1. Introduction

If there are two things that are most common to all people in the world, they are intelligence and emotion. Intelligence allows us to make complex decisions in a complex world; it governs everything that we do, big or small, and it does so without our having to understand why it works. Emotion is very much the opposite. It "clouds our judgment", drives people to make irrational choices, and makes us constantly wonder how we can control that which seems uncontrollable. Or at least this is how these two mental forces have been viewed and portrayed since the Romantic Era [7,8]. Recently, however, this view has changed quite dramatically.

Despite intelligence and emotion being viewed as almost opposites for most of history, they are now viewed more as partners in the intricate dance that is human cognition [3,6,7]. Just as in a dance, one of the two may typically lead the other, as many would suggest is true of intelligence generally leading emotion, but no dance is possible without both. Genuine human cognition is only possible when you have both intelligence analyzing the relevant information and emotion helping guide the way when intelligence is not enough [7]. Although there is not a clear need for emotion to co-occur with intelligence, it is clearly so in human beings and is therefore the only way that we can explore both.

In this paper I will very briefly survey the current work in the study of emotion and the different answers that people have thought of to answer the questions of "how does emotion work?" and "what does emotion mean?". I will also give a brief history of the pursuit for artificial intelligence as background. After surveying these areas, I will give a brief review of the current applications of emotion to artificial intelligence research and make general comments regarding issues with this work; these comments are not meant to be at all comprehensive critiques of these works since each critique could be a paper in and of itself. The purpose of these brief comments is meant to highlight the flaws of the "how?" and "what?" approach to understanding emotion as an addition to intelligence. The paper will wrap up with a brief discussion of the problems that are evident in all current work mixing emotion with cognition and issues that I believe need to be addressed in this area in general.

[…]

3.2.5 Improved computation with improved understanding

The final group of papers utilizes both emotional understanding to improve computational systems and also makes an effort to improve knowledge about emotions and what part they play in intelligence. The first of these papers uses an emotional extension to a learning system as a means of helping decide between various possible responses to different situations. The second paper takes a computational stance in order to explore possible advantages that emotions give for adaptability and autonomy.

Sandra Gadanho's work, [5], is a good example of attempting to improve computational systems and also makes an effort to improve knowledge about emotions and what part they play in intelligence. The system discussed in this paper uses a rule-

based learning system with a neural network at its core. The system has a set of predefined actions that it can undertake and a set of predefined goals. Emotional modification of this learning system is done by adding homeostatic variables that give rise to emotion based on the difference of these variables from their optima. The purpose of these emotional responses to change in the homeostatic variables is to give an indication of the "goodness" of the decisions that were made which lead to that change. This is very similar to the first work discussed in 3.2.4 which used "emotion-like" analyses of program behavior to store "goodness" of choices that lead to that computation. The use of homeostatic variables as the means of generating emotion is based in the neurophysiological work done by Damasio [1,2]. The use of emotion in this model attempts to improve the understanding of how emotion may affect cognitive decision-making by comparing several versions of the existing cognitive system with and without the addition of an emotional component.

The experiment discussed in this paper is similar to experiments run on other systems that have been discussed here and appears to be the quintessential way of justifying the use of emotion with cognition; provide an agent with a problem and give them many possible solutions or conflicting goals and see how the introduction of the "emotion" to the cognitive system improves performance [7]. Here, the task is for a robot to find light sources within a maze so that it can replenish energy that is constantly draining. The robot must make choices between avoiding obstacles, following the walls, or seeing out light sources[1] while it tries to maintain three homeostatic variables: energy, welfare, and activity. The results that come out of the introduction of emotion into the existing system gives rise to the same questions that have come up with the other work discussed here: what justification is there for the decisions that were made relating to the "emotion" used in this system?

The biggest issue relating to justification is based on the way in which the emotional system was added onto the learning system. Qualitatively, the system uses emotion as a means of gauging the quality of action choices. However, functionally the system uses the emotional system as a means of interrupting and guiding the cognitive system. This seems quite backwards based on the evolutionary view of how emotion and intelligence developed. Since we see things that are very reminiscent of emotion in a wide variety of animals that have significantly less advanced cognitive abilities than humans do, it is not unreasonable to argue that the emotional system we have is either driven by or derived from an instinctive system that developed a long time ago. Given this, wouldn't it make more sense to develop a system that uses cognitive processes to modify decisions that originate in an emotional system? Or perhaps this common view creates a false dichotomy between which of the two components modifies the other. Perhaps there is a third component, instinct, that is the prime originator of motivation that is then fed into both emotion and cognition which then influence each other? This is all speculation with the point of suggesting that the existing view of emotional modification of plans originating in cognitive processes appears to be inconsistent with explanations of the evolutionary development of the brain.

[1] Why these actions are to be considered mutually exclusive is unclear but is an important part of the justification for introducing “emotion” to this system.

Even if you ignore the issues given about the order of interaction between emotion and cognition, there is still an issue in the choice of homeostatic variables and their lack of grounding. The argument behind homeostatic variables being the driving force behind emotions is that it was evolutionarily advantageous to be able to sense and respond to changes in important, internal variables[2]. However, the variables and the responses to their changes have been developed over millions of years of evolution. Making unjustified decisions about what homeostatic variables to give an agent and what responses they are allowed to make in response to changes in those variables removes this grounding. When the grounding is removed, the computation done by this "emotion" system (and the systems in the other papers as well) seems no different than any other computation done by a purely cognitive system. As such, it is not clear from the results of the experiment or from the basis of the model itself why this model has any emotional components.

[…]

Conclusion

Although done in different ways, each research program discussed in this paper made some claim at improving the performance of an existing system by introducing an additional system that they claimed was emotion. But is it enough to use emotion-based words such as "happy" and "afraid" to justify that a process duplicates emotion? It is quite apparent from the existing literature that there is no real consensus as to what emotion inherently is. It is not apparent where it emanates from in the brain, what role it plays in cognition, and what makes it so easy for other humans to identify but so hard for them to describe; it isn't even clear that the things we label as emotion are actual a coherent group [3,4]. Given all of these problems, the idea that emotion can be duplicated outside of a human being seems naive at best and arrogant at worst. The one exception to this is the use of computational models of emotion for improving theories of emotion. However, all that these systems can hope to do is emulate certain responses that are consistent with responses that are attributed to emotion in human beings. Until there is a better idea of the biology behind emotion, it seems unlikely that there will be any significant models of emotion that can be extrapolated to computational systems.

References

[1] Damasio, A. (1994). Descartes Error – Emotion, Reason and the Human Brain. New York: Gosset/Putnam Press.

[2] Damasio, A. (1999). The Feeling of What Happens: Body and Emotion in the Making of Consciousness. New York: Harcourt Brace.

[3] DeLancey, C. (2002). Passionate Engines. Oxford: Oxford University Press.

[4] De Sousa, R. (1987). The Rationality of Emotions. Cambridge: MIT Press.

[5] Gadanho, S. (2003). Learning Behavior-Selection by Emotions and Cognition in a Multi-Goal Robot Task. *Journal of Machine Learning Research, 4, 385-412.*

[2] It is well supported by biology and other scientific fields that some notion of grounding is required in any attempt to duplicate or understand human emotion [3].

[6] Gratch, J. (1999). Why You Should Buy an Emotional Planner. *Proceedings of the Agents '99 Workshop on Emotion-based Agent Architectures*.

[7] Minsky, M. (1986). The Society of Mind. New York: Simon & Schuster.

[8] Pinker, S. (1997). How the Mind Works. New York: W. W. Norton & Company.