depending upon what is available, what special offers there are, what time is available for shopping, etc., the decisions are made and the dinner is composed. Hence, apart from navigating in a physical space (the actual mall, the shop, the web pages shown on the screen), there is navigation in a space with product information and a space containing constraints that need to be considered in order to make decisions. In addition we can talk about a social space in which a buyer operates. A sales person can be part of this social space, but so can other customers or people looking for information. We can talk to them in order to get what we want (direct social navigation) but we can also notice and follow their behavior (indirect social navigation). In a virtual world we can have a social space inhabited by live agents, virtual agents and visitors.

3 What Will Be Influenced by Interaction Characteristics?

Now that we've looked at a virtual environment from the visitor's point of view, we now want to add some observations from the point of view of the environment's owner, the company that wants to sell and wants to make profits, while keeping good relations with its customers. Clearly, offering facilities as mentioned above helps in being a 'good' company. However, there are other aspects, which should be given attention. In particular, the virtual mall owner wants to know about the client's preferences. Not only concerning products, information and advice, but also concerning interaction behavior with the environment, with the domain agents made available in this environment and with other clients visiting this environment. The following, observations, clearly not independent from each other, can be given.

- *Consumer Buying Behavior (CBB)*. There exist several models that describe consumer-buying behavior. In Guttman et al. [2] some of them are related to the use of agent technologies. The main stages that are distinguished are need identification, product brokering, merchant brokering, negotiation, purchase and delivery, and service and evaluation. In particular product and merchant brokering and negotiation are considered to be the agent-centric stages of the models.
- *Trust and security*. Their users will not trust non-secure web systems. However, it is not necessarily the case that technically secure systems will be perceived as trustworthy. The concept of trust in e-commerce systems has been studied by, among others, Soltesz [8]. We adhere to his definition of trust: *The users positive expectations on the e-commerce systems ability to function accordingly to the user intentions expressed through his actions, under conditions of risk*. The paper by Soltesz is interesting because of his avocation of the use of the social psychological

concept of 'action space' in e-commerce research. Action space refers to the possibility of the user to control his environment. Obviously, different tasks and different user groups may require different action spaces. The design of the action space influences the users degree of trust.

- *Commitment*. Trust may be a prerequisite, but it does not necessarily lead to commitment. How can we stimulate the users commitment, his or her sympathy or even love for the system? When we are able to evoke such emotions towards the system, it certainly will help the users willingness to accept (to a certain degree) slow, wrong and non-robust feedback from the system. In a spoken language dialogue system a careful design, anticipating a user's errors and the system's shortcomings, of the prompts of the system is the main non-intimidating tool for achieving commitment.

- *Presence.* 'Presence' (the experience of being in one place or environment, even when one is physically situated in another; see Lombard & Ditton [5]) is an issue in the design of virtual environments that can be very much influenced (using sound, natural interactions, visual effects, etc.). Presence very much influences the user's well being, his or her patience, his or her willingness to buy and explore and to take advice seriously. Causes and effects of presence have been studied in experiments, but the number of parameters involved, e.g., task and user characteristics, makes systematic research difficult.

Among the design issues of 'action space' are: information needed to increase users trust (transparency of the system, who are the owners and trusted third parties, what are the customers rights, which privacy policies are maintained), communication to increase users trust (with company representatives, other users and trusted third parties), which interaction decisions are available (to cancel an order, to delete an electronic trail, etc.) and which information about the system status is available. These issues also influence the commitment the user shows to the system.

Needless to say that for many tasks and activities one may not expect that presence will necessarily enhance performance. On the contrary, presence can be noise which distracts a user from performing a task and which throws away the advantages of individual access from a home PC to, e.g., an information service. There is no need for a virtual waiting queue or the sound of a leaving plane when all we want to know is when the next flight to Amsterdam leaves.

Also in real life an environment may be designed to exploit the 'positive' feelings with respect to the environment, to stimulate trust, commitment and community feeling. In Kooijman [4] a rather comprehensive survey of design concepts for shops and shopping malls can be found. Hardly any attempt has been made to translate such concepts to virtual environments. We expect that in the near future much work will be devoted to the transition of knowledge of design, building, use and user appreciation of physical spaces to knowledge of similar issues in virtual environments.

