

Emotions and Multimodal Interface-Agents: A Sociological View

Daniel Moldt

University of Hamburg, Computer Science Department

Moldt@informatik.uni-hamburg.de

Christian von Scheve

University of Hamburg, Institute for Sociology

CvScheve@gmx.net

Abstract

Designing human-computer interfaces that are easy and intuitive to use is important for the use of computer technology in general. Due to the growing complexity of information systems, more and more elaborated concepts of interface design can be found. One of the most advanced concepts is agent technology. Within this conceptional approach *emotional agents* are of increasing importance because emotions have a strong influence on interactions. At the same time communication channels between humans and computers become more powerful. This can be seen in an intensified research in multimodal interfaces. Applying the concept of agents to multimodal interface design allows to combine the advantages of both approaches. Until now, this combination mostly relies on cognitive psychological approaches to emotion, although emotions are also a fundamentally social process. This is why sociology has to offer interesting perspectives. Within the DFG-research project “Sozionik”, sociological theory is used to enhance computational systems, in this case hybrid societies. We will show in which way sociological theories of emotion can be used to design emotional agents. This is also of importance regarding the use of multimodality since the emotional consequences of applied modalities are subject to social norms and rules. Consequently we call for a sociologically founded user model and a model of an agent’s “social self”.

1.1 Introduction

As computers and information systems are becoming more and more widespread within the general public, an increasing number of untrained and inexperienced users seek access to these technologies. It is therefore necessary to design systems in such a way that also people who are less familiar with high technology are able to use it. This is of utmost importance when the ability to deal with these systems becomes a precondition for the participation in public social life.

Facing these circumstances, users should be able to fall back on familiar mechanisms and modalities when interacting with computational systems. To accomplish this task, human-computer interaction should be modeled in a way that makes it comparable to interpersonal interactions, thus allowing users to rely on skills obtained from human to human and, ideally, face-to-face interactions. In this respect, multimodal interfaces are a suitable approach, since they have the potential to apply dimensions and modalities of human interactions to the field of human-computer interaction. Using multiple modalities the

insufficiencies of specific modalities could be compensated by employing appropriate alternatives.

When designing multimodal interfaces with the aim to model social interpersonal interactions, one has to bear in mind that the choice of an adequate modality has to be made according to the task, the user, and the situation in question. It is therefore necessary that multimodal interfaces allow for a dynamic task-, user- and situation-specific adaptation.

When modeling the characteristics and modalities of interpersonal interactions (e.g. mimics, gestures, voice-intonation or tactile behaviors), the meanings and effects of the employed modality with respect to an individual actor should be taken into consideration. We want to take a closer look at the emotional components of modalities in general since they play an important role in interpersonal interactions. Furthermore, emotions have the potential to transcend the technical, rationality based context of human-computer interaction in favor of a more interpersonal level (Bente/Otto 1996, 224f).

Approaches from the field of cognitive psychology that deal with these issues, suppose that multimodality can be used to support and to control human-computer interaction in a way that elicits certain user-emotions or supports emotional system feedback. Verbal communication for example may be enhanced by nonverbal clues to clarify intentions and the contents of the communication. That means specific modalities are used as a mediator of information that is of relevance in interaction or communication.

Without a doubt this is an important and widely accepted view on multimodality - nevertheless we suppose that employing a specific modality may *itself* evoke emotions in an actor. This is because the utilization of modalities in an interaction often is subject to social norms and rules just as emotions are. Sociological theory outlines a framework that shows how emotions and emotion expressions are bound to social norms and consequently to other actors' expectations:

- Specific norms and rules ("feeling rules") directly influence the emotions of an actor by establishing what is socially expected and by forcing him/her to perform "emotion work".
- Social norms and rules in general are indirectly responsible for the elicitation of certain emotions. For example negative emotions (e.g. guilt or shame) can be evoked when an actor realizes that his/her behavior is considered deviant or inadequate.
- Deviant behavior does not only lead to specific emotions in deviant individuals but also in other individuals participating in an interaction.

