A Concern-Centric Society-Of-Mind Approach To Mind Design

Steve Allen[1]

Multi-Agent Systems Group,
German Research Centre for Artificial Intelligence,
Stuhlsatzenhausweg 3, D-66123 Saarbrücken, Germany
Email: Steve.Allen@dfki.de

Abstract

In this poster summary, we argue that mental concern-processing mechanisms are amenable to a society-of-mind approach to mind design. We illustrate our case with an information-level analysis of the emotion process, relating the different classes of emotional state to the different layers of our motivated agent framework. We describe how a society-of-mind design-based implementation strategy allows us to add depth to our agent architecture, and incrementally account for more and more of the phenomena of interest. Finally, we report on the results of recent research into the design of cognitively-inspired emotional agent architectures.

1 Introduction

Concerns are broadly defined as dispositions to desire the occurrence, or non-occurrence, of a given kind of situation [Frijda 86, page 335].

Not all concern processing mechanisms need explicit representational forms or structures (as some are emergent), but they do need a systematic framework within which they can be described and operate. In this summary we will use our motivated agent framework (Figure 1) to briefly elucidate the concern-processing mechanisms inherent in the human emotion process.

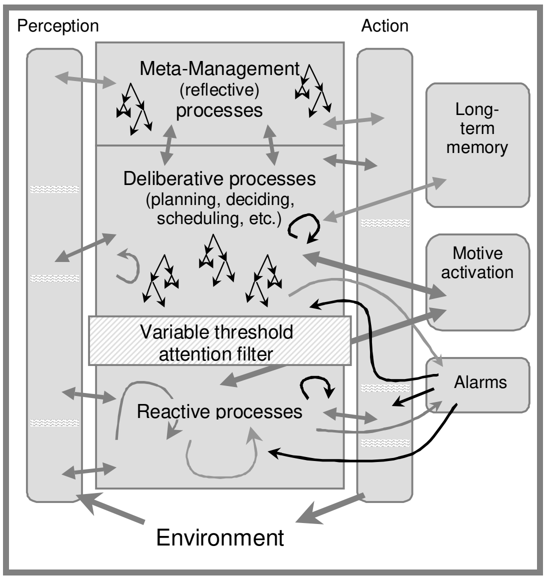

Figure 1 Motivated Agent Framework [Sloman 99]

2 Emotional States

By referring different definitions and theories of emotion to the different layers of the motivated agent framework, we can identify three main classes of emotional state [Sloman 99] – *primary*, *secondary*, and *tertiary*.

Primary emotional states: such as being startled, terrified, or sexually stimulated, are typically triggered by patterns in the early sensory input and detected by a global alarm system.

Secondary emotional states: such as being anxious, apprehensive, or relieved, depend on the existence of a deliberative layer in which plans can be created and executed with relevant risks noticed, progress assessed, and success detected. An alarm system capable of detecting features in theses cognitively generated patterns is still able to produce global reactions to significant events in the thought process [see also Damasio 94 and Picard 97].

Tertiary emotional states: such as feeling humiliated, ashamed, or guilty, can be further characterised by a difficulty to focus attention on urgent or important tasks. These emotions cannot occur unless there is a meta-management layer to which the concept of “losing control” becomes relevant.

The three different classes of emotional state should be seen as orthogonal to the common emotion type labels used in everyday language. For example, fear can take the form of a primary, secondary, or tertiary emotion. Each class of emotional state has its own physiological characteristics and hedonistic tone, further underlining the futility of talking about emotional states as active states of a discrete “emotion” system (or systems).

[1] In collaboration with the Cognition and Affect project at Birmingham University.

3 Society-of-Mind

Emotional states are best viewed as an emergent phenomena arising from the interaction of a number of different systems and cognitive processes (only some of which are specific to the generation of emotional states). We can start to make these systems/processes more explicit by mapping their abstract information-processing representations onto our motivated agent framework. This mapping process is performed within the context of Frijda's [86] emotion process, resulting in a generalised design for an emotional agent (see Figure 2).

