What Emotions Are Necessary for HCI?

Dolores Cañamero
Artificial Intelligence Research Institute, Spanish Scientific Research Council

Introduction

The Human-Computer Interaction (HCI) community is showing increasing interest in the integration of affective computing in their technology. Particular attention is being paid to research on emotion recognition, since computer systems should be able to recognize human emotions in order to interact with humans in a more adaptive and natural, human-centered way. However, other aspects of emotion might be equally important, depending on the kind of interaction we are aiming at. My research has focused on two different aspects of affective computing. On the one hand, the generation of affect-driven expressive musical performances that can elicit different emotional states in humans, and that can also contribute to the understanding of how listeners perceive the expressive aspects that the musician intends to communicate. On the other hand, my work focuses on emotion synthesis for action selection and behavior modulation in autonomous agents. Both topics raise many open questions that are equally relevant within the HCI community.

Expressing and Eliciting Emotion with Music

One of the main problems with the automatic generation of music is to achieve the degree of expressiveness that characterizes human performances. The main difficulty arises from the fact that performance knowledge—including not only the sensible use of musical resources, but also the ability to convey emotion—is largely tacit, difficult for musicians to generalize and verbalize. Humans acquire it through a long process of observation, imitation, and experimentation (Dowling and Hardwood 1986). It seems thus natural to adopt a similar approach for the automatic generation of expressive music. This was the main motivation to develop SaxEx, a Case-Based Reasoning system that turns inexpressive perfomances of jazz ballads into expressive ones by imitating human performances of other jazz ballads stored in the case base of the system.

In its current version (Arcos et al. 1999) SaxEx uses, in addition to expressive musical parameters (dynamics, rubato, vibrato, articulation, and attack), information concerning the affective dimensions of music in order to generate expressive performances. The system is given as input: a score (a MIDI file), which provides the melodic and the harmonic information of the musical phrase; a sound file containing an inexpressive performance of the phrase; and qualitative labels along three affective dimensions (sad-joyful, tender-aggressive, and calm-restless), through which the user can specify the expressiveness desired in the output. Values for affective dimensions guide the search in the memory of cases so that the system gives priority to performances with similar affective qualities. The selected performances are used to modify the inexpressive input file with the help of background musical knowledge (theories of music perception and understanding). The output of SaxEx is a new expressive performance (a sound file), obtained by transformations of the original inexpressive sound through imitation of the expressive resources of the selected sample cases (Cañamero et al. 1999). Readers can visit the SaxEx web site for sound examples: http://www.iiia.csic.es/Projects/music/Saxex.html. As for the uses of SaxEx in HCI, we envisage three main applications: as a pedagogical tool, as an experimentation tool for musicians, and as a module in multimedia applications to generate expressive music coupled with video images or animated cartoons.

Although the approach adopted seems to be very successful in grasping performance knowledge—listeners agree that the music generated by SaxEx sounds very natural and human-like, and they can easily identify the affective intention of the piece—a number of issues are still to be solved. The affective space is characterized in SaxEx along three dimensions (sad-joyful, tender-aggressive, and calm-restless), instead of using isolated adjectives, for two main reasons. First, the use of affective adjectives for musical purposes presents problems related with linguistic ambiguity. Second, the selected dimensions reflect three dimensions of affective meaning which seem to be culturally universal (Osgood et al. 1975): "evaluation" (good vs. bad), "potency" (powerfulness vs. powerlessness), and "activity" (activation vs. tranquillity). It is not clear, however, that these dimensions allow to express the full range of emotions that music could in principle convey. Much experimentation using very different musical styles would be needed in order to look for possible counterexamples. A major difficulty here is the fact that not much theory is available to guide experimentation. Surprisingly, as Peretz. et al. (1998) point out, although music is often characterized as the language of emotions, very few studies in psychology have been devoted to the study of music as an emotional language. Another issue that remains open is how musical elements relate to emotions. In SaxEx, musical analysis of the score and its partition into affective regions are currently performed separately. One of the extensions envisaged is to have the system learn associations between affective labels and expressive musical parameters for situations appearing recurrently in the phrases, and to use this knowledge to help assess when performers introduce expressive variations due to the logical structure of the score vs. other motives (e.g. their own expressive intention). Again, little psychological research has been devoted to this issue. It would seem that the joyful-sad dimension is to a big extent conveyed by structural properties of the music—mode and tempo of the piece (Crowder 1984)—rather than by human interpretation, but to our knowledge, almost nothing is known about the other dimensions.

Does HCI Need Emotion Synthesis?

The other research line that I have undertaken in affective computing is emotion generation or synthesis—computers (or in general artifacts) that "have" emotion, to put it in Picard's words (Picard 1997). My work in emotion synthesis focuses on the use of an emotional system for action selection and behavior control in autonomous agents living in a highly dynamic world, where different elements, external and internal, can threaten their lives. Emotions are in this case adaptive mechanisms related with the agent's survival, and are elicited by the presence of events that are significant with respect to survival-related goals. In my physiological model of emotions (Cañamero 1997), a motivational system drives behavior selection and organization based on the notions of arousal and satiation. Basic emotions (anger, boredom, fear, happiness, interest, and sadness) exert further control of the agent's behavior through the action of hormones that may affect the agent's perceptive, attentional, and motivational states, also modifying the intensity and execution of the selected behavior. Emotions are therefore a key element in determining what the resulting behavior of the agents, including their interactions with other agents, whether artificial or human, is going to be.

