Natural Forms of Communication and Adaptive Behaviour in Human-Computer-Interaction

Madlen Wuttke[1,*] and Kai-Uwe Martin[2,*]

[1] Institute for Media Research, Chemnitz University of Technology, Chemnitz, Germany
madlen.wuttke@phil.tu-chemnitz.de
[2] Chair of Computer Engineering, Chemnitz University of Technology, Chemnitz, Germany
kai-uwe.martin@informatik.tu-chemnitz.de

Abstract. Scientific research over the last two decades imputes a beneficial effect on human-computer interaction by depicting a virtual communication partner onscreen due to the persona effect and the media equation theory. On the other hand, looking back at the historic component of human-computer interactions, the burden of adaptation has always been on humans to understand the machine and to communicate in accordance with its standards. This paper describes natural communication and interaction strategies of humans and computers as well as their importance to scientific research.

Keywords: Intelligent and agent systems, Pedagogical Agents, Natural Forms of Communication, Adaptive, Mobile Learning.

1 Human-Computer-Interaction

This paper discusses the possibilities of HCI from two perspectives: the social sciences, which aspire to develop a natural, human-like multi-channel communication, and the computer sciences, which engages with technology's capabilities and limitations. Historically, the burden of accommodation has always been on the side of humans [1]. Since machines are being built by humans and we already possess the evolutionary advantage of being able to adapt our behaviour, it seems logical to assume this division of labour. The problem is the hereby artificially limited pool of humans capable of interacting with computers. A broader distribution and application of computer systems in the general population was therefore bound to a more user accessible form of interaction.

Following in this line of thought, the development of different kinds of interfaces and the intuitive knowledge about their functions and usage is key to any successful human-computer interaction. Examples of such are already visible in the market today.

The Samsung Galaxy S4 [2] mobile phone has a feature installed which enables the phone's camera to accurately track the gaze of a user. Although the built-in front camera is not capable of tracking eye-movements, this quit simple gaze-tracking allows for a wide array of helpful and user friendly features. For example in a biometric

M. Kurosu (Ed.): Human-Computer Interaction, Part II, HCII 2014, LNCS 8511, pp. 641–647, 2014.

way which allows to unlock the phone, otherwise it enables the video player to pause a presentation once the user physically diverts his attention away from the phone. Another example is the scrolling of long text documents or websites. Once the gaze of the user is reaching the bottom of the screen, the gadget automatically begins to scroll down – allowing for a more intuitive way of reading text on a hand-held device like a smartphone.

TV manufacturers are currently implementing cameras into their latest products which, obviously, allow for an implementation of video-chat functionalities. Furthermore, it enables the devices to check for people looking at the screen. Devices capable of displaying 3D videos without the help of wearing goggles, by adjusting the surface of the screen, use the camera to identify a user's head and position within the visible space in front of the TV [3].

And, of course, car manufacturers are installing additional ways of interaction into their cars. Examples are vibrating steering wheels once diverting across lanes accidentally, cameras, checking the eyes for signs of fatigue or simply microphones to be able to react to voiced commands.

But these cases of application are examples of specific devices while the implementation of new interactive methods in the case of everyday computer systems is stagnating. The same is the case for the establishment of additional Human-Computer-Interaction-Methods in software with an educational purpose.

2 Designing, Development and Evaluation

When designing, developing and evaluating human computer interactions, it is helpful to use a multi-disciplinary approach to facilitate a broader insight of its processes and requirements. This holistic approach facilitates a more profound approach to the challenges of future HCI research. As a result, this paper describes natural communication and interaction strategies of humans and computers, inspects the technical requirements as well as developments and deals with future possibilities and improvements.

While the direction of interactional development has not changed much since its early stages, the challenge remains to establish a natural form of HCI, meaning to enable computer systems to both acknowledge and interpret multi-facetted user generated input and to create adequate responses.

Especially during the last decade natural input methods like gesture based controls using range cameras, voice recognition based on cloud computing and face recognition based approaches were developed, advanced and established. All these methods share a common base: Whenever either side is trying to influence the other by an interaction, an adapted response is expected. This results in the simple conclusion that every form of interaction demands at least some form of adaptive capability [4].