It is obvious that not only the modalities used to express emotions but also the use of modalities in general are part of this rule system. When utilizing multiple modalities in human-computer interfaces in order to support or to control interactions by means of corresponding emotions, the sociological perspective could be helpful. To be situational- and user-adequate, modalities should be applied according to valid social norms and rules and not in an ad hoc fashion. These rules and norms refer to gender-, cultural- or class-specific differences amongst individuals. Underlying this approach is the assumption that the same rules that guide interpersonal interactions are also valid in human-computer interaction. This assumption is backed up by several studies by Nass and others (Nass et al. 1994; Nass et al. 1994a; Bellamy/Hanewicz 1999)

To make our point of view clear once again: We do not want to examine differences in the symbols that are communicated by utilizing specific or multiple modalities (e.g. who uses what gesture when, with which intention, to reach what goal), instead we want to investigate the emotional effects that result from the utilization of specific modalities

themselves with respect to the participating actors and the social situation in which the interaction takes place. In most cases, it is not the mere use of a modality that elicits an emotion like a stimulus-response reaction, rather it is an actor's interpretation in view of a situational context that leads to an emotion.

Multimodality and the use of separate modalities are dependent on a system of social norms and rules and almost always have socio-emotional consequences on an actor, even if these consequences are unintended.

To illustrate in which way multimodality depends on sociological factors, we will refer to symbolic-interactionist perspectives on emotion (Hochschild 1979; Fiehler 1990). The cognitive appraisal process that is responsible for emotion elicitation does not only recur on psychological factors like beliefs, desires or intentions (Ortony et al. 1988), it also takes into account sociological categorization-schemes. Internalized social norms and rules determine how an actor has to behave in specific situations, that means what kinds of emotions and emotional expressions are appropriate, socially expected and considered to be adequate in an interaction. These internalized social norms and rules also constitute, which modalities are to be used in specific interaction situations or which of them are considered to be inadequate.

One fundamental concept in realizing user adaptive interfaces is the use of agent technology. Already at the stage of the specification of an agent, key characteristics of modern architectures like autonomy, intelligence, flexibility, mobility, or adaptivity are needed. In view of the above stated requirements for multimodal interface design, these characteristics qualify agent technology to be especially suitable for our approach. Within complex socio-technical environments agents could be designed in a way that enables them to take the role of human interaction partners, thus substituting them in part. This approach seems to be especially suited because agents are considered to bear an enormous potential regarding processes of sociomorphic and anthropomorphic attribution (Nass et al. 1993; de Angeli et al. 1999).

Taking into consideration that emotions are vital for interpersonal interactions and - as a mechanism of reducing complexity - increase the effectiveness of interactions and communications (Gerhards 1988, 88f), it appears obvious to equip interface agents with emotional components. In this case a sociological foundation seems to be appropriate for those emotions that are social in nature. One basic assumption of sociological approaches is that some emotions are subject to a process of adaptation to other actors' expectations and specific situational conditions. Therefore it is essential to provide emotional agents with a sociologically founded "social self", a corresponding user model and the ability to define interaction situations according to sociological factors (Moldt/von Scheve 2000).

Since multimodality is becoming more and more popular in human-computer interface design, it is important to observe closely the emotional components and consequences that result from utilizing multiple modalities. Dealing with this issue by using sociological emotion theories could lead to predictability and adequacy of the resulting emotions. We consider sociologically founded emotional agents as a conceptual basis for the approach proposed here.

In the following sections we will illustrate the potential of sociological emotion theories in this respect. The goal is to enable agents to choose appropriate and adequate modalities.

1.2 The Sociological Potential

In this section we discuss how sociological theory can be used to model emotional human-agent interaction. The main question within this approach is how emotions are elicited by “social facts” and how emotional expressions of agents can be adapted to user expectations resulting in lifelike, interpersonal interactions. We do not want to participate in the discussion on the interrelation between cognition and emotion here.¹ As already stated above, symbolic-interactionist theories of emotion are suited for our approach. They deal with questions about the social construction of emotions, how emotions are socially treated, coded, and judged and how “social facts”, that means norms, rules, standards, etc. determine how actors cope with and express their emotions (see also Tritt 1991). These theories are of interest in the following ways:

With them it should be possible to set up generalized emotion expectations. These expectations are foremost of interest in user-adapted interactions. Depending on the application an emotional agent may embody a certain role and/or personality. In analogy to interpersonal relationships a user will develop specific expectations not only concerning the agent’s task performance but also its social behavior and emotional expressions. Depending on how well an agent can meet and fulfill these expectations its behavior will elicit corresponding user emotions. Generally, it can be said that meeting the expectations will elicit positive emotions and failing to meet expectations will result in negative emotions. Also specific modalities are expected to be used in certain situations. These are either the modalities used to express emotions (they depend on “feeling rules”) or the use of modalities in general. Using an inappropriate modality in a specific situation may lead to emotions in those who interpret the modality as being inadequate and in those who utilized the inadequate modality.

