AFFECTIVE COMPUTING: CHALLENGES

Rosalind W. Picard
MIT Media Laboratory
Cambridge, MA USA
E-mail: picard@media.mit.edu

1. INTRODUCTION:

Many people think of Spock, the half-Vulcan, half-human character of Star Trek, as the patron saint of computer science. He was highly intelligent, highly rational, highly unemotional, and (at least at the university where I work) attractive to women. A popular image is that Spock didn't have emotions: after all he almost never expressed emotion, excepting his characteristic pronouncement of the word "fascinating" upon pondering something new. In fact, as the actor Leonard Nimoy describes in his book [Nimoy 1995] the character Spock did have emotion; he was just very good at suppressing its expression. But most people think of Spock as not having emotion. When someone never expresses emotion, it is tempting to think that emotion is not there.

In affective computing, we can separately examine functions that are not so easily separated in humans. For example, the Macintosh has been displaying a smile for years upon successful boot-up. But few people would confuse its smile – albeit an emotional expression – with a genuine emotional feeling. Machines can fake the appearance of an emotion quite well, without having any feelings similar to those we would have: They can separate expression from feeling. With a machine it is easy to see how emotion expression does not imply "having" the underlying feeling.

Machines that might actually "have" feelings is the key area of affective computing that I expressed serious doubt about in my 1997 book Affective Computing. I think the discussions there, and in a later book chapter (Picard 2003) on this topic are still timely and I will not plan to add to them here. Researchers in the last decade have obtained dozens of scientific findings illuminating important roles of emotion in intelligent human functioning, even when it looks like a person is showing no emotion. These findings have reshaped scientific understanding of emotion and have inspired a number of researchers to consider that emotional mechanisms might be more valuable than previously believed. Consequently, a number of researchers have charged ahead with building machines that have several affective abilities, especially: recognizing, expressing, modelling, communicating, and responding to emotion. And, within these areas, a number of new criticisms and challenges have arisen. The rest of this paper addresses such matters.

2. SENSING AND RECOGNIZING EMOTION

Criticism 1: The range of means and modalities of emotion expression is so broad, with many of these modalities being inaccessible (e.g., blood chemistry, brain activity, *neurotransmitters), any many others being too non-differentiated. This makes it unlikely that collecting the necessary data will be possible or feasible in the near future.*

Criticism 2: People's expression of emotion is so idiosyncratic and variable, that there is little hope of accurately recognizing an individual's emotional state from the available data.

In the book *Affective Computing*, I gave reasons why I thought many computers will need the ability to recognize human emotion, ideally at the same level that people can. This level is variable: On the one hand, three-quarters of computer users admit to swearing at computers [Mori survey, 1999], a verbal form of emotional expression that most people have no difficulty recognizing. On the other hand, if you were asked to jot the emotional state of the next person you see (feel free to try this) the challenge grows significantly. In fact, many people don't know how to articulate their own feeling state. Think about it: you have better access to your innermost feelings than anyone, but you still do not always know how to "recognize" or label what you are feeling. In particular it can be hard to assign a single word to label a feeling state. Nonetheless, we have feelings essentially all the time: it is common to hear somebody say, "Sorry, I wasn't thinking", but not "Sorry, I wasn't feeling."

One of the first questions my students and I investigated was "could a wearable computer be designed to recognize a person's emotion over an extended period of time?" We used four skin-surface sensors, freshly mounted in the same way each day, for several weeks of data collection. (The sensors included: electromyogram, skin conductance, blood volume pulse, and respiration. See Picard et al. (2001) for details.) The idea was that a wearable computer naturally affords skin-surface sensing, and long-term monitoring. The sensors measured physiological parameters related to autonomic nervous system changes. We were interested in how daily physiological variations would influence results, and in particular, we wanted to remove the variations we expected to find from one person to another. Thus, we gathered a huge amount of data from one person over many weeks, in contrast with most experiments, which gather from many people over a short single session, then look for averages.

