Augmented Metacognition Addressing Dynamic Allocation of Tasks Requiring Visual Attention

Tibor Bosse[1], Willem van Doesburg[2], Peter-Paul van Maanen[1,2], and Jan Treur[1]

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

Abstract. This paper discusses the use of cognitive models as augmented metacognition on task allocation for tasks requiring visual attention. In the domain of naval warfare, the complex and dynamic nature of the environment makes that one has to deal with a large number of tasks in parallel. Therefore, humans are often supported by software agents that take over part of these tasks. However, a problem is how to determine an appropriate allocation of tasks. Due to the rapidly changing environment, such a work division cannot be fixed beforehand: dynamic task allocation at runtime is needed. Unfortunately, in alarming situations the human does not have the time for this coordination. Therefore system-triggered dynamic task allocation is desirable. The paper discusses the possibilities of such a system for tasks requiring visual attention.

Keywords: Visual attention, cognitive modeling, augmented metacognition.

1 Introduction

The term *augmented cognition* [6, 10] was used by Eric Horvitz at the ISAT Woods Hole meeting in the summer of 2000 to define a potentially fruitful endeavor of research that would explore opportunities for developing principles and computational systems that support and extend human cognition by taking into explicit consideration well-characterized limitations in human cognition, spanning attention, memory, problem solving, and decision making. This paper focuses on extending human cognition by the development of principles and computational systems addressing task allocation of tasks requiring visual attention. In previous work [2], cognitive models of visual attention were part of the design of a software agent that supports a naval warfare officer in its task to compile a tactical picture of the situation in the field. In the domain of naval warfare, the complex and dynamic nature of the environment makes that the warfare officer has to deal with a large number of tasks in parallel. Therefore, in practice, (s)he is often supported by software agents that take over part of these tasks. However, a problem is how to determine an appropriate allocation of tasks: due to the rapidly changing environment, such a work division cannot be fixed beforehand [1]. Task allocation has to take place at runtime, dynamically. For this purpose, two approaches exist, i.e. *human-triggered* and *system-triggered* dynamic task allocation [3]. In the former case, the user can decide up to what level the software agent should assist her. But especially in alarming situations the user does not have the time to think about such task allocation [7]. In these situations it would be better if a software agent augments the user's metacognitive capabilities by means of system-triggered dynamic task allocation. This paper discusses the usage of cognitive models of visual attention that can be incorporated within assisting software agents offering augmented metacognition in order to obtain such a system-triggered dynamic task allocation.

In Section 2 a further elaboration on the motivational background for augmented metacognition is given. In Section 3 a generic design of augmented metacogntion based on cognitive models of visual attention is described. In Section 4 some applications of the framework are introduced and discussed. The paper is concluded with a general discussion and some future research.

2 Augmented Metacognition: Motivational Background

Support of humans in critical tasks may involve a number of aspects. First, a software agent can have knowledge about the task or some of its subtasks and, based on this knowledge, contribute to task execution. Usually, performing this will also require that the software agent has knowledge about the environment. This situation can be interpreted as a specific form of augmented cognition: *task-content-focused augmented cognition*. This means that the cognitive capabilities to do the task partly reside within the software agent, external to the human, and may extend the human's cognitive capabilities and limitations. For example, if incoming signals require a very fast but relatively simple response, in speed beyond the cognitive capabilities of a human, a software agent can contribute to this task, thus augmenting the human's limited reaction capabilities. Another example is handling many incoming stimuli at the same time, which also may easily be beyond human capabilities, whereas a software agent can take care of it.

If the software agent provides task-content-focused augmented cognition, like in the above two examples, it may not have any knowledge about the coordination of the subtasks and the process of cooperation with the human. For example, task allocation may completely reside at the human's side. However, as discussed in the introduction, when the human is occupied with a highly demanding task, the aspect of coordination may easily slip away. For example, while working under time pressure, humans tend to spend less attention to reflection on their functioning. If the software agent detects and adapts to those situations it will have a beneficial effect; e.g., [8]. This type of reflection is a form of metacognition: cognitive processes addressing other cognitive processes. A specific type of support of a human from an augmented cognition perspective can also address such reflective aspects: *augmented metacognition*. This is the form of augmented cognition that, in contrast to task-content-focused augmented cognition, addresses the support or augmentation of a human's limitations in metacognitive capabilities. The type of augmented metacognition discussed in this paper focuses on dynamic task allocation.

