

Evaluation of an Autonomous DVD Sales Agent

Christopher Mumme, Niels Pinkwart

Department of Informatics, Clausthal University of Technology,
38678 Clausthal-Zellerfeld, Germany

1 Introduction

Virtual Environments (VEs) can open up new possibilities for eBusiness applications: Using 3D representations, products can be presented realistically and can be arranged next to each other as is done in real shops, making use spatial contexts [1]. Furthermore, in the real world, most shops engage salesclerks to increase the shop turnover because these employees simplify the decisions of buyers. In order to achieve this positive effect in traditional online web shops, some companies like Amazon work with recommender systems. In this paper we investigate whether similar approaches are also possible and effective in 3D shops and if the customers like them. We describe the evaluation of an autonomous virtual sales agent [2] that is able to provide customers in a virtual video store with advice.

2 Study

Study design: In order to evaluate the salesclerk agent, we investigated the following research hypotheses and questions: [H1] The agent gives advice to the human customers in such way that their purchases fit the requirements of the customer better, and [H2] human users accept and follow the recommendations of the virtual agent. [Q1] Do users like the way the agent interacts with them? [Q2] Do they buy a considerable amount of goods that they do not know yet? [Q3] Based on their experience with our agent, can users imagine being counselled by a virtual agent?

To investigate these questions, we conducted a study (36 students took part in) at TU Clausthal. The participants were assigned randomly to groups of sizes between 3 and 5 who then used the virtual shop synchronously. In each study session, each participant was given a profile of a fictitious person. These contained between 4 and 8 pieces of information about which films or genres the persons like and some information about the person (such as age or sex). The participant's task was to buy suitable DVDs for this person as a birthday present. We assigned the participant groups to two study conditions (with and without agent) randomly. Independently of condition, all participants went through two study runs: one with and one without our sales agent being available in the shop. In each run, the participants stayed in the virtual store for 30 minutes (20 to explore and 10 to purchase) and had 30 virtual Euros to purchase. After the two runs, the participants had to fill out a questionnaire.

Results: To examine H1, we evaluated how suitable the purchases of the customers were with respect to the profiles of the fictitious persons they had to buy a

present for. Our goal was to determine the difference between the actual purchase of the customer and the best possible purchase for the fictitious person. To measure the suitability of a DVD for a given profile description, we compared the full person profile information with the DVD set in the store. This way, we calculated the best possible combination of DVDs that could be bought for a given profile and then compared the actual purchases of the participants to this maximum score. For this analysis, we excluded four participants who were not counselled by the agent. Our results are (“with agent” runs in boldface): Run 1 (without agent): 58.2%; Run 2 (with agent): **80.3%**; Run 1 (with agent): **75.8%**; Run 2 (without agent): 61.5%; This results clearly show that in both study conditions, the purchases of the users were much more suitable when they had the virtual agent available in the shop. A paired-samples t-test confirmed that this difference is statistically significant ($p < .001$). H1 is thus confirmed.

To examine H2, we compared the buying behavior of the participants to the recommendations they received from the agent, again excluding the four persons from the analysis who did not interact with the agent. In total, the sales agent recommended 121 DVDs to the 32 customers. 47.1% (57) of these recommendations were followed and the DVDs were bought. Another view on this data: 88 DVDs were bought in total. 64.8% of these films were recommended by the agent before. H2 is thus clearly confirmed: the users frequently followed the recommendations of the virtual agent.

To answer research question Q1, we asked the participants whether they perceived the dialogue with the agent as natural. About 28% (28.2 natural; 21.9 neutral; 50% unnatural) of the participants answered in a positive way. Taking into account the general difficulty of implementing natural language into a computer program and the fact that this was the first time we tested our agent, we consider these numbers as very good and promising results.

To answer research question Q2, we asked the participants in the questionnaire whether they knew the films they bought or. The result was that the 36 participants bought a total of 185 DVDs of which they did not know 47% (87 DVDs). Overall, while the participants bought slightly more known than unknown films, the share of unknown movies is still considerable. Partially, the high amount of known films can be explained by our shop’s range of products. We also asked the participants about their TV consumption in hours per day. The result was an average of two hours per day, i.e. several hundred movies per participant per year. It is therefore realistic to assume that many movies in the shop were already known to the users.

Finally, we were able to answer Q3 with the questionnaire item whether the participants could imagine to be counselled by a virtual agent in a virtual store. Here, 64% of the participants answered with “yes”, supporting the general idea of virtual salesclerk agents based on the prototype they saw.

References

1. Herder, J., Jaensch, K., Horst, B., Novotny, T.: Testmärkte in einer Virtuellen Umgebung. Workshop on AR/VR in der Produktentstehung, pp. 97 -- 110, Germany (2004)
2. Mumme, C., Pinkwart, N., Loll, F.: Design and Implementation of a Virtual Salesclerk. In Proceedings of IVA 2009, 379 – 385. Springer, Heidelberg (2009)