Research regarding new forms of human-computer interaction is invested in developing either intuitive ways of operating or about implementing new and yet untouched forms of dialogue based communication channels. For example Krämer [5] refers to an increased demand for user-centered behavior by computer systems, since they have the computing power and the capability for computer scientists to produce

such systems. She pertains further to the growing number of users which are less and less educated in a certain way about how to engage with computer systems, therefore the natural and intuitive ways of interaction will have to increase. The most sophisticated way of a human computer interaction is based on the principles of human to human communication, face-to-face, as humans are evolutionized to do.

The challenge for establishing a real-life-communication between these two entities is to implement all the audible, tactile, olfactoric, visible and invisible channels used in an everyday conversation. While the human part of HCI is equipped with, and used to, all the communication possibilities, the expected interpretation and equally capable distribution of these non-verbal forms of communication on the side of the computer system demands a daring effort.

Since the 1990s, scientific research is focused on empirical evidence for beneficial effects regarding human-computer interaction. Most experimental setups rely on depicting an ever so lifelike but still virtual communicational agent onscreen. The idea behind it is the evidence shown by the persona effect and the postulations of media equation theory. Ideally this would be a person-like robot, but the depiction of a person onscreen has been proven appropriate for facilitating a connection between user and computer [6], [7]. But while the rendering of a person's face, its mimic abilities and lifelike animations have progressed quite rapidly over the last decade [8], [9], the development in the field of text to speech software is not progressing quickly enough to keep up with the computers abilities to graphically display itself as a photo-realistic communication partner [10], [11].

This discrepancy hinders any dynamic approach of interactive communication to pre-recorded voice samples of another human, talking on behalf of the computer. Besides developing real-life appearances and realistic voice effects, research regarding the other non-verbal cues of communication is even more prolongated.

The specified approaches, from simple Input/Output to adaptive forms of interaction, new forms of interaction and finally natural ways of communication deserve a clear exploration and evaluation matching the socio-cultural requirements with the potential of the current available technology to provide a steady progress in human computer interaction. Since this is a convergence of interdisciplinary scientific research traditions, social sciences for explanation of human behavior and applied information sciences to develop software in accordance with user expectations, the potential for future research is extensive.

3 Input and Output

In order to establish new communication channels between the user and the computer, adequate and accessible hardware has to be implemented. For instance the early printout feedback of machines have been replaced by an electronic monitor to display information, while the previously used machine assembler code has been replaced by an attempted natural form of language which is still used today in the form of 'if', 'then' and 'else'. This usage of a more accessible form of communication led to the formation of a dialogue oriented attempt of interacting with machines [5].

Due to this development computers became more accessible, were easier to use for specific tasks and allowed for their embedded use today. To promote this development further, the goal should be to enable computers to be more refined and possess improved forms of interactive input and output. Current input methods are basically improvised forms of communication, although processor power and software development are conceptually ready for a more natural and capable of real-world representation of a communicational counterpart.

In the case of an educational system, this would enable a pedagogical agent to perceive and respond to individual issues of knowledge acquisition. Drastically changing the way of presenting material to a learner who, for the first time, is enabled to react appropriately and naturally once being distracted or unable to follow the presentation [12].

4 Adaptive Forms of Interaction

Looking at human-computer interactions, it is important to state that the idea behind any form of interaction is a palpable result. Humans want to interact with a computer in order to make it compute something. And whenever a computer is prompting an error-message it attempts to communicate to a user to change something regarding the previously attempted interaction. So whenever either side is trying to influence the other by an interaction, an adapted response is expected.

This results in the simple conclusion that every form of interaction demands at least some form of adaptive capability [4]. Even the push of a radio-button on a website results at least in the visual depiction of a selected item. But the attempted effect of an adaptive interaction is most probably to facilitate an open form of dialogue between the user and a machine where one influences the other due to reciprocal communication – being in the shape of a natural language or by selecting metaphorical icons which represent an intuitive function like the recycle-bin.

