Leah E. Steinberg
THEA 209
Final paper

Theater for the Emotions

During the course of this term, we explored how theater and science interact. It seemed to me that a challenge facing us at the beginning of the class was to find out how science, a stereotypically dry and factual area, could be turned into a show that would capture the attention and emotion of an audience. It was no surprise to me that we found ourselves gripped in the passions of the science plays that we read, but it did pique my interest. How do these shows make us feel so much about simple facts of the physical world? If we approach the question scientifically, we end up looking at how the physical emotion-maker (also known as the body, specifically, the brain) works. It is beyond the scope of this course to get into the neuroscience of emotions, but we did take a brief tour of cognitive science and I think we're qualified to explore emotion in the context of several different theories of mind. Steven Pinker gave an explanation of emotion from an evolutionary standpoint,[1] but he left the mental structure that implements emotion unspecified. The early part of the class focused on the life, work, and philosophies of Alan Turing, which form a theoretical basis for modern computational cognitive science. It seems that if we could write down a list of rules for a universal Turing

[1]See chapter 6 of Pinker.

machine complicated enough, we could produce rational thought. However, humans don't always (or even often) act rationally, and if cognitive science is truly interested in understanding the way humans think, it cannot afford to ignore the all-pervasive phenomenon of emotion, as I will argue below. Yet it turns out that ". . . emotions remain the mental phenomena most neglected by contemporary philosophy of mind and the sciences of mind . . . for the traditional interests of philosophers of mind, cognitive scientists, artificial intelligence researches, and many others, emotions remain peripheral, sometimes even a seemingly irrelevant issue." (DeLancey, vii). In this paper I set out to explore what happens when three different schools of cognitive science attempt to include emotions in their explanations of the mind, and why it is important for them to do so.

In order to continue, we must take the time to determine what we mean when we talk about emotions. What are these mysterious feelings that affect us so deeply? Craig DeLancey would argue that "affect" is the key word. He claims that "*[a]ffects are body states that are motivational*. . . the principle feature of these motivations is that they are internal physical states of an organism that cause it to perform an action if the organism is not inhibited by different motivations or otherwise constrained." (5-6, [emphasis his]). This definition seems to paint emotions, which DeLancey classifies as a subcategory of affects, as purely physical and causal. Antonio R. Damasio notes that, while emotions can be culturally influenced, they have a neurological basis, and can occur without conscious thought.[2] Bringing cultural and social influence into the picture is significant, for doing so introduces an element that clearly separates emotions from basic reflexes such as withdrawing one's hand from a flame, over which culture has little to no influence. Yet at the same time, these emotions can occur "*without conscious thought*," making them unlikely candidates for cognitive processes. DeLancey uses the term "subcognitive" to describe situations where mental processes are "not propositional attitudes or complex symbolic representations." (35). From all of this I draw the definition of an emotion to be a bodily state (including – but especially not limited to – a neural state) that is subcognitive in nature, yet can affect conscious thoughts, processes, and social interactions, and in turn be affected by these same influences.

Why should emotions be important to cognitive science in the first place if they are not themselves cognitive processes? I make the case that emotions are a universal, integral part of the human mind, and thus indispensable to the field. A simple appeal to common sense shows how intuitive this claim is: "Emotions, after all, are the threads that hold mental life together. They define who we are in our own mind's eye as well as in the eyes of others." (LeDoux, 11). However, as we know from many class discussions, science is not founded on the basis of intuitions and hunches, so I must produce more concrete evidence of the importance of emotions in cognition to make

[2] See Damasio 1999, 51-52 for his complete definition.

a convincing argument. Fortunately, this evidence abounds.

Most people have experienced times when they feel that their judgment is clouded by emotion, but this is typically believed to happen only when the emotion is particularly strong (e.g., a normally sober person drinking to excess after the loss of a loved one). It is therefore somewhat surprising to learn that ". . . a growing body of research indicates that even mild and even positive affective states can markedly influence everyday thought processes, and do so regularly." (Isen , 261). As this is the case, emotion is surely relevant to the study of cognition. Emotion has been shown to affect memory, creativity, social judgment, rationality, and reason [3] – to name a few important cognitive functions – leaving me puzzled as to how it can possibly be ignored when studying human thought and behavior. So why is emotion left out of major theories of mind? Or is it? If it isn't, how does it fit into each of them?

