User Affect, Emotions and Personality in HCI
How interaction can be improved by utilizing user's affect

Theodoros Foteinakis
School of Computer Science
University of Birmingham
Birmingham, United Kingdom
theofote@gmail.com

Abstract—In recent years the concept of designing applications has become user-centred and affect-driven. This survey describes the basic background of affective HCI, the main psychological aspects of the user and how emotions are created. Furthermore, the ways that machines can sense the user's psychological state and the specificities of each method are analysed along with some key factors about manipulating it to achieve an optimum outcome. Finally, the ways machines respond to users' emotions and related constraints are explained and examples of affective computing cases are provided along with suggestions for future designing of interactive applications utilizing the most effective techniques and patterns.

Keywords—Affective HCI; Affective Computing; Affect Recognition; Personalized Human-Computer Interaction; Emotions; Human-Centred Design

I. INTRODUCTION

Using a machine like a personal computer does not always produce the same results in terms of quality. While a computer is actually soulless materials, combined to form a functional piece of technology capable of accomplishing a huge variety of tasks, making it adaptable to the user's emotional state can make a great difference in the outcome of its interaction with him (Brave and Nass, 2003). Brave and Nass argue that emotions used to be considered irrelevant with Human Computer Interaction and that the user should discard his own as much as possible to achieve effective use of technology. Recent psychology studies proved that emotional state plays a critical role to every kind of interaction with computers and since then, much research has been done on this field in order to improve interaction and achieve maximum effectiveness and enjoyability on computer usability (Calvo and D'Mello, 2010). “More often than not, the user is now the central component of system design and user needs drive both the nature of the user interface, and the function allocation of tasks between the user and the machine” (Hudlicka, 2003).

The main element of Human Computer Interaction and associated affect and personality research is the emotional state of the user. “Emotion is at the heart of any human experience and an essential component of user-product interactions and user experience” (Forlizzi and Battarbee, 2004). This essay will describe some basic themes in the emotion creation process as related parts of the human brain and emotion processing in them.

There are many parameters needed to be taken under consideration to effectively detect a user's psychological state. The human brain is complicated for a machine to fully understand it and while a human can correctly interpret the feelings of another by observing the facial expressions, gestures and tone of the other person's voice apart from processing the actual words that are being used, a machine can find it very difficult to do so (Pantic and Rothkrantz, 2003). On the contrary, there are cases where a computer can easily detect signals not visible for humans that indicate a particular emotion, mainly in cases such as autism (Kuriakose et al., 2012). This is a very important aspect of Human-Computer Interaction and will be discussed in the following sections of this paper.
Affect detection and sensing can lead a machine to change its state so that it corresponds to the user's mood. There are countless fields that this can be applied to and in many cases the benefits are extremely valuable such as medical or high risk and responsibility tasks that need technology assistance to guarantee error prevention (Fragopanagos and Taylor, 2005; Kuriakose et al., 2012). Apart from cases of high importance, technology and most specifically HCI, can provide the luxury of machines that respond to the user's needs without being necessary that the user will take any actions. In this survey such tasks are going to be analysed in order to make clear which are the aspects where affect recognition can make the difference.

Apart from making machines fully automatic, the way that they interact with the user is changing. Being able to recognise his needs and current psychological state means that they are going to respond accordingly, attempting to change the user's mood if it is negative and adapt to it during all this procedure so that the best result out of this process will be produced (Hudlicka, 2003). In order for this to be possible, some terms and aspects of machines have to be investigated so that the details that can affect the user's attitude and emotional state can be manipulated resulting the optimal outcome (Brave and Nass, 2003).

The rest of the essay is organised as follows: Section 2 provides some basic information about emotions and their nature, section 3 describes the most common methods for machines to sense human emotions, section 4 is about the way emotions can emerge during the interaction process (probable stimulus) and section 5 briefly explains what happens after a machine has successfully detected the emotional state of the user. Finally, in section 6 possible future work is provided and in section 7 the conclusions of this research paper are explained in total.

