AI Magazine Volume 20 Number 1 (1999) (© AAAI)

Book Review

Response to Sloman's Review of *Affective Computing*

Rosalind W. Picard

Let me begin by expressing appreciation to Aaron Sloman for his compliments, insights, and lengthy effort to elaborate on some of the ideas in my book. Sloman was one of the first in the AI community to write about the role of emotion in computing (Sloman and Croucher 1981), and I value his insight into theories of emotional and intelligent systems. Alas, Sloman's review dwells largely on some details related to unknown features of human emotion; hence, I don't think the review captures the flavor of the book. However, he does raise interesting points, as well as potential misunderstandings, both of which I am grateful for the opportunity to comment on.

Use of Affective Computing

Sloman writes that I "welcome emotion detectors in a wide range of contexts and relationships, for example, teacher and pupil." This might sound innocuous, but its presumption of the existence of emotion detectors is not. Many users have fearfully expressed to me that the thought of a machine that could "recognize their emotions" strikes terror into their already anxious hearts. These users tend to not understand the limits of the technology; they are already so amazed at what computers can do, that they almost expect computers will be able to do anything. One user, with grave concern on her face, asked me in a hushed voice as she turned her back to her computer, "Does it know that I don't like it?" At one time, I would have discounted such remarks, but now that nontechnical users are in the majority, their feelings and fears demand not just our attention but also our careful choice of language.

I use the expression *emotion recognition* only when established as shorthand for the unwieldy but more accurate description "inference of an emotional state from observations of expressions and from reasoning about an emotion-generating situation." The computer cannot directly read internal thoughts or feelings, and therefore, there is no "emotion detector" as such. It can detect certain expressions that arise in conjunction with an internal state: pressure profiles of banging on a mouse, video signals of posture and facial expressions, audio signals of vocal expressions, physiological signals such as respiration, and more. The machine tries to recognize patterns of physical signals and associated behaviors that might be correlates or expressions of emotion. The best of what I think computers could detect in the foreseeable future is described in the book.

I urge a responsible focus on applications where the user perceives a benefit with the communication of emotion and where the user remains free to hide this information if he/she desires. In contexts where humans interact with computers naturally and socially (Reeves and Nass 1996), we might benefit by having the computer recognize expressions such as boredom, confusion, pleasure, or frustration and then adjust its behavior appropriately. For example, the computer might speed up if we seem bored, offer an alternate explanation if we appear confused, and try to improve its knowledge of our preferences by watching for signs of pleasure or frustration. We send affective cues almost effortlessly, and such cues can be loaded with helpful information. Affective cues are a natural way that humans give feedback to learning systems. Recognition of our natural forms of expression is a means of giving machines more human-centered behavior.

My students and I currently use tools of expression recognition to gather data to hone the abilities of our research systems, always with the consent of those involved. However, Sloman's remarks imply that I favor inserting "emotion detectors" willy-nilly into situations. In fact, I decry even the relatively benign intrusions, such as emotional agents that jiggle about on the screen, smiling at you in an annoying and inappropriate fashion, costing you precious time while you search for a way to disable them. Although inappropriate use of affect might be the most common affront with this technology, there are also potentially more serious problems (chapter 4.)

Sloman writes that in lieu of being hooked up to emotion-sensing devices, he would prefer us all to become more sensitive. I heartily agree and am disappointed that he would emerge from my book with any other impression. What Sloman misses in his remark is the flip side: As people interact increasingly with and through nonaffective computers, then their emotion sensitivity can actually be diminished. The aphorism "if you don't use it, you'll lose it" might just as well apply to social and emotional skills, which are able to be learned and to be improved with use. The fact that many high-functioning autistics love communicating over the net because it "levels the playing field" should be a wake-up call to us: Current forms of computer-mediated interaction limit affective communication.

What Is Sentic Modulation?

Sloman's review might seem confusing in places whether or not you've read my book. Part of the confusion is trace-

Sloman's primary complaint revolves around his claim that Damasio and I believe "that secondary emotions always trigger primary mechanisms." The insertion of the always *is Sloman's because I would not write this indefensible statement ….*

able to a misunderstanding about what I mean by *sentic modulation* (sentic comes from the Latin *sentire,* the root of the words *sentiment* and *sensation*.) Sentic modulation, such as voice inflection, facial expression, and posture, is the physical means by which an emotional state is typically expressed.

