Attention Aware Systems

Claudia Roda

American University of Paris

Computer Science Department

147, rue de Grenelle

75007 Paris

c.roda@ac.aup.fr

Julie Thomas

American University of Paris

International Communication Department

6, rue du Colonel Combes

75007 Paris

Thomas@aup.fr

Attention Aware Systems

ABSTRACT

Attention-aware systems (AAS) are systems capable of adapting to, and supporting, human attentional processes especially in situations of multi-tasking, frequent interactions with other users, and highly dynamic environments. In this paper we discuss the rationale for AASs and their role within current HCI research, we review current research and applications of AASs, and we highlight the issues that will need to be addressed in the future for their design.

KEYWORDS

Attention aware systems, adaptive systems, notification systems, attentive user interfaces, transparent systems, visual attention, tracking.

INTRODUCTION

Much information science research has focussed on the design of systems enabling users to access, communicate, and use information quickly and efficiently. However the users' ability to exploit this information is seriously limited by finite human cognitive resources. In cognitive psychology the role of attentional processes in allocating cognitive resources has been demonstrated to be crucial. Attention is often defined as the set of processes guiding the selection of the environmental stimuli to be attended. Access to information therefore is not only regulated by its availability but also by the users' choice to attend the information - this choice being governed by attentional processes. Recently several researchers and practitioners in Human Computer Interaction (HCI) have concentrated on the design of systems capable of adapting to, and supporting, human attentional processes. These systems, that often rely on very different technologies and theories, and that are designed for a range of applications, are called *attention-aware* systems (AAS). In the literature these systems have also been referred to as *Attentive User Interfaces* (Vertegaal, 2003). However, we prefer using the former name as it stresses the fact that issues related to attention are relevant to the design of the system as a whole rather than limited to the interface. The recent interest in this field is testified by the publication of special issues in academic journals (e.g. Communication of the ACM 46(3) 2003, International Journal of Human-Computer Studies 58(5) 2003) and by the organisation of specialised fora of discussion (e.g. the workshop on "Designing for Attention" (Roda & Thomas, 2004)).

In this paper we discuss the rationale for AASs and their role within current HCI research, we briefly review current research in AASs, and we highlight some open questions for their design.

BACKGROUND - RATIONALE FOR AND ROLE OF ATTENTION-AWARE SYSTEMS

In this section we analyze the rationale for AASs and we discuss their role in HCI research.

Why attention-aware systems

Information overload is one of the most often mentioned problems of working, studying, playing, and generally living in a networked society. One of the consequences of information overload is the fast shift of attention from one subject to another or one activity to another. In certain situations the ability to quickly access several information sources, to switch activities, or to change context is advantageous. In other situations it would be more fruitful to create and maintain a focus whilst offering the possibility to switch attention to other contents or activities only as a background low-noise open choice. System awareness about the cost/benefits of attentional shifts with respect to the user's goals is essential in environments where (1) attentional switches are very often solicited, or (2) where the users' lack of experience with the environment makes it harder for them to select the appropriate attentional focus, or (3) where an inappropriate selection of attentional focus may cause serious damage to the system, its users, or third parties. Systems relying highly on multi-user interaction, such as virtual communities and certain systems supporting cooperative work, are examples of environments where attentional switches are often solicited. Online educational systems are example of environments where the lack of knowledge and experience of users with the subject at hand makes it harder for them to select the appropriate attentional focus and may easily cause a loss of focus. Life critical systems are examples of environments where an inappropriate selection of attentional focus may cause serious damage. The need for *AAS*s is quite widespread especially if one considers that assessing, supporting, and maintaining users' attention may be desirable in other environments such as entertainment and e-commerce.

Attention-aware systems in HCI research

A large portion of research on human attention in digital environments is based on the findings of cognitive psychology. For example Raskin (2000) analyses how single locus of attention, and habit formation have important consequences on human ability to interact with computers. He proposes that habit creation is a mechanism that can be used to shift the focus of users from the interface to the specific target task.

The study above follows the classic "direct manipulation" school (Shneiderman, 1992; Shneiderman, 1997) which aims at supporting the attentional choices of the user by making the device "transparent" so that the user can focus on the task rather than on the interface. The wide range of systems designed with this aim is often referred to as *transparent systems*, a term also employed in ubiquitous computing (Abowd, 1999; Weiser, 1991).

