Virtual Humans Growing up: From Primary Toward Secondary Emotions

Christian Becker-Asano, Stefan Kopp, Nadine Pfeiffer-Leßmann, Ipke Wachsmuth

In order to understand and model the role of emotion in cognitive processes we attempt to integrate theoretical approaches originating from different disciplines in an implemented cognitive architecture for embodied agents. Our virtual humanoid agent Max employs this architecture to generate believable human-like behaviors in a variety of situational contexts. In this article, we describe how we go about endowing Max's architecture with increasingly elaborated kinds of emotions – from primary emotions like happiness and fear, toward secondary emotions like hope and relief.

1 Introduction

In recent years the integration of "emotion-driven" behaviors has become prominent in the field of virtual humans [7] for at least two reasons: First, the integration of emotions—or at least emotional expressions—is supposed to support the believability of the artificial interlocutor in Human-Computer Interaction. Second, from a more theoretical point of view it is argued that without the integration of non-rational concepts, such as emotion, the ultimate step toward human intelligence might never be accomplished by Artificial Intelligence.

The virtual human Max, developed at Bielefeld University's AI Group, is a testbed for studying human-like behavior in natural face-to-face interactions [9]. We describe here how we incrementally endow Max's cognitive architecture with simulated emotions. In this we follow Damasio's [4] distinction of "primary" and "secondary" emotions. Primary emotions are ontogenetically older types of emotions and they lead to basic behavioral response tendencies like "flight-or-fight" behaviors. They are elicited in immediate response to stimuli that might also originate from internal, bodily processes. In contrast, secondary emotions like "relief" or "hope" are assumed to arise from higher cognitive processes, based on an ability to evaluate preferences over outcomes and expectations.

In Section 2 we present theoretical approaches in psychology to shed light on the somewhat fuzzy concept of emotion. After introducing an integrated architecture in Section 3, we will describe in Section 4 how we started out with simulating primary emotions. Comparable to human children, Max directly reveals his primary emotions by modulated involuntary behaviors and changing facial expressions. In human adults, however, a smile may as well be connected to a secondary emotion as, e.g., relief. In Section 5 we present how we have started to model the higher-order cognitive processes that underlie the simulation of secondary emotions. We will thereby show how our growing up agent is becoming able to experience increasingly elaborated kinds of emotions and how these extensions advance his abilities to interact in a gaming scenario.

2 Background

The concept of emotion has long been subject to controversy in psychology. Two major strands of theories can be distinguished. *Cognitive* emotion theories are focusing on the cognitive appraisal processes [20] and structures [12] necessary to elicit the full range of emotions in adult humans. On the other hand (and not to be treated completely in separation as we will see), *dimensional* emotion theories [5] are based on the idea to classify emotions along a varying number of dimensions of connotative meaning [14]. In the following, we will present these strands of theories in more detail and report on computational systems that integrate emotions into their architectures.

2.1 Emotion theories

The emotion model proposed by Ortony, Clore and Collins [12] (in short, OCC) has often been the basis for modeling emotions in cognitive architectures of embodied characters. In this conceptual model, a total of 22 emotion categories are differentiated that can be deduced logically from events, agents and objects. Although explicitly designed to be applied computationally, it was frequently criticized for major methodological drawbacks. Recently, Ortony et al. [13] have argued for distinguishing three levels of information processing that give rise to different classes of affective states. "Proto-affect" is considered a product of lower-level, hard-wired, reactive processes. "Primitive and unconscious emotions" arise from routine level processing characterized by awareness without self-awareness. At last, the reflective level gives rise to "full-fledged, cognitively elaborated emotions" originating from higher-order cognitive functions.

In contrast to the conceptual OCC approach, Scherer's "component process model" identifies different functional sources of emotion elicitation and proposes five distinct components, in which the regulatory functions reside [20]. Notably, these components can also be associated with corresponding parts of the human nervous system. As these components are assumed to process "stimulus evaluation checks" (SEC, in short) in parallel, timing and synchronization are central to this theory.

With regard to the different degrees of awareness and conscious processing, Damasio's [4] primary emotions are understood as basic, more automatic behavioral response tendencies. Secondary emotions occur when the individual starts to form "systematic connections between categories of objects and situations, on the one hand, and primary emotions, on the other" [4, p. 134]. The ability to express verbally secondary emotions such as hope or relief also furthers the impression of a interlocutor as a sensible human-like social partner [21].

