Available online at www.sciencedirect.com
ScienceDirect
Cognitive Systems Research 10 (2009) 1–5
Cognitive Systems RESEARCH
www.elsevier.com/locate/cogsys

Editorial

Modeling the cognitive antecedents and consequences of emotion

Action editor: Ron Sun

Jonathan Gratch [a,*], Stacy Marsella [b], Paolo Petta [c]

[a] Institute for Creative Technologies, University of Southern California, 13274 Fiji Way, Marina del Rey, CA 90292, United States
[b] Information Sciences Institute, University of Southern California, 4676 Admiralty Way, Marina del Rey, CA 90292, United States
[c] Center for Brain Research, Medical University of Vienna, Freyung 6/2, A 1010 Vienna, Austria

Available online 25 June 2008

1. Introduction

The last decade has seen an explosion of interest in emotion in both the social and computational sciences. Emotions arise from interactions with both people and technology. They color human perception and decision making and shape a person's moment-to-moment responses to their social and physical environment. Emotions are expressed through changes in speech, facial expression, posture and physiological processes, and these changes provide essential clues to a person's beliefs, desires, intentions and likely future behavior. Recognizing and exploiting such influences can have broad impact across a variety of disciplines: Incorporating the influence of emotion increases explanatory power of models of human decision making (Loewenstein & Lerner, 2003); Responding to a student's emotions can enhance the effectiveness of human or computer tutors (Conati & MacLaren, 2004; Graesser et al., 2008; Lepper, 1988), and modeling emotional influences can enhance the fidelity of social simulations, including how crowds react in disasters (Lyell, Flo, & Mejia-Tellez, 2006; Silverman, Johns, O'Brien, Weaver, & Cornwell, 2002), how military units respond to the stress of battle (Gratch & Marsella, 2003), and even large social situations as when modeling the economic impact of traumatic events such as 9/11 or modeling inter-group conflicts (Marsella, Pynadath, & Read, 2004).

More generally, an understanding of the cognitive and social function of human emotion complements the rational, individualistic and disembodied view of cognition that underlies most artificial intelligence and cognitive system research. Emotional influences that seem irrational on the surface may have important social and cognitive functions that would be required by any intelligent system. For example, Herb Simon (1967) theorized that emotions serve to interrupt normal cognition when unattended goals require servicing. Robert Frank argues that social emotions such as anger and guilt reflect a mechanism that improves group utility by minimizing social conflicts, and thereby explains people's "irrational" choices to cooperate in social games such as the prisoner's dilemma (Frank, 1988). Similarly, Alfred Mele (2001) claims that "emotional biases" such as wishful thinking reflect a rational mechanism that more accurately accounts for social costs, such as the cost of betrayal when a parent defends a child despite strong evidence of their guilt in a crime (see also Ito, Pynadath, & Marsella, 2008). At the same time, findings on non-conscious judgments (e.g., Barrett, Ochsner, & Gross, 2007; Moors, De Houwer, Hermans, & Eelen, 2005) have enriched our understanding of how cognitive style is shaped by the socio-emotional context, often in adaptive ways. More broadly, appraisal theorists such as Lazarus (1991), Frijda (1987) and Scherer (2001) have argued that emotions are intimately connected with how organisms sense events, relate them to internal needs (e.g., is this an opportunity or a threat?), characterize appropriate responses (e.g., fight, flight or plan) and recruit the cognitive, physical and social resources needed to adaptively respond. Thus, an understanding of emotion's function can inform the design of cognitive systems that must survive in a dynamic, semi-predictable and social world.

This special issue of Cognitive Systems Research gives a cross-section of contemporary psychological and computational research on the interplay of cognition and emotion. The articles arise from a recent interdisciplinary symposium on Modeling the Cognitive Antecedents and Consequences

* Corresponding author.
E-mail addresses: gratch@ict.usc.edu (J. Gratch), marsella@isi.edu (S. Marsella), paolo.petta@meduniwien.ac.at (P. Petta).

1389-0417/$ - see front matter
doi:10.1016/j.cogsys.2008.06.001

of Emotion that brought together leaders in psychological and computational approaches to emotion for three days of intense discussion. The articles represent the current state of an ongoing discussing to bridge the divide between computational and psychological perspectives on emotion, illustrating both that theories on the function of emotion in human cognition can yield key insights into the design and control of intelligent entities in general, and that computational models of human mental processes can inform psychological theories through the exercise of concretizing them into working and testable systems.

