A Primitive Emotion and Its Cooperative Function Simulated in Neural Networks: Towards a Theory of Emotions as Cognitive Functions

2006/03/12

N. Nagataki, H. Tsukimoto, M. Shibata, H. Hattori, H. Ito, T. Mino, A. Shimojima, N. Shinohara, and T. Kashiwabata

Emotions are usually considered to be states with specific qualia resulting from various physical or mental conditions, without any particular functions other than that of disturbing normal cognition. However, as recent studies in many aspects of cognition show, emotions have their own unique functions contributing to survival of cognitive agents. We have engaged in the research of clarifying what causal roles emotions play and how they relate to other cognitive mechanisms when agents tackle various intellectual and affective tasks. For that purpose, we have deployed two working hypotheses: multi-dimensionality and computational intractability of emotions.

As Frijda rightly argues, the core of an emotion is readiness to act in a certain way. It should then be noted that there are many ways of achieving readiness to act in various situations. For emotions have many layers operating differently in different cognitive tasks, which we call multi-dimensionality of emotion. In this research we focused on a function of emotion which seems to belong to the most primitive layer of emotional cognition, *i.e.* a cooperative function, and we simulated its working in a hypothetical situation of people hunting dangerous animals together. The cooperative function, being among rather elementary ones, is, as we take it, a basis of many higher-order emotional states, such as joy, sadness, anxiety, fear, disgust, or even moral sentiments. The aim of our research in this regard is to verify that agents having acquired the cooperative function have a primitive form of emotion by showing that they can solve certain cognitive problems more easily than those without acquiring it.

Our second hypothesis is that emotional functions cannot be computationally tractable, that is, that there are no effective ways of programming them in classical AI systems. Therefore we employed instead a simulation model by means of neural networks, and besides that, we used genetic algorithms dealing with weight parameters of neural networks. The cooperative function was confirmed to emerge through evolution on this setup. More specifically, through the games of hunting dangerous animals together, the networks at a certain generation acquired the cooperative function. The point of simulating emotions on neural networks is that it can provide biochemical and causal solutions to many computationally insoluble cognitive problems. Then as a paradigmatic problem, we posed to our agents (*i.e.* our neural networks) the prisoner's dilemma, which we believe does not allow any simple rational solutions other than those based on the altruist practical considerations from an evolutionary perspective.

We will show how well our agents can cope with the prisoner's dilemma, and make some philosophical and psychological remarks about what emotions are in general from our functional viewpoint.