Another area of research focuses instead on designing interfaces and systems capable of guiding the users in the choice of attentional focus. The system is seen as proactive, visible, and capable of supporting the users in their choices. This type of systems are often designed as artificial agents (Bradshaw, 1997; Huhns & Singh, 1997) acting as proactive helpers for the user (Maes, 1994; Negroponte, 1997) and they are frequently referred to as *proactive/adaptive systems*.

The two approaches are often regarded as divergent: (1) responding to different needs and (2) requiring different design choices. However this is not necessarily the case, as it should become apparent from the following discussion of these two alleged differences on users' needs and design choices. Concerning the ability to respond to user needs, consider for example, one of the metaphors most often used for proactive systems: Negroponte's English butler (Negroponte, 1995). "The best metaphor I can conceive of for a human-computer interface is that of a well-trained English butler. The "agent" answers the phone, recognizes the callers, disturbs you when appropriate, and may even tell a white lie on your behalf. The same agent is well trained in timing, versed in finding the opportune moments, and respectful of idiosyncrasies. People who know the butler enjoy considerable advantage over a total stranger. That is just fine." (Ibid. p.150). Isn't this proactive/adaptive system an exquisite example of transparent system? The English butler certainly knows to disappear when it is the case, but he is there when required and he is capable of proactive behavior such as selecting the calls you may want to receive or even telling a joke if appropriate! Concerning the design choices a few considerations should be made. First of all, any system needs to be proactive in certain situations (e.g. reporting errors) and transparent in others. Secondly, certain applications, in particular those where the user has a good knowledge of the most effective attentional focus, require mostly transparent interfaces, whilst certain others, where the user is more in need of guidance, require more proactive interfaces. Also the user's needs, the system's functionality, and the use that is made of the system, may change with time. Therefore it may be desirable for a system, that is initially very proactive, to slowly become transparent, or vice-versa. Finally, applications exist where the user is expected to focus on the system/interface itself, e.g. digital art. As a consequence, just as proactive adaptive behaviors may not always be desirable, transparency itself may, under certain conditions, not be desirable.

This brings us to another reason for studies related to AASs. In the last two decades there has been a shift on the use and market of Information and Communication Technologies (ICT) from strictly task oriented (work related) to more a pervasive personal and social, use of these technologies. Performing a task or achieving a goal may not be the main target of the user who instead may turn to ICT artifacts for their symbolic or affective value, entertainment, or pleasure in general - see for example Lowgren's arguments for Interactive Design versus classic HCI in (Löwgren, 2002). Capturing and maintaining user attention may then actually be the ultimate goal of the system.

The real challenge of modern interface design is therefore at the meta-level. We should not aim at designing transparent or proactive systems. Rather we should aim at designing systems capable of reasoning about users' attention, and consequently decide how best to disappear or to gain and guide user's attention. Focusing on attentional mechanisms also provides a framework that reconciles the direct manipulation user interfaces approach and the interface agents approach as clearly presented and exemplified by Horvitz (1999).

HUMAN ATTENTION AND SYSTEMS CAPABLE OF SUPPORTING IT

This section briefly reviews the work done so far in AASs - for a more extensive review see (Roda & Thomas, 2005). It should be noted that attention has not often been prioritised as a specific subject of research in HCI (with some notable exceptions including the *Attentional User Interface* project at Microsoft research (Horvitz et al., 2003)). As a consequence, much of the work relevant to the development of AASs appears in the context of other research frames. This is especially the case as attention processes are related to, and necessary for, the successful accomplishment of many diverse activities.

Human attention has been widely researched in cognitive psychology and, more recently, in neuropsychology. Although there is no common agreement on a definition of "attention", attention is generally understood as the set of processes allowing humans to cope with the, otherwise overwhelming, stimuli in the environment. Attention therefore refers to the set of processes by which we select information (Driver, 2001; Uttal, 2000). These processes are mainly of two types: endogenous - i.e. guided by volition - and exogenous - i.e. guided by reaction to external stimuli. Given this view of attention as selection of external stimuli it is obvious that attention is somehow related to human sensory mechanisms. *Visual attention*, for example, has been widely studied in cognitive psychology and it is particularly relevant to HCI since the current predominant modality for computer-to-human communication is screen display. Using the results of psychological studies in visual attention, some authors have proposed visual techniques for notification displays that aim at easy detection whilst minimising distraction (Bartram et al., 2003). Attention on modalities other than visual, as well as attention across modalities, have not been investigated to the same extent as visual attention. However, Bearne and his colleagues (Bearne et al., 1994) propose guidelines for the design of multimedia systems grounded in attentional mechanisms.

