Igor Napolskikh

Professor Paul Thagard

PHIL 256

March 15, 2012

THE ROLE OF EMOTIONS IN HUMAN COGNITION AND ROBOTICS

The Issue

Emotions have puzzled and fascinated thinkers throughout history, from Plato and Aristotle to modern cognitive psychologists. Until relatively recently, the widely accepted view on emotions has been that they are only an interference on the neutral behaviour of a person. However, there has been a rapid increase in the study of the effects of emotion on cognition, popularizing the view that an adequate theory of cognition that ignores emotion is likely to prove inadequate (Eysenck, Keane, 2009). In this paper, the question of the importance of emotions in human cognition and the potential of application of emotions to robotics are discussed.

Alternatives

To understand the full impact of emotions on cognition, one must first look at the nature of the concept to define what it is to have emotions. The main theories among cognitive psychologists on emotional phenomena can be generally categorized into two groups, one focusing on the physiological aspects of emotions and their effect on bodily reactions (James-Lange theory)(Damasio, 1994), and the other based on a person's general state, that emotions are primarily judgments, evaluations, and thoughts (Lazarus' appraisal theory) (Thagard, 2005). A third recently emerging category is a hybrid of the two perspectives, arguing that bodily responses are central to emotions but emphasizes the importance of concepts and judgments in emotional experience. Other views include a focus on the social account of emotions, their importance in the workings of a society and on social interactions. Further separation of perspectives involves the abstraction levels at which emotional experience is defined as, how much molecular interaction plays in the more general experience of emotions and whether a more functionalist perspective of emotions is adequate.

A common conclusion that can be reached from the physiological perspective of emotions, tracing its origins as far back as Plato, is that emotions are unnecessary and insignificant. Since they often result in a conflict between logical reasoning, they must be forcibly channeled through reason and logic to result in any beneficial value. This view extends further into a more negative approach to emotions, that their unpredictable nature often interferes with people's way of life (Gadanho, 1999). For example, those with clinical depression, anxiety disorder, or bipolar disorder are hampered by their emotions in performing everyday tasks. For this reason, they must be treated with drugs to force them out of such heightened emotional states into a more neutral state; with drugs that may have their own side effects and health consequences. This view implies that human computation has little to nothing to do with emotions and that they are mainly an obstruction to reason. From this perspective, it would be unnecessary to include any sort of emotions in the design of robotics, as it would only be a hindrance to the overall performance of the robot.

On the opposing side, supported by the appraisal theory, is a perspective that link emotions to goal achievement. This perspective argues that emotions have a significant influence on decision-making, by providing the necessary focus and readiness to action required for this behavior (Thagard, 2005). Moreover, emotions provide a summary appraisal for a given situation which acts as an integral part of solving complex problems, which could include multiple potential goals (2005). From a functionalist perspective, this implies that the study of the roles emotions play in cognitive processes could assist the fields of artificial intelligence and robotics in the creation of more robust systems. Problems where such systems could benefit in terms of efficiency and effectiveness include dealing with multiple goals and constraints, rapidly changing environments, and social interactions.

There are two other two positions on the involvement of emotions in computation and the applicability of emotions to robotics (Thagard, 2005). The first states that emotions are not computational, and therefore can only be modeled, thus having no place in robotics. The other focuses on the uniqueness of the brain, and states that emotions arise from a particular computation of the brain, thus, cannot be replicated on other systems. The second view is much more likely than the first, and has more in common with a functionalist view of appraisal theory mentioned above. The difference can be seen as the required similarity between another system in the brain for the production of emotions.

Evidence

From the appraisal and hybrid theory of emotions, emotions are very important in human cognition as they are heavily involved in the way we interpret information and make decisions towards a chosen goal. Different emotions serve as guides to the way we should manage our actions, given the uncertainty which surrounds the consequences of such actions (Oatley, 1992). By investigating the positive functions of emotions, we can hypothesize the reason for the evolution of these mental features and the potential survival value of emotions that have contributed to our cognitive behaviour. There is a substantial amount of evidence in support of the view that basic emotions are related to adaptive biological, motivational, and social processes that have functional significance for an organism (LaFreniere, 2011). In fact, emotions should be viewed as adaptive mechanisms that are often excessively relied upon as a function of basic learning processes, rather than as a result of random or disordered biochemical processes (Rasmussen, 2010).

