Attribution and Adaptation: The Case of Social Norms and Emotion in Human-Agent Interaction

Daniel Moldt
University of Hamburg, Computer Science Dept.
Vogt-Koelln-Str. 30, D-22527 Hamburg, Germany
moldt@informatik.uni-hamburg.de

Christian von Scheve
University of Hamburg, Institute of Sociology
Vogt-Koelln-Str. 30, D-22527 Hamburg, Germany
xscheve@informatik.uni-hamburg.de

Abstract
Adapting interface agents' behaviors to cognitive user states and processes is a difficult task. It is briefly outlined how other properties of the user can additionally be utilized to make human-agent interaction more social. Social-structural implications, categorizations, and social norms are crucial factors in interpersonal interactions. Sociological theory analyzes and provides conceptualizations of social environments. We propose to use these theories in human-computer interaction in order to account for the peculiarities of specific social worlds. We illustrate the potential of sociological theory by giving an example of the interrelation between social norms and emotion. To do this, an existing emotional assistant agent is described.

Keywords
Social Norms; Attribution; Emotion; Human-Agent Interaction; Sociological Theory

INTRODUCTION
Due to the rapidly increasing deployment of computational systems in everyday life, efforts toward designing systems that are capable of interacting with users in an interpersonal style are intensified. These systems require a minimum set of seemingly intelligent behaviors, the possibility to acquire a certain amount of information about the user, and the ability to adapt to a specific user according to the information gathered. In this area, research is often focused on the cognitive processes underlying interaction: the properties of a user are often tried to be modeled from the "inside out". That means those properties that are most difficult to obtain and to realize, because they are often located in the innermost recesses of the mind, are to be modeled. Many individuals would even find it hard for themselves to put these properties, e.g. goals, beliefs, intentions, preferences, or emotional state, in concrete terms.

Furthermore, a user-model is usually considered to be an explicit representation of the properties of a *particular* user. But, in everyday social interactions human actors merely have any explicit information about each other, regarding the mentioned properties. Therefore, we propose to make use of those mechanisms in human-computer interaction that make up a great deal of social, interpersonal interactions. These mechanisms are extensively analyzed by sociological theory, which, unfortunately, has not had a great impact on human-computer interaction research until now. We will briefly introduce sociological concepts here, that depict the interrelation between emergent social phenomena (such as norms and rules) and emotion. To illustrate in which way these concepts can be used in emotional agents design, we will describe the "CyMON" architecture, that has been developed by Agentscape AG and is deployed in a real-world application.

ACTOR MODELING IN SOCIAL INTERACTIONS
In this section, we will analyze everyday, non-conflict social interactions. These interactions make up a substantial part of human affairs, where there is usually no need for (conscious) deliberation, negotiation, or extensive cooperation. Interaction sequences of this kind often are highly standardized, institutionalized, and even ritualized and constitute what Collins has called "interaction ritual chains" (Collins 1981). For actors engaging in such an interaction ritual, there is not necessarily a need (1) to reason about specific goals, motifs, beliefs, etc. of another actor or (2) to use large parts of cognitive resources e.g. for situation-specific short term planning, decision-making or action selection, unless the course of an interaction deviates from actors' expectations (for an illustration of everyday-action see Schütz (1944)). What, then, if not conscious cognitive processing and reasoning facilitates such an interaction?

A prerequisite for successful social interaction of the described kind is a shared amount of knowledge about significant symbols and social facts, e.g. codes of conduct, hierarchies, roles, social status, power, and the like. This knowledge does not only vary between different cultures and societies but also between smaller forms of social aggregation, e.g. groups or communities. Knowledge about symbols and social facts allows to attach (social) percep-

tions with a socially situated meaning, which is *shared* and bilaterally accepted (in dyadic interactions).

According to one's knowledge about meaningful symbols of a specific social environment on the one hand, and perceptions of social stimuli on the other hand, actors attribute specific motifs, desires or intentions towards each other. These attributions generally depend on three factors: (1) the situation an interaction takes place in, (2) the relationship actors maintain with each other, and (3) the concepts of self (or self-attributions) maintained by participating actors. Since these factors are dynamic in nature and may be updated during an interaction, the premise of an explicit representation of the user's properties is hardly manageable in real-time (Bianchi-Berthouze et al. 1999).