II. RESEARCH ON EMOTIONS

Emotions in general used to be considered unmeasurable phenomena, providing information only to philosophers since they were not considered to be connected with mediation of behaviour and motivation (Hudlicka, 2003). After it was found that emotions actually have a huge impact on people by “controlling us”, altering our perspectives and our reactions, emotion research gained extreme popularity and was strongly associated with HCI and Artificial Intelligence in the technology field. Rational aspects of behaviour like decision making, learning, planning, perception and action selection were found that are highly affected by emotions for important but also for insignificant decisions in daily life.

The human brain contains three key regions that handle the emotion recognition process. The “Thalamus”, the “Limbic System” and the “Cortex” (LeDoux, 1995). While the Thalamus gets as an input information from the external environment and does the basic emotion processing, it sends information to the Cortex and to the Limbic System. The Limbic system is used for processing primitive emotions and it can either send signals to the body resulting psychological response or to the cortex for cognitive processing of more complicated ones (Brave and Nass, 2003). It is worth to mention that there can be stimuli that is internally generated and then processed like when someone is thinking about a had task he may be obligated to complete and thus stress and anxiety may emerge.

Neuroscience and psychology has proven that emotions can alter the processing of information in the human brain by making it even bypass the cortex, often causing unexpected behaviour as also that there are sensors in the human body that are directly connected to the emotional circuitry in the human brain (Hudlicka, 2003). The part of the human brain which is responsible for such processing is called “the Amygdala” and is a key term in emotion research (LeDoux, 1992). “The Amygdala” is defined as “a structure in the limbic system that is linked to emotions and aggression. The amygdala functions to control fear responses, the secretion of hormones, arousal and the formation of emotional memories” (Cherry, n.d.). Damasio argues that cognitive processing is directly linked with emotions to produce what is widely known as rational decision-making (Damasio, 1994). However, most people ignore the presence of emotions to every single decision they might take, unless the current situation is an extreme one, full of emotional tension.
One very important aspect of emotions is their impact on social relations, by constructing, maintaining, altering or even destroying social interactions (Hudlicka, 2003). The strong connection between cognition and emotion causes unexpected behaviour in many occasions as it can also be handled in a positive way (e.g. someone calms down by breathing slowly and deeply before he performs an act in order to be concentrated and relaxed instead of nervous). This connection is being investigated by researchers to be able to analyse psychological issues and find a way to correct them.

Human Computer Interaction is based on emotion theory to enhance user experience and system effectiveness. People cannot think or act without at least unconsciously involve emotional processing in it (Picard, 1997). This applies to everything including using a machine like a computer. As Brave and Nass argue “It is now understood that a wide range of emotions plays a critical role in every computer-related, goal directed activity” (Brave and Nass, 2003).

Since not all emotions are characterized by the same level of complexity, they can be categorized as “primary” and “secondary”. Primary emotions are primitive reactions to sound and image like fear, disturbance and pleasing feelings and can be caused by the tone of certain sounds, the colours and brightness of an image, its motion (Detenber and Reeves, 1996) and the way they are introduced to the user etc. (Brave and Nass, 2003). Brave and Nass argue that emotions belonging to the “secondary” category are the ones that HCI is targeting mainly. Such emotions are satisfaction, frustration and pride. It is not rare for the Limbic System and the Cortex to collaborate like in cases where the initial emotion of fear can be identified as something harmless or something that triggers an action that needs to be done immediately. It is worth to mention that researchers struggle to categorize emotions in a universally acceptable way (Lee and Narayanan, 2005).

The ultimate goal for an effective machine which senses the user’s emotions would be to make it correspond exactly as a human being (Pantic and Rothkrantz, 2003). Pantic and Rothkrantz argue that a machine cannot interpret emotions correctly and with ease in every occasion since people are not the same and their reactions for a specific emotional state may differ. Cultural differences are a strong example of large groups of people reacting in a different way under the same circumstances. However, despite the fact that the main (primary) emotions tend to be expressed universally in the same way and this could be taken into advantage by researchers to create a universal model for people’s reactions in defined situations, some of the
secondary ones like happiness and surprise tend to be expressed in different ways (Ekman et al. 1987).

Designing systems for different cultures needs to be done with caution and knowledge of who is going to use them (Bourges-Waldegg and Scrivener, 1998). As people from different cultures are characterised by different ways of interpreting events and unique ways to express themselves, a machine specially designed for them can follow particular patterns to match those specifications. However, there may be machines or applications that are designed to be used in a multicultural environment such as a large international company or even a service on the internet. Bourges-Waldegg and Scrivener argue that in these cases, design must follow commonly accepted rules, as close as possible to general preferences ignoring cultural differences as long as it is possible.