In the field of robotics, artificial intelligence plays a vital role in determining a robot's effectiveness, especially in dynamic and uncharted environments. A robot must be able to efficiently gather and process information to make conclusions about its current state and available options, and then choose the most suitable option that will further its progression towards its goal. These tasks involve trade-offs that limit the robot's robustness as a sacrifice for more reliable performance. Allocated memory is at odds with the speed at which that memory can be processed, and engineers and researches often find themselves having to make difficult decisions in planning for the unknown by choosing one over the other. In most practical cases, the robot has limited amounts of processing power, memory, and electrical power that can be spared for accomplishing certain action-planning and execution within a given time frame. The most efficient method to manage all of these actions is therefore highly desirable and as a result, a vast amount of research has gone into optimizing the performance of robots in various circumstances.

Evidently, it would be advantageous for a robot to have a higher level of abstraction akin to the human layer of emotions in order to manage the actions of a sophisticated system in dynamic environments (Gadanho, 1999)(Parsi & Petrosino, 2010). The successful management of these actions would depend on the current state of the robot (e.g. available power status, where its joints may be located, how far it is from its current goal destination) and the inputs it receives to describe its environment. Although there are a variety of methods to manage these tasks, the most common being Bayesian theory, they often fall short of the complexity, adaptability and efficiency of their human counterpart. Explorations of the potential of implementing emotions into robotics have already taken place for a few decades (Gadanho, 1999)(Fellous & Arbib, 2005)(Hanson, 2009)(Morse, Herrera, Clowes, Montebelli & Ziemke, 2011).

Supporting appraisal theory, numerous studies have been performed that relate to the parts of the brain most likely responsible for the phenomena of emotions to a participant's ability to make decisions (Gupta, Koscik, Bechara & Tranel, 2011). The amygdala has been shown to perform a primary role in the processing and memory of emotions, and is especially important for decision-making, by triggering autonomic responses to emotional stimuli, including monetary reward and punishment (2011). Specifically, patients who have suffered damage to their amygdala lack autonomic responses to reward and punishment, and cannot utilize so called "somatic marker" type cues to guide future decision-making (Gupta, Koscik, Bechara & Tranel, 2011). Numerous studies have been done to analyze the cognitive consequences of deficits in emotional signaling (somatic states). Lesions in the ventromedial prefrontal cortex or amygdala, have been shown to lead to poor judgment in decision-making in personal and social realms (Bar-On Tranel, Denburg & Bechara) and adaptive decision-making for gains and losses (Weller, Levin, Shiv & Bechara, 2007). Experiments involving a functional MRI indicate that emotion and higher cognition can be truly integrated in brain functional organization (Gray, Braver, Raichle, 2002).

Conclusion

Emotion and cognition are commonly thought of as two independent processes that frequently, but not always, beneficially interact. The subjects of emotions, feelings, and cognitive processes are often ill-defined in discussions of a general social setting. However, it is much more valuable to study them as a coherent, functional concept that can be talked about in ordinary language (Oatley, 1992). Such a point of view could provide conclusions that benefit the understanding of emotions from a psychological and philosophical perspective, and applications that promote the fields of artificial intelligence and by extension, robotics.

From a functionalist perspective, a sophisticated robot could (and perhaps should) be able to exhibit system patterns that resemble some of those discussed above. An argument could be made that this would be a necessity in order to have a highly efficient autonomous robot able to survive with limited resources in an ill-defined environment. As robots become more sophisticated and begin to traverse more unpredictable environments with greater independence, they must be designed to better adapt to new situations and quickly make decisions in such an environment.

A more specific example involves robots placed in a situation where their survival hinges on their ability to act independently from tele-operators. These robots must be able to adapt to their environment quickly and effectively, changing the way they motivate their actions towards a given goal and likely altering the sub-goals of a given overall goal or constraint as well. A robot encountering danger might disregard its previous goals of conserving battery life and give priority to evacuating the area, thus satisfying the overall goal of staying alive. Consider the hostile environment of the surface of Mars, where autonomous mobile robotic exploration is taking place. The entire process of sending a single robot to the surface of Mars on a space transport consumes an extraordinary amount of time and resources, thus, the robot's survival and its ability to function properly are high priority goals. The robot chosen for the mission must be well equipped with tools that would aid it in dealing with the many unpredictable situations that could arise during the mission.