Concerning the connotative meaning it has been claimed that any emotion can be characterized as a continuous progression in a three-dimensional space [5]; see also [22]. The three dimensions are commonly labeled *Pleasure/Valence (P)*, representing the overall valence information, *Arousal (A)*, accounting for the degree of activeness of an emotion, and *Dominance/Power (D)*, describing the experienced "control" over the emotion itself or the situational context it originated from. The degree of *Dominance* can further be seen as a general aspect of an agent's personality traits. Given equal negative *Pleasure* and medium *Arousal*, dominant individuals have a much stronger tendency to get angry, whereas individuals with a submissive personality are more likely to feel sad in the very same situational context. The three-dimensional abstract space spanned by these dimensions is referred to as *PAD-space*.

Scherer and colleagues [19] argue for combining dimensional affect representation and cognitive appraisal theory. They emphasize that cognitive processes leading to the elicitation of emotions might remain partly non-conscious. In their discussion the authors assume that it "seems likely that there is a strong and immediate effect of appraisal (and the consequent physiological and expressive changes) on feeling, which is then followed by a weaker but more continuous tuning of appraisal by feeling in the course of emotion regulation" [19, p. 108]. Exactly this complex interaction of *cognitive appraisal* and "bodily feeling" is key in our approach to computational emotion simulation.

2.2 Computational models

Based on the OCC model, Reilly [17] used emotion simulation to increase the believability of artificial actors in a drama scenario, but had to admit that the system was only "moderately successful". Especially the necessity to implement rules specific to a single domain, let researchers soon start to think of alternative ways to integrate emotions into cognitive architectures. For example, Gratch and Marsella [6] describe a framework for modeling emotions to integrate appraisal and coping processes, which is based on the Belief-Desire-Intention (BDI) approach to modeling rational agents. Central to their concept are "appraisal frames and variables" to capture the emotional value of external and internal processes and events. Making use of the agent's reasoning capabilities and concepts such as likelihood and desirability, individual instances of emotion are aggregated into a current emotional state and an overall mood. Mood, defined as a more long-term indicator of the agent's well-being, is considered beneficial because it has been shown to impact "a range of cognitive, perceptual and behavioral processes, such as memory recall (mood-congruent recall), learning, psychological disorders (depression) and decision-making" [6, p. 18]. Furthermore, following the idea of mood-congruent emotions, the mood value is used to weight otherwise equally active emotional states (such as fear and hope at the same time).

Remarkably, Gratch and Marsella's framework has been the first fully implemented and truly domain-independent architecture for emotional conversational agents. However, it seems questionable whether it is sufficient to rely on purely rule-based systems when simulating mental states like emotions. Recent neuropsychological research suggests that many non-conscious processes are heavily involved in the elicitation of emotions. Also, bodily feedback is considered responsible for a variety of emotions and should not be neglected in the attempt to simulate them. Thus, we propose a combination of the two major strands of theories introduced above: cognitive and dimensional theories of emotion.

3 An integrated architecture

Max is based on a cognitive architecture [10] in which deliberative processes are interpreting perceptions to decide which action to take next, and to compose behaviors that realize it. This is implemented following the BDI approach [16] of modeling rational behavior. Knowledge is represented as distinct beliefs kept and updated in memory. Deliberative processing allows the agent to plan ahead by internally simulating and evaluating actions against the background of these beliefs. Additionally, internal feedback information originating from the agent's physis layer is continuously fed back to the cognition layer.

Reflexes, immediate responses to external events, are realized by direct connections between perception and action at the physis layer. Reactive behaviors include gaze tracking and focusing the interlocutor in response to prompting signals. Likewise, fast-running stimulus-response loops react to internal events such as a changing emotional state. In addition, involuntary behaviors pertaining to Max's virtual physis like eye blink, breathing, and body sway are simulated at this layer.