4 The Case: Twente Virtual Theatre

We discuss a virtual world (Nijholt & Hulstijn [6]) for presenting information and allowing natural interactions about performances, associated artists and groups, availability of tickets, etc., for some existing theatres in the city of Enschede, the

Netherlands. Our virtual theatre1[1] has been built according to the design drawings made by the architects of our local theatre. Part of the building has been realized by converting AutoCAD drawings to VRML97. Video recordings and photographs have been used to add 'textures' to walls, floors, etc. Sensor nodes in the virtual environment activate animations (opening doors) 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. and they can enter the theatre and walk around, visit the concert hall, admire the paintings on the walls, go to the balconies and, take a seat in order to get a view of the stage from that particular location. When the performance hall is entered, the lights dim, spotlights are moving over the stage and some music starts playing. Information about today's performances is available on an information board that is automatically updated using information from the database with performances. In addition, as may be expected, visitors may go to the information desk in the theatre, see previews of performances and start a dialogue with an information and transaction agent called 'Karen' (Figure 1). Karen has a 3D animated talking face. On Karen's desk is a monitor on which visitors can see pictures or video previews of performances. Unfortunately, for most performances no video preview is available yet, so we cannot display them for every performance that is in the database.

Fig. 1. Karen at the Information Desk

Domain-dependent tasks are assigned to agents in our environment. It can be useful to visualize such agents using talking faces and animated 3D avatars. From several studies it has become clear that people engage in social behavior toward machines. It is also well known that users respond differently to different 'computer personalities'. It is possible to influence the user's willingness to continue working even if the system's performance does not seem to be adequate. Users can be made to enjoy the interaction and they can be made to perform better, all depending on the way the interface and the interaction strategy have been designed.

---

[1] The system is accessible using the Cosmo Player browser (http://parlevink.cs.utwente.nl)

4.1 An Agent Platform in the Virtual Environment

In the current prototype version of the virtual theatre we have an information and transaction agent, we have a navigation agent and there are some agents under development. An agent platform has been developed in JAVA to allow the definition and creation of intelligent agents. Users can communicate with agents using speech and natural language keyboard input. Any agent can start up other agents and receive and carry out orders of other agents. Questions of users can be communicated to other agents and agents can be informed about each other's internal state. Both the information & transaction agent and the navigation agent are in the platform. But also the information board, presenting today's performances, has been introduced as an agent. And so can other objects in the environment.

4.2   Information, Transaction, and Navigation Services

Karen, the information & transaction agent, allows a natural language dialogue with the system about performances, artists, dates, prices, etc. Karen wants to give information and to sell tickets. Karen is fed from a database that contains all the information about performances in the (existing) theatre.

In our current version of the dialogue system of which Karen is the face user utterances are simplified using a great number of rewrite rules. The resulting simple sentences are parsed. The output can be interpreted as a request of a certain type. System response actions are coded as procedures that need certain arguments. Missing arguments are subsequently asked for. The system is modular, where each 'module' corresponds to a topic in the task domain. For example, a module has to take care of a date a user is referring to (next Wednesday, over two weeks, tomorrow).

Users need help in order to be able to get the information they want. Karen knows about theatre performances and artists and the user knows that she should address Karen in order to get such information. However, when we visualize information (that is, we introduce information boards, posters, doors to rooms containing particular information, embodied agents with particular knowledge, the building and geographically and semantically associated information, etc.), the user expects to get answers that show that the system knows about what the visitor sees and explores and is at least able to initiate a dialogue about the topics the user introduces and to return believable (not necessarily satisfactory) answers. Presently, our navigation agent knows about a geometrically described environment. Where can we find the keyboard, where is Karen, where is the entrance, etc.? When a user addresses the navigation agent with a question that is about performances, the navigation agent makes an attempt to give control to Karen, the information and transaction agent.

4.3 From Environment to Community

Today there are examples of virtual spaces that are visited and inhabited by people sharing common interests. These spaces can represent offices, shops, classrooms, companies, etc. 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, recreational, educational) interests, but share common conditions (driving a car, being in hospital for some period, having the same therapy, belonging to the same political party, etc.).