Modern technology is evolving to allow autonomous robots to emulate the processing of information and make decisions in ways that are increasingly similar to those of human brains. Examples include Field Programmable Gate Arrays (FPGA), which are able to process information in parallel (similar to how the brain functions) rather than serial processing (how most average computers function)(Maxfield, 2004). Various hardware designs can be uploaded to these units that will “rewire” the internal components for specific tasks. Similarities can be drawn between the operability of these units and the neural networks of the brain. In fact, there are already studies being done on using parallel hardware, including FPGA’s, to simulate the computational, representational and dynamics aspects of neural activity (Eliasmith, Anderson, 2004).

To pursue this parallel into practical applications, one could implement FPGAs or similar configurable parallel processing units on a robot to have a more sophisticated system of hormonal adjustments based on emotions. This could be done by allowing the hardware of the robot to be rewired by a supervisory unit when required, based on past states of emotions, the associated actions and outcomes, and the environment. The changes on the micro scale that occur over time due to certain hormones being exhibited more frequently would therefore be analogous between the robot’s FPGA gates and a human’s neural network neurons. Furthermore, this method has the potential to account for a wide range of phenomena that been used to support physiological and appraisal accounts of emotions (Dalgleish & Bramham, 1999). The method involves a supervisory process which is able to judge which FPGA design configuration is most appropriate in the current situation, how to modify a design to make it more appropriate in their situation, and reflect on the effectiveness of previous designs. Judgments of the design to be used would be made based on what state the robot is currently in at the lowest abstraction level (e.g sensory inputs, such as camera images, temperature sensors, GPS location, battery level, etc.) and the interpreted state from these inputs (e.g. robot is safe from potential harm, robot has enough power to complete a goal). A robotic operating system such as the Robotic Operating System (ROS) developed by Willow Garage would make this implementation seamless, by allowing a more modular approach to the organization of the system (Quigley et al, 2009). The communicative framework of ROS would further allow for easy communication of emotional states between the robots for easier management and coordination of the robots. The modular and open source approach to ROS would also allow for a widespread implementation of algorithms developed for its utilization. This could pave the path for a standard emotional framework of robotics, used by both industry standard robots and lower budget recreational robots.

To demonstrate the idea behind such a system, let's go back to the hypothetical example of robots operating on the surface of Mars. A colony of robots have been sent for a mining expedition, to autonomously gather and study minerals, asynchronously reporting results of their findings back to Earth at a telecommunications base a few times a day. The robots are equipped with overall management system, tracking the goals and constraints of the robots, such as which minerals should be sought after, excavated and analyzed, and the operational state of the robot based on sensory input (such as battery level and robot GPS location). An emotional framework on each robot determines which emotional state should be currently active and determines the types of actions a robot will execute based on its emotional state and the current circumstances. If one of the robots were to suddenly venture into an area that is well-known to have dangerous sandstorms, it would focus more processing power on inputs from the visual systems to detect an incoming storm. The power management and distribution of processing units would be based on the design of an FPGA, at the lowest hardware level of the robot. The robot would switch its emotional state from enjoyment/interest to surprise if certain sensor thresholds were reached to indicate sudden environmental changes, such as increasing wind speeds and visual gusts, causing more processing and electrical power to be reactionarily distributed from analysis tools to the motors and visual sensors. If analysis of the patterns of the sensory changes by the robot resulted in a judgment of the situation as dangerous, the robot would then go into an emotional state of terror. Immediately, the robot's reaction would be to signal out to the other robots of the danger and begin the planning and execution of an escape route of the area. Once out of reach of its potential danger, the robot's emotional system would then reflect on the effectiveness of its actions. Analysis would be done on the data to determine whether the conditions were really dangerous (i.e. whether a storm was really present or if this was a false-positive) and whether the system's hardware reactions were appropriate (e.g. did the robot become damaged because it didn't react quickly enough, had the robot overreacted and lost too much battery power in the process). The results of this analysis would then be used to determine whether to modify the hardware design responsible for the emotional reaction (e.g. an FPGA board) and how to optimally modify it for future situations. A robot who always cried wolf in a situation and wasted other robots battery power when forced out of an area due to potential danger might be reprimanded into shame for being too cowardly in the given circumstances. The emotional system of that robot would then change the way judgments by the robot are made based on visual sensory data of the storm at the software level, rather than at the lower abstraction layer of hardware. This example demonstrates the potential emotional system resembling the hybrid theory of human emotions as implemented on robotics systems. On a physiological level, the robot's change in the arrangement of the logic gates in its FPGA implementation is similar to a human's hormonal change in the release of various neurochemicals (i.e. noradrenaline, dopamine, and serotin) when switching between emotional states. On a judgmental level, the robot's change in confidence thresholds for detecting a sandstorm is akin to a person changing their perception of danger based on previous experiences.