Systems capable of supporting and guiding user attention must, in general, be able to: (1) assess the current user focus, (2) and make predictions on the cost/benefits of attention shifts (interruptions). We conclude this section with a review of the work done so far in these two directions.

Several sensory-based mechanisms for the detection of users' attention have been employed, including *gaze tracking* (Hyrskykari et al., 2000; Vertegaal, 1999; Zhai, 2003), gesture tracking (Hinckley et al., 2000), head pose and acoustic tracking (Stiefelhagen, 2002). Horvitz and his colleagues (Horvitz et al., 2003) propose that sensory-based mechanisms could be integrated with other cues about the current user's focus. These cues could be extracted from users' scheduled activities (e.g. using online calendars), users' interaction with software and devices, and information about the users and their patterns of activity and attention. In any case, even when employing mechanisms capable of taking into account all these cues, a certain level of uncertainty about users' focus, activities, goals, and best future actions will always remain and will have to be dealt with within the system (Horvitz et al., 2003).

The problem of evaluating the cost/benefit of *interruptions* has been researched mostly in the context of *notification systems* (Brush et al., 2001; Carroll et al., 2003; Czerwinski et al., 2000; Hudson et al., 2003; McCrickard et al., 2003b; McCrickard et al., 2003c). This research aims at defining the factors determining the likely utility of a given information, for a given user, in a given context and the costs associated with presenting the information in a certain manner, to the user, in that context. McCrickard and Chewar (2003) integrate much of the research in this direction and propose an attention-utility trade-off model.

FUTURE TRENDS

AASs will be crucial for the development of applications in a wide variety of domains including education, life critical systems (e.g. air traffic control), support to monitor and diagnosis, knowledge management, simulation of human-like characters, games, and e-commerce. In order to unleash the whole potential of these systems however, there are many fundamental aspects of attention, of the mechanisms that humans use to manage it, and of their application in digital environments that require further exploration. As it will result obvious from the description below, this exploration would greatly benefit from a more interdisciplinary approach to the design of AASs.

First, although a very significant amount of research on human attention has been undertaken in psychology, several HCI researchers agree that the reported theories are often too far removed from the specific issues relevant to human computer interaction to be easily applied to this field of research (McCrickard et al., 2003c) and that more focussed research in this direction is needed (Horvitz et al., 2003).

A second important issue in the design of AASs is the definition of parameters against which, one could measure their efficiency. In their work on notification systems, McCrickard and his colleagues (McCrickard et al., 2003a) advance a proposal in this direction; however further discussion is needed in order to achieve an agreement on parameters that are generally accepted.

Third, although the visual modality has been extensively researched in cognitive psychology and HCI, this work is mostly focussed on still images. How would the principles apply to moving images?

Fourth, much work remains to be done on modalities other than visual. In particular, research on attention in speech (from phonetics to semantics and rhetoric) (Argyle & Cook, 1976; Clark, 1996; Grosz & Sidner, 1990) could be fruitfully applied to HCI research in AASs. Distribution of attention over several modalities is a field that also deserves further research.

Fifth, most of the work on the evaluation of the cost/benefits of interruptions has been done taking the point of view of the user being interrupted; such analysis however, should also take into account the cost/benefit to the interrupter, and the joint cost/benefit (Hudson et al., 2002; O'Conaill & Frohlich, 1995).

Sixth, certain aspects of human attention related to social and aesthetic processes have been largely disregarded in current research. How could these processes be taken into consideration? Furthermore, most of the target applications in AASs assume that the user is in a 'work" / task-oriented situation. How would AAS design apply to different situations (play, entertainment)?

CONCLUSION

AASs are systems capable of reasoning about user attention. In a task-oriented environment such systems address the problem of information overload by striving to select and present information in a manner that optimizes the cost/benefit associated with users' shifts of attentional focus between contexts and tasks. In this article we have reviewed the work done so far in this direction. We have also indicated some issues related to the future development of AASs. Amongst these the most significant ones are the need to further investigate the application of AASs in environments that are not task-oriented, and the need to take into account collaborative situations when evaluating the cost/benefit of attentional shifts.

TERMS AND THEIR DEFINITION

Locus of attention

Amongst all sensory input, the locus of attention is the input to which one allocates mental resources. Input that falls outside the locus of attention may go absolutely unnoticed. An example of locus of attention is a specific section of a computer screen.

Direct manipulation user interfaces

Interfaces that aim at making objects and actions in the systems visible by [graphical] representation. They were originally proposed as an alternative to command line interfaces. The system's objects and actions are often represented by metaphorical icons on screen (e.g. dragging a file to the recycle bin for deleting a file). Designers of direct manipulation user interface strive to provide incremental reversible operations and visible effects.