In the previous subsections we have looked at possibilities for theatre visitors to access information, to communicate with agents designed by the provider of the information system and to explore an environment with the goal to find information or to find possibilities to enter into some transaction. Hence, we have a community of people interested in theatre, in music, in performers, and their environment has been modeled along the lines of an existing theatre. We need to investigate how we can allow communication between users or visitors of this web-based information and transaction system. Users can help each other to find certain information, they can inform each other (especially when they know about the other's interests), they can have conversations about common interests and they can have domain-related collaboration (e.g., in our case, they can decide to perform a certain play where the actors are distributed among different web sites but sharing the same virtual stage).

It is an important question how to integrate the human visitors of our environment with our models of agent interaction, with our models of multi-modal interaction and multi-media presentation, with models of non-verbal agent behavior (associated with verbal behavior) and with models of agent movements. Hardly any research results are available and no experiments have been performed from which we can learn how humans behave in such agent-rich environments. See [7] for the embedding of our environment in a multi-user environment where users can share scenes, can see each other in the environment and can chat with each other about their interests.

5   Conclusions and Future Research

Presently, our agents are not 'believable', i.e., they have very specific and detailed domain knowledge but they are not able to use some common-sense knowledge concerning family, relatives, environment in which they are, etc. Simple questions concerning these issues should be understood, at least in the sense that it becomes possible to recognize them and to generate an utterance that can at least be considered as an attempt to get the dialogue going. Examples are Jennifer, a virtual saleswoman, and Erin, a virtual bartender, both designed by Extempo Systems Inc. [3]. Jennifer, e.g., has detailed knowledge about cars and racing, but questions on family or country music can be dealt with in a believable way. Jennifer has personality, customers feel confident to talk with her and a company that employs Jennifer will build relationships of affection, trust and loyalty with its customers. The information that is elicited from a customer is stored in marketing and customer databases.

Other issues that need to be studied are user modeling and personalization. Currently, our agents do not employ a model of a user or of user groups. No user model has been 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/her preferences and habits and to the history of interaction with the system. Personalization techniques [1] should be studied and they should be applied to the dynamic generation of presentations in virtual environments and the customization of interaction and presentation styles to different users.

Finally, when considering our specific application domain, we should (now that we have made clear that there are many similarities) study what makes this domain different from others, what properties CBB models for this specific domain have. Similarly, we can look at the usual criteria for selling products via the Internet: technology fit, logistic fit and culture fit. Providing information about cultural events and selling tickets for such events scores high on these 'fitness' scales. There is no need to test the product before buying it, to touch or smell it, to find out how it works and how easy it is to use, etc. Visitors of our theatre will be interested in seeing previews, hearing audio fragments and read additional information about performances, artists and directors, but this is all possible using WWW.

References

1. Ardissono, L. & A. Goy. Tailoring the interaction with users in electronic shops. Proc. *7th Conference on User Modeling*. Banff, Canada, Springer-verlag, 35-44, 1999.
2. Guttman, R.H., A.G. Moukas & P. Maes. Agent-mediated electronic commerce: a survey. *Knowledge Engineering Review*, June 1998.
3. Isbister, K. & B. Hayes-Roth. Social implications of using synthetic characters: An examination of a role-specific intelligent agent. KSL Report No. 98-01, January 1998, Department of Computer Science, Stanford University,
4. Kooijman, D. *Machine en theater. Ontwerpconcepten van winkelgebouwen*. Uitgeverij 010, Rotterdam, 1999.
5. Lombard, M. & T. Ditton. At the heart of it all: The concept of presence. *Journal of Mediated Communication* 3, Nr.2, September 1997.
6. Nijholt, A. & J. Hulstijn. Multimodal Interactions with Agents in Virtual Worlds. In *Future Directions for Intelligent Systems and Information Science*, N. Kasabov (ed.), Physica-Verlag: Studies in Fuzziness and Soft Computing, 1999, to appear.
7. Nijholt, A. & H. Hondorp. Towards communicating agents and avatars in virtual worlds. *EUROGRAPHICS 2000*, A. de Sousa & J.C. Torres (eds.), August 2000, to appear.
8. Soltesz, T. Action space as a mean to establish trust in electronic commerce systems. In: A. Paiva (ed.), *Affect in Interactions: Towards a New Generation of Interfaces.* Proceedings workshop of the I3 Annual Meeting *Community of the Future*, Siena, Italy, October 1999.