A problem that still remains unsolved, also within sociology, is the specification of the factors that limit the generalization of emotion-response-expectations. What diversifies these expectations? Social factors could be culture, religion, milieu, class or gender, mainly aspects that are supposed to lead to different valid social norms and rules.

In an explorative way we put the above mentioned aspects in concrete terms. We do not try to define emotions in sociological terms, neither do we want to classify emotions nor emotion-words or modes of emotion expression. It is our aim to lay theoretical foundations for designing multimodal emotional agents.

1.2.1 Emotion Work and Feeling Rules

Hochschild’s concepts of “emotion work” and “feeling rules” refer to the cognitive work an actor performs when modifying emotions and social rules and norms that determine what emotions and emotion expressions are adequate and expected in social situations (Hochschild 1979). When performing emotion work an actor expresses specific emotions according to valid situational feeling rules and expectations of interaction partners, although he/she may not necessarily *feel* this emotion. Emotion work can be an individual as well as an interactional process: “[...] emotion work can be done by the self upon the self, by the self upon others, and by others upon oneself” (Hochschild 1979, 562).

Hochschild illustrates this phenomenon using the example of flight-attendants and bill-conductors. Flight-attendants are supposed to express positive emotions all the time they

¹ For a discussion of this issue see Lazarus (1984).

are on board of an aircraft. This is to establish a comfortable and friendly atmosphere and to mediate a feeling of sympathy toward the passengers. On the other hand, bill-collectors are supposed to induce an atmosphere of fright, shame and fear, the debtors shall feel uncomfortable and pay their bills (Hochschild 1983, 137ff).

But emotion work is not only performed at work, also in private life emotion work is a common phenomenon. Emotion work is not limited to the expression of emotions (“surface acting”). By means of cognitive work it is also possible to re-interpret and to modify ones own emotions in a way that the desired emotion is at last really felt (“deep acting”) (Hochschild 1979, 558). Feeling rules are social norms and rules that regulate respectively indicate emotion work in social situations. For example it is expected to show grief on funerals.

Feeling rules can be seen as a frame guiding emotion work and determining what emotions are adequate in specific situations. A question that is left unanswered by symbolic-interactionist theories of emotion is what the primary (social) elicitors of emotions are. The concepts of feeling rules and emotion work generally presuppose existing emotional stimuli that are modified and adapted by means of emotion work. Emotions generated by emotion work are certainly not primary emotions.

1.2.2 Defining Social Situations

Fiehler suggests to divide emotion elicitation into two domains (Fiehler 1990, 64f): The first domain is that of emotion elicitation in exceptional situations, for example emergency situations. In such situations actors are overwhelmed by their feelings, there is simply no time for a cognitive interpretation of the event and emotions in such situations often have an intensity that leads to an immediate automatic behavior.

But since our approach is not concerned with more the less automatic behavior but with meaningful social action, the second domain is of interest to us. That is the domain of familiar, regular situations where emotional intensity is mostly not as high as in emergency situations thus leaving enough time for a cognitive interpretation of the situation (Fiehler 1990, 65).²

These situations can be structured in three ways by an actor: First an actor has to define the situation. Is it a known, regular, familiar, unfamiliar, routine, etc. situation? Second it has to become clear if it is a situation of social interaction, and if this is the case, the type of interaction has to be identified (conflict, discussion, ritual, argument, etc.). Third the interaction partner(s) have to be analyzed. That means an actor has to determine what his/her relationship with the interaction partner is like (is it a friendship, an acquaintance, romantic love, a secretary, an assistant, a neighbor, etc.) and what role the interaction partner takes. Depending on how the situation has been defined specific feeling rules apply that enable an actor to interpret interaction relevant emotion expressions of other actors and indicate what actions and emotions are adequate in that situation. This way feeling rules constitute a means of reducing complexity since they make the actors' situational interpretations routine. They set up an interconnection between types of situations

² We will neglect here, that internalized social norms may also result in a more or less automatic emotional response on external stimuli. In this case, the norms and rules in question are internalized to a degree that lets us refer to them as a part of the personality, thus shifting an examination to a psychological level.

and expectable emotions.