5 New Forms of Interaction

Research regarding new forms of human-computer interaction is invested in developing either intuitive ways of operating or about implementing new and yet untouched forms of dialogue based communication channels. And due to the success of social network sites, the development which begun with the already mentioned metaphorical icons via apparent impasses like the data glove, is now picking up speed mostly because of the possibility to finance new and creative ideas via social-online-crowdsourcing platforms [13], due to which the different ways of interaction might increase quite rapidly over the next couple of years.

This is also true for the formation of new communication channels which is dependent on progress and the development of better and faster hardware. But a simple webcam to track gaze and infer attention can be enough to enhance the way we interact with a computer system, since it opens up a visual channel for the system to act on behaviour instead of relying on the keyboard and mouse interface to simply react to the user.

Fig. 1. Educational software with a gaze perceptive pro-active pedagogical agent

6 Educational Settings

An educational setting is per se a program which relies heavily on interactive processes between a learner and some form of teaching software. Since the implementation of pedagogical agents seems to have a beneficial effect on the retention and transfer of knowledge [6] the enhancement of the dialogue between learner and software is paramount. So called 'conversational agents', have a high impact on research efforts. And again, a weak point in human-machine interaction is the limitation to the few already implemented channels of communication. Compared to a real-world face-to-face, teacher to student communication setting, additional information is being conveyed than currently possible.

But as in the rest of HCI research, most of empirical studies is focused on checking for beneficial or hindering aspects in appearance or social behaviour. Instead, research should be broadened to include not simply methods of instructional design but also to include real-world expectations of behaviour. For example if the noise level of the environment might be distracting to the learner. Or if a student's gaze is directed at the correct area of the screen.

As described by Wuttke [12] the pro-active educational aspects might best be dealt with inside a distinct module, called the Electronic Educational Instance (EEI).

The EEI is basically an add-in component capable of being implemented in a variety of information presenting devices like educational software or even recreational devices like Smart-TVs or even cars. Inside the EEI all the sensory information is

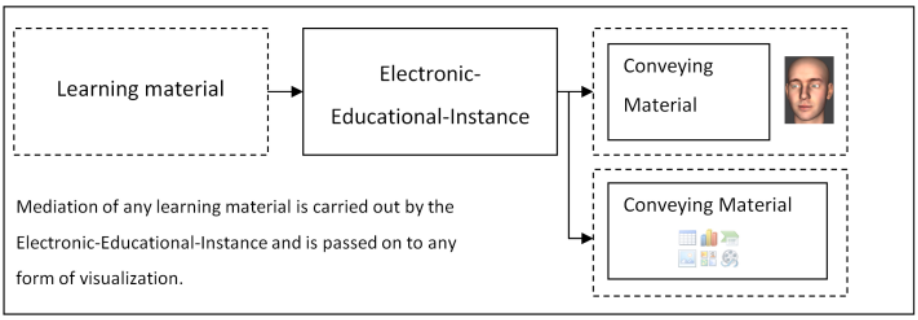

Fig. 2. Electronic Educational Instance (EEI)

gathered and situational context is created. Allowing for the device to be aware of potential disturbances and enabling it to react adequately.

7 Mobile Systems

The presented forms of interactions and communication have mainly been discussed within the background of traditional computer systems. With the ongoing development and growing distribution of mobile devices like smartphones and tablets, the research is continually focused on developing explicitly intuitive forms of communicating with the device. Traditional forms of providing input like a keyboard and mouse have to be represented onscreen in order of providing these new devices with the advantage of being easy to transport and easy to handle. Touch-based input is as intuitive as a form of input can get – you see, you touch (or push) – whatever reaction one wants to evoke.

Speech-recognition and optical character recognition (OCR) are other forms, which essentially attempt to emulate human perception systems. As an intermediate step, the ubiquitous distribution of Quick-Response-Code (QR-Code) is another example how human communication is adapted to the lower level of computer recognition. OCR still is simply too inaccurate to reliably implement it as a form of input, but due to the QR-Code, letter recognition is not any longer a prerequisite for transferring textual information from the real world into the computer system. While augmented reality applications enable users to perceive the real world and the additional layer of supplementary information, accessible only by using a computer system.