We knew from Schachter and Singer's work in the 1960's that many psychologists believe that physiological signals would not distinguish more than arousal level. We found, however, that eight emotions could be distinguished at levels significantly higher than chance: we developed pattern recognition algorithms that attained 81% classification accuracy instead of the predicted 12.5% of a random classifier. Furthermore, these distinctions were not limited to arousal, but also included discrimination of emotions having similar arousal and varying valence – positive or negative – characteristics. We also found that the variations from day to day, even for the same emotion, could be larger than variations among different emotions on the same day; we developed new algorithms to accommodate these variations.

In the end, we concluded that recognizable physiological differentiation does appear with the eight emotions we investigated. However, it would be wrong to conclude that a computer can recognize people's emotions with 80% accuracy. Although this issue is more fully addressed in the paper reporting this experiment and its results [Picard et al., 2001], the main limitation is that our experiment forced a choice among eight kinds of emotion. Thus, the results are not unlike those in the early days of speech recognition when a person spoke the digits, "one, two, …., eight" into a microphone, and the computer recognized which of the eight digits was spoken. Although we managed to capture the equivalent of day-to-day variations in voice that might arise with head-colds, drowsiness, caffeine, and moods, the task was still a limited one: recognize eight discrete categories. Researchers have a long way to go before solving the emotional equivalent of continuous, speaker-independent natural language recognition.

Although comparing affect recognition to speech recognition is useful for highlighting how nascent and challenging the research is, the comparison breaks down when considering how words differ from emotion. On the one hand, a word is discrete, and you can usually know if you recognized it correctly or not. However, emotions can be thought of as both discrete and continuous, and it is hard sometimes to know how to describe them. I find the weather metaphor particularly useful in trying to think about emotion:

> 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 analogous 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... [Jerome Kagan, 1984]

With emotion, as with weather, one can build sensors for measuring the physical equivalents of temperature, pressure, humidity, etc. One can also build successful algorithms for combining patterns of such measures, and thus recognize the emotional equivalents of a tornado or a blizzard. At this level, I expect emotion recognition to be solvable by machine, at least as well as people can label such patterns. However, I do not expect researchers will have success matching human labels when such labels may not exist: just like we don't have special names for most of the states of weather – but rather only for its extreme states. In the in-between cases, we tend to use adjectives like, "ok day" or "partly cloudy" referring only to some quality of the weather. Similarly, not all aspects of affective state will be relevant or useful to observers. Sometimes just reporting the quality of the state – it's "ok" or it's "not so great" will suffice. Not that computers couldn't label any state that they could represent, only that such labels might not matter much unless they are in service of some greater purpose.

And this latter service is what motivates much of affective computing –how can we enable computers to better serve people's needs – adapting to you, vs. treating you like some fictional idealized user, and recognizing that humans are powerfully influenced by emotion, even when they aren't showing any emotion? More recent work on affect recognition at MIT has focused not on recognizing eight affective states with names like "joy" or "anger", but on recognizing things like "the state you're in when all is going well with the computer" vs. "the state you're in when encountering annoying usability problems." Each of these states may be a complex mix of emotions; however, if the pattern of their expression is distinctive, then we can try to have the machine recognize it, and respond to it. Our belief is that such discrimination is of value in designing systems that reduce user frustration.

There's an old saying in the business world: if you can't measure you can't manage it. Affect, like weather, is hard to measure; and like weather, it probably can't be predicted or controlled with perfect reliability. But, if we can do significantly better than random, then people will at least be less likely to get caught in a thunder storm without an umbrella. If computers can at least measure affect that is clearly expressed to them, say by irate users, then such measures can be useful in comparing product and interface designs.

It is likely to be the case that we can get good measures of some affective states but not others. Imagine if we could accurately measure how frustrating a certain kind of interaction is for a class of users: then you can say that making the interface one way is more frustrating than making it another way. I would like to see that added to price-performance curves: for the same price/performance, this interface is less frustrating. That kind of measurement can impact sales, not to mention user health and productivity. Similarly, one might try to measure the positive affect benefits of a particular design or interaction: Alice Isen has shown beneficial effects of positive affect on even very rational decision making (for an overview of several studies supporting this, see Isen (2000)).