2. Background

Emotion research spans an enormous body of work across a wide range of scientific disciplines and contains within it a diversity of competing theoretical perspectives. This special issue emphasizes appraisal theories of emotion (Ellsworth & Scherer, 2003), the dominant psychological theory over the last twenty years and, with its emphasis on cognitive processes, the most congenial to cognitive systems research. Before describing the contributions to the special issue we first review differing theoretical perspectives and recent progress on computational models of emotional processes.

Although there is no consensus on the definition of the term *emotion,* the following proposal by Scherer is illustrative: emotions are episodes of synchronized recruitment of mental and somatic resources allowing an organism to adapt to or cope with a stimulus event subjectively appraised as being highly pertinent to its needs, goals, and values (Scherer, 2004). This definition emphasizes the alignment of several distinct components including cognitive processes (e.g., appraisal), physiological processes (e.g., ANS arousal), behavioral tendencies and responses (e.g., facial expressions). Other theorists also emphasize the importance of awareness (e.g., "I feel mad!"). Most research treats emotion as short-term and changeable in contrast with longer term moods or dispositional tendencies (e.g., personality). Some theorists also emphasize that emotions are intentional in the sense that they make reference to a specific entity or situation. For example, one may be in an angry mood (a non-intentional state) but to experience the emotion of anger, one must be angry at something.

Theories differ in which components are intrinsic to an emotion (e.g., cognitions, somatic processes, behavioral tendencies and responses), the relationship between components (e.g., do cognitions precede or follow somatic processes), and representational distinctions (e.g., is anger a prototype or a natural kind). For example, *discrete emotion* theories argue that emotions are best viewed as a set of discrete sensory–motor programs (Ekman, 1992; LeDoux, 1996; Öhman & Wiens, 2004). Each of these programs consists of a coherent brain circuit that links eliciting cognitions and somatic responses into a single neural system. At the other extreme, *dimensional* theories (e.g., Russell, 2003) argue emotions are simply cognitive labels we apply retrospectively to sensed physiological activation, which, rather than consisting of discrete motor programs, is characterized in terms of broad bipolar dimensions such as valence and arousal (e.g., I feel negative arousal in a context where I have been wronged, therefore I must be angry). Each of these perspectives has merit and its own body of empirical support and it remains an open challenge to reach an overall synthesis (see Parkinson's article in this special issue).

Since the 1980s, appraisal theories have become a major theoretical perspective in the study of emotion and the dominant contemporary theory underlying computational models of emotional processes. Appraisal theories emphasize the cognitive antecedents of emotional experience. The central tenant of appraisal theories is that the organism's evaluation of its circumstances plays the primary role in eliciting and differentiating emotional responses. Appraisal theories posit a set of discrete judgments, called appraisal variables, which characterize the impact of events (real or imagined) on the organism's beliefs and desires. Some of these proposed variables include pleasantness, expectedness and coping potential. According to appraisal theories, these judgments largely determine the organism's emotions and behavioral responses. For example, an unexpected negative event may provoke fear and a tendency to freeze or run away. In this sense, appraisal theories resemble the discrete emotion perspective in proposing a coherent linkage between elicitation (in terms of appraisal) and somatic response, but they differ in claiming a far richer and more flexible mapping between elicitation and response that better captures the subtlety and richness of human emotion.

Although individual appraisal theories differ in terms of their posited appraisal dimensions and their process assumptions, computational models of emotion have been most influenced by the appraisal theory of Ortony, Clore, and Collins (1988), chiefly as it is described with a clarity that can be readily translated into a computer program. Clark Elliott's Affective Reasoner was the first attempt to realize this theory (Elliott, 1992). Most subsequent computational approaches have focused on the appraisal component of emotion, proposing more general and comprehensive techniques for deriving appraisal variables from a representation of perceptions, knowledge and goals. For example, whereas Affective Reasoner used hand-crafted rules (e.g., during a football match, a goal scored by my opponent is *undesirable*), the subsequent EM system (Neal Reilly, 1996) divided appraisal into general reasoning mechanisms that operated over domain-specific knowledge structures.

One unfortunate consequence of this reliance on Ortony, Clore and Collins' model is that it focused most concretely on the cognitive *structure* of emotions (i.e., appraisal dimensions) but not the overall emotion process, and the resulting computational models reflect this narrowness. Other appraisal theorists, such as Richard Lazarus (1991), Nico Frijda (1987) and Klaus Scherer (2001), proposed more comprehensive theories that not only encompassed a wider