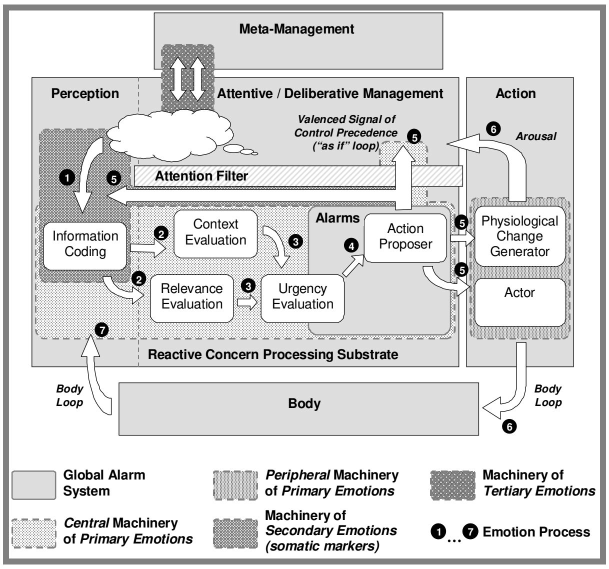

Figure 2 – An Information-Level View of the Emotion Process

Having established an abstract design for an emotional agent (noticeably devoid of an "emotion" module), we can now start to refine the architecture through our design-based research methodology – building a series of complete broad-but-shallow implementations of our design to incrementally cover more and more of the phenomena of interest.

We are able to capitalise on the society-of-mind design philosophy by adding depth to our agent designs through the addition of new specialist members within the existing society-of-mind architecture. Furthermore, drawing inspiration from the fields of neurology, we started to map these information-level agents onto regions of the human brain [LeDoux 96, Damasio 94]. For example, Allen [2000] describes the design for an emotional society-of-mind agent architecture, based on earlier work by Cañamero [97] and members of the Cognition and Affect project at Birmingham University [Beaudoin 94 and Wright 97].

4 Conclusions

In this brief poster summary, we have tried to give a flavour of the concern-centric society-of-mind approach we advocate for mind design. Although we have focussed on a single aspect of mind, that of the emotion process, our approach is general enough to apply to other mental phenomena.

Acknowledgements

The author would like to thank Aaron Sloman and members of the Cognition and Affect project at Birmingham University, and the Multi-Agent Systems and Intelligent User Interfaces groups at the German Research Centre for Artificial Intelligence, for their active support of this research.

5 References

Allen, S. (2000). *Concern Processing in Autonomous Agents.* Submitted PhD Thesis. School of Computer Science, University of Birmingham.

Allen, S. (1999). Control States and Motivated Agency. In E. André (Ed.) *Behavior Planning for Life-Like Characters and Avatars: Proceeding of the i3 Spring Days '99 Workshop*. pages 43-61.

Beaudoin, L. (1994). *Goal Processing in Autonomous Agents*. PhD Thesis, School of Computer Science, University of Birmingham.

Cañamero, D. (1997). Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In *Proceedings of the First International Symposium on Autonomous Agents, AA'97*, Marina del Rey, CA, February 5-8, The ACM Press.

Damasio, A. R. (1994, 96). *Descartes' Error: Emotion, Reason and the Human Brain*. London: Papermac. (first published 1994, New York: G. P. Putman's Sons.)

Frijda, N. H. (1986). *The Emotions*. Cambridge: Cambridge University Press.

LeDoux, J. E. (1996). *The Emotional Brain: The Mysterious Underpinnings of Emotional Life*. New York: Simon and Schuster.

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

Sloman, A. (1999). Architectural Requirements for Human-like Agents Both Natural and Artificial. (What sorts of machines can love?). In K. Dautenhahn (Ed.) *Human Cognition And Social Agent Technology*, John Benjamins Publishing.

Wright, I. P. (1997). *Emotional Agents*. PhD Thesis, School of Computer Science, University of Birmingham.