Precisely when a physical event is or is not sentic modulation is not something I lingered on in the book, but certainly it is the case that not all physical events are sentic modulation, only those that carry emotional information. Facial expression and vocal intonation are the best-known examples, but sentic modulation can affect most any behavior or physical motion: Some people type differently when angry versus when in a neutral state.

It is possible in the future that sentic modulation will be able to be sensed not only using wearable (skin-surface or clothing-layer) computers but also with miniature implantable and inhalable devices. The inhalables and implantables might be used to interpret biochemical and bioelectrical changes that carry emotional information. My current research with sentic modulation focuses on recognizing patterns gathered by noninvasive sensing: viewing behaviors, listening to sounds, and sensing signals from skin-surface contact. However, sentic modulation refers to all physical changes that communicate emotional information, whether sensed noninvasively or otherwise.

Sloman dwells on the limits of sentic modulation, which indeed exist, especially given an incomplete understanding of the phenomena. However, I do not hide the limits, and readers can see that I have also written about verbal communication and the important influence of thoughts and context on emotions. Sloman's disagreement is with my emphasis on nonverbal forms; he claims that verbal forms of emotion are the most primary and natural. He supports the claim by citing the long history of letter writing, the outpourings of people in news groups, and examples of e-mail and Marc Antony's soliloquy, pieces that he claims carry a certain clear affective message. In contrast, I have no trouble reading these pieces with an entirely different tonality, such as a wry and jesting voice, as well as with the voices he presumes. Unfortunately, you can't hear my tone of voice in this text right now, no matter how carefully I choose my words.

The problem of communicating affect linguistically is an important one. When computers solve it, it will no doubt be useful for alerting hasty e-mail senders to the unintentionally nasty tone of their mail and other such applications. However, I suspect that no amount of linguistic brilliance will ever suffice to unambiguously communicate tone. My relative emphasis on the nonverbal is based on findings that show that nonverbal cues (body language, behavior, and so forth) typically form somewhere between 65 and 95 percent of human communication (Wolfgang 1979; Mehrabian 1972; Birdwhistell 1970).

Emotions: Simple and Complex

Sloman and I have both fallen prey to sometimes writing in an oversimplified way about emotions. The athlete's story (which both he and I write about) deserves more careful analysis than space permits here, but let me suggest a key issue: Emotional self-report is fraught with problems. When the athlete rattles off her list of feelings to the public eye, she rattles off not just **what she thinks she feels** but also **what she thinks she should say that she feels**. In this flurry of thoughts and feelings, she anticipates an event and concludes, "The thought of that makes me sad." At this moment, she might or might not experience a flicker of sadness, and even if she does, it might or might not make an observable dent in her sentic modulation. However, later when the event happens and she genuinely **feels** sad, the sadness is likely to become apparent. Although her words are unquestionably important, it would be fascinating to watch her face and gestures, and to hear her voice during this discussion of her feelings, because these signals communicate much more than her words alone.

The phenomena of trying to recognize emotion is too often misinterpreted as one of recognizing a small set of simply labeled states, such as sadness. Jerome Kagan's (1984) comparison of emotion to weather provides a nice metaphor:

> The term *emotion* refers to relations among external incentives, thoughts, and changes in internal feelings, as weather is a superordinate term for the changing relations among wind velocity, humidity, temperature, barometric pressure, and form of precipitation. Occasionally, a unique combination of these meteorological qualities creates a storm, a tornado, a blizzard, or a hurricane—events that are anal-

Sloman and I have both fallen prey to sometimes writing in an oversimplified way about emotions.

> ogous to the temporary but intense emotions of fear, joy, excitement, disgust, or anger. But wind, temperature, and humidity vary continually without producing such extreme combinations. Thus meteorologists do not ask what weather means, but determine the relations among the measurable qualities and later name whatever coherences they discover… (p. 155).