Social situations are often characterized by a lack of informational clues about interaction partners (Ego and Alter Ego). Beliefs, desires or intentions of an Alter Ego are seldom easily accessible from Ego's point of view (even Ego often does not know what his own intentions and beliefs are). Therefore, in social interactions, actors may fall back on social categorizations such as roles, social class, gender, cultural background, religious affinity, etc. to attribute the intentions that could motivate behaviors and actions of Alter Ego, thus giving them a socially situated meaning and making them understandable and to some extent also predictable (Moldt/v. Scheve 2001).

Due to several constraints (time, uncertainty, imperfect information, etc.) these categorizations often cannot be obtained by means of discourse or extensive verbal exchange. Thus, actors use perceptions of social stimuli which are easily accessible, e.g. physical appearance, status-symbols, occupation, group affiliations or prosody to categorize an Alter Ego. The power and advantage of these attributions and categorizations is, that they reduce complexity generated by symbolic cognition and also often determine (reciprocally) the way social stimuli are perceived, processed and adapted (as one aspect of social cognition) (Festinger 1957; Forgas 2000).

In everyday interactions, options and alternatives for decision making and action generation decrease to a degree that allows for quick, although often quite unspecific decisions and inferences. This can best be seen regarding unpleasant side-effects that can occur from these mechanisms, like stereotypization, discrimination, prejudice or even racism.

According to categorizations and attributions made during an interaction, specific social norms and rules apply and guide actors' behaviors and actions toward each other. They help actors to deal with contingency and reciprocity problems. We argue that these aspects (social categorizations, norms and rules) should be taken into account when modeling the behavior of (personified, anthropomorphic) intelligent interface agents. This way agents could be enabled to adapt to a user, respectively to groups of users, without having to gather information about every individual's cognitive representations in the first place.

For instance, demographic data obtainable from the user could be used to generate a user-model consisting of information about social status, prestige, cultural and symbolic capital, class, gender, etc. Based upon this model and a variety of prototype situations that can occur in an interaction, rule-based conclusions could be drawn from matching situation, user-model and prevailing norms (Moldt/von Scheve 2001a).

This will lead to a very shallow model providing guidelines for emotional reactions which are not necessarily deeply cognitively grounded. But as we will illustrate, shallow models of emotion and emotional interaction may be sufficient for specific applications such as persuasive computing or e-commerce applications.

INTERACTING WITH AGENTS

The basic assumption of our approach is, that intelligent agents are not intelligent in a way that is comparable to human intelligence, rather they are able to show behaviors as if they had human intelligence. On the other hand, users generally know that agents are inanimate objects rather than intelligent living beings. Nevertheless they tend to attribute characteristics of interpersonal subjectivity, personality, emotionality or intelligence toward these agents (anthropomorphism or "intentional stance") (Dautenhahn 2000; Nass et al. 1993). They behave as if the agent was an intelligent and intentional entity with human-like qualities. We consider this a prerequisite to apply sociological models of interaction in this respect.

When approaching human-agent interaction from a sociological point of view, it has to be assured that human action directed toward an interface agent as "Alter Ego" qualifies as some kind of *social* action, thus legitimating the use of an interactionist approach. Regarding the role of emotion in this process, Geser (1989: 233) notes that to constitute social interaction it is sufficient to have *one* entity that acts socially. Another entity (in this case an interface agent as the addressee respectively recipient of social action) is only of interest as a source of bodily or verbally manifested behaviors, e.g. speech acts, body movements, gestures or mimics which are being perceived and processed by the socially acting entity (in this case the user), and lead to alterations of the user's state of mind (e.g. by evoking respect, sympathy, pity, etc.).

This assumption is backed up by studies carried out by Nass and associates who found out that users tend to perceive human-computer interaction in "self-" and "other-" dimensions just like in interpersonal relationships (Nass et al. 1994/1994a). Similarly, users often assign sociomorphic and anthropomorphic attributes and behavioral roles toward computers. These processes are normally limited to human interaction partners (Turkle 1984: 147).

Until now, the process of attributional reciprocity (or "as if" behaviors) is often neglected, although it bears the potential to improve human-agent interaction. It should be possible to shape this process in a way that allows for more interpersonal and meaningful interactions. Since this process of attributional reciprocity is virtually already a simulation of social interaction (in a sociological sense) it seems reasonable to use sociological theories of interaction to analyze how these "as if" behaviors may be connected and related to one another. Using the exemplary concept of "emotional action" we briefly sketch how this could be achieved.