I will take this opportunity to briefly sketch the three philosophies of mind I wish to consider in this paper. Cognitive science is by no means limited to these three approaches, but they form the basis of the major schools of thought in the field at present. In class, we studied the underpinnings of a representational theory of mind when we learned about Turing machines. The view of minds as formal symbol manipulators, where the symbols being

[3] See Alice M. Isen's chapter "Positive Affect and Decision Making" in Haviland, chapter 11 in DeLancey, and Damasio 1994 for examples.

manipulated represent the world, is often known as classic artificial intelligence, or classical AI.

An alternative to explicit rule-following can be found in connectionist theory. Again, we were introduced to connectionism in class by Pinker[4]; he explained the basics of connectionist networks with their learning ability due to varying connection weights between nodes. He pointed out some of the benefits of connectionism, such as easy generalization and the ability to work with incomplete data. A similar approach to connectionism asks how a neural net actually acquires knowledge and intelligence. The connectionist regime depends on accurate feedback to modify the connection weights, but the real world does not often hand back an explicitly graded version of life's test questions. Embodied, embedded, or extended cognition focuses on the interactions between the intelligent agent and its environment. Proponents of embedded cognition view cognitive phenomena as emergent behavior derived from interactions between the agent and its surroundings, driven and shaped by the agent's environment.[5]

Now that we have an outline of our three theories of mind, let's see how they attempt to explain the mystery of emotions.

It soon becomes apparent, when looking at classical AI, why the emotion question is often ignored. In Hubert L. Dreyfus' criticism of classical AI he

[4] See chapter 2 of Pinker.

[5] See Andy Clark's book for a more detailed description.

points out that "[t]here is no reason to suppose that moods, mattering, and embodied skills can be captured in any formal web of belief . . . all AI workers and cognitive psychologists are committed, more or less lucidly, to the view that such noncognitive aspects of the mind can simply be ignored." (177). He brings up the point in order to argue that these subcognitive abilities are vital for establishing the context necessary for cognition; we can see from the above arguments that much more than context is lost without them. DeLancey further spells out the problem [6]: "If emotions and cognitive states can be subcognitive but can influence cognition, then they both fall outside of the symbolic computational model, and they influence the very kind of things it aims to explain."(194). DeLancey also points out that many of the classical AI algorithms were derived from having people report the steps of their cognitive processes, yet there is evidence that rationality is affected by emotion, as noted above, and these reports often lead to decisions that are not typical of daily life, (198 and endnote 59, page 232), so the entire basis for many AI programs may be rooted in misinformation. It seems fairly clear, simply on the basis of examining classical AI with respect to emotion, that it is an inadequate explanation of how the human mind works.

Connectionism would appear to be a better candidate for dealing with emotional effects on cognition, as emotions may be represented by the nodes and weights inside a connectionist network. Placing emotions within the net-

[6] See chapter 11 of DeLancey for multiple arguments dismissing classical AI and connectionism.

work certainly does make them ". . . an inseparable part of how we see and represent the world around us, the way we select, store and retrieve information, and the way we use stored knowledge structures" (Forgas, 593), which seems like a step forward from the classical AI position of complete exclusion. DeLancey is less optimistic in pointing out that connectionist models which attempt to encode emotions ". . . place emotion and the other nodes on the same level [as such cognitive categories as 'game' or 'uninteresting']," (195), but we have established that emotions are subcognitive, making the connectionist solution appear to give perhaps too much weight to emotions. However, because DeLancey only considers a semantic connectionist model, he may be overlooking all the possibilities of connectionism. Perhaps there are other ways of accounting for emotion in connectionist networks that would place it at an appropriate level with respect to traditionally cognitive processes. Incorporating emotion into connectionist nets may be a matter of focusing on a suitable level of representation and finding an emotion-eliciting environment with which the network can interact.