Augmented metacognition can be provided by the same software agent that provides task-content-focused augmented cognition, or by a second software agent that specialises on metacognition, for example the task allocation task. The former case results in a reflective software agent that has two levels of internal processing: it can reason both about the task content (object-level process) and about the task coordination (meta-level process); e.g., [9]. The latter case amounts to a specific case of a reflective multi-agent system: a multi-agent system in which some of the agents process at the object level and others at the meta-level.

The distinction made between task-content-focused augmented cognition and augmented metacognition provides a designer with indications for structuring a design in a transparent manner, either by the multi-agent system design, or by the design of a reflective software agent's internal structure. This paper focuses on the latter, the design of the reflective internal structure of the software agent. An implementation of such an agent has been evaluated for two case studies.

3 Augmented Metacognition Design

In this section, first the generic design of the proposed augmented metacognition is presented in Section 3.1. After that, Section 3.2 describes how principles of Signal Detection Theory (SDT) can be applied within this design.

3.1 Prescriptive and Descriptive Models

The present design is based on the idea that the software agent's internal structure augments the user's metacognitive capabilities. This structure is composed of two maintained models of the user's attention. The first is called a *descriptive model*, which os a model that estimates the user's actual attentional dynamics. The second is called a *prescriptive model*, which prescribes the way these dynamics should be. In Figure 1 a conceptual design of such a software agent is shown. Depending on the user's and the agent's own attentional levels, the agent decides whether the user (or the agent itself) is paying enough attention to the right tasks at the right time. This is determined by checking whether the difference between described attention and prescribed attention is below a certain threshold. In Figure 1 this comparison is depicted in the middle as the *compare* process. Based on this, the agent either adapts its support or it does not, i.e. the *adapt* process in Figure 1.

From the perspective of the agent, the runtime decision whether to allocate a task to itself or to the user comes down to the decision whether to support this task or not. The question remains what the agent could use as a basis for deciding to take over responsibility of a task, i.e. by exceedance of a certain threshold, using both descriptive and prescriptive models of user attention. An answer to this is that the agent's decision to support can be based on several performance indications: (PI1) a performance indication of the user concerning her ability to appropriately allocate attention (to the right tasks at the right time), (PI2) a performance indication of the agent concerning its ability to soundly prescribe the allocation of attention to tasks, (PI3) a performance indication of the system concerning its ability to soundly describe the user dynamics of the allocation of attention to tasks, and (PI4) a performance indication of the agent concerning its ability to soundly decide to support the user in her task to allocate attention to tasks.

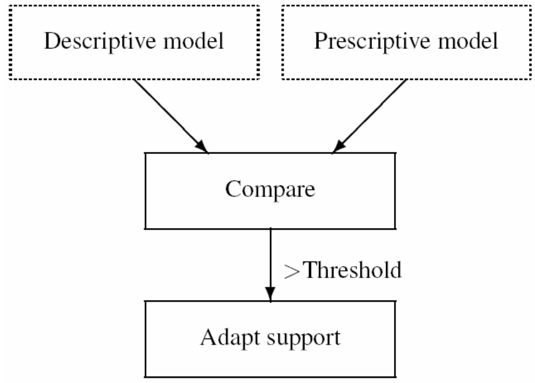

Fig. 1. Conceptual model of the attention allocation system.

3.2 Some Principles of SDT

On of the ways to let the agent estimate the performances of the user and the agent itself from the previous paragraph is by using the principles of *Signal Detection Theory*, or simply SDT [5]. In this subsection a theoretical framework based on SDT is defined, including a method that constitutes a means for identifying when to trigger attention allocation support.