EMOTIONAL AGENTS

An exemplary illustration of a sociologically founded model of social interaction is provided by means of "emotional action". Emotions are considered to be an increasingly relevant factor in interface agents design, either to convey intentionality, to influence a user, to communicate meaning or simply to make interactions more "comfortable" (Picard 1997; Bates 1994). The sociological concept of "emotional action" describes in a wider sense to what extent emotional behaviors (felt emotions and emotion expressions) are subject to social norms and rules. Emotions in social interactions are not solely dependent on cognitive appraisal but also on a system of social norms and rules ("feeling rules") that directly influence the elicitation of emotions and the way actors deal with their emotions (coping, mood-joining, emotion work) (Hochschild 1979). "Emotional action" thus is an actor's intentional behavior directed to regulate and adapt an emotional state or expression to meet the expectations of other actors.

In order to model lifelike social interactions, it is important to consider the ability to reflexively deal with one's emotions. Enabling emotional expressive agents to adapt emotion expressions according to prevailing norms may encourage users to attribute roles, personality, other social qualities or even intelligence toward an agent. That means, a user will be (unconsciously) willing to behave as if the agent was a human-like entity. On the other hand, acting and behaving in conformity with prevailing norms is an important part of social intelligence and supports the overall "as if" intelligent behavior of an agent.

These abilities and strategies are often neglected in contemporary emotional agents design. But this approach is not limited to what is often referred to as "emotional intelligence". There are many other behaviors which are perceived as being "intelligent" although they are not more than social match-making.

We are aware of the fact that these "shallow models of emotion", as Sloman (1999) has called them, are in no way sufficient representations or models of the natural phenomenon. Although progress in "deeply rooted" cognitive science emotion research is made rapidly nowadays, we hardly have any "deep" models or architectures that are applicable to real world interfaces. In view of what we have argued in the preceding section, it seems appropriate to use these "shallow" models, for example in a way that is explained in the following section.

An Exemplary Application

In this section we will introduce the "Flirtmaschine" as an example of applied emotional agents design. The "Flirtmaschine" is a social match-making website with the primary goal of arranging romantic love relationships amongst its users. It is developed by Agentscape AG.[1] Because of existing contacts between our departments and Agentscape AG we have the possibility to briefly introduce the "Flirtmaschine" here. Due to copyright-reasons we are unable to present the architecture as a whole; we will therefore leave out technical details and focus on aspects which are most relevant for our approach.

Deployed on the "Flirtmaschine" website is the "Cyb" (Create Your Bot), an emotional personal assistant agent which is based on the "CyMON" architecture (Create your Match and Organizing Netware).[2] The Cyb is personified and graphically represented in form of a cartoon-character. It is comparable to a "believable social and emotional agent" as described by Reilly (Reilly 1996) and also bears characteristics of a synthetic-character as described by Elliot and Brzezinski (Elliot/Brzezinski 1998). To represent emotion expression up to 250 Flash-animations may be used. In the context of the "Flirtmaschine" the Cyb has to deal with three fundamental tasks:

(1) The agent takes the role of a matchmaker between the partner-seeking users. The user will be asked several questions to obtain personal user data and to generate psychological and sociological user profiles which are compared and matched, resulting in a partner-suggestion.

(2) The Cyb assists users in navigating through editorial contents, which are part of the "Flirtmaschine". It can suggest specific articles of potential interest to the user. To make proposals, the agent can fall back upon the information stored in the user-profiles or on content related user preferences that become accessible through aggregated click streams.

(3) The agent should encourage the user to visit the site for successful match-making and for commercial reasons. In this respect, emotionality is considered to be a crucial means since it is constitutional in establishing social relationships. Furthermore it is capable of binding users to a social relationship because of the emotional feedback they get out of it. Besides the promise to find an appropriate

---

[1] The "Flirtmaschine" is located at: http://www.flirtmaschine.de See also http://www.agentscape.de

[2] "Cyb", "CyMON", and "Flirtmaschine" are registered trademarks of Agentscape AG.

partner, the interaction with the agent itself should be encouraging enough to visit the site ("Tamagotchi Effect").

The Emotion State Machine (ESM) of the CyMON architecture is following the OCC-model of emotion (Ortony et al. 1988). Emotional states result from an appraisal process, i.e. the interpretation of situational conditions or user actions according to the agent's internal beliefs, desires or intentions. Internal states are related to a goal hierarchy, that means depending on which goal is currently the most active (thus being pursued with greatest effort), specific intentions will result and guide the agent's behavior and emotional states/reactions. An emotional state will conversely influence the agent's desires and intentions by guiding and regulating action-selections.