The explicit inclusion of emotional mechanisms that have direct feedback on the behavior of the system—in this case, agents—seems natural in this kind of application, but it is less clear whether every system that has to interact with humans must "have" emotions as well. Work on believable agents shows that humans easily attribute the illusion of life, personalities, and emotions, to characters that do not have explicit mechanisms for them. For some applications, one may think that the ability to recognize the emotional state of the user and to reason about it would suffice for the system to respond in a way that is adapted to the user's state. I will argue, however, that endowing the system with some mechanism that allows it to generate and show emotional behavior—putting some sort of emotions into it—can be very beneficial for most applications. Indeed, for a natural and human-like interaction with humans, HCI systems must be adaptive themselves. Emotions are essential mechanisms for adaptation, not only as far as survival is concerned, but most importantly to be able to properly interact in the social world. In social interactions, emotions are mechanisms for communication and negotiation, and they emerge from (and have an impact on) the interactions among the different individuals involved. If we devoid HCI systems of the ability to *have* emotions, we may end up developing autistic HCI systems. My claim is thus: Emotion synthesis is necessary to close the human-computer loop in HCI. Many tough problems have to be solved first, regarding for instance the choice of the emotional repertoire needed for the different applications, as well as many design choices; the solution of these problems requires the collaboration and mutual feedback of both the affective computing and HCI communities.

Conclusion: What Does Anthropomorphism Buy Us?

Since HCI is essentially human-centered, it seems natural to take a human-centered approach when integrating affective components in this technology. However, it is far from clear what "human-centered" means as far as affective computing is concerned. In what could be called the weak sense of the term, human-centered affective computing is intended to be adapted to the users' emotional world in order to enhance their cognitive capabilities and interactions, hiding the computer in the human-computer loop. The strong sense of the term seems to imply, in addition, some form of anthropomorphism, i.e. the attempt to integrate elements that mimic human emotion. Is the second sense of the term necessary to achieve the first one? Does affective computing need to take into account the full complexity of the human emotional apparatus? What does it mean for a computer to have human-like emotion?

I argued somewhere else (Cañamero 1998) that, if we have an engineering purpose in mind, it is neither desirable nor possible (at least for now!) to endow artificial systems with emotions of full human-level complexity (see (Picard 1997) for a list of components of human emotional systems). I also advocated for the adoption of a functional approach (as opposed to a "componential" one) to guide the design of the system. Following Frijda (1995), from a functional point of view, we have to pay attention to the properties of the structure of humans and their environment which are relevant for the understanding of emotions, and that can be transposed to a structurally different context to give rise to the functions or roles of emotions we are interested in. Not all the functions of emotions are relevant for every application—their choice, as well as of the mechanisms used to give rise to behavior fulfilling these roles, depend on the kind of interactions needed in the application. In other words, the design of the emotional system must be guided by the requirements on the system and on the interactions with the human users and their environment, so that they can show human-adapted (and in some sense human-like) emotional behavior. The other side of the coin could be that human users might tend to attribute to a system showing human-like emotional reactions the full complexity of their own emotions. This may result in the user being sometimes worried about the system, disappointed, or frustrated when s/he realizes that the system is not a human. Or perhaps this danger will simply not exist, as humans will become used to interacting with affective systems and get to know them (including their limitations). In any case, we as designers should always make clear what is "inside the system" in order to avoid promoting false illusions in the users. This is not in contradiction with the fact of trying to make emerge not only human-centered, but also human-like emotional interactions in HCI.

References

Arcos, J.L., Cañamero, D., López de Mántaras, R. (1999). Affect-Driven CBR to Generate Expressive Jazz Ballads. In *Proceedings of the 3rd Intl. Conf. on Case-Based Reasoning*. Berlin-Heidelberg: Springer-Verlag, LNAI (in press).

Cañamero, D. (1997). Modeling Motivations and Emotions as a Basis for Intelligent Behavior. In W. Lewis Johnson (Ed.), *Proceedings of the 1st Intl. Conference on Autonomous Agents*, pp. 148–155. New York: The ACM Press.

Cañamero. D. (1998). Issues in the Design of Emotional Agents. In *Emotional and Intelligent: The Tangled Knot of Cognition. Papers from the 1998 AAAI Fall Symposium*, pp. 49–54. TR FS–98–03. Menlo Park, CA: AAAI Press.

Cañamero, D., Arcos, J.L., López de Mántaras, R. (1999). Imitating Human Performances to Automatically Generate Expressive Jazz Ballads. In *Proceedings of the AISB'99 Symposium on Imitation in Animals and Artifacts* (AISB'99, Edimburgh, UK, April 7–9, 1999) pp. 115–120.

Crowder, R.G. 1984. Perception of the Major/Minor Distinction: I. Historical and Theoretical Foundations. *Psychomusichology*, 4, 3–12.

Dowling, W.J., Harwood, D. 1986. *Music Cognition*. Academic Press, 1990.

Frijda, N.H. 1995. Emotions in Robots. In H.L. Roitblat, J.-A. Meyer (Eds.), *Comparative Approaches to Cognitive Science*, 501–516. The MIT Press.

Osgood, C.H., May, V.H., Miron, M.S., 1975. *Cross-Cultural Universals of Affective Meaning*. Urbana, IL: University of Illinois Press.

Peretz, I., Gagnon, L., Bouchard, B. (1998). Music and Emotion: Perceptual Determinants, Immedicay, and Isolation after Brain Damage. *Cognition*, 68, 111–141.

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