The emotion recognition complexity goes even further when looking at what signals a machine is detecting and how it interprets them. Emotions can be detected through many elements of interaction between a human and a machine such as text, facial recognition, gesture recognition, vocal intonations, words used in vocal commands etc. Facial expressions and vocal intonations are considered the most important ways a machine interacts with a human and they are the main parts of interaction that are being investigated (Pantic and Rothkrantz, 2003). Pantic and Rothkrantz argue that a multi-modal analyser will need multiple channels to correctly interpret an emotion because it has to take into consideration multiple factors at the same time such as analysing a smile which may not always indicate happiness as in many occasions it is accompanied by tears and it expresses grief.

Facial expressions and tone of speech are considered the most important ways of expressing emotions in terms of targeting them to be sensed by a machine. The direct connection of tone of speech and facial expressions with primitive feelings (anger, fear, sadness, happiness, surprise disgust) makes them very helpful when designing for HCI (Busso et al., 2004). Not all emotions are optimally detected using the same senses, for instance happiness and anger are better identified though eyesight while sadness and fear through audio (voice) (De Silva et al., 1997). Buso et al. argue that a fusion of both facial expressions and vocal intonations recognition can offer the best results when the appropriate approach of the process is used depending on the circumstances (Busso et al., 2004).

Recognition of emotion from text as user input is one the most important kinds of interaction since “text is still the main communication tool on the Internet” (Wu et al., 2006). Text can provide information on whether it is part of a “happy” or a “sad” document. However, recognizing user emotions through text has been proven very difficult because of the big variety of languages used and the complexity of them (e.g. at some cases people may not even use their native language properly). If the recognition is correct then the system is able to change its state according to the psychological state of the user and provide a much more effective and even pleasant type of interaction (Lee et al., 2002). In order for a machine to be able to understand the user's emotions, several methods has been adapted with the most popular one to be using emotional keywords and classifying them in categories to calculate the emotion level of each case while other methods also exist as pragmatic intent and paragraph structure (Dijkstra et al., 1995). Lee et al. introduced a new method of identifying user affect from text by extracting semantic information out of it and combined with research on psychology models and custom algorithms, the probabilities of each emotional state could be calculated for each situation. The experiment with this method showed that it produced better results than most of already known and used methodologies for extracting information about the psychological state of the user through text (Lee et al., 2002). However, text is considered an outdated method of interaction and as Cassell et al. argue, “Multimodal interfaces that include voices, faces and bodies can now manifest a much wider and more nuanced range of emotions than was possible in purely textual interfaces” (Cassell et al., 2000).

Finally, using voice commands and gestures are important methods of emotion recognition. The public is familiar with the former, mainly through the wide use of voice commands in most mobile phones, even some of the second generation ones (2G) had a few basic voice commands available like
calling someone by saying his name (assuming he is stored as a contact in the telephone catalogue of the device). Microsoft's Kinect for Xbox 360 is a very popular application of gesture-recognition interaction. Gesture interaction is the evolution of an earlier form of interacting with a machine like a mouse or a joystick of a personal computer, well known types of interaction that however required some hardware along with the gestures to be possible for the machine to understand user input (Freeman and Roth, 1995). Freeman and Roth argue that there are two types of gesture interaction, the static and the dynamic. While the static is able to recognize an image of a gesture made by the user, dynamic gesture interaction can recognize movement as a sequence of images. “Gestures are expressive, meaningful body motions involving physical movements of the fingers, hands, arms, head, face, or body” (Mitra and Acharya, 2007). Since gestures are very difficult for a machine to detect because they are ambiguous and not completely specified, much caution is needed in the interpretation of them to avoid errors.

IV. HOW ARE EMOTIONS CREATED DURING THE INTERACTION WITH A MACHINE

What could possibly be the cause of specific emotions that are raised during the interaction with a machine? Do these emotions pre-exist, are they created the exact moment when the interaction takes place or could they emerge after an amount of time after the interaction is finished? What is the role of a user's personality in this and what other interaction characteristics heavily affect this process? Many questions regarding emotions and how they appear are raised when trying to analyse interaction between a human and a machine.