To let a software agent reason about the performance of the user concerning her ability to appropriately allocate attention (PI1), a formal framework in SDT terms is needed in which it can describe it. These terms are mainly based on a mathematical description of the following situations:

1) The descriptive model of user attention indicates that attention is paid (A) to the tasks that are required by the prescriptive model (R). This situation is also called a *hit* (HIT).
2) The descriptive model of user attention indicates that attention is not paid (not A) to the tasks that are not required by the prescriptive model (not R). This situation is also called a *correct rejection* (CR).
3) The descriptive model of user attention indicates that attention is paid (A) to the tasks that are not required by the prescriptive model (not R). This situation is also called a *false alarm* (FA).
4) The descriptive model of user attention indicates that attention is not paid (not A) to the tasks that are required by the prescriptive model (R). This situation is also called a *miss* (MISS).

The task to discriminate the above situations can be set out in a table as a 2-class classification task. The specific rates of HITs, FAs, MISSs, and CRs, are calculated by means of probabilities of the form P(X|Y), where X is the estimate of certain behaviour and Y is the estimate of the type of situation at hand. The descriptive and prescriptive models mentioned earlier can be seen as the user's attentional behaviour (A or not A) in a specific situation that either requires attention (R) or does not (not R). A HIT, for example, would be in this case P(A|R), and a FA would be P(A|not R), etc. This classification task is shown in Table 1. A similar task can be defined for the other performance indicators, i.e. PI2, PI3, and PI4.

Table 1. A 2-class classification task based on a descriptive (attention allocated) and prescriptive (attention required) model of user attention.

		Attention required?	
		Yes	No
Attention allocated?	Yes	HIT = P(A | R)	FA = P(A | not R)
	No	MISS = P(not A | R)	CR = P(not A | not R)

In SDT, the measure of sensitivity (d') is commonly used as an indicator for various kinds of performances. The measure is a means to compare two models, in this case descriptive and prescriptive models. Hence the calculation of such sensitivities can be used by the agent to determine whether to support the user or not. For instance. low sensitivities (< threshold) may result in the decision to adapt support. The calculation of sensitivity in terms of the above mentioned HIT, FA, MISS, and CR, can be done by using the following formula:

d' = HIT – FA = CR – MISS

As can be seen in the formula, to calculate sensitivity, the measurement of HIT and FA are sufficient. No estimates of CR or MISS are needed, since HIT – FA is equal to CR – MISS.[1] Furthermore, sensitivity is dependent on both HIT and FA, rather than on HIT or FA alone. A user that has a high sensitivity as a result of attending to all tasks all the time (high HIT rate), is not only impossible due to the maximum capacity of human attention, but also very inefficient. Think of the very limited attention each task probably will receive due to unneeded FAs. The other way around, a low FA rate as a result of attending to nothing, is obviously not desired as well.

4 Applications

This section discusses two applications of the presented framework for task allocation based on visual attention. In Section 4.1, a pilot study is described, of which the main aim was to establish a (descriptive) model of a person's visual attention in the execution of a simple task in the warfare domain. For this pilot study, a simplified version of an Air Traffic Control (ATC) task was used. Next, Section 4.2 addresses a

[1] This is due to the fact that HIT = 1 – MISS and therefore a high HIT results in a low MISS, and vice versa (the same holds for CR = 1 – FA).

more realistic case: the task of Tactical Picture Compilation (TPC) by a naval warfare officer. For both cases, it is explained how descriptive models of visual attention may be used for task allocation, using the design introduced in Section 3.

4.1 Multitask

In order to test the ideas presented in the previous sections, a pilot study has been performed. The setup of this pilot study consisted of a human participant executing a simple warfare officer-like task [2]. To create such a setup, the software Multitask [4] was used (and slightly altered in order to have it output the proper data). Multitask was originally meant to be a low fidelity ATC simulation. In this study, it is considered to be an abstraction of the cognitive tasks concerning the compilation of the tactical picture, i.e. a warfare officer-like task. A screenshot of the task is shown in Figure 2.