Endogenous attentional processes

Refers to the set of processes of voluntary (conscious) control of attention. These processes are also referred to as top-down, or goal-driven. An example of endogenous attentional mechanism is the attention you are paying at this page as you are reading. Endogenous attention is voluntary, it requires explicit effort, and it is normally meant to last.

Exogenous attentional processes

Refers to the set of processes by which attention is captured by some external event. These processes are also referred to as bottom-up, or stimulus-driven. An example of this mechanism would be the attention shift from your reading due to a sudden noise. Exogenous attention is triggered automatically and it normally lasts a short time before it is either shifted or becomes controlled by endogenous processes.

Visual attention

Visual attention is the process by which we select the visual information most relevant to our current behaviour. In general, of all the visual stimuli we receive, we only attend to a few, this determines what we "see". Visual attention controls the selection of appropriate visual stimuli both by pruning irrelevant ones and by guiding the seeking of relevant ones. Research in visual attention aims at understanding the mechanisms by which human sensory and cognitive systems regulate what we see.

Gaze tracking

The set of mechanisms allowing to record and analyse human eye-gaze. Gaze tracking is normally motivated by the assumption that the locus of eye-gaze may, to some extent, correspond to the locus of attention or it can help capturing user interests. Several techniques exist for eye tracking varying in their level of intrusion (from requiring the user to wear special lenses, to just having camera-like devices installed on the computer), their accuracy, and ease to use. Normally devices need to be calibrated before use (some systems allow to memorise calibrations for specific users).

Gesture tracking

The set of mechanisms allowing to record and analyse human motion. Gesture may be tracked either in 2D or 3D. Gesture tracking ranges from the recording and analysis of postures (e.g. head, body) to that of more detailed elements such as hand fine movement, or facial expression. The aims of gesture tracking in HCI span from recognising the user's current activity (or lack of), to recognising emotional states. Gesture tracking is often used in combination with gaze tracking.

REFERENCES

Abowd, G. D. (1999, 16-22 May 1999). Software Engineering Issues for Ubiquitous Computing. Proceedings 21st International Conference on Software Engineering (ICSE'99), Los Angeles, CA, USA. 75 - 84

Argyle, M., & Cook, M. (1976). Gaze and Mutual Gaze: Cambridge University Press.

Bartram, L., Ware, C., & Calvert, T. (2003). Moticons:: detection, distraction and task. International Journal of Human-Computer Studies, 58(5), 515-545.

Bearne, M., Jones, S., & Sapsford-Francis, J. (1994). Towards usability guidelines for multimedia systems. Proceedings second ACM international conference on Multimedia, San Francisco, California, United States. 105 - 110

Bradshaw, J. (1997). Software Agents: AAAI Press/The MIT Press.

Brush, B., Bargeron, D., Gupta, A., & Grudin, J. (2001). Notification for Shared Annotation of Digital Documents (Technical Report MSR-TR-2001-87): Microsoft Research Microsoft Corporation.

Carroll, J. M., Neale, D. C., Isenhour, P. L., Rosson, M. B., & McCrickard, D. S. (2003). Notification and awareness: synchronizing task-oriented collaborative activity. International Journal of Human-Computer Studies, 58(5), 605-632.

Clark, H. H. (1996). Using Language. Cambridge: Cambridge University Press.

Czerwinski, M., Cutrell, E., & Horvitz, E. (2000). Instant messaging: Effects of relevance and time. Proceedings HCI 2000 - 14th British HCI group Annual Conference. 71–76

Driver, J. (2001). A selective review of selective attention research from the past century. British Journal of Psychology, 92, 53-78.

Grosz, B. J., & Sidner, C. L. (1990). Plans for discourse. In P. R. Cohen & J. Morgan & M. Pollack (Eds.), Intentions in communication (pp. 417–444). Cambridge, MA: MIT Press.

Hinckley, K., Pierce, J., Sinclair, M., & Horvitz, E. (2000, Nov. 2000). Sensing techniques for mobile interaction. Proceedings ACM UIST 2000 Symposium on User Interface Software and Technology, San Diego, CA. 91–100

Horvitz, E. (1999, May 1999). Principles of mixed-initiative user interfaces. Proceedings ACM SIGCHI 1999, Pittsburgh, PA. 159-166

Horvitz, E., Kadie, C., Paek, T., & Hovel, D. (2003). Models of attention in computing and communication: from principles to applications. Communications of the ACM, 46(3), 52-59.