In Figure 1 some common emotions of users are displayed, grouped according to similarities to briefly indicate the complexity of a user's psychology. It is apparent that except their plurality, an additional difficulty in distinguishing them is the fact that some are very similar, yet different like the emotion of feeling pleased and the one of feeling glad. It is extremely difficult for a machine to discriminate between such cases.

There are several factors that play a key role in the creation of emotions of a user. At this stage it is important to distinguish emotions from other terms that describe psychological aspects like mood and sentiments (Brave and Nass, 2003). An emotion is always about a relationship with a particular object and not a general feeling, a user gets angry with something or feels happy about something (Frijda, 1994). Something has to interact with the user for the emotion to be created, emotions cannot be evoked without the interference of an object. A mood on the other hand, is a more general feeling that is not dependent on an object and it could be a continual situation lasting from a small period of time to weeks, months etc. Depression is an extreme case of mood. Lastly, sentiments are characteristics of an object (or an interface etc.) that the user associates with an emotional state, in some cases due to previous experience with such characteristics or subconsciously, without cognitive processing of them. A good example would be the black colour, which is a social stereotype to be associated with sadness and grief.

In many cases, moods can affect the kind of emotions that are generated towards a particular stimulus (Brave and Nass, 2003). Being in a positive mood may produce different emotions and better

Figure 1 - Baumgartner's multidimensional scaling for 28 basic experiential emotions
performance towards interaction with an interface than being in a negative one, thus positive stereotypes should be encouraged (e.g. an application for learning could be presented as entertainment, achieving much better results due to the positive perception of it by the user). Also sentiments tend to be easily placed in stereotypes which means that people who may learn about them can interpret a characteristic of an application according to general beliefs and not essentially how they would do it according to their own personality.

Generating positive emotions has been proven by researchers that enhances the experience a user has with a machine, making him frame it as entertainment rather than work (Brave and Nass, 2003). Brave and Nass argue that for this reason interaction should be handled in a way that it does not disturb the users but creates positive emotions, according to some patterns to avoid causing unexpected or inappropriate behaviour (like laughter at work). Better performance at work can also be achieved through producing positive user affect (Isen et al., 1991) and additionally, combined with a positive mood, users can evaluate an interface or even a web service more positively no matter what emotion may be generated through the interaction process (Brave and Nass, 2003). Additionally, emotions “play a critical role in what have traditionally been considered 'rational' aspects of behaviour: perception, decision-making, learning, planning and action selection” (Hudlicka, 2003). Considering this finding, along with the ability of positive emotions to improve performance and perception of an interaction, it is apparent that focusing on emotions and trying to manipulate the way they are evoked is a very promising strategy for enhancing the interaction experience between a machine and its user.

An important key aspect of users' psychology is that negative experiences tend to be remembered better than positive ones (Brave and Nass, 2003). Thus, avoiding user frustration is usually a priority over improving the current form of interaction to generally make it more advanced by designers, since negative results will cost more and will be harder to reverse. Moreover, users seem to strongly remember occasions where their current mood has been congruent with the stimuli of the situation rather than others with incongruent mood-stimuli combination (e.g. a user will remember more facts about a cheerful interface or an enjoyable application when he is in a positive mood than a formal, strict application) (Ellis and Moore, 1999).

Another important specificity of Human Computer Interaction is that users want to enjoy the interaction process, they want it to be smart, playful and rewarding (Petersen et al., 2004). An easy and rewarding experience is always a more welcomed one by the users, hence the effort needed to learn how to use an interface should be as little as possible and the learning process should be rewarding from the start so that the user will not lose interest in it (Forlizzi and Battarbee, 2004). Petersen et al. argue that “beyond rational and functional requirements”, aesthetic interaction is becoming a more and more popular concern of designers. According to Petersen et al. “Aesthetic Interaction is not about conveying meaning and direction through uniform models; it is about triggering imagination, it is thought-provoking and encourages people to think differently about the encountered interactive systems, what they do and how they might be used differently to serve differentiated goals”. Aesthetic design is “rooted in the socio-cultural context of people's everyday life” and is becoming an important element for the future of a system, interface etc., however more research on this field is required for it to be established in every production life-cycle of future systems (Petersen et al., 2004).