1.2.3 Extended Feeling Rules

Fiehler has extended and specified Hochschild's concepts of feeling rules and emotion work. He introduces the additional concepts of manifestation rules, correspondence rules and coding rules that are described in the form of if-then explanations (Fiehler 1990, 77-87), a format that is also interesting from a pragmatic point of view (operationalization).

1. Following these additional aspects feeling rules codify which emotions, from Ego's point of view, are situational adequate and, from Alter's point of view, are expected. Is a situation being described as a type *X* situation it is adequate and expected to show emotional expressions of the type *Y*.
2. Manifestation rules determine how intense an emotion is expressed. If a situation is interpreted as type *X*, it is adequate and socially expected to show emotion manifestations of type *Y* in an interaction-relevant way (Fiehler 1990, 78). The manifestation rules, also called "display rules" or "expression management" (Gordon 1981), mainly come into effect when a feeling, as it is expected according to valid feeling rules, is not being felt.
3. Correspondence rules depict how interaction partners should react on perceived emotional expressions. They codify which corresponding emotions respectively emotion manifestations are adequate and socially expected in accordance to a perceived emotion manifestation of an interaction partner. Is the emotional state of an interaction partner described as type *X* then it is adequate and socially expected to have or at least to express an emotion *Y*. This phenomenon is also called "mood-sharing" or "mood-joining" (Denzin 1980, 257-256).
4. Coding rules state what kind of behavior is considered to be an expression of which emotion. These specifications may be mimics, gestures, language, intonation, etc. In general one could formulate an if-then explanation as follows: If under the conditions of *X1-Xn* the behavioral patterns of *Y* respectively *Y1-Yn* are observed, they can be considered to be a manifestation of emotion *Z*.

These rules constitute a complex system that depicts which emotions and which emotional expressions (manifestations) are situational adequate and expected. What they do not show is how an actor actually deals with emotions and the emotional expressions and what techniques he/she can employ to regulate them in case a difference between situational expectations and actually felt/expressed emotions comes to the mind of an actor. Hochschild's concept of emotion work or emotion regulation is following this point.

1.2.4 Dual-Monitoring and Emotion Work

Emotion work can be looked at from two different perspectives: First an emotion itself can be focused by an actor. This means that in a specific situation an actor feels a specific emotion that does not fit the situation or the applying feeling rules. It is perceived as being inadequate; by the self and also by others. Consequently an actor will try to regulate this emotion to fit the situation.

The second perspective is that of the situation and the corresponding feeling rules being focused. The applying feeling rules coerce actors to evoke or to manifest specific emotions to meet situational expectations. This perspective does not rely on a feeling that has

already been elicited. Instead it is an anticipatory view which is why emotion work can not only be seen as reactive-action but also as anticipatory-action (Fiehler 1990, 87-89).

Consequently social situations call for a process of dual-monitoring: On the one hand the social situation itself has to be monitored and on the other hand a steady surveillance of one's emotions and emotion expressions is necessary. In case this dual-monitoring reveals discrepancies between a social situation and a currently felt/expressed emotion, there are two possibilities of regulation:

An actor may hold on to the definition and interpretation of a situation, thus accepting the appropriate feeling rules, and trying to regulate his/her feelings according to the rules. Another possibility is to cognitively re-interpret the situation so that the feeling rules applying to this re-definition are congruent with and legitimate actually felt emotions.

In view of the first possibility, resulting emotions or emotion expressions may differ in the degree of performed emotion work. An actor may try to regulate the emotion itself ("deep acting") or merely the emotion manifestation ("surface acting"). The successful regulation of an emotion mostly is accompanied by a corresponding and adequate manifestation. Techniques to perform emotion regulation can be cognitive, gestural, mimic, bodily, verbal-communicative, etc. and pursue the goal to suppress, evoke or transform emotions and emotion expressions (Fiehler 1990, 91).