Most of the agent's behavior is adaptable to the actions a user actually performs and to the user-model that contains sufficient demographic data to categorize a user in terms of social class, occupation, gender, income, status, and the like. But, until now, it is quite difficult to perceive an emotional user state in order to adapt an agent's *emotional* behavior according to the perceived state (at least for web-based applications) (Picard 1998).

The Cyb's sole possibility to perceive an emotional user-state is by questioning (in a natural-language dialogue) the user at login. Adaptation therefore only occurs as a direct, yet appropriate natural-language and graphical response to the perceived state. Despite the problems of modeling emotional user states, the concept of "emotional action" could help to improve this adaptation. The agent's knowledge about emotional user states, be it as vague as it is, could have ongoing consequences on forthcoming agent-user interactions.

As we have argued, "feeling rules" and "emotion work" are constitutional to the construction of emotion in a social context. According to his emotional state and the situation an interaction takes place in, the user will expect specific behaviors and emotion expressions of "Alter Ego" (in this case the Cyb). These expectations are not universal, but related to the user's "social self", e.g. the class or milieu he belongs to.

According to the data stored in the user-profile, the Cyb could categorize the user in a way described earlier and match these categorizations with the interaction situation and the prevailing norms and rules for emotion expression. This would allow for a socially adequate adaptation of the agent's emotion expressions. The emotional state originally generated by the ESM could then be modified and regulated in order to fit the user's expectations.

An example: We start from the initialization of an agent-user relationship. When the Cyb has introduced itself and described what it will do for the user, specific expectations will result on the side of the user, probably corresponding to expectations from a comparable relationship in the real world. So the question is: what is a corresponding relationship in the "real" world? We would suggest it is that of a personal assistant that helps performing specific tasks but who is also addressable when it comes to private concerns.

Since the personal assistant and the user perform role actions to some extent, specific expectations result from this role-behavior. This relationship is also characterized by a specific relational structure: The user will certainly hold more power than the assistant but on the other hand he is expected to grant status and appreciation to the assistant for accomplishing tasks in a satisfactory way. If the assistant has performed a task of normal difficulty level it will expect to be granted with an adequate amount of status (this could be praise or appreciation of some kind). Should the status grant excess expectations, thus being interpreted as inadequately high, embarrassment may result on the side of the secretary. If the status grant is considered to be insufficient, anger or depression could be the outcome (see Kemper 1978).

But now, let us consider the appropriate feeling rules for this situation: The assistant knows that the relationship is in the first case an occupational relationship, private concerns are secondary. Consequently the secretary will avoid expressing his emotions, because they are considered to be inadequate in this situation. Instead he may try to re-interpret the situation as follows: Maybe the user is very busy at the moment, he simply forgot to say, for example: "Thank you, well done." This cognitive re-interpretation could negate the initial emotion, thus keeping the social relationship on a constant level. Also, no consequences in view of the assistant's action-selection would occur.

This example exactly describes the concept of "emotional action" and in which way it can be applied to emotional agents design. As long as there are technical limitations preventing a dynamic recognition of emotional user states, sociological theory and the concept of "emotional action" can be used as a fall-back position. But, the views expressed here are also necessary to obtain a more complete picture of natural emotional phenomena in addition to "deep" cognitive models.

OUTLOOK

As we have shown, sociological theory combined with cognitive models of emotion can help to answer urging questions in emotion-based human-computer interaction. Furthermore, emotions are considered to have strong effects on phenomena like social change, social structural dynamics or the emergence of norms (Staller/Petta 2000; Moldt/v. Scheve 2002). These aspects are also of importance for all applications dealing with larger aggregates of agents (multi-agent systems, artificial societies) or groups of users (e.g. computer supported cooperative work environments). To take into account these social structural aspects of emotion and relate them to existing cognitive

theories of emotion is one further aim of our research. Without a unified approach it is hardly possible to get a clear picture of natural emotional phenomena and to make them usable for improved technical systems.

To model the interrelation between emotion and social structure (as described by sociological theory), we will draw upon results of the interdisciplinary project "ASKO" ("Acting in Social Contexts") at University of Hamburg, that has successfully modeled theories of organizational choice by means of Petri nets (Heitsch et al. 2000).