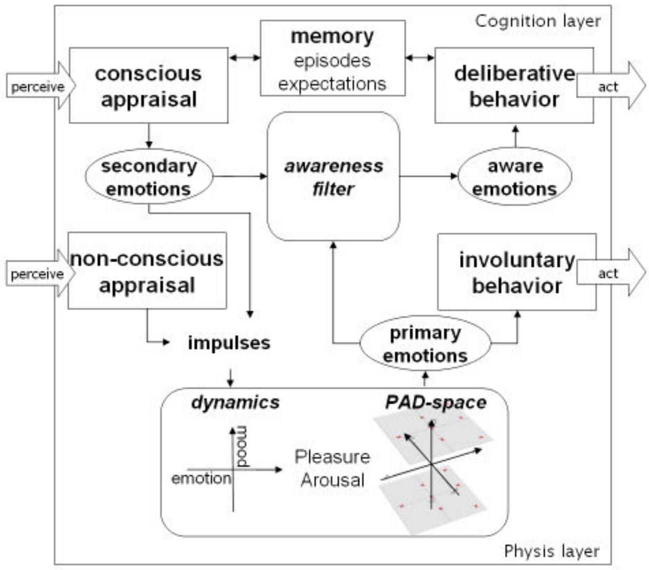

Figure 1: An outline of the emotion system leading to mood-congruent primary and secondary emotions.

Max's architecture has been extended to integrate the simulation of primary and secondary emotions as shown in Figure 1, and as explained in the next sections. The idea of a coherent emotion dynamics, the bottom part of Fig. 1, is realized in a three-dimensional emotion space. As shown in the upper part, conscious and non-conscious appraisal are distinguished, separately giving rise to primary and secondary emotions. Both kinds of emotions are fused and filtered to ensure the mood-congruency of aware emotions.

To validate Max's architecture with respect to the simulation of emotions, he was applied in various scenarios. In the following, we will first explain the simulation of primary emotions in the context of a cooperative scenario in which Max is employed as a virtual guide in a public museum [8]. Then, we will present the aforementioned extensions toward the simulation of secondary emotions in the context of a competitive gaming scenario [3].

4 Primary emotions

Max's architecture was enhanced in several ways in order to model primary emotions. Indispensable to the simulation of any kind of emotion is the appraisal of external and internal events (see Section 2.1). In contrast to the original OCC-model [12], at this level, we do not distinguish between perceptions of agents, actions, or objects as being more or less affective events. Instead, we started with the idea of "intrinsic pleasantness or unpleasantness" [20] as the most basic aspect of an emotionally relevant event, giving rise to primary emotions [4]. These basic events are assumed to originate already from non-conscious appraisal at the agent's physis layer (including the perceptual system) and may only subsequently be followed by cognitive elaboration (see Figure 1).

The emotion dynamics is based on the assumption that emotions have a fortifying or alleviating effect on the mood of an individual. An emotion is commonly understood as a shortlived phenomenon, whereas a mood is a longer lasting, valenced state. According to findings of empirical psychology [11], a mood creates a predisposition to experience certain changes of emotion. For example humans in a positive mood are more susceptible to positive than negative emotions, and vice versa.

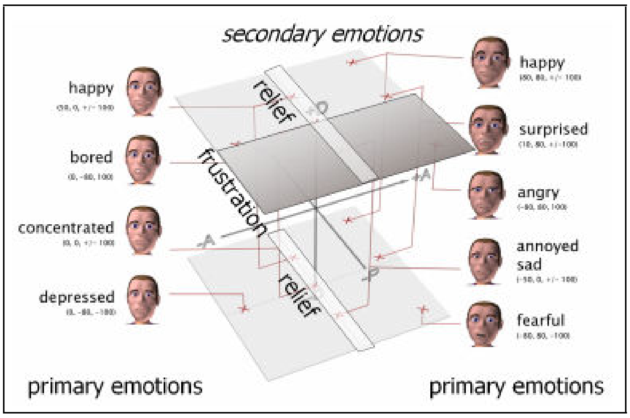

Figure 2: Nine primary emotions located in PAD-space (secondary emotions are represented in the same space, see Sect. 5).

The next step of emotion processing consists of a categorization of these valences onto primary emotions in PAD space as introduced in Section 2.1. This categorization is based on the assumption that every primary (as well as secondary) emotion can be located in PAD-space with respect to its inherent degree of *Pleasure*, *Arousal* and *Dominance* (cf. Figure 2). In case of primary emotions such as anger, fear or happiness we argue that such a representation is already sufficient to capture their respective connotative meanings. Consequently, the elicitation of primary emotions takes place directly in PAD-space on the basis of a distance metric. First, a continuous, mathematical mapping from emotion dynamics to the Pleasure-Arousal subspace is applied. Next, the subjective state of Dominance is continuously updated drawing on conscious appraisal of situational parameters in the cognition layer. Finally, a distance metric is used to calculate the activation levels of the primary emotions located in PAD-space; for details cf. [1].