Getting computers to recognize the analogs of humidity, temperature, tornadoes, and other useful coherences is the easier (although not easy) part of emotion recognition. In contrast, getting computers to understand what gave rise to the variables, and what they mean, seems to require solving the hard AI problem. I wrote about both and wrote more about the former because I think it is likely to appear soonest in machines. Solving language understanding, situation understanding, and computer common sense will no doubt help make computers better at interpreting human emotions. However, we must not forget that a cat can usefully discriminate important human expressions such as approval and disapproval without having rich and subtle linguistic abilities or a deep understanding of the structure of human minds. Computers can achieve several useful and important affective abilities without solving hard AI.

Sloman's resistance to my ideas on mixed emotions puzzled me initially, but now I think that it can be explained by the difference in our research slants. Sloman's approach over the decades has been primarily one of proposing architectures that would **give rise to** cognitive and affective behavior, which contrasts with my recent focus on designing and building systems that can **recognize and respond intelligently to** human emotions. He has focused on sources of emotion, and I have focused on communication of emotion. The problem is not parallelism for sources; the problem is on the recognition side when multiple emotions modulate the same output channel. Given an output channel, how do we find which emotions are modulating it? In this case, we need to ask how the multiple emotions have been combined, which is where I hope my metaphors are useful.

Secondary Emotions and Sloman's New Category

Sloman's primary complaint revolves around his claim that Damasio and I believe "that secondary emotions always trigger primary mechanisms." The insertion of the *always* is Sloman's because I would not write this indefensible statement (consider, for example, when the triggering mechanism is damaged.)

I prefer not to attempt a detailed comparison of Sloman's versus Damasio's viewpoints but will comment on Sloman's request for a new category of "central secondary emotions." I find his arguments sound good at first but on further inspection are too weak to win my support. Sloman says there might be ways for redirecting thought processes other than signals caused by an emotional system. This statement is perfectly reasonable; even noise might redirect thoughts, but this is not the heart of Sloman's claim.

The real issue is that Sloman argues that there are emotions that cause changes in mental events without any corresponding sentic modulation. He suggests that possible examples are guilt, infatuation, and anxiety. One of the problems with Sloman's argument is that the three examples he offers—guilt, infatuation, and anxiety—are all discussed in the literature as having corresponding biochemical, physiological, or other behavioral changes, that is, sentic modulation. For example, the emotion of guilt has been argued to be distinguished by behavior where an individual tends to move in space as if trying to repair an action (Lewis 1993).

Sloman's claim seems to be based primarily on his personal experience, which he tells me includes an existence without a rich set of feelings. He is not the first I've encountered who has said this, confirming that there is a wide range of emotional experience, even among apparently healthy and high-functioning humans. However, it is premature to say that there is no sentic modulation with certain emotions, at least not until there is concrete evidence to support this. One might not be aware of sentic modulation as it happens, and it might not be obviously visible, but this does not mean that it does not exist.

If there were no sentic modulation for certain emotions, then what would distinguish these "cold" emotions from mere thoughts? I wrestled with this in chapter 2 of my book, where I argue that not only can we give "purely cognitive" (cold) emotions to machines, but some machines already have these. Such purely cognitive emotions are currently believed to exist in humans with various kinds of brain damage or neurological abnormality, where the ordinary "feelings" no longer co-occur with the thoughts that would otherwise produce them. These patients remark after looking at a horrific scene, for example, that it no longer makes them "feel" horror, like it would have done before their brain damage, even though they know cognitively that it should make them feel bad.

I agree with Sloman that secondary

emotions might have different ways in which they signal cognitive changes, but I think that the creation of the class of "central secondary emotions" that Sloman proposes is not only not justified but is not the correct granularity to account for the variety of experiences we find in the literature. I would suggest instead that there is a small set of analog signaling mechanisms, which include the pathways that are damaged in Damasio's patients, and that a variety of means, including various kinds of developmental abnormality and brain damage, can modify the normal range of "feelings" in an emotional experience by modifying these mechanisms.

Sloman tries to bolster his argument by pointing out that emotion is not necessary for searching but that other nonemotional techniques can provide the functions hypothesized for emotion in the decision-making process of narrowing down large sets of possibilities. His arguments here are reasonable, but they miss a much bigger point. The point is not that emotion biases decision making (the focus in Damasio's book) but that increasingly, neurologists, cognitive scientists, and psychologists are finding that human emotion biases memory formation, which, in turn, appears to bias decision making, creativity, planning, perception, judgment, mood-congruent memory retrieval, and a host of specific phenomena. There are many findings other than Damasio's, and one must look at all of them. The simplest explanation that addresses all these findings right now is emotion. Sloman's explanations fit searching but not the many other phenomena.

