Feeling Fabricated: Artificial Emotion

By: Andrew Paulin

For: Dr. Anna Greenspan
MMEDIA 3B03

November 23, 2003

From robot companions such as the Furby to recent films like A.I., the concept of machines that operate on an emotional level has become increasingly popular in modern culture. The combination of the artificial and something intrinsically natural such as human emotion challenges traditional machinistic views that still persist today. But, machines that act like humans are not a new idea. Computing pioneer Alan Turing states the focus of his 1950 article “Computing Machinery and Intelligence,” in the first line: “I propose to consider the question ‘Can machines think?’” The Turing Test is meant to determine if a program has intelligence. This is how it works: An interrogator is connected to one person and one machine via a terminal, and therefore can't see her counterparts. Her task is to find out which of the two is the machine, and which is the human, only by asking them questions. If the machine can "fool" the interrogator, it should be considered intelligent (Turing).

This test has been subject to much criticism, and has been at the heart of many debates in artificial intelligence, philosophy and cognitive science for over 50 years. Regardless of its validity, the Turing Test illuminates some relevant insight to the discussion at hand. On one level, machines that know human behaviour and can socially interact with humans in a natural manner are of considered superior intelligence. Machines that are able to understand, respond to and generate emotions in humans, and are able to behave according to social conventions, will be able to communicate more effectively, and will be perceived as intelligent. Artificial emotion is a promising approach currently being used to develop and improve artificially intelligent machines. It is equally significant that the Turing Test shows that the nature of development of artificial intelligence is bound to knowledge of human intelligence. The computer must appear to be a human in the mind of a human. If artificial intelligence is designed to interact with humans, then it is imperative that it is in some way modeled after human intelligence. This is also true of artificial emotion. In order to successfully implement artificial emotional systems into artificial intelligence, existing knowledge of human emotion must be used as a foundation.

There is still much to know about human emotion. Many theories have surfaced over the years discussing the cause and origin of emotion. Most modern theories emphasize the functional aspects of emotions. According to this view, one function of emotions is to prepare and motivate us to deal with challenges in our environment. In a fight or flight situation, anger prepares and motivates one to get rid of an obstacle. Fear, on the other hand, provides motivation to avoid or escape danger. Emotions are somewhat like instincts, prompting a response to environmental changes. But they are more flexible than instincts; they allow for fine tuning of responses, and rely on past experiences and situational knowledge (Smith, 344). Dietrich Dorner, a psychologist working with AI researcher Joscha Bach, is currently devising methods of incorporating modern psychology regarding emotion, motivation, and affectivity into artificially intelligent agents. He is able to test these agents by providing them with a virtual environment, and 'emotional' stimuli.   Dorner's agents react to their environment by forming memories, expectations and immediate evaluations. They are also given a set of varied reflexive parameters which serve as "emotional cues" that influence their behaviour and perception (Bach, 3). The agent's resulting emotional behaviour is complex, and it is nearly impossible to classify a specific emotion that is occurring within the program. Even so, the change in their behaviour over a short time of emotional experiences is obvious. When agents perceive danger, they reflexively respond in an emotional manner, which contributes to a self-developed fight or flight reflex. They no longer examine their environment closely, but try for obvious escapes. Depending on the internal self-assessment of the agents, they may also attempt to confront or pursue the perceived threat (Bach, 5). While it is uncertain why the agent chooses fear above anger, or vice versa in any given situation, it is obvious that it has developed an adaptive emotional structure conceptually based on that of humans.

Emotion also serves a communicative role. Emotions serve as signals to other people, revealing how one is feeling, and thus how they are likely to behave. These signals manifest themselves physiologically; they are reflexive and uncontrollable. The Cannon-Bard theory of emotion holds that the central nervous system is innately linked with emotional signals and responses. The thalamus receives emotion-arousing stimuli from the senses, and serves as a relay hub. It simultaneously transmits information to the cerebral cortex, which results in a subjective experience of emotion, and down through the autonomic nervous system to the body's organs. This results in bodily reactions, which occur with the feelings of emotions in one's mind (Cannon, 108, 121). This notion is very similar to Rodney Brooks' belief that artificial intelligence requires a physical body. "Having a body provides a natural grounding. All computations that might be done on a robot were in service of the robot in the reality of the physical world" (Brooks, 143). IBM's artificial emotion project, currently known as Project Joshua Blue, incorporates an emotional model derived from current emotion theories built on the Cannon-Bard theory. The system includes valence - the degree of attraction or aversion that is evoked from events or situations - and arousal, homeostasis, and drive states. It also includes multiple point physical proprioception and a pain/pleasure system, which are all affected by the robot's 'emotions.' The system architecture responsible for the creation and transmission of emotion is based on "a semantic network of nodes connected by wires along which activation spreads" (Alvaredo, 3). In this activation model, the length of wires captures semantic distance. The conductance of wires is adjusted dynamically based on the emotional context. This design permits cognitive processes, physical systems and mental representations to be continuously influenced by emotions. The system is motivated and guided by affect to navigate its environment and acquire meaning through principles of learning (Alvaredo, 5). Although Joshua Blue is still being developed, it is a good example of an artificial system that incorporates a physical manifestation of emotion.