1.2.5 Summary

Considering emotional agents it has become clear that it is necessary to enable agents to perform emotion work respectively emotion regulation. This is especially important when modeling emotional interface agents whose tasks have to be seen in interactional contexts. To perform emotion regulation successfully a rule-based system of emotion work should be implemented. Agents must have knowledge about feeling rules, correspondence rules, coding rules, manifestation rules and the interrelations between these rules. Furthermore an agent has to analyze situations in the form of:

- Is this situation an interaction situation?
- What kind of interaction is this situation?
- What kind of relationship is maintained by the interaction partners?

To analyze situations successfully an agent has to hold or to build up a database containing information on prototype situations. Matching database contents and situational conditions will show, which feeling rules are valid for the situation in question and which techniques of emotion regulation are at an agent's disposal and may be used. Corresponding to the feeling rules and the prototype situations there should be sets of modalities an agent can choose from. This way it is ensured that only adequate modalities are used.

Furthermore an agent should contain a model of its "social self", that means knowledge about its social and task-related role, its position in the social space and maybe the class or milieu it belongs to. According to this conception of the self and the applying feeling rules the degree of the socially expected emotion regulation can be seen. On the other side, the agent has to build up a user model containing sufficient personal information to characterize the user in sociological terms. This would enable the agent to categorize the user according to the above mentioned factors and to predict user-expectations.

As can be seen from these requirements, using symbolic-interactionist theories of emotion in emotional agents design is strongly related to the adaptation to user expectations. This is why we consider this approach to be especially suited for user-adaptive and multimodal

systems.

1.3 Conclusion

In the case of user centered interface design it has become clear that multimodality, emotionality and agent based interactions deserve a special attention. In the near future it will be important to ensure the incorporation of these approaches into a unified theoretical framework and conception of human-computer interface design. Sociological theories of emotion seem to be a valuable and suited approach to theoretically support the realization of this conceptional framework and to extend (not to substitute) existing models inspired by cognitive psychology.

Multimodality as an aspect of interpersonal interaction can be considered as an emotion elicitor as well as a part of a norm-based rule system of emotional expressions. This is why multimodality cannot be looked at without taking into account the emotional components that accompany the utilization of multiple or separate modalities. Consequently, applying multimodality to human-computer interfaces calls for appropriate mechanisms and concepts that consider these social norms and rules.

Sociologically founded emotional agents are one step in this direction because they are able to simulate the differentiated processes of human emotional behavior. They ensure that also the socio-emotional components of multimodality can be modeled and applied adequately.

Because social interaction is based on intersubjective reciprocity, a model of the “social self” of an agent, an elaborated user model, and the definition of an interaction situation are necessary to simulate this process as far as possible in an agent-user relationship.

Looking at this approach in a more general way, it becomes obvious that not only the components that are being represented to the user but also other parts of an agent architecture could be modeled with the help of sociological theory. Our suggested architecture is that of a distributed (multi-agent) system that is being represented toward the user by a single interface agent. In contrast to contemporary architectures, that try to implement all capabilities within a single entity, an emotional agent could be considered a multi-agent system that is responsible for an adequate utilization of multiple modalities.³

We think that in the foreseeable future sociological theories will gain more and more relevance for the design of man-machine interfaces. This is based on two assumptions: First there is an increasing interest and a growing need for an understanding of socially coordinated distributed problem solving. In this respect emotions (and their social nature) are one important component. Second, *social* interaction is not possible without the interpretation of the interaction partner (this includes psychological as well as sociological aspects). But since actors often only have insufficient personal information about an interaction partner, they have to rely on sociological categorizations to make a meaningful inter-

³ We suggest that every emotional agent is considered a multi agent system. Within this system there are one or more agents that are directly or indirectly responsible for the different aspects of emotionality. It is the interrelation between emotional and non-emotional, sociological and psychological/neuro-scientific (and other) components of the distributed system that is important when dealing with the questions and issues presented in this article.

action possible. This is especially true for human-agent interactions, since an agent's knowledge about personal user information usually is very limited. Therefore, a transformation of sociological concepts of interaction seems to be a desirable and effective solution.

We think that an intensive interdisciplinary exchange between sociologists and computer scientists could be very productive to provide both parties either with means to build life-like, failure tolerant and performative systems or to specify and improve theoretical concepts of social interaction, coordination and problem solving. In our opinion the paradigm of agents respectively actors is the best suited metaphor to use in this context, since both scientific fields aim at understanding these principles which we consider to be strongly related with each other.

