

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.

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.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2002, April 20-25, 2002, Minneapolis, Minnesota, USA.
Copyright 2001 ACM 1-58113-453-3/02/0004...\$5.00.

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.

ACTOR MODELING IN SOCIAL INTERACTIONS

The kinds of social interactions to be analyzed here are those of everyday, usually non-problematic and non-conflict 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) perceptions 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 attrib-

ute 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 realtime (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 (see Moldt/von Scheve 2001). The power and advantage of these attributions and categorizations is, that they reduce complexity generated by symbolic cognition and also often determine 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 behaviors 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") (see Dautenhahn 2000). 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.

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 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.

Furthermore, sociological theory in combination with cognitive models of emotion can help to answer urging questions concerning the role of social structure in the emotion process. On the other hand, emotions are considered to have strong effects on phenomena like social change, social structural dynamics or the emergence of norms (see Staller/Petta 2000). These aspects are 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 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.
- 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.
- 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.
- 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.
- Picard, R.W. (1997): *Affective Computing*. Cambridge: The MIT Press.
- Schütz, A. (1944): The Stranger: An Essay in Social Psychology. In: *American Journal of Sociology*, 49: 499-507.
- 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