The Affective Computing Research Lab at MIT is also interested in the communicative aspect of emotion. Their goal is “to create personal computational systems endowed with the ability to sense, recognize and understand human emotions, together with the skills to respond in an intelligent, sensitive, and respectful manner toward the user and his/her emotions” (Picard, 11). The Affective Computing Research Lab is currently working on the development of computers that aid in communicating human emotions, computers that assist and support people in development of their skills of social-emotional intelligence, and computers that have emotional mechanisms, as well as the intelligence and ethics to appropriately manage, express, and otherwise utilize these emotions. One such project that concentrates on emotional communication is the “Computer Response to User Frustration Project.” The goal of the project is to design human-computer interaction systems to actively support human users in their ability to regulate, manage, and recover from their own negative emotional states, particularly frustration, which occur during computer use. A study designed to test this agent was conducted: A system was built that elicits frustration in human subjects. The interaction agent then initiated several social, emotional-content feedback strategies with some of the subjects, in an effort to help relieve their emotional state. These strategies were designed to provide many of the same cues that human counselors employ when helping relieve strong, negative emotions in others. Two control groups were exposed to the same frustrating stimuli, one of which was given no emotional support at all; the other enabled subjects to report problems and "vent" at the computer. Subsequent behaviour was then observed, and self-report data was collected (Picard, 63-65).

Behavioural results showed the agent was significantly more effective than the two controls in helping relieve frustration levels in subjects. The project shows that strategic, social, emotional interaction with a computer by users who are experiencing frustration can help alleviate this negative state. They also provide evidence that humans may benefit emotionally in the short term from computers that respond in socially appropriate ways toward their emotions. The implications of this work suggest a new role for computers in human life (Picard, 66).

Another modern psychological theory focuses on the role of physical facial expressions in emotion. The Facial Feedback hypothesis holds that facial expressions play a key role in initializing, or at least modulating, the experience of emotion. “Facial expressions are not merely the outward expression of emotion, they contribute to the feeling itself” (Izard, 494-495). It is not entirely understood how this process works, but many researchers believe that there is a two-way feedback loop between facial muscles and brain centres responsible for emotional experience. They are hard-wired to each other and act reflexively together. This line of human emotion research in many ways parallels the work of Cynthia Breazeal and the Sociable Humanoid Robotics Group at MIT. Breazeal believes that a social interface is a truly universal interface for human-computer interaction. Her work centres around building robots that share a similar morphology to humans, and can communicate in a manner that supports the natural communication modalities of humans. This includes facial expression, as well as body posture, gesture, gaze direction, and voice (Breazeal, 18-20). One robot in particular, Kismet, is especially relevant to the discussion at hand. Kismet is an expressive, anthropomorphic robot that engages people in natural and expressive face-to-face interaction. The robot is about 1.5 times the size of an adult human head and has a total of 21 degrees of freedom. Three direct the robot’s gaze, another three control the orientation of its head, and the remaining 15 move its facial features (eyelids, eyebrows, lips, and ears). To visually perceive the person who interacts with it, Kismet is equipped with four colour cameras. In addition, Kismet has two small microphones (one mounted on each ear). A lavalier microphone worn by the person who is interacting with Kismet is used to process voice. Inspired by infant social development, psychology, and evolution, this work integrates theories and concepts from these diverse viewpoints to enable Kismet to enter into natural and intuitive social interaction with a human and to eventually learn from them. This is based on human parent–infant exchanges. To do this, Kismet perceives a variety of natural social cues from visual and auditory channels, and delivers social signals to the human through gaze direction, facial expression, body posture, and vocal babbling. The robot has been designed to support several social cues and skills that could ultimately play an important role in socially situated learning with a human instructor. These capabilities are evaluated with respect to the ability of naive subjects to read and interpret the robot's social cues, the robot's ability to perceive and appropriately respond to human social cues, the human's willingness to support the robot's learning, and how this produces a "rich, flexible, dynamic interaction that is physical, affective, social, and affords a rich opportunity for learning" (Breazeal, 74). Kismet will someday be able to reflexively detect emotional content in the actions and voice of humans that it communicates with. It is also able to display recognizable emotions to people.