In closing, emotions have thus been shown to be an integral aspect of human cognitive processing and the link between the two should be studied further to allow for development of more robust robotic systems. The seemingly distant gap between reality and the science fiction worlds of emotional robotics like *Blade Runner* and *WALL-E* seem to be closing. Although there is still a long way to go before robots have acquired a full range of human-like emotions, there has already been progress made with great opportunity for future research. Robotics engineers and cognitive psychologists will undoubtedly be soon joining forces in greater numbers to create the future of emotionally governed frameworks of reasoning.

References

Bar-On, R., Tranel, D., Denburg, N. L., & Bechara, A. (2003). Exploring the neurological substrate of emotional and social intelligence. *Brain: A Journal of Neurology*, *126*(8), 1790-1800.

Dalgleish, T. & Bramham, J. (1999). Cognitive Perspective. In David Levinson, James J. Ponzetti, Jr., & Peter F. Jorgensen (Eds.), Encyclopedia of Human Emotions, pp. 118-121. New York: Macmillan.

Damasio, A. R. (1994). Descartes' error: Emotion, reason, and the human brain. New York: G. P. Putnam's Sons.

Eliasmith, C., & Anderson, C. H. (2004). *Neural Engineering: Computation, Representation, and Dynamics in Neurobiological Systems*. Cambridge, Mass.: MIT Press.

Fellous J., Arbib M. A. (Eds.), (2005). *Who needs emotions?: The brain meets the robot*. New York, NY, US: Oxford University Press.

Gadanho, S. C. (1999). "Reinforcement Learning in Autonomous Robots: An Empirical Investigation of the Role of Emotions." University of Edinburgh, College of Science and Engineering, School of Informatics.

Gray J.R, Braver T.S, Raichle M.E. (2002) *Integration of emotion and cognition in the lateral prefrontal cortex.* Proc Natl Acad Sci USA. 2002;99: 4115–4120.

Gupta, R., Koscik, T. R., Bechara, A., & Tranel, D. (2011). The amygdala and decision-making. *Neuropsychologia, 49*(4), 760-766.

Hanson, D. (2009). David Hanson: Robots that "show emotion". [Video file]. Retrieved from http://www.ted.com/talks/lang/en/david_hanson_robots_that_relate_to_you.html

LaFreniere, Peter. (2011) *Adaptive Origins: Evolution and Human Development*. New York: Psychology Press

Maxfield, C. (2004). *The design warrior's guide to FPGAs: devices, tools and flows*. Boston: Newnes

Morse, A. F., Herrera, C., Clowes, R., Montebelli, A., & Ziemke, T. (2011). The role of robotic modelling in cognitive science. *New Ideas in Psychology, 29*(3), 312-324.

Oatley , K . (1992). The Structure of Emotions. Best Laid Schemes : The Psychology of Emotions. Cambridge, England: Cambridge,University Press.

Parisi, D., & Petrosino, G. (2010). Robots that have emotions. *Adaptive Behavior, 18*(6), 453-469.

Quigley, M., Gerkey, B., Conley, K., Faust, J., Foote, T., Leibs, J., Berger, E., Wheeler, R., & Ng, A. (no date) “ROS: an open-source Robot Operating System”. Stanford University, University of Southern California. Computer Science Department.

Rasmussen, P. R. (2010). *The quest to feel good*. New York, NY, US: Routledge/Taylor & Francis Group.

Sakaki, M, Gorlick, M. A., & Mather, M. (2011). Differential interference effects of negative emotional states on subsequent semantic and perceptual processing. *Emotion, 11*(6), 1263-1278.

Thagard, P. (2005). *Mind: Introduction to Cognitive Science* (2nd ed.). Cambridge, Mass.: MIT Press.

Weller, J. A., Levin, I. P., Shiv, B., & Bechara, A. (2007). Neural correlates of adaptive decision making for risky gains and losses. *Psychological Science, 18*(11), 958-964.