Sloman writes that effects such as rapid redirection of thought processes can happen without primary emotions being involved and that there might be some higher system that controls both of these: Both the lack of emotion in Damasio's patients and the corresponding impaired cognitive processes might be caused by something else. Sloman's hypothesis that another "higher" mechanism could cause all these effects is logical. However, Sloman's hypothesis has yet to identify any nonvirtual regions of our brain's architecture that could account for the mechanism he proposes. Damasio's hypothesis is appealing for its overall simplicity and grounding in neuroscience, which is not to say Damasio is right and Sloman is wrong; however, there is a principle in science to favor the simplest explanation that fits the most evidence, and currently, this argument does not favor Sloman's proposal. The challenge remains for scientists to determine precisely how the many regulatory signaling mechanisms of the cognitive and affective systems work their marvelous and manifold influences.

On Emotion and Intelligence

Sloman and I are mostly in agreement on this topic, recognizing that emotional skills are important for intelligent human interaction. Sloman's casual mention of Spock, however, invites misunderstanding. Sloman writes of *Star Trek*'s Spock as if Spock had no emotions, a common misconception. The half-human Spock not only had emotions, but actor Leonard Nimoy writes wonderfully of how Spock developed his abilities to suppress the expression of his emotions, an important value in the Vulcan culture (Nimoy 1995).

Sloman and I both think that mechanisms to help overcome resource limits are important for intelligent systems; however, Sloman calls emotions "side effects" of these mechanisms, and I point to a growing collection of studies from cognitive science and neuroscience, suggesting that the emotional system includes or otherwise directly influences these mechanisms. I find the literature's studies compelling and suggest that instead of debating "is emotion a main effect or side effect," we put more energy into trying to understand the mechanisms that make up the emotional system and its influences.

Ultimately, we might be able to imitate all the known "emotional" phenomena in machines without invoking the "E-word," without precisely duplicating any of the mechanisms of the human emotional system. However, I would argue that once we have duplicated in the machine all the important biasing, regulatory, motivational, behavioral, and other expressive phenomena associated with human emotions, and done so in a sufficiently general, flexible, and efficient way, that we will have given the machine mechanisms that are essentially emotions.

Every living example of general intelligence that we know has emotion, which does not mean there might not be another path to intelligence without such mechanisms. However, why take time to search for an alien mechanism when we already have a set of mechanisms—emotion—that appears to be able to do the job? If we emulate human affect abilities, versus alien mechanisms, then (1) we stand a better chance of understanding the resulting behaviors of the computer and (2) the process of figuring out how to build the emotions can help us in figuring out how the human emotional system works, one of the most important potential benefits. Understanding and emulating human emotion might or might not hold the key to solving AI, but we are far from complete in our efforts to understand intelligence if we do not learn more about emotion.

References

Birdwhistell, R. L. 1970. *Kinesics and Context*. Philadelphia: University of Pennsylvania Press.

Kagan, J. 1984. *The Nature of the Child*. New York: Basic.

Lewis, M. 1993. Self-Conscious Emotions: Embarrassment, Pride, Shame, and Guilt. In *Handbook of Emotions,* eds. M. Lewis and J. Haviland, 563–573. New York: Guilford.

Mehrabian, A. 1972. *Nonverbal Behavior.* Chicago: Aldine-Atherton.

Nimoy, L. 1995. *I Am Spock*. New York: Hyperion.

Reeves, B., and Nass, C. 1996. *The Media Equation*. Cambridge, U.K.: Cambridge University Press Center for the Study of Language and Information.

Sloman, A., and Croucher, M. 1981. Why Robots Will Have Emotions. In Proceedings of the Seventh International Joint Conference on Artificial Intelligence, 197–202. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Wolfgang, A., ed. 1979. *Nonverbal Behavior: Applications and Cultural Implications*. San Diego, Calif.: Academic.