V. RESPONDING TO USER'S AFFECTIVE STATE

The main aim of recognising user affect and emotions is to make the machine adaptable according to them by changing it's state in real-time. This way, the machine will always try to get the best result during the interaction with the user and respond in appropriate ways to address user affect depending on the situation. However, things are not very simple when making a machine adaptable to the user, there are several constraints and decisions to be made in order to avoid exaggerated behaviour and prevent errors (Hudlicka, 2003).

Solutions in defining the exact state a machine should adapt are not explicit. In general, level of stress is considered something that should be handled in a way to be maintained as low as possible
but in some cases stress is essential to achieve maximum performance (Yerkes and Dodson, 1908) such as high responsibility or high risk tasks like surveillance of pilots while flying an aeroplane (Matthews, 2000; Pantic and Rothkrantz, 2003). Hence, not all situations are the same and should be treated accordingly, some of them do not even require a change of the machine's state as it would be useless (e.g. most users would prefer an ATM machine to remain as we know it) while others are heavily counting on the machine's change of state to successfully complete a task with the maximum performance and preserve the user's stress to low levels (Hudlicka, 2003).

There are several different cases of user affect machines are called to deal with. The main question is whether a machine should “maintain a particular state for a particular task, induce a particular state or simply make the user's experience enjoyable?” (Hudlicka, 2003). In general, the level of adaptability is dependant on the context. Hudlicka argues that adaptive user interfaces are slowly making progress in establishing themselves in the industry, yet they are still a minority. In most cases machines still assume normative performance ignoring user affect and personal characteristics. However, this kind of machines tend to become less popular compared to affect-driven ones which are gaining reputation and in some cases are becoming irreplaceable.

It is worth to mention that the higher the responsibility of the user and sensitivity of a situation (e.g. air traffic control), the most severe the consequences in case of error. Sensing human psychological state before the user is about to perform can prevent many unwanted effects, the user could be tired, bored, depressed etc. which means not able to focus properly on his task. Apart from extreme cases though, there is room for attempting to help the user perform even in a not ideal psychological state by changing the state of the machine to encourage him, raise his interest, guide him when he may be unable to concentrate. In each situation there are parameters that need to be taken under consideration in order to achieve maximum performance and avoid errors.

One of the most well-known uses of interactive systems, sensing changes in the human body such as heart rate, blood pressure and electro-dermal activity is the lie detector (Figure 2). Medical diagnosis is another field where affective computing offers great help such as in cases of schizophrenia. Doctors used to identify the patient's psychological state by the expressions and self-description provided by the patients which would often be problematic or misleading (Fragopanagos and Taylor, 2005). An emotion-sensing machine does not rely on individuals' perception of the emotions and can be more accurate sensing signals that humans are not capable of doing so. Cases of autism are also a good example, where individuals struggle to express their

Figure 2 – A lie detector can sense changes in the human body such as heart rate, blood pressure and electro-dermal activity

Figure 3 – Microsoft's Xbox 360 Kinect sensor, a very popular case of HCI based on gesture recognition
emotions but the physiological signals humans create but not sense, can easily be sensed by a machine (Kuriakose et al., 2012). As an example, a well known technology that has been practised in the United States described as “Telemedicine” was named Tele-Home Health Care (Tele-HCC) and it constantly provides doctors with information about the patients' state by collecting data as heart rate, blood pressure and oxygen saturation (Nasoz et al., 2004).

Gaming is also benefited by interactive systems like Microsoft's Xbox 360 Kinect sensor (Figure 3) which senses movement without any other hardware required. This way, the user can control the character in the game or steer the car depending on what kind of game he is playing. Gesture and movement interaction in this case provide a unique gaming experience, making the user feel as part of the game rather than just playing it. Other relevant examples are Sony's PlayStation 3 Move and Nintendo’s Wii system with the difference that these systems require the user holding controllers and the movement is sensed through them.

