Technical and Privacy Challenges of Multimodal Dynamic Adaptive Systems

Lisa Anthony
UMBC
Information Systems Department
1000 Hilltop Circle
Baltimore, MD 21250 USA
lanthony@umbc.edu

CHI 2011, May 7–12, 2011, Vancouver, BC, Canada.
ACM 978-1-4503-0268-5/11/05.

Abstract
This paper describes one vision of an accessible future, in which a user's interactions with technology are dynamically and automatically adapted to the specific context of use in each new interaction episode, taking into account user needs, task elements, and environmental factors. Lessons from human-computer interaction and multimodal input tell us that considering these components of context will favorably impact how well such systems can adapt to new users and situations. This vision brings great promise for supporting natural and transparent interfaces, allowing users to switch fluidly between interaction modes as appropriate; it also brings new research challenges and requires balancing issues of privacy and ethics with supporting increased convenience for users.

Keywords
Dynamic accessibility, multimodal interaction, context, privacy, ethics, human-computer interaction.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

General Terms
Human Factors.

Introduction

Whether a user is interacting with handheld mobile devices, using a desktop computer, or experiencing smart or virtual environments, *context* determines what interactions are desirable and suitable. Future interactive technologies must be able to serve a wide variety of users, in terms of physical and cognitive characteristics, and in terms of the tasks for which, and environment in which, they are using the technology.

This author anticipates a future in which interaction technologies can dynamically adapt to the specific context of use with each new interaction episode, taking into account user needs, task elements and environmental factors. Supporting multimodal interaction, allowing the user to fluidly change modalities to suit the changing context, becomes critical. For example, the user could shake a device to cancel instead of touching a button when wearing gloves. Prior work in human-computer interaction and multimodal input [2] has shown that performance of a system can improve with the consideration of context, forming a complementary relationship: what's good for the user is also good for the system, and vice versa.

However, collecting this personal and contextual information brings with it issues related to privacy and ethics in terms of the way such data is handled. This paper outlines a research vision for dynamically accessible, transparent interfaces that maintain users' control of their own privacy without overburdening them to micromanage contextual adaptations.

Transparent, Multimodal Interfaces

"Transparency" as applied to interfaces or interactions is used to describe the removal of the "physical interface as a barrier between the user and the work [he or] she wishes to accomplish" [1]. An (imperfect) synonym might be "*invisible*"[1]. The user does not have to alter his or her behaviors when performing a task in order to be supported by an automated system. The system does not impose extraneous cognitive load on the user by forcing the user to adapt to the system, but in order to provide a truly transparent interface, the system must be able to adapt to the user seamlessly.

Allowing the user to communicate with the system in the same ways that he or she might communicate with another person (e.g., multimodal interactions such as speech, gesture, facial expressions, etc), improves the transparency of the interaction by eliminating the requirement to learn a new language to communicate with the system. Truly multimodal interaction implies that the user can switch fluidly between supported input modalities at any time, perhaps even mid-interaction. Most multimodal systems to date have focused on offering "bimodal" interfaces such as speech and pen, or speech and lip reading [7]; offering trimodal interfaces, and beyond, will be required to push into the future of natural multimodal support.

Adapting to Context

The term "context" has been used in many ways and to cover many different ranges of variables. Let us adapt a very broad definition from Beyer and Holtzblatt's *Contextual Design* [5], and define context to be the whole of the set of factors having to do with the potential *user population* of a system (e.g., education

[1] Compare this definition to the one usually given in automated systems to describe revealing a system's underlying algorithms and state to the user; a synonym there might be "exposed".

level, impairment, native language, etc), the potential *environments* in which the system could be used (e.g., outdoors, on-the-go, noisy situations, etc), and the potential *tasks* for which the system could be used (e.g., flow of information, domain knowledge, etc). There has been much prior work on the detection of context through the use of environmental or user-worn sensors, and classification of sensor readings into a representation of the context. See [3] for a survey.

Once the system is able to detect all the elements of the current context, it can then adapt to it. For example, it might provide the user with a digital sketchpad when the user types or says he or she is working on a conceptual design problem; or it might alter its spoken language understanding approach to one designed for stronger waveforms when it detects that the user is speaking in frustrated, emphatic tones.

In addition, the user might consciously choose an adaptation, such as preferring to use voice input on a mobile device while walking instead of touch input, or preferring to use touch input while at the opera instead of voice. The system that is truly multimodal will be able to recognize the streams of input, whether they are synchronous or not, and synthesize them into one representation of the desired user action or request.