REFERENCES

Bates, J. (1994): The Role of Emotions in Believable Agents. In: Communications of the ACM 37 (7): 122-125.

Bianchi-Berthouze, N.; Berthouze, L.; Kato, T. (1999): Understanding Subjectivity: An Interactionist View. In: Kay, J. (Ed.): User Modeling. Proceedings of UM'99. Wien: Springer, 3-12.

Collins, R. (1989): On the Microfoundations of Macrosociology. In: American Journal of Sociology 86 (5): 984-1014.

Dautenhahn, K. (2000): Socially Intelligent Agents and The Primate Social Brain - Towards a Science of Social Minds. In: Proceedings of the AAAI Fall Symposium "Socially Intelligent Agents - The Human in the Loop". Technical Report FS-00-04. Menlo Park: AAAI Press, 35-51.

Elliot, C.D.; Brzezinski, J. (1998): Autonomous Agents as Synthetic Characters. In: AI Magazine, 19 (2): 13-30.

Festinger, L. (1957): A Theory of Cognitive Dissonance. Palo Alto: Stanford University Press.

Forgas, J.P. (2000)(Ed.): Feeling and Thinking. The Role of Affect in Social Cognition. Cambridge: Cambridge University Press.

Geser, H. (1989): Der PC als Interaktionspartner. In: Zeitschrift für Soziologie, 18 (3): 230-243.

Heitsch, S.; Hinck, D.; Martens, M. (2000): A New Look into Garbage Cans – Petri Nets and Organisational Choice. In: Time for AI and Society. Proc. Of the AISB'00 symposium on Starting from Society: The Application of Social Analogies to Computational Systems. Birmingham/UK, 17th-20th April 2000.

Hochschild, A.R. (1979): Emotion Work, Feeling Rules, and Social Structure. In: American Journal of Sociology, 85 (3): 551-575.

Kemper, T.D. (1978): A Social Interactional Theory of Emotions. New York: Wiley-Interscience.

Moldt, D.; v. Scheve, C. (2001): Emotional Actions for Emotional Agents. In: Agents & Cognition. Proc. of the AISB'01 Symposium on Emotion, Cognition, and Affective Computing. York: SSAISB, 121-128.

Moldt, D.; v. Scheve, C. (2001a): Emotions and Multimodal Interface-Agents: A Sociological View. In: Oberquelle, H.; Oppermann, R.; Krause, J. (Eds.): Mensch & Computer 2001. Tagungsband der 1. fachübergreifenden Konferenz. Stuttgart: Teubner, 287-295.

Moldt, D.; v. Scheve, C. (2002)[under review]: Emotions in Hybrid Social Aggregates. In: Mensch & Computer 2002. Proceedings of the Conference Mensch & Computer '02. Stuttgart: Teubner.

Nass, C.; Steuer, J.; Tauber, E.R. and H. Reeder (1993): Anthropomorphism, Agency, & Ethopoeia: Computers as social actors. In: INTERCHI'93 (INTERACT and CHI) adjunct proceedings of the Conference on Human Factors in Computing Systems. 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 (3): 543-559.

Nass, C.; Steuer, J.; Tauber, E.R. (1994a): Computers are social actors. In: Human Factors in Computing Systems, CHI'94 Conference Proceedings, 72-78.

Picard, R.W. (1997): Affective Computing. Cambridge: The MIT Press.

Picard, R. (1998): Toward Agents that recognize Emotion. MIT Media Lab Perceptual Computing Section, Technical report #515.

Reilly, W.S. (1996): Believable Social and Emotional Agents. Ph.D. thesis, (CMU-CS-96-138), Carnegie Mellon University. Pittsburgh.

Schütz, A. (1944): The Stranger: An Essay in Social Psychology. In: American Journal of Sociology, 49: 499-507.

Sloman, A. (1999): Beyond Shallow Models of Emotion. Paper for I3 Spring Days Workshop on Behavior Planning for life-like characters and avatars. Sitges, Spain.

Staller, A.; Petta, P. (2000): Introducing Emotions into the Computational Study of Social Norms. In Proceedings of the 2000 Convention of the Society for the Study of Artificial Intelligence and the Simulation of Behaviour (AISB'00), Birmingham, UK, April 17-20, 2000.

Turkle, S. (1984): The second Self. Computers and the Human Spirit. London: Granada.