Figure 3: Max in a smalltalk conversation with the visitor.

In the museum guide scenario Max is employed as a conversational guide in a public computer museum; see [8]. By means of a video camera Max can perceive the presence of museum visitors, and he engages them in conversations in which he provides information about the museum, the exhibition, and other topics (see Figure 3). Non-conscious appraisal in this scenario is based on the analysis for skin-colored regions in the incoming visual information. In case of success a small positive emotional impulse is sent to the emotion dynamics and a reactive, gaze following behavior is triggered concurrently. In effect, Max's mood increases when people are around. In the absence of visitors the emotion dynamics is generating the emotional state of boredom and special behaviors are triggered such as leaning back and yawning. This physical exertion is modeled to have an arousing effect by automatically setting the boredom value (and thus also the arousal value) to zero.

Concurrently, the interpretation module sends a positive (or negative) emotional impulses if the visitor's utterance has been understood as a compliment (or an insult, resp.). Likewise, the achievement of a desired discourse goal, e.g., coming to know the visitor's age after having asked for it, causes the dialog manager to send positive impulses to the emotion system.

In this scenario Max always feels dominant, independent of who is having the initiative in dialog. The elicited primary emotions, once categorized, are asserted as beliefs of the agent and influence the agent's deliberative reasoning. An appropriate facial expression is concurrently being triggered on a reactive level in accordance to the primary emotion letting Max express his emotional state directly.

A detailed example of an interaction together with the generated emotion dynamics can be found in [2]. A first evaluation of this collaborative scenario revealed that Max evokes natural communication strategies in the visitors, who ascribe a certain degree of sociality to the agent [8].

5 Toward secondary emotions

According to [18], desires and actual beliefs can be reinterpreted to generate expectations and newly acquired beliefs may lead to a match or mismatch with these expectations. This ability is an essential prerequisite in the process of cognitive appraisal leading to secondary emotions such as relief or frustration. For example, "relief" may be specified as "the disconfirmation of the prospect of an undesirable event" [12, p.121].

We currently extend our agent's reasoning abilities enabling him to first generate expectations, then appraise these expectations based on past experiences and finally evaluate current events in the light of former expectations. Whenever this conscious appraisal process gives rise to a secondary emotion, its inherent valence information is turned into an emotional impulse (see Figure 1). This impulse affects the emotion dynamics in the same way, and at the same time, as the outcomes of the concurrently running non-conscious appraisal. That way the valence of the agent's mood is also influenced by the higher-order, cognitive emotions. Additionally, some aspects of the connotative meanings of secondary emotions are represented in PAD-space by the two areas labeled "relief" and "frustration" in Figure 2. The agent is only able to become aware of emotions that pass an "awareness filter", which guarantees mood-congruency of emotion elicitation. This filter takes as input a set of candidate secondary emotions together with the currently active state of emotion dynamics in PAD coordinates. If the reference point falls into a region of one or more secondary emotions in PAD-space (see Figure 2), these are added to the output set of aware emotions. The resulting set of aware emotions consists of primary as well as secondary emotions and is subject to deliberation that might give rise to coping behavior.

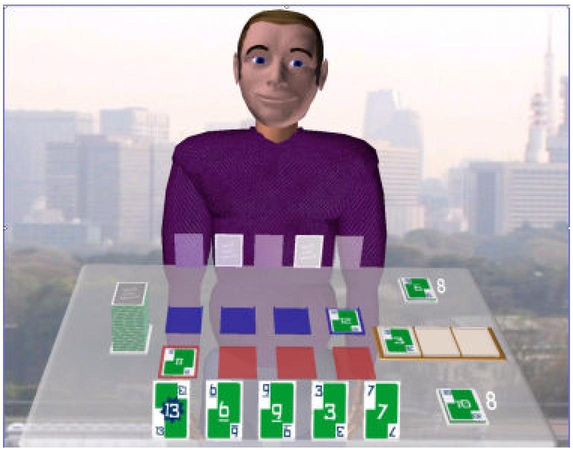

Figure 4: Max playing Skip-Bo against a human [2].