Hudson, J. M., Christensen, J., Kellogg, W. A., & Erickson, T. (2002). "I'd be overwhelmed, but it's just one more thing to do": availability and interruption in research management. Proceedings SIGCHI conference on Human factors in computing systems, Minneapolis, Minnesota, USA. 97 - 104

Hudson, S. E., Fogarty, J., Atkeson, C. G., Avrahami, D., Forlizzi, J., Kiesler, S., Lee, J. C., & Yang, J. (2003, April 05 - 10, 2003). Predicting Human Interruptibility with Sensors: A Wizard of Oz Feasibility Study. Proceedings CHI 2003, Ft. Lauderdale, Florida, USA. 257 - 264

Huhns, M. N., & Singh, M. P. (1997). Readings in Agents: Morgan Kaufmann.

Hyrskykari, A., Majaranta, P., Aaltonen, A., & Räihä, K.-J. (2000, November 2000). Design issues of iDict: A gaze-assisted translation aid. Proceedings ETRA 2000, Eye Tracking Research and Applications Symposium, Palm Beach Gardens, FL. 9-14

Löwgren, J. (2002). Just How Far Beyond HCI is Interaction Design? Boxes and Arrows. Retrieved 20/11/04, 2004, from the World Wide Web: http://www.boxesandarrows.com/archives/just_how_far_beyond_hci_is_interaction_design.php

Maes, P. (1994). Agents that reduce work and information overload. Communications of the ACM, 37(7), 30 - 40.

McCrickard, D. S., Catrambone, R., Chewar, C. M., & Stasko, J. T. (2003a). Establishing tradeoffs that leverage attention for utility: empirically evaluating information display in notification systems. International Journal of Human-Computer Studies, 58(5), 547-582.

McCrickard, D. S., & Chewar, C. M. (2003). Attuning Notification Design to User Goals and Attention Costs. Communications of the ACM, 46(3), 67-72.

McCrickard, D. S., Chewar, C. M., Somervell, J. P., & Ndiwalana, A. (2003b). A model for notification systems evaluation—assessing user goals for multitasking activity. ACM Transactions on Computer-Human Interaction (TOCHI), 10(4), 312 - 338.

McCrickard, D. S., Czerwinski, M., & Bartram, L. (2003c). Introduction: design and evaluation of notification user interfaces. International Journal of Human-Computer Studies, 58(5), 509-514.

Negroponte, N. (1995). Being digital. New York: Vintage Books, Random House Inc.

Negroponte, N. (1997). Agents: from direct manipulation to delegation. In J. Bradshaw (Ed.), Software Agents (pp. 57-66): AAAI Press/The MIT Press.

O'Conaill, B., & Frohlich, D. (1995). Timespace in the Workplace: Dealing with Interruptions. Proceedings CHI '95 Conference Companion, Denver, Colorado, United States. 262-263

Raskin, J. (2000). The Humane Interface: New Directions for Designing Interactive Systems: Addison-Wesley.

Roda, C., & Thomas, J. (2004). Designing for attention. Proceedings HCI2004 Designing for Life - 18th British HCI group Annual Conference, Leeds, UK. Vol. 2 249-250

Roda, C., & Thomas, J. (2005). Attention Aware Systems: Theory, Application, and Research Agenda. Computers in Human Behavior, Forthcoming.

Shneiderman, B. (1992). Designing the User Interface: Strategies for Effective Human-Computer Interaction: ACM Press.

Shneiderman, B. (1997). Direct manipulation versus agents: path to predictable, controllable, and comprehensible interfaces. In J. Bradshaw (Ed.), Software Agents (pp. 97-106): AAAI Press / MIT Press.

Stiefelhagen, R. (2002, October 14 - 16). Tracking Focus of Attention in Meetings. Proceedings 4th IEEE International Conference on Multimodal Interfaces, Pittsburgh, Pennsylvania. 273-280

Uttal, W. R. (2000). Summary: Let's pay attention to attention. The Journal of General Psychology, 127(1), 100- 112.

Vertegaal, R. (1999). The GAZE groupware system: mediating joint attention in multiparty communication and collaboration. Proceedings SIGCHI conference on Human factors in computing systems, Pittsburgh, Pennsylvania, United States. 294--301

Vertegaal, R. (2003). Introduction. Communications of the ACM, 46(3), 30-33.

Weiser, M. (1991). The computer of the 21st century. Scientific American, 265(3), 66-75.

Zhai, S. (2003). What's in the eyes for attentive input. Communications of the ACM, 46(3), 34-39.