The main contribution of this paper is to enlarge the awareness of a successful implementation of sociologically founded concepts of social interaction. Traditionally, computer science that is working on multimodal interfaces is focussing on "hard" facts and at most on research results from the field of cognitive psychology. Nevertheless there may be some interesting sociological theories to look at, even if they seem to pose more questions than answers at a first glance.

1.4 References

- Bellamy, A.; Hanewicz, C. (1999): Social Psychological Dimensions of Electronic Communication. In: *Electronic Journal of Sociology* 4/1999 (1). URL: <http://www.icaap.org/iuicode?100.4.1.2>
- Bente, G.; Otto, I. (1996): Virtuelle Realität und parasoziale Interaktion. In: *Medienpsychologie* 8/1996 (3), 217-242.
- De Angeli, A.; Gerbino, W.; Nodari, E. und D. Petrelli (1999): From tools to friends: where is the borderline? In: *Online Proceedings of the Workshop on Attitude, Personality and Emotions in User-Adapted Interaction*. URL: <http://aos2.uniba.it:8080/pr-ws-um.html>
- Denzin, N.K. (1980): A Phenomenology of Emotion and Deviance. In: *Zeitschrift für Soziologie* 9/1980 (3), 251-261.
- Fiehler, R. (1990): *Kommunikation und Emotion. Theoretische und empirische Untersuchungen zur Rolle von Emotionen in der verbalen Interaktion*. Berlin: de Gruyter.
- Gerhards, J. (1988): *Soziologie der Emotionen: Fragestellungen, Systematik u. Perspektiven*. Weinheim: Juventa.
- Gordon, S.L. (1981): The Sociology of Sentiments and Emotion. In: Rosenberg, M.; Turner, R.H. (Eds.): *Social Psychology. Sociological Perspectives*. New Brunswick: Transaction, 562-592.
- Hochschild, A.R. (1979): Emotion Work, Feeling Rules, and Social Structure. In: *American Journal of Sociology* 85/1979 (3), 551-575.
- Hochschild, A.R. (1983): *The Managed Heart. Commercialization of Human Feeling*. Berkeley: University of California Press
- Lazarus, R.S. (1984): Thoughts on the Relations Between Emotion and Cognition. In: Scherer, K.R.; Ekman, P. (Eds.): *Approaches to Emotion*. Hillsdale: Lawrence Erlbaum, 247-257.
- Lenzmann, B. (1998): *Benutzeradaptive und multimodale Interface Agenten*. Sankt Augustin: Infix
- Maes, P. (1994): Agents that reduce work and information overload. In: *Communications of the ACM* 37/1994 (7), 31-40.
- Moldt, D.; von Scheve, C. (2000): Soziologisch adäquate Modellierung emotionaler Agenten. In: Müller, M. (Ed.): *Benutzermodellierung: Zwischen Kognition und Maschinellern Lernen*. 8. GI-Workshop ABIS'00. Osnabrück: Institut für Semantische Informationsverarbeitung, 117-131.
- Nass, C.; Steuer, J.; Tauber, E. and H. Reeder (1993): Anthropomorphism, Agency, & Ethopoeia: Computers as Social Actors. In: *Adjunct Proceedings of the Conference on Human Factors in Computing Systems, INTERCHI'93 (INTERACT and CHI)*, Amsterdam, 111-112.
- Nass, C.; Steuer, J.; Henriksen, L. and D.C. Dryer (1994): Machines, social attributions, and ethopoeia: performance assessments of computers subsequent to "self-" or "other-" evaluations. In: *International Journal of Human-Computer-Studies* 40/1994 (3), 543-559.
- Nass, C.; Steuer, J.; Tauber, E.R. (1994a): Computers are Social Actors. In: *Proceedings of the Conference on Human Factors in Computing Systems (CHI'94), "Celebrating Interdependence"*, Boston, 1994. New York: ACM Press, 72-78.
- Ortony, A.; Clore, G.L.; A. Collins (1988): *The Cognitive Structure of Emotions*. Cambridge: University Press.
- Tritt, K. (1991): *Emotionen und ihre soziale Konstruktion: Vorarbeiten zu einem wissenssoziologischen, handlungstheoretischen Zugang zu Emotionen*. Frankfurt am Main: Lang.