As a competitive face-to-face interaction between a human and Max we implemented a cards game called "Skip-Bo" [3]. In this game, both players have the conflicting goals of getting rid of their cards (see Figure 5). As Max always has the control over the game in a sense that he corrects the human player in case of a false move, his emotion system is initialized to reflect the state of high *Dominance*. However, when the human player is at least two cards ahead to win the game, the *Dominance* value is changed to reflect a state of *Submissiveness*. Consequently, when Max is highly aroused and in a state of negative pleasure, he sometimes expresses the primary emotion *fear* instead of *anger* within this scenario.

An empirical study revealed a desirable effect of the direct expression of primary emotions [15]. However, as the ability to express secondary emotions furthers the impression that Max is a sensible human-like social partner [21], we integrate means to elicit secondary emotions within this scenario as follows. When Max comes to believe that the opponent may play a card hindering him to fulfill one of his goals, the expectation of an undesirable event is generated. However, upon perceiving and interpreting the opponent's actions, Max realizes that the card the opponent has played is not as bad as expected. In result the secondary emotion relief is proposed and, at the same time, a positive emotional impulse is sent to the emotion dynamics. If the corresponding PAD values fall into the *relief* region (cf. Figure 2), Max will become aware of the secondary emotion *relief*. In addition, one of the primary emotions happiness, annoyance or sadness is passed on to deliberation, depending on the dominance and pleasure values. Other secondary emotions will be integrated similarly letting the agent experience and express increasingly complex emotional states.

6 Conclusion

In this paper, we described our approach toward the integration of emotions into a virtual human's cognitive architecture. Along the lines of Damasio [4] we started with the simulation of "primary" emotions as a product of a concurrently running emotion dynamics. In contrast to the classical attempts of emotion simulation, which are mainly based on conceptional emotion theories and implemented as an extension of existing reasoning processes, our emotion dynamics is realized based on dimensional theories of emotions and the idea of emotional impulses derived from intrinsic pleasantness and goal-conduciveness [19]. Thus, we at first started to simulate more infant-like, primary emotions. In first evaluations, where our agent was situated in two different interaction scenarios, we found that the direct expression of primary emotions—including negative ones—yields positive effects on the acceptance of Max as a coequal social partner.

In ongoing work, we extend the emotion simulation capabilities of our agent toward secondary emotions as outlined in Section 5. To guarantee mood-congruency of emotions, we suggest a combination of higher-order, cognitive appraisal with bodily-grounded emotion dynamics, which lets Max not only reason about emotions but also "have" emotions of his own. The higher-order cognitive processes extend the classical BDI-approach to generate expectations based on short-term memory. The competitive game scenario is being extended to include secondary emotions, enabling us to study and validate their effects. We expect that combining the two classes of emotion theories, as proposed here, will significantly contribute in the attempt to increase the believability and acceptance of virtual humans.

7 Acknowledgement

This research has been partially supported by the DFG in the Collaborative Research Center 673, and by the Japan Society for the Promotion of Science.

References

[1] C. Becker, S. Kopp, and I. Wachsmuth. Simulating the emotion dynamics of a multimodal conversational agent. In *Workshop on Affective Dialogue Systems*, LNAI 3068, pages 154–165. Springer, 2004.

[2] C. Becker, S. Kopp, and I. Wachsmuth. Why emotions should be integrated into conversational agents. In T. Nishida, editor, *Conversational Informatics: An Engineering Approach*, chapter 3. Wiley, November 2007, pages 49–68.

[3] C. Becker, H. Prendinger, M. Ishizuka, and I. Wachsmuth. Evaluating Affective Feedback of the 3D Agent Max in a Competitive Cards Game. In *Affective Computing and Intelligent Interaction*, LNCS 3784, pages 466–473. Springer, 2005.

[4] A. Damasio. *Descartes' Error, Emotion Reason and the Human Brain*. Grosset/Putnam, 1994.

[5] T. L. Gehm and K. R. Scherer. Factors determining the dimensions of subjective emotional space. In K. R. Scherer, editor, *Facets of Emotion*, chapter 5. Lawrence Erlbaum Associates, 1988.

