Synthetic Emotion

Rosalind W. Picard

Imagine the most important interview of your life, with the head of the company that you have always wanted to work for. Your potential boss asks you tough questions about problems you have solved, challenges you have faced, and why you want to leave your present job. At the end of this grueling meeting, you are told that you were too nervous-sounding, had unusually short pauses in your speech, were evasive with eye contact, and had cold clammy hands. Fortunately, this was not the real thing.

You have just had a practice interview in front of your trusted computer. The computer interviewing agent, displaying the face and nuanced expressions of the CEO, asked you questions while listening to changes in your voice and discourse parameters. It watched your facial expressions and body language, sensing changes in physiological parameters such as your skin conductivity, temperature, and color. It observed your affective responses to see where and how they differed from what it usually senses from you in your day-to-day interaction. The computer recognized not only *what* you said and did but also *how* you said and did it. It noticed both overt and subtle emotional expression, much like a person with keen emotional intelligence. Then it carefully constructed visual and verbal responses that were appropriate to what it recognized from you.

With a camera, microphone, and various sensors, the computer "affective mirror" just described is more advanced than the reflective glass people have traditionally practiced in front of, even though it is not as insightful as a skilled human listener. Nonetheless, there are times when a computer's availability, patience, and nonpartisan judgment cannot be beat. Being able to "try out" that important talk in front of your computer has a certain convenience and privacy that some find encouraging and comfortable. But, can a computer really recognize what you are feeling, and respond in an intelligent way?

New research in "affective computing," computing that relates to, arises from, or deliberately influences emotion, aims to give computers skills of emotional intelligence, including the abilities just illustrated. These skills are crucial for learning and for savvy social interaction. For example, recognizing someone's emotional response is key to sensing if what you have just done or said is met with by approval or disapproval, interest or boredom, confusion or understanding. If you are a healthy person, you naturally began to acquire this ability as an infant or toddler. The youngest members of our society can see if they have pleased or displeased their caregivers long before they master verbal and computational skills. If you are a computer, or an autistic person, then these abilities do not come naturally. Computers are increasingly verbal, and have formidable computational abilities, but still they cannot see if the graphics they are displaying are delighting or boring the viewer.

How might computers get better at recognizing emotion? Machines are

presently being equipped with sensors – cameras, pressure indicators, muscle tension sensors, and more – that can detect patterns of expression from the user. These patterns can be assembled for clues to the user's emotional state. Machines will combine facial expression recognition, vocal inflection analysis, and other physical responses with other information you might share, such as your goals, standards, and preferences. The machine will reason about your expressions jointly with what it can sense regarding the situation at hand, and try to infer how you might feel.

Within a couple decades, we will see computers that get to know your typical emotional expressions: how you look and act when you really like something vs. when you are not interested or you feel disliking, what you tend to do when all is going smoothly vs. what you do when something is increasingly frustrating, and so forth. The desktop computer might recognize if you are behaving the way you usually behave late afternoons when you are in the mood to play a video game, and pre-load the latest one for you, just in case you decide to play. Your wearable computer might recognize that you are walking in a particularly lilting and probably joyful way, and ask you if you'd like to hear some old favorite tunes similar to those you previously requested when you walked that way. Your car computer might notice that your driving parameters are a bit different than usual, and guess that maybe your unusually loud singing is distracting you; it might relay cell phone calls to your voicemail and display a small increase in a risk meter, if you approve of that option. If you don't like something any of these computers does, and you frown or curse at it, it could offer to stop doing that which irritates you, or at least make sure you know how to turn off that irritating feature.

Researchers have recently taught a computer to recognize which of eight emotional states a user is expressing over a long period of time, with over eighty-percent recognition accuracy, based on facial expressions or on physiological signs such as respiration and muscle tension. However, the current results place the state of the art in affect recognition at about where speech recognition was in the 50's. The "emotions" the computer can recognize now are a small set of pre-segmented and carefully expressed states, analogous to the early days of speech recognition where the computer could recognize the digits "1," (pause) "2," (pause) ..., "10," but not recognize the same digits occurring as part of a conversation, much less recognize the other words in the conversation. The problem is more than one of pattern recognition. It is not clear if even in the next millennium the computer will know what the user *truly* feels, since one's innermost feelings remain mysterious in form, and are unable to be sensed by any known methods. Often people do not even know what they feel, and words are insufficient for such description. Because emotions are comprised of both thoughts and feelings, and because neither of these is perfectly observable, the emotion recognition process is mostly one of inference based on partial observations.

An outsider can see a change in how you are moving – with upward, joyous

motions, with lumbering, heavy, slow ones, or with your characteristic neutral ease, all which can reflect emotional state, but none of which declare it. What the computer can potentially erceive is similar to what another person can perceive – external visible or audible changes in how you move, talk, gesture, choose words, and make facial expressions. Nonetheless, the clues that people use to infer emotional information, although imperfect, are powerful communicators. Computers could make use of these cues to better learn how to please and serve users.

Artists have a long history of communicating emotion in graphics, but the process has been largely intuitive. Only recently have researchers tried to mechanize the dynamics of emotion, designing computational programs that can automatically change, for example, the state of an animated character from happy to sad. An emotional state change effects not just face, voice, and posture, but the entire spatio-temporal form of the character, physically and cognitively. The character will pick up an object differently when happy than when sad. The character will walk differently. Even the way that the character listens will be affected – people who are in a good mood are more likely to hear "presents" than "presence," and those in a bad mood are more likely to hear "banned" than "band." Human emotion involves a constellation of interacting bodily signals, influencing thought and behavior, modulating every action and movement.

The next millennium will bring a balance to scientists' understanding of cognition and emotion, and a more authentic model of human behavior to graphics and animation. Researchers who try to synthesize mechanisms of human emotion in graphical systems will ask questions that emotion researchers have yet to contemplate, and this will help advance knowledge. In the late nineteenth century, Freud and James argued for the importance of understanding emotion in human behavior; but, the twentieth century has downplayed emotion, modeling man as an unimpassioned cognitive machine. Such a model describes only in part, as seeing in a mirror dimly. The new century will lead to the use of computing for more than making a perfect image; computing will be used to illuminate the nature of human emotion.