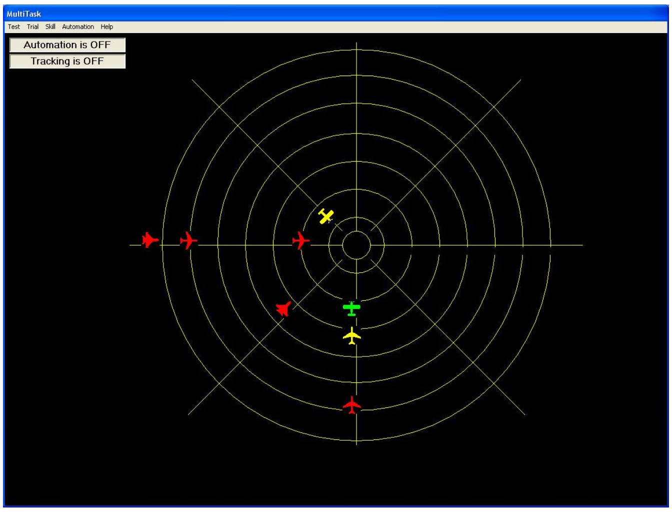

Fig. 2. The interface of the used environment based on MultiTask [4].

In the pilot case study, the participant (controller) had to manage an airspace by identifying aircrafts that all are approaching the centre of a radarscope. The centre contained a high value unit (HVU) that had to be protected. In order to do this, airplanes needed to be cleared and identified to be either hostile or friendly to the HVU. The participant had to click on the aircraft according to a particular procedure depending on the status of the aircraft. Within the conducted pilot study, three different aircraft types were used, which resulted in different intervals of speed of the aircrafts. The above dynamic properties of the environment were stimuli that resulted in constant change of the participant's attention. The data that were collected consist of all locations, distances from the centre, speeds, types, and states (i.e., colours).

Additionally, data from a Tobii x50 eye-tracker[2] were extracted while the participant was executing the task. All data were retrieved several times per second (10-50 Hz).

Based on such data, a cognitive model has been implemented that estimates the distribution of the user's attention over the locations of the screen at any moment during the experiment [2]. This model uses two types of input, i.e., *user-input* and *context-input*. The user-input is provided by the eye-tracker, and consists of the (x, y)-coordinates of the gaze of the user over time. The context-input is provided by the Multitask environment, and consists of the variables speed, distance to the centre, type of aircraft, and aircraft status. The output of the model is represented in the form of an dynamically changing 3D image. An example screenshot of this is shown in Figure 3 at an arbitrary time point.[3] The x- and y-axis denote the x- and y-coordinates of the grid, and the z-axis denotes the level of attention. In addition, the locations of all tracks, the status of the tracks, the location of the gaze, and the mouse clicks are indicated in the figure by small dots, colour, a star, and a big dot, respectively. Figure 3 clearly shows that at this time point there are two peaks of attention (locations (12,10) and (16,9)). Moreover, a mouse click is performed at location (16,9), and the gaze of the subject is also directed towards that location.

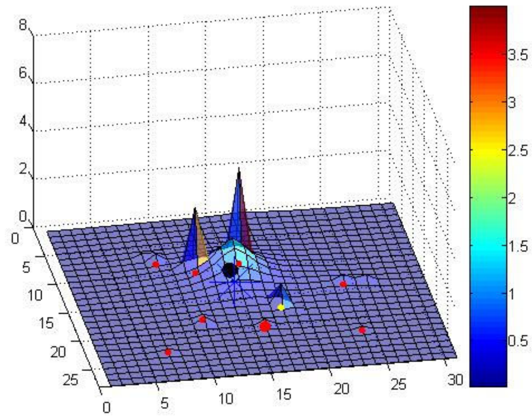

Fig. 3. Example Output of the Cognitive Model of Visual Attention [2].

In terms of Section 3, the presented model is a descriptive model of the task. If, in addition to this a prescriptive model is created, both models can be used for dynamic task allocation, using the principles of Signal Detection Theory. Hence the presented model of visual attention can be used for augmented metacognition purposes: the system maintains a cognitive model of the attentional dynamics of an user, and accordingly, extends the user's metacognitive capabilites. By introducing a threshold, a binary decision mechanism can be established, which decides for each location whether it receives (enough) attention or not ("A" or "not A" in Table 1). The idea is to use such a mechanism for dynamic task allocation for the type of tasks in the naval

[2] http://www.tobii.se.

