Automated Visual Attention Manipulation (Extended Abstract)*

Tibor Bosse [a] Rianne van Lambalgen [a] Peter-Paul van Maanen [a,b]
Jan Treur [a]

[a] *Department of Artificial Intelligence, Vrije Universiteit Amsterdam, De Boelelaan 1081a, 1081 HV Amsterdam, The Netherlands*
{tbosse, rm.vanlambalgen, treur}@few.vu.nl
[b] *TNO Human Factors, P.O. Box 23, 3769 ZG Soesterberg, The Netherlands*
peter-paul.vanmaanen@tno.nl

In the domain of naval warfare, it is crucial for the crew of the vessels involved to be aware of the situation in the field. Examples of important questions that should be addressed continuously are "in which direction are we heading?", "are we currently under attack?", "are there any friendly vessels around?", and so on. To assess such issues, one of the crew members is usually assigned the Tactical Picture Compilation Task (TPCT): the task to identify and classify all entities in the environment. This is done by monitoring a radar screen for radar contacts, and reasoning with the available information in order to determine the type and intent of the contacts on the screen. However, due to the complex and dynamic nature of the environment, this person has to deal with a large number of tasks in parallel. Often the radar contacts are simply too numerous and dynamic to be adequately monitored by a single human, which compromises the performance of the task.

For these reasons, it may be useful to offer the human some support from an intelligent ambient system, consisting of software agents that assist him in the execution of the Tactical Picture Compilation Task. For example, in case the human is directing its attention on the left part of a radar screen, but ignores an important contact that just entered the radar screen from the right, such an agent may alert him about the arrival of that new contact. To be able to provide this kind of intelligent support, the system somehow needs to maintain a model of the cognitive state of the human: in this case the humans focus of attention. It should have the capability to attribute mental, and in particular attentional states to the human, and to reason about these. In psychology and philosophy this characteristic is often referred to as Theory of Mind (or ToM). According to [2], agents, both human and software, can exploit a Theory of Mind for two purposes: to anticipate the behaviour of other agents (e.g., preparing for the consequences of certain actions that the other will probably perform), and to manipulate it (e.g., trying to influence the actions that the other will perform). In case of an intelligent system to support naval crew members, both purposes are relevant, but require a different type of support. This study is related to the latter type, the type that tries to manipulate the focus of attention.

A number of approaches in the literature address the development of software agents with a Theory of Mind. Usually, such agents maintain, in one way or the other, a model of the epistemic (e.g., beliefs) and/or motivational states (e.g., desires, intentions) of other agents. However, for the situation sketched above, such agents ideally also have insight in another agents attentional states. After all, if a supportive agent is to find out whether the human is ignoring some contact, it needs to have some knowledge about which contacts the person is paying attention to. This idea is in line with the theories of cognitive scientists like Gärdenfors, who claims that humans have a Theory of Mind that is not only about beliefs, desires, and intentions, but also about other mental states like attentional, emotional, and awareness states.

The current paper is the result of a project that aims to develop intelligent agents to support naval crew members in the Tactical Picture Compilation Task, based on the ideas described above. To this end, four models have been developed. First, a dynamical model of human attention is needed, which estimates where the persons attention is, based on information about features of objects on the screen and the persons

*This paper is an extended abstract of [3].

gaze. Second, a reasoning model is needed to reason through the first model in order to generate beliefs on attentional states at any point in time. Third, a model is needed that compares the output of the second model with some normative attention distribution, and determines whether there is a discrepancy. Finally, a model is needed that uses the output of the third model to determine how to alert the human that he is ignoring something important. An initial version of the first two models has already been developed and were adopted from this earlier work ([4], [5] and [1], respectively). The current paper has its focus on the development of the other two models.

Concerning future work, an important challenge would be to perform a more elaborated validation of the supportive system. This can be done is several steps. First, to obtain more data, the experiment introduced in this paper will be performed with a larger number of participants. The resulting data can then be used to check (possibly using automated analysis tools) whether the supporting agent is successful in various situations. As part of this validation, also different strategies and parameter settings will be tested. For example, does adapting the shape of an object provide better results than adapting its luminance, or adapting multiple features? Similarly, in addition to manipulation of bottom-up attention, is it useful to manipulate top-down attention as well? Furthermore, in a later stage of the project, it is planned to evaluate whether the software agent indeed improves the task performance of the user.

Ackowledgments

This research was partly funded by the Royal Netherlands Navy (program number V524).

References

[1] T. Bosse, F. Both, C. Gerritsen, M. Hoogendoorn, and J. Treur. Model-based reasoning methods within an ambient intelligent agent model. In M. Mühlhäuser, A. Ferscha, and E. Aitenbichler, editors, *Proceedings of the First International Workshop on Human Aspects in Ambient Intelligence, Constructing Ambient Intelligence: AmI 2007 Workshops Proceedings*, LNCS. Springer-Verlag Berlin Heidelberg, 2009.

[2] T. Bosse, Z.A. Memon, and J. Treur. A two-level bdi-agent model for theory of mind and its use in social manipulation. In *Proceedings of the AISB 2007 Workshop on Mindful Environments*, pages 335–342, 2007.

[3] T. Bosse, R. van Lambalgen, P.-P. van Maanen, and J. Treur. Automated visual attention manipulation. In L. Paletta and J.K. Tsotsos, editors, *Proceedings of WAPCV 2008*, volume 5395 of *LNAI*, pages 262–277. Springer-Verlag Berlin Heidelberg, 2009.

[4] T. Bosse, P.-P. van Maanen, and J. Treur. A cognitive model for visual attention and its application. In T. Nishida, editor, *Proceedings of the 2006 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT-06)*, pages 255–262, Hong Kong, P.R. China, 2006. IEEE Computer Society Press.

[5] T. Bosse, P.-P. van Maanen, and J. Treur. Simulation and formal analysis of visual attention in cognitive systems. In L. Paletta and E. Rome, editors, *Proceedings of WAPCV 2007*, volume 4840 of *LNAI*, pages 463–480. Springer-Verlag Berlin Heidelberg, 2007.