In the case of mobile learning applications, the sensory equipment which is presently installed into a handheld device enables a time- and space-dependent variation of displaying the teaching materials. For example the GPS would allow the application to accurately predict the location of a learner. In the case of the situation ‘waiting for the bus’ a non-verbal and easy to understand explanation could be offered in the form of a small knowledge-nugget. While in the case of being at the library it would show a textual but more elaborated chunk of knowledge. And once the device’s GPS and the wireless network at home is detected, it will produce an auditory chunk of knowledge like teaching vocabularies.

8 Conclusion

A wide array of issues has been discussed regarding the ways and capabilities of human computer interactions. While the direction of interactional development has not changed much since its early stages, the challenge remains to establish a natural form of HCI, meaning to enable computer systems to both acknowledge and interpret multi-facetted user generated input and to adequately respond in kind. Being able to use all the aforementioned channels human would use in a common conversational setting as well – both verbal and non-verbal.

Since this is a convergence of interdisciplinary scientific research traditions, social sciences for explanation of human behaviour and applied information sciences to develop software in accordance with user expectancies, the potential for future research is extensive. To facilitate this research, a theoretical discussion of the essential aspects of expected components and behaviour has been presented.

References

1. Draper, S.W., Norman, D.A.: Introduction. In: Norman, D.A., Draper, S.W. (eds.) User Centered System Design: New Perspectives on Human-Computer Interaction, pp. 1–6. Erlbaum, Hillsdale (1986)
2. Galaxy S4, http://www.samsung.com/us/topic/our-galaxy-smartphones (retrieved on January 31, 2014)
3. Toshiba 55zl2, http://www.toshiba.de/contents/de_DE/PRODUCT_DESC/files/1134731.pdf (retrieved on January 31, 2014)
4. Niegemann, H.M., Domagk, S., Hessel, S., Hein, A., Hupfer, M., Zobel, A.: Kompendium multimediales Lernen. Springer, Heidelberg (2008)
5. Krämer, N.C.: Soziale Wirkungen virtueller Helfer: Gestaltung und Evaluation von Mensch-Computer-Interaktion. Kohlhammer, Stuttgart (2008)
6. Lester, J.C., Converse, S.A., Kahler, S.E., Barlow, S.T., Stone, B.A., Bhogal, R.S.: The persona effect: Affective impact of animated pedagogical agents. In: Pemberton, S. (ed.) Human Factors in Computing Systems: CHI 1997 Conference Proceedings, pp. 359–366. ACM Press, New York (1997)
7. Reeves, B., Nass, C.: The Media Equation: How people treat computers, televisions, and new media like real people and places. Cambridge University Press, New York (1996)
8. Kim, Y.: Learners' expectations of the desirable characteristics of virtual learning companions. International Journal of Artificial Intelligence in Education 17(4), 371–388 (2007)
9. Agada, R., Yan, J.: Research to Improve Communication by Animated Pedagogical Agents. Journal of Next Generation Information Technology 3(1), 58–69 (2012)
10. Atkinson, R.K., Mayer, R.E., Merril, M.M.: Fostering social agency in multimedia learning: Examining the impact of an agent's voice. Contemporary Educational Psychology 30, 117–139 (2005)
11. Veletsianos, G.: The impact and implications of virtual character expressiveness on learning and agent–learner interactions. Journal of Computer Assisted Learning 25, 345–357 (2009)
12. Wuttke, M. (2013). Pro-Active Pedagogical Agents. In: Fakultät für Informatik (Ed.). Proceedings of International Summerworkshop Computer Science, pp. 59-62 (July 2013)
13. Marjanovic, S., Fry, C., Chataway, J.: Crowdsourcing based business models. In: Search of Evidence for Innovation 2.0. Science and Public Policy 39, pp. 318-332 (2012)