[3] See http://www.few.vu.nl/~pp/attention for a complete animation.

domain as considered in this paper. For example, in case an user is already allocated to some task, it may be better to leave that task for him or her, and allocate tasks to the system for which there is less or no commitment from the user (yet).

4.2 Tactical Picture Compilation Simulator

The characterizations of different attentional states in relation with adaptive task allocation was investigated in another case study, namely one in the naval surface warfare domain. In Figure 4 a snapshot of the interface of the Tactical Picture Compilation (TPC) Simulator, used in this study, is shown. Similar to the case study in the previous section, this study was also conducted in an effort to augment the metacognitive capabilities of the user in naval operations in naval operations and to leverage cognitive models of attention. However, the present study focused on a more realistic domain: its goal was to establish an implicit work division between a naval officer and a supportive agent based on a cognitive model of the TPC task. TPC is a critical task in naval surface warfare. It is continuously performed by naval warfare officers during operations at sea. The main goal is to create an accurate tactical picture of the immediate surroundings of the ship using the ship's sensors. Changes in the tactical picture can occur either extremely slow or very rapidly depending on the traffic density at sea.

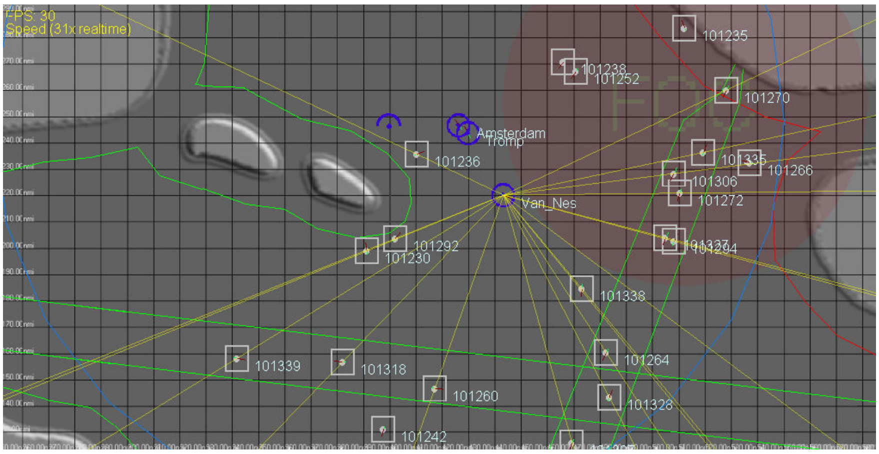

Fig. 4. Tactical picture compilation simulation that was used to implement a dynamic task allocator based on a cognitive model of attention.

The application contained a TPC agent that supported the officer's interpretation of the behaviour of ships in the area. The officer observes the behaviour of the ships in the area over time in order to deduce if they have hostile intent or not. The attentional capacity of the officer may come under stress during transitional periods from low traffic densities to high traffic densities or during intense tactical action. During these periods of high cognitive workload, the officer is supported by the TPC agent. The question is how to determine the optimal division of workload without adding to the workload of the officer by letting him or her decide what the division should be. Instead, the main region of interest on the radar screen is determined by using a cognitive model of attention. The model is similar to the one used for the study mentioned in Section 4.1. It deduces the region of interest based on an interpretation of eye gaze behaviour of the officer as well as information about the properties of the objects and spaces visible on the radar screen. The various symbols and lines are not equally visually salient and the model must correct for this. The regions that are not covered by the modeled attention of the officer are then assigned to the TPC agent for processing. The officer therefore *implicitly* communicates the metacognitive decision which radar tracks are desired to be handled by the TPC agent.

This implicit communication is based on the assumption that the user's cognitive performance in determining threat is better than that of the TPC agent. This means that no prescriptive variant of the TPC model was used to determine the work division between the TPC agent and the officer. A prescriptive model of the TPC task, as described in Section 3, may be used to inform the work division decision, by checking if the officer pays attention to the radar tracks that represent the greatest threat. If not, then the optimal work division may need to be re-assessed. The TPC agent will automatically be assigned those tracks that fall outside the scope of attention of the officer. In future versions of this application, a prescriptive model can be implemented to enhance performance in situations where threat is ambiguous or situations that contains multiple threats from different directions. In those situations, the most optimal work division may consist of the officer covering one threat and the TPC agent covering another, instead of just keeping track of nominal radar tracks. A prescriptive TPC model is then required to detect threats outside the scope of attention of the officer.