Sometimes, when confronted with a really difficult problem like affect recognition, it is helpful to remember that partial solutions can still be of value. It is known that infants recognize some kinds of affect in speech, apparently long before they recognize what is said. Also, dogs, who presumably have no sophisticated models of affect, often give the appearance of being able to recognize it (for example, putting their ears back and tail down when their owner looks like he's had a bad day, as if to show a kind of empathy.) Such examples of recognition remind us that there are many applications even for systems that do not fully solve the problem. And, when a computer is capable of conversing with a person, the computer can augment its imperfect non-verbal sensing with an occasional verbal inquiry – "hey, how's it going?

As part of his doctoral research, Tim Bickmore crafted a "Relational Agent," a conversational character designed to build and maintain a long-term social-emotional relationship with users who were undergoing a month-long program to increase their exercise levels [Bickmore, 2003]. One of the factors believed to contribute to the success of this agent (vs. a control agent, which was also conversational about the task and goals involved in the interaction but didn't practice any deliberate social-emotional relationship skills) was its ability to converse about the user's feelings and show occasional empathetic caring and concern with both text and bodily expression. Even though users (rightfully) doubted the character had any feelings or really cared, and the computer character was up front about its limited abilities; nonetheless, users rated the relational agent significantly higher on likeability, trust, respect, feelings it cared for them, and willingness to continue interacting with it. These ratings were significantly higher on both days 7 and 27 of the long-term interaction. Thus, appropriate use of even limited affective abilities may lead to improved quality of experience for users, especially over long-term interaction.

3. AFFECT MODELING

Criticism 3: Little real progress has been made with cognitive modeling. How can we hope to make progress in modeling the much more subtle, multi-modal, and idiosyncratic processes that characterize emotional responses.

While some data exist regarding the effects of emotion on decision making and behavior, the majority of these data were collected in highly artificial lab environments, and both their robustness and their generalizability is questionable. IN addition, within the emotion research community, much disagreement exists regarding the types of mechanisms that mediate the effects of emotions, to the extent that attempts have even been made to characterize these mechanisms.

Existing models of emotion use highly stylized stereotypes of personality types and emotional responsiveness, which do not correspond to real behavior in real people.

The role of situational factors in emotion expression is poorly understood.

Together, these factors make it unlikely that any valid models of emotional processing will be constructed in the near future.

The machine vision community knows and appreciates many of these same modelling problems, which are not necessarily particular to emotion. In that community, researchers who try to understand and model human vision interact with those who want to build machines that see. One might think that the latter group couldn't make much progress without better answers, and better data, from the former group. But, in fact, both groups have learned from each other, and the latter group has built many useful systems, despite that neither group understands fully how the human vision system works. I expect that affect modelling research will proceed similarly: learning from human and animal models of affect, and in turn informing the development of such models by trying to build computationally working systems. I also expect attempts to model affect to result in computational solutions that are useful, even if they turn out to not follow nature's mechanisms.

With any complex modelling problem, there are ways to handle complexity and contribute to progress in understanding. If we want to model affect at the neuropeptide signalling level, then we are limited indeed, because we know only a small part about how such molecules communicate among the organs in our body [Pert 1997] and real-time sensing of local molecular activity is not easily accomplished; computer modelling is likely to push the human understanding in such an area. But there are other levels where we know more, and can describe a set of requirements, at least for high-level functions to be met. For example, it is known that emotions contribute to regulating and guiding attention, and to helping make decisions, generally biasing one's selection of next moves away from negative or harmful choices. A practical computational case for implementing such features can be seen in the work of Prof. Cynthia Breazeal at the MIT Media Lab. Breazeal is working to develop a "brain/controller" for a very complex one-of-a-kind robot made by Stan Winston Studios. This robot has 61 degrees of freedom; consequently, there are many ways it can physically move, including undesirable ways. Moreover, it is a very expensive machine, and cannot be easily repaired. Puppeteers who operate it have to be careful not to harm it. One of Cynthia's priorities in giving this machine "smarts" (to alleviate some of the tedious work of the puppeteers) is to give it mechanisms of self-protection: internal warning and regulatory systems that function, like emotions, to alert it to avoid certain potentially harmful behaviours. Although this is a small part of what an emotion system does, attempts to build these aspects of the system should not only prove useful – saving money and time in costly repairs -- but should also help clarify what is and is not known about corresponding human and animal mechanisms of emotion related to self-protection.