Prior work in human-computer interaction and multimodal input has shown that the use of context can help improve results of this recognition and synthesis process (e.g., [2], [7]). This evidence supports pursuit of the vision of dynamic accessibility described in this paper. A complementary relationship is formed by the two halves: as the system utilizes context from multimodal input and other sources, it can more successfully comprehend the user's intent, enabling it to provide more situation-appropriate interactions to the user, which in turn encourages the user to continue to use multimodal input and provide this rich context.

This author anticipates several key technical challenges involved in supporting context-sensitive adaptations:

- How to enable fluid switching between manual control and automatic adaptation;
- How to manage conflicting adaptations suggested by separate but co-occurring contextual cues;
- What kinds of adaptations are well-received by users, and which are not and should be avoided;
- How to map the contextual cues to specific adaptations for one's domain; and,
- How to improve the appropriateness and success of adaptations over time.

While not a comprehensive list, it sets the stage for the types of research challenges to be faced in this field.

Issues of Privacy and Ethics

Consider that the types of contextual information gathered by a dynamic accessible system may contain highly personal and identifiable information, such as one's most common routes or locales (e.g., home to work), cognitive or physical abilities or impairments, Internet search habits, and so on. Certain users may be sensitive about having specific information tracked and logged over time. For example, in the frustration example above, the user might not want his or employer to be aware of his frustration at work. While studies on privacy attitudes and behaviors typically indicate that users do desire to preserve their privacy,

their actions seem to belie this, as they continue to invest a high amount of trust in Internet services [6].

Also, it has been found that when using websites that collect personal information in a social setting, such as Facebook[2], which has very permissive privacy settings by default, users do not understand the privacy settings available and the implications of their choices [4]. As conscientious researchers and designers of dynamically accessible systems, it is our responsibility to ensure that users' privacy and security is maintained, even when users themselves may not insist upon it.

It is this author's view that, equal to the technical challenges involved in supporting context-sensitive adaptations, are concerns of privacy and ethics. Key issues to be addressed are outlined here, but new ones will certainly arise as more data is available:

- How to balance users' desire for convenience with concerns of ethics and privacy;
- How to respect social and interpersonal norms when providing adaptations designed to make a task easier, without implying that the user needs extra help;
- How to maintain transparency of what data is being collected, how it is being used, and the system's underlying algorithms, without being too complex; and,
- How to ensure the system operates within users' conceptual models of privacy-preserving data use.

Dynamic accessibility is a viable and powerful goal for the future of interaction, but researchers and designers have an obligation to maintain the privacy of their users and to use their information ethically.

[2] http://www.facebook.com/

Conclusions

In sum, this paper has presented a vision for an accessible future in which the user may interact transparently with a system that can dynamically adapt to the user's situation. Multimodal interaction enables the fluid context-relevant switching between modalities that would be required in such a system. Prior work in multimodal interaction indicates this is a viable future goal. However, as context-adaptive systems collect more information about their users, we must consider issues of system transparency, privacy and ethics.

Acknowledgments

The author would like to thank the workshop organizers for helpful comments on earlier drafts of this paper.

References

[1] Abowd, G.D. Software Engineering Issues for Ubiquitous Computing. In *Proc. ICSE 1999*, 75-84.

[2] Anthony, L. *Developing Handwriting-based Intelligent Tutors to Enhance Mathematics Learning.* PhD Thesis, Carnegie Mellon University, 2008.

[3] Baldauf, M., Dustdar, S. & Rosenberg, F. A Survey on Context-Aware Systems. *Intl J Ad Hoc Ubiquitous Computing 2*, 4 (2007), 263-277.

[4] Besmer, A. & Lipford, H.R. Users' (Mis) Conceptions of Social Applications. In *Proc. GI 2010*, 63-70.

[5] Beyer, H. & Holtzblatt, K. *Contextual Design*. Academic Press, San Diego, CA, USA, 1998.

[6] Fox, S. Trust and Privacy Online. *Pew Internet & American Life Project*, Aug 20 2000, http://www.pewinternet.org/Reports/2000/Trust-and-Privacy-Online.aspx, accessed on Feb 5 2011.

[7] Oviatt, S. Breaking the Robustness Barrier: Recent Progress on the *Design* of Robust Multimodal Systems. In *Adv. in Computers 56*, Elsevier (2002), 305-341.