[6] J. Gratch and S. Marsella. A domain-independent framework for modeling emotion. *Cognitive Science Research*, 5:269–306, 2004.

[7] J. Gratch, J. Rickel, E. André, J. Cassell, E. Petajan, and N. Badler. Creating interactive virtual humans: Some assembly required. *IEEE Intelligent Systems*, 17:54–63, 2002.

[8] S. Kopp, L. Gesellensetter, N. Krämer, and I. Wachsmuth. A conversational agent as museum guide – design and evaluation of a real-world application. In *Intelligent Virtual Agents*, LNAI 3661, pages 329–343. Springer, 2005.

[9] S. Kopp, B. Jung, N. Lessmann, and I. Wachsmuth. Max - A Multimodal Assistant in Virtual Reality Construction. *KI-Künstliche Intelligenz*, 4/03:11–17, 2003.

[10] N. Lessmann, S. Kopp, and I. Wachsmuth. Situated interaction with a virtual human - perception, action, and cognition. In G. Rickheit and I. Wachsmuth, editors, *Situated Communication*, pages 287–323. Mouton de Gruyter, Berlin, 2006.

[11] R. Neumann, B. Seibt, and F. Strack. The influence of mood on the intensity of emotional responses: Disentangling feeling and knowing. *Cognition & Emotion*, 15:725–747, 2001.

[12] A. Ortony, G. L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, 1988.

[13] A. Ortony, D. Norman, and W. Revelle. Affect and proto-affect in effective functioning. In *Who needs emotions: The brain meets the machine*. Oxford University Press, 2005.

[14] C. E. Osgood, G. J. Suci, and P. H. Tannenbaum. *The Measurement of Meaning*. University of Illinois Press, 1957.

[15] H. Prendinger, C. Becker, and M. Ishizuka. A study in users' physiological response to an empathic interface agent. *International Journal of Humanoid Robotics*, 3(3):371–391, 2006.

[16] A. Rao and M. Georgeff. Modeling Rational Agents within a BDI-architecture. In *Proc. of the Intl. Conference on Principles of Knowledge Representation and Planning*, 1991.

[17] W. S. N. Reilly. *Believable Social and Emotional Agents*. PhD thesis, Carnegie Mellon University, 1996. CMU-CS-96-138.

[18] R. Reisenzein. Emotions as metarepresentational states of mind. *Cybernetics and Systems*, 2:649–653, 2006.

[19] K. Scherer, E. Dan, and A. Flykt. What determines a feeling's position in affective space? A case for appraisal. *Cognition and Emotion*, 20:92–113, 2006.

[20] K. R. Scherer. Appraisal considered as a process of multilevel sequential checking. In K. R. Scherer, A. Schorr, and T. Johnstone, editors, *Appraisal Processes in Emotion*, chapter 5. Oxford University Press, 2001.

[21] J. Vaes, M.-P. Paladino, and J.-P. Leyens. The lost e-mail: Prosocial reactions induced by uniquely human emotions. *British Journal of Social Psychology*, 41:521–534, 2002.

[22] W. Wundt. *Vorlesung über die Menschen- und Tierseele*. Voss Verlag, Leipzig, 1922.

Contact

Dipl.-Inform. Christian Werner Becker-Asano
Universität Bielefeld
Technische Fakultät, AG Wissenbasierte Systeme (KI)
Postfach 100131, 33501 Bielefeld
Tel.: (0521) 106-2919
E-mail: cbecker@techfak.uni-bielefeld.de

Bild

Christian Werner Becker-Asano is a Ph.D. student and research assistant in the AI Group at Bielefeld University. In his research, he concentrates on the simulation of complex emotions for embodied agents.

Bild

Stefan Kopp is Assistant Prof. at the AI Group and member of the CRC 673 "Alignment in Communication" at Bielefeld University. His current research interests include cognitive models of conversational behavior in embodied agents.

Bild

Nadine Pfeiffer-Leßmann is a Ph.D. student and research assistant at the CRC 673 "Alignment in Communication" at Bielefeld University, with a special interest in cognitive architectures for cooperative interaction partners.

Bild

Ipke Wachsmuth is Director of the Center for Interdisciplinary Research (ZiF) and Chair of Artificial Intelligence at Bielefeld University, Germany. He is co-initiator and member of the executive committee of the DFG Collaborative Research Center 673 "Alignment in Communication".