An Exploration of User Engagement in HCI

Christopher Peters
Department of Engineering and Computing
Coventry University
United Kingdom
c.peters@coventry.ac.uk

Ginevra Castellano
Dept. of Computer Science
School of EECS
Queen Mary University of London, United Kingdom
ginevra@dcs.qmul.ac.uk

Sara de Freitas
Serious Games Institute
Coventry
United Kingdom
s.defreitas@coventry.ac.uk

ABSTRACT

Engagement is a concept of the utmost importance in human computer interaction, not only for informing the design and implementation of interfaces, but also for enabling more sophisticated interfaces capable of adapting to users. While the notion of engagement is actively being studied in a diverse set of domains, the term has been used to refer to a number of related, but different concepts. This paper represents a first attempt at exploring a number of important concepts that the term has been used to refer to, of relevance to both human-human and human-machine interaction modelling.

1. INTRODUCTION

Investigations of the way in which people may engage and interact with machines play a significant role in the design and implementation of intelligent interfaces for a variety of applications, from learning to recreation. Such interfaces should be capable of adapting to the human user and act in an appropriate manner according to the context of the situation [14]. Being able to monitor, interpret and react appropriately to users' interest and engagement, is an important step towards achieving this. The notion of engagement is being considered in a number of diverse research domains, scientific and commercial. However, there remains great variability, overlap and often vagueness with respect to what exactly the term is. For example, a number of related concepts, such as interest, sustained attention, immersion and involvement, are sometimes used interchangeably and relationships are often unclear. This paper represents a first, exploratory attempt at considering some of the broad number of descriptions of engagement and related concepts from across several domains broadly related to HCI. We describe consistent fundamentals behind the concept of engagement as it is presented by researchers working in domains such as robotics, affect recognition, computer games and conversational characters. We do this, with a mind to modelling engagement capabilities for machines, over two Sections: in the first, we describe fundamental factors in engagement (Section 2) and in the second, we attempt to relate different views on engagement to the perception-cognition-action loop (Section 3).

2. FUNDAMENTAL FACTORS

When consulting a dictionary in the English language, the term *engagement* appears to be used in at least two common senses. First of all, it can be used in the sense of *starting*, referring to an initiation of contact. For example, a user may engage in interaction with a machine by moving into a specific range upon which the machine responds. This does not seem to imply any necessary duration to the interaction, but rather a phase (Section 2.1) - the user may engage briefly with the machine and then decide to move away. In a longer term sense, engagement also refers to the concept of *being occupied with*. In this respect, the engagement seems to imply a more sustained involvement. In the literature, engagement is referred to in a number of different ways: as a process; as a stage in a process, or the overall process; as an experience; as a cognitive state of mind; an empathic connection; as a perceived or theorised indicator describing the overall state of an interaction. Nonetheless, in most studies relating to engagement, two underlying fundamentals are apparent: attentional and emotional involvement. Selective attention to a stimulus seems necessary in order for a most basic form of engagement to occur with it; this form of engagement may be limited to a quick glance at a potentially relevance stimulus that proves to be of no further interest. A more sustained form of attention provides a more elaborate requirement for engagement and also allows the possibility of affective involvement [16].

2.1 Phases of engagement

Engagement, when referred to in the sense of a process, can be regarded as being composed of a number of discrete stages or phases through which it may progress. These may relate to the intensity or degree of involvement of participants with respect to the focus of engagement. In the study of engagement with robots, Sidner et al. [17] refer to engagement as a process by which \individuals in an interaction start, maintain and end their perceived connection to one another". It is a natural starting point to consider engagement as consisting of at least these three broad phases.

O'Brien and Toms [11] refer to four phases of engagement: a point of engagement, sustained engagement, disengagement, and re-engagement. The concept of re-engagement raises the important issue of when an engagement can be considered to end. In some circumstances, it may be clear: a person may get up out of a seat and walk away from an interaction with the computer. However, in many other cases, it is much harder to determine. For example, if the user looks away briefly, it may just mean that they have been temporarily distracted. In some cases, looking away may actually signal engagement, such as during shared attention scenarios, when looking at an object under mutual consideration [12]. Social space has also been used in [9], among other indicators, to help categorise engagement according to the stages of *present*, *attending/interested*, *engaged* and *interacting*. This is an important consideration during mobile scenarios, where robots and users are free to move around the environment, highlighting the important role of the context of the interaction.

2.2 Focus of engagement

An important factor is consider is what exactly it is that the user is engaged with. This is often difficult to measure with confidence, particularly for more sophisticated forms of engagement. For example, the fact that a user is in the vicinity of a screen or looking at one does not mean that they are paying attention to it (they may be day-dreaming for example) or that they are paying attention to important aspects. One way to alleviate this situation is to consider only attention towards currently relevant aspects of the scene. For example, in [12] during interaction with a virtual character, three qualities of engagement are defined, relating to the user (1) not looking at the screen at all, (2) looking at irrelevant aspects of the scene, and (3) looking at relevant aspects of the scene with respect to the ongoing interaction. In this situation, attention towards relevant aspects of the scene at appropriate times is regarded as signalling the highest quality of interaction engagement between the user and the character.

2.3 Interest

Another related concept to the study of engagement is that of *interest*. Interest has been described as being functionally important, for motivating interaction and learning, as a mechanism of selective attention \that keeps the creature's attention focused on a particular object, person or situation" [2]. It is particularly noteworthy that many works regard interest as an affective state (for example [10]). This can be explained by the many complex overlaps between attention and affect: attention is deemed to be required for varying degrees of affective processing while, conversely, emotional stimuli capture, maintain and may modulate attention.

3. CONCEPTS

The term *engagement* has been mentioned in many different senses throughout the literature, referring to a number of diverse, although often related concepts. Here, we relate engagement to the action-cognition-perception loop