4. EMOTION EXPRESSION

C*riticism 4: The sine qua non of emotion expression is the physical body. To the extent that computers do not have physical bodies, they cannot reliably and believably express emotion. Existing attempts at expressing emotions in robots, which are at times referred to as embodied, are unrealistic and therefore unconvincing, and unable to generate the type of affective responsiveness characterizing human-human interaction.*

The opening statement of this criticism rings true; but in a superficial way: it is also true for all human forms of expression, not just emotion. Can you express even a cold mathematical solution, say the derivative of sin(x), without using your body? The last century has so separated the mind and body that people forget that cognitive thought involves the brain, with is part of the body. A thought involves biochemical and electrical signalling mechanisms: physical changes in the body. Emotions also involve such changes. Both thoughts and emotions can subsequently trigger actions visible to others: gestures, facial movements, head nods, shifts in posture, and other physical forms relating both to cognition and affect.

Computers also have bodies, albeit very different ones from ours. And we know computers can give the appearance of having emotion – expressing it with a facial display or a cute noise (think R2D2 in Star Wars). Disney animators have also demonstrated that emotional expression can be communicated effectively through characters, such as Aladdin's magic carpet, even though such characters don't have humanoid bodies. Thus, the issue is not one of the body, or of the body's capabilities to express, but the issue is a much harder one: how to get machines to express emotion in an appropriate way?

Embodied emotional expression is a challenge, but it is not as challenging as designing systems for appropriate expression. Somebody who smiles at the wrong time is considered stupid; machines don't yet know when the wrong time is. The same can be said for many other social and cultural display rules for emotion. Microsoft's animated paperclip office assistant has been particularly egregious in this regard: exhibiting joyful expressions right after annoying people, perhaps even while they are searching for how to turn it off. Imagine a human who expressed smiles and did a little happy dance after you cursed at him, declared him unhelpful, and showed him the door? Would you want him to visit you again? A lot of efforts to give machines expressive abilities have grossly neglected fundamental rules of human affect expression, and have consequently failed in accord with this neglect.

5. ETHICS

Criticism 5: Emotions, perhaps more so than thoughts, are ultimately personal and private. Providing information about the most intimate motivational factors and reactions. Any attempts to detect, recognize, not to mention manipulate, a user's emotions thus constitutes the ultimate breach of ethics and will never be acceptable to computer users. Attempts to endow computers with these abilities will lead to widespread rejection of such computer systems and will help promote an attitude of distrust to computers in general.

If your boss yells at you, is it wrong to detect his angry voice, or to recognize he is angry? Is it unethical, once you've recognized his anger, to try to take steps to alleviate his anger (or to "manipulate" it, perhaps by sharing new information with him, so that he is no longer angry?) One can think of situations where the foregoing answers are "no": e.g., he is yelling at you directly, and clearly wants you to recognize it and take steps in response. And, one can imagine the answers might be more complex if you surreptitiously detected his anger, and had nefarious purposes in mind by attempting to change it.

Humans routinely detect, recognize, and respond to emotions or manipulate them in ways that most would consider highly ethical and desirable. Playing music to cheer up a friend's mood, eating chocolate, exercising to perk oneself up, and other manipulations count among many that can be perfectly acceptable. That said, I do agree that there can be unscrupulous uses (by people and by people via machine) of affect detection, recognition, expression, and manipulation. Some of these, including ways affective machines might mislead customers, assuage productive emotional states, and violate privacy norms, are discussed in Picard (1997) and in Picard and Klein (2002).

6. UTILITY OF CONSIDERING AFFECT IN HCI

Criticism 6: Airplanes do not flap their wings. Just because humans use emotional abilities, and just because human-human interaction is replete with affective reactions, does not mean that computers should aspire to emulate what may in fact not be the best way for their uses. Emotions and passions tend to be more problematic than helpful in human-human interaction. There is not need to contaminate purely logical computers with emotional reactiveness.