This form of decision support is implicit in the sense that there is no explicit communication between officer and agent about the decisions supported. Both the work division decision and task related decisions happen automatically. An interesting question is to determine whether the acceptance of this form of decision support by naval officers can be enhanced if the system communicates to the officer the reasons on which it bases its decisions. This might cause the officer to trust the system more, because the reasons on which decisions are made inspire confidence. On the other hand, situations might develop in which the officer generates expectations about the decisions made by the supporting agent. This might lead to a form of paranoia in which the officer is distracted from the main task (TPC) because of the desire to check the decisions of the supporting agent.

5 Discussion

The Augmented Cognition International Society defines augmented cognition as 'an emerging field of science that seeks to extend a user's abilities via computational technologies, which are explicitly designed to address bottlenecks, limitations, and biases in cognition and to improve decision making capabilities.' The Society also formulated a goal: '.. to develop computational methods and neurotech tools that can account for and accommodate information processing bottlenecks inherent in human-system interaction (e.g., limitations in attention, memory, learning, comprehension, visualization abilities, and decision making).' Augmented cognition is a wide area, that is applicable to various types of cognitive processes. As the area develops further, it may be useful to differentiate the field a bit more, for example, by distinguishing more specific classes of application.

In this paper, such a distinction is put forward: augmented cognition focusing on task content versus augmented cognition focusing on task coordination. As the latter is considered a form of metacognition, this suggests augmented metacognition as an interesting subarea of augmented cognition. The paper discussed applications to the metacognition used for dynamic task allocation within this area. It has been pointed out how functioning of human-computer systems can be improved by incorporating augmented metacognition in them. Especially in tasks involving multiple stimuli that require fast responses, this concept is expected to provide a substantial gain in effectiveness of the combined system.

Acknowledgments. This research was partly funded by the Royal Netherlands Navy under program number V524.

References

1. Bainbridge, L.: Ironies of automation, *Automatica* 19 (1983) 775-779
2. Bosse, T., Maanen, P.-P. van, and Treur, J.: A Cognitive Model for Visual Attention and its Application, In: *Proceedings of the 2006 IEEE/WIC/ACM Int. Conf. on Intelligent Agent Technology (IAT-06)*, IEEE Computer Society Press, Hong Kong (2006)
3. Campbell, G., Cannon-Bowers, J., Glenn, F., Zachary, W., Laughery, R., and Klein, G.: *Dynamic function allocation in the SC-21 Manning Initiative Program*, Naval Air Warfare Center Training Systems Division, Orlando, SC-21/ONRS&T Manning Affordability Initiative (1997)
4. Clamann, M. P., Wright, M. C. and Kaber, D. B.: Comparison of performance effects of adaptive automation applied to various stages of human-machine system information processing, In: *Proceedings of the 46th Ann. Meeting of the Human Factors and Ergonomics Soc.* (2002) 342-346
5. Green D. and Swets J: *Signal detection: theory and psychophysics,* Wiley: New York (1966)
6. Horvitz, E., Pavel, M., and Schmorrow, D. D.: Foundations of Augmented Cognition. Washington, DC, National Academy of Sciences (2001)
7. Inagaki, T.: Adaptive automation: Sharing and trading of control, *Handbook of Cognitive Task Design* (2003) 147-169
8. Kaber, D.B. and Endsley, M.R.: The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task, *Theoretical Issues in Ergonomics Science* 5(2) (2004) 113-153
9. Maes, P. and Nardi, D.: *Meta-level Architectures and Reflection*, North-Holland, Amsterdam (1988)
10. Schmorrow, D. D. and Kruse, A. A.: Augmented Cognition. In W.S. Bainbridge (Ed.), *Berkshire Encyclopedia of Human-Computer Interaction.* Great Barrington, MA: Berkshire Publishing Group (2004) 54-99