Enhancing connectionist nets by placing them in rich environments brings us to the embedded theory of mind. Perhaps these nets could feel emotions if placed in the proper setting, without specifically coding for affects. DeLancey is not enthusiastic about embodied cognition, but he does not reject it as vehemently as he does classical AI and connectionism. His most forceful argument against embodiment is that it requires ". . . a proper account of the body of the organism . . . to explain the functions of affects and other mental states." (200). I believe his pessimism springs from the notion that emotions will not be evident in a body much less complicated than our own, which would rob the embodied viewpoint of its explanatory powers since it would not differ enough from the original to be of any use in bringing the interactions to a level we can easily understand. This is a point to keep in mind, after all, what is the value of an extremely complicated model? Michael R. W. Dawson discusses this question and concludes that models are used for prediction, so a predictive model, even if complicated, has value.[7] Dawson praises the synthesis, emergence, and analysis approach to understanding models[8], which is extremely applicable to embodied models. DeLancey combats the possibility of emotions as emergent properties by stating that ". . . the inherited and biologically based, and not the emergentist, view is the more supported by the findings of biology and other relevant sciences [when considering the basic emotions]," (204). When emotions are considered to be physical states, as we have defined them to be, it seems difficult to understand how they could arise unexpectedly if one considers them on the level of chemical interactions in the brain, or physical interactions on the atomic level of those chemicals. If, however, one looks at the situation from a more abstract level – which, as Daniel C. Dennett argues,

[7] See chapter 2 of Dawson, especially page 6.

[8] See chapter 6 of Dawson.

has its explanatory and predictive benefits[9] – embodied cognition seems the most likely alternative for gracefully and successfully accepting emotion into its explanation of mind.

Looking at our three different theories of mind, we see that classical AI is severely incapable of accounting for emotions. Connectionism fares better in appearing to account for the effects of emotion, and the embodied viewpoint does better, yet, with the possibility of not only representing, but even creating emotions. None of the theories I've explored have come up with a perfect answer to why we cry when we know that the person who appears to be lying dead on the stage is really an actor and is actually alive. Still, with a scientific challenge like the machinery of emotion in front of us, how can we resist the attempt to solve the question? If nothing else, a play specifically about emotion would have fascinating possibilities.

[9] See chapter 3 "True Believers: The Intentional Strategy and Why It Works" by Daniel C. Dennett in Haugeland.

References

Clark, Andy. Being There: Putting Brain, Body, and World Together Again. Cambridge: The MIT Press, 1997.

Damasio, Antonio R. Descartes' Error: Emotion, Reason, and the Human Brain. New York: G. P. Putnam's Sons, 1994.

Damasio , Antonio R. The Feeling of What Happens: Body and Emotion in the Making of Consciousness. New York: Harcourt Brace & Company, 1999.

Dawson, Michael R. W. Minds and Machines: Connectionism and Psychological Modeling. Malden: Blackwell Publishing, 2004.

DeLancey, Craig. Passionate Engines: What Emotions Reveal About Mind and Artificial Intelligence. New York: Oxford University Press, 2002.

Dennett, Daniel C. "True Believers: The Intentional Strategy and Why It Works." Mind Design II: Philosophy, Psychology, Artificial Intelligence. Ed. John Haugeland. Cambridge: The MIT Press, 1997. 57-79.

Dreyfus, Hubert L. "From Micro-Worlds to Knowledge Representation: AI at an Impasse." Mind Design II: Philosophy, Psychology, Artificial Intelligence. Ed. John Haugeland. Cambridge: The MIT Press, 1997. 143-182.

Forgas, Joseph P. "Network Theories and Beyond." Handbook of Cognition and Emotion. Ed. Tim Dalgleish and Mick Power. Chichester: John Wiley & Sons Ltd, 1999. 591-611.

Isen, Alice M. "Positive Affect and Decision Making." Handbook of Emotions. Ed. Michael Lewis and Jeannette M. Haviland. New York: The Guildford Press, 1993. 261-277.

LeDoux, Joseph. The Emotional Brain. New York: Simon & Schuster, 1996.

Pinker, Steven. How the Mind Works. New York: W. W. Norton & Company, 1997.