Breazeal acknowledges that Kismet has not been programmed for every emotion. For example, she has been saving 'surprise' for the day when she installs the learning algorithms she is working on. Then learning will become Kismet's desired stimulus; learning will make the robot happy. As Kismet begins to learn, Breazeal says, "it will slowly become more socially sophisticated, like an infant becoming a small child" (Breazeal, 68-69). This emotion-based learning is conceptually based on human processes, especially infant cognitive development. It also supports the idea that emotional robots are intelligent robots.

The capabilities of displaying and detecting emotion seem to be critical components of creating intelligent systems with whom humans can comfortably relate and communicate. The emotional aspect distinguishes a ‘dead’ machine from one who is believable, ‘alive,’ and trustworthy - and able to pass Turing’s test. But does this mean that the ultimate goal of AI is to mimic human intelligence to a point of simulation? Should artificial emotion aspire to be a replica of human emotion? In many ways, the Turing Test is a testament to the egocentric nature of human intelligence. It assumes that human thought is the highest form of intelligence against which all others should be measured. Yet, modern computers can outperform the human brain in areas of complex mathematics, counting, and memory. In “On Computational Wings,” Ford and Hayes use the early development of artificial flight as an analogy to the development of artificial intelligence. The Wright brothers were inspired to build their flight machine by watching birds, but they did not try to imitate a bird’s flight process. “The Turing Test should be relegated to the history of science, in the same way that the aim of imitating a bird was eventually abandoned by the pioneers of flight” (Ford, 8). Ford and Hayes extend this idea to modern aeronautic engineering. Its main goal is obviously not to make machines fly exactly like pigeons, to the point where they can fool other pigeons. They also point out that no aircraft could ever be mistaken for a bird, because it lacks the precision that a bird has. Aircraft cannot perch in trees, scoop fish, or hover motionless in a warm draft of air. However, no bird can fly faster than sound, or at 40 000 feet (Ford, 10).

Modern AI research, then, should be inspired, but not limited, by human intelligence. Most of the AI research previously discussed supports this idea: they are not attempts at true recreations of human emotion, but rather the projects are largely based on evoking and/or detecting emotion in the humans they are interacting with. They have their foundations in existing knowledge of human emotion and social interaction, but are not limited by it. Dorner’s Psi agents experience ‘emotions’ that are unintelligible to humans. Kismet’s facial features are decidedly non-human. But they are meant to interact with humans, in the hope that humans will gain a better understanding of intelligence, both human and nonhuman. By experimenting with artificial emotion, there will be much knowledge gained about human and nonhuman emotion, and how it interacts with other cognitive systems to produce intelligent behaviour.

Works Cited

Alvaredo et al. (2001). “Project Joshua Blue: Design Considerations for Evolving an Emotional Mind in a Simulated Environment.” Paper from the 2001 AAAI Fall Symposium, AAAI Technical Report, FS-01-02. AAAI Press, Menlo Park, CA,2001. <http://www.csupomona.edu/~nalvarado/PDFs/AAAI.pdf>

Bach, J. “Enhancing Perception and Planning of Software Agents with Emotion and Acquired Hierarchical Categories,” Proceedings of MASHO 02, German Conference on AI. Karlsruhe, Germany. 2002. <http://www.informatik.hu-berlin.de/~bach/publication/BachMasho2002.pdf>

Breazeal, C. Designing Sociable Robots, MIT Press, Cambridge, MA. 2002.

Brooks, R.A. “It’s 2001 Already.” Flesh and Machines: How Robots Will Change Us, Pantheon, 2002. (version referenced: Thinking Machines. Ed. A. Greenspan. Hamilton, ON. McMaster University CCPS, 2003. pp 141-160.)

Cannon, W. B. (1927). The James-Lange theory of emotions: a critical examination and alternative theory. American Journal of Psychology, 39, 106-124.

Ford, K.M., Hayes, P.J. (1998). “On Computational Wings.” Understanding Artificial Intelligence. Ed. S. Fritz. New York, NY. Warner Books, 2002. pp 5-18.

Izard, C. E. (1990). Facial expressions and the regulation of emotions. Journal of Personality and Social Psychology, 58, 487-498.

Picard, R. W. (1997). Affective Computing, MIT Press, Cambridge, 1997.

Smith, C. A. (1989). Dimensions of appraisal and physiological response in emotion. Journal of Personality and Social Psychology, 56, 339-353.

Turing, A. M. (1950). Computing Machinery and Intelligence. Mind*, 49,* 433-460. <http://cogprints.ecs.soton.ac.uk/archive/00000499/00/turing.html>