Commentary on Eva Hudlicka, "Guildelines for designing computational models of emotions."

# ON GUIDING THE DESIGN OF AN ILL-DEFINED PHENOMENON

Christian Becker-Asano (Freiburg Institute for Advanced Studies, Freiburg University, 79108 Freiburg; christian@becker-asano.de; http://www.becker-asano.de)

**Abstract**: The quest of modeling "emotions" computationally has spurred the interest of a number of researchers from such fields as human-machine interaction, artificial intelligence, and even emotion psychology itself. As Hudlicka correctly points out, over the last 15 years quite many different approaches to modeling emotions have been followed that are based on diverging conceptions which are derived from a number of different theoretical findings and assumptions. Her aim to provide guidance in this complicated endeavor is as timely as it is potentially helpful for the Affective Computing community—at least in principle. In this commentary it will be argued that our conceptual knowledge might still be too limited to allow for providing a technical "how-to guide" for the general task of modeling a phenomenon that is inherently ill-defined. On the contrary, the multitude of existing computational models, which rightly claim to realize this or that aspect of emotions, should encourage scientists of the computing sciences to further explore alternative architectures of emotions, as long as they are derived from or at least relate to findings and theories of the neighboring disciplines such as philosophy, neuro-biology, and emotion psychology. In this respect, the author follows Hudlicka's final argument that "computational modeling [..] has the potential to contribute to the elucidation of the underlying mechanisms of affective processes."

Why is any kind of computational simulation of emotions inside an agent's architecture necessary at all for a virtual or robotic agent? Indeed, following Hudlicka's distinction of "Two Categories of Emotion Roles" (cp. Figure 2 in the original article), it might be sufficient for a roboticist to focus on the interpersonal role of emotions alone. Developing sophisticated means for a robot's multimodal emotion expression is already challenging enough and by realizing this ability appropriately long-term human-robot interaction could already be improved significantly. Of course, the question remains when a robot should express which kinds of emotions?

An engineer's response to this question would most likely be: It depends on the particular task the (humanoid) robot has to fulfill; see also (Hudlicka, et al., 2009). In particular, why would a robot need to express emotions, if it were never to encounter humans in fulfillment of its task? And if it were to serve humans directly, why should it ever need to express negative emotions? Why don't we just supply the robot with a rule set of stimulus-response patterns that matches a number of inputs relevant to the current context onto a predefined set of (multimodal) emotional expressions?

As argued before (Becker, Kopp, & Wachsmuth, 2007), only focusing on a subset of emotional expressions might be sufficient for short-term technical solutions, but if humanoid robots are to fulfill adequate roles in human society not being limited to specific application contexts, a more complete picture of emotional competence needs to be integrated into their computational architectures. Regardless of whether or not such vision of a future, in which robots share our social domain, might ever become reality, a second motivation for integrating deep models of emotions into agent architectures is important: the experimental-theoretical motive. From this point of view, virtual and robotic agents can be used as tools to gain a deeper understanding of inherently human phenomena,

e.g., the fluent combination of verbal and non-verbal expressivity, joint attention as a basis for social learning, and, of course, the affective dimension in social interaction.

Consequently, Hudlicka's aim to provide "guidelines" on this matter should be taken with a grain of salt. For example, the root of her "taxonomy of affective states and traits" (cp. Figure 1 of the original article) is labeled "affective factors", which are supposed to be different from "affective states" further below. The latter, however, don't subsume "emotions" and "moods" in the same figure, although the last paragraph of Section 2.2 suggest exactly this to be the case by stating that "[e]motions, moods, affect and feelings all refer to affective states [..]." In effect, it is not surprising that the proposed taxonomy has no influence on the guidelines later on. At least when aiming at an object-oriented implementation of an agent's emotional cognitive architecture, however, a clear taxonomy needs to be established; see (Becker-Asano & Wachsmuth, 2010; Becker-Asano C. , 2008), for example.

The requests stated in terms of questions to be answered by emotion theories in Section 3.3 are indeed helpful for emotion modeling and should be considered by anyone, who is new to this field. This one-way approach of only using theoretical evidence and theories to guide the design of computational models can be criticized, because it falls short of the high expectations that some researchers of emotion psychology have concerning the potential of devising computational models of their very own theories. In fact, Hudlicka's valuable overview and knowledgeable advice concerning the "Affective Requirements Analysis" (cp. Table 12 of the original article) shows above all, how many slightly different realizations of computational emotion architectures still have to be explored. In this respect, finding a useful method for evaluating the different emotion architectures is another very important challenge for Affective Computing researchers, because at the end of the day computational modeling promises to provide exactly that new means of practical testability.

## References

Becker, C., Kopp, S., & Wachsmuth, I. (2007). Why emotions should be integrated into conversational agents. In T. Nishida (Ed.), *Conversational Informatics: an Engineering Approach* (pp. 49-68). London: John Wiley & Sons Ltd.

Becker-Asano, C. (2008). *WASABI: Affect Simulation for Agents with Believable Interactivity.* Amsterdam: IOS Press.

Becker-Asano, C., & Wachsmuth, I. (2010). Affective computing with primary and secondary emotions in a virtual human. *Autonomous Agents and Multi-Agent Systems, 20*(1), pp. 32-49.

Hudlicka, E., Payr, S., Ventura, R., Becker-Asano, C., Fischer, K., Leite, I., et al. (2009). Social interaction with robots and agents: Where do we stand, Where do we go? *Affective Compting and Intelligent Interaction* (pp. 698-703). Amsterdam: IEEE.