First, let me address the last two lines of this criticism, which I believe have been solidly refuted by a number of bodies of work over the last few decades – for example, from the research of Damasio (1994), Isen (2000), Ledoux (1996), Salovey and Mayer (1990), their teams, and many others. Their studies have supported vital roles for emotion in many background processes: perception, decision-making, creativity, empathic understanding, memory, as well as in social interaction. Too many people, especially those who don't know the latest literature on emotion, think that emotion is only there when it jumps up and grabs their attention. Like with a blizzard, they don't notice the weather until it inconveniences them. But the latest findings suggest that emotion is always there, and usually it is helping regulate and bias processes in a helpful way, a way that contributes to intelligent functioning. Affective computing has emphasized such a role from its start: it has never been about making machines that would look "more emotional"; instead, it has been about making machines that would be more effective.

That said, cars do not have legs and airplanes do not flap their wings. Just because every living intelligent system we know of has emotion does not mean that intelligence requires emotion. Although people are the most intelligent systems we know, and people's emotion appears to play a vital role in regulating and guiding intelligence, it does not mean there might not be a better way to implement some of these goals in machines. It may be possible that there is something like the wheel, which has no precise human or animal equivalent, but which provides for some of the same locomotion goals. There may exist a kind of alien intelligent living system, something we've never encountered, which achieves its intelligence without having anything like emotion. Although humans are the most marvellous example of intelligence we have, and we wish to build systems that are natural for humans to understand, these reasons for building human-like systems should not limit us to thinking only of human abilities.

Finally, affective computing has always emphasized the need for a balance. Some machines won't need any emotional abilities; while others might be improved with some subset of them. There is a time to express emotion, and a time to forbear; a time to sense what others are feeling and a time to ignore feelings. In every time, we need a balance, and this balance is missing in computing. Designers of future computing can continue with the development of computers that ignore emotions, or they can take the risk of making machines that recognize emotions, communicate them, and perhaps even "have" them, at least in the ways in which emotions aid in intelligent interaction and decision making.

7. REFERENCES

Bickmore, T. (2003) *Relational Agents: Effecting Change through Human-Computer Relationships.* Ph.D. thesis, MIT Media Arts and Science, 2003.

Breazeal, C. (2003) http://robotic.media.mit.edu/projects/Leonardo/Leo-robot.html

Damasio A. (1994). *Descartes' Error: Emotion, Reason, and the Human Brain.* Gosset/Putnam Press, New York, NY 1994.

Isen, A. M. (2000). "Positive Affect and Decision Making", in *Handbook of Emotions*, Eds., M. Lewis and J. Haviland. Guilford, New York, 2000.

Kagan, J. (1984). *The Nature of the Child*. Basic Books, New York, 1984.

LeDoux, J. E. (1996).*The Emotional Brain*, Simon & Schuster, New York, 1996.

Mori Survey (1999). Rage Against the Machine, 1250 users in the United Kingdom.

Nimoy, L. (1995). *I am Spock*. Hyperion, New York, 1995.

Pert, C. (1997). *Molecules of Emotion: Why You Feel the Way You Feel.* Touchstone, 1997.

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

Picard, R. W. (2003), "What does it mean for a computer to "have" emotions?" in "Emotions in Humans and Artifacts," ed. by R. Trappl, P. Petta and S. Payr. MIT Press.

Picard, R. W. and Klein, J. (2002) "Computers that Recognise and Respond to User Emotion: Theoretical and Practical Implications," *Interacting with Computers*, 14, 2 (2002), 141-169.

Picard, R. W., Vyzas, E. and Healey, J. (2001). "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," *IEEE Transactions Pattern Analysis and Machine Intelligence*, Vol. 23, No. 10, October 2001.

Salovey, P. and Mayer, J. D. (1990). "Emotional Intelligence", *Imagination, Cognition and Personality*, Vol. 9, No. 3, 1990, 185-211.

Schachter, S. & Singer, J. E. (1962). Cognitive, social and physiological determinants of emotional state. *Psychological Review*, *69*, 379-399.