An emotion-sensitive automatic tutor could sense the user's frustration, excitement and feelings such as boredom or even irritation and adjust the tutoring process accordingly (Fragopanagos and Taylor, 2005). This way the session may be slowed down if the user struggles to follow, it may become faster if it is proven too comfortable for him to understand it or stop if his ability to pay attention may decrease below a certain level. Fragopanagos and Taylor argue that “understanding and developing automatic systems for emotion recognition can assist in generating faces and/or voices endowed with convincingly human-like emotional qualities”.

Surveillance has also huge gains from emotion sensing systems like cases of air traffic surveillance where the users need to be perfectly focused and psychologically capable of accomplishing their tasks. In such cases there is no room for mistakes and an affect-sensing system can help the user concentrate or even prevent him from doing his job if there is a probability that he will fail to do so due to tiredness, boredom or any other emotional reasons. Surveillance can also be applied to tasks such as flying an aeroplane or driving a car.

The possible applications of emotion-sensing machines can be countless. A personal computer could sense the user's tiredness and adjust the brightness of the screen to help him focus. A car could be connected with sensors that detect whether the driver has consumed alcohol above a certain level and prevent him to drive or sense his level of anger and stress and control the maximum speed he can reach and stabilize the wheel beyond the usual level to avoid accidents due to sudden uncontrolled moves. In cases where the driver may be novice, an electronic tutor could be available giving some basic advice for the first few months. Placing sensors for facial recognition in a TV could be used in cases where children may be watching something inappropriate like scenes of violence or things that a child is not interested to, like political shows and the system could change the program to something else attempting to change the child's emotions to positive.

VI. FutuRe Work

Systems that can sense user affect and are able to change their state are already being designed rather than normative performance ones, yet adaptive systems are not welcomed in every case depending on the context of the service they are about to provide. In the future, some patterns should be defined to effectively distinguish the requirements of each situation to avoid exaggerations or bad kinds of service.

Furthermore, cultural differences that can have a huge impact on the ways a user may express himself compared to another user belonging to a different culture, should be defined and grouped in a way that each group will be characterised by common expression techniques. This way, designing for each group will be easier and more culture-specific so that it will properly address user affect. Considering how complicated expression of emotions can be due to differences of individuals, cultural influence can make the situation even more complex if not handled properly. An exception should be made for systems that will be globally used, those systems should adapt some universal ways of people to express themselves, to aim addressing at least the most primitive emotions that tend to be the same for every culture.
Interactive systems are mostly constructed to serve particular needs rather than providing a playful, enjoyable experience to the user. Since the outcome of cases where the user does not treat the interaction as a chore is highly positive compared to old fashioned systems with unfriendly interfaces and much effort needed to get familiar with them, in the future some rules could be defined to effectively categorize cases where a system needs to be strict and formal (possibly advanced professional high responsibility applications) so that the rest of the cases can set as a priority the satisfaction of the user and the positive effect on his general mood. Sentiments of a system is one of the keys to this procedure and a deeper understanding of sentiments could be extremely beneficial for the future of interactive systems.

VII. CONCLUSIONS

Emotions play a crucial role in everyday life because of their direct connection with cognitive processing and they have an impact in every decision from the most important to the most insignificant one. That is why there has been an explosion of research on the field of emotions in HCI in recent years, a field that was created recently and led to an entirely new approach of technology providing fundamental findings and innovations.

A machine, in order to sense the user's emotions can use several tactics depending on the purpose it is designed to serve and also to the kind of input that is necessary. Vocal intonations and facial expressions are the most popular methods because of their ability to predict a huge variety of emotions with satisfying precision. Nowadays, machines that are designed ignoring user affect and emotional state are not probable to last since the levels of efficiency and also user satisfaction are much higher, turning the interaction process as a playful experience. It is very important for designers of such systems to take into consideration the sentiments that those systems will produce, setting as a priority to not disturb or annoy the user in any way rather than putting enhanced features and additions to it.

Finally, when designing systems providing the services needed in a framework that follows some appropriate aesthetic patterns, along with the usage of knowledge about when and where the machine should respond in particular ways can provide a complete, satisfactory experience. The risks and responsibilities of the system depending on the context of the service it provides can be high in several cases, hence the designing process should be done according to all available findings about how a machine should change its state to respond to user affect as also without excluding any new type of information that may be presented as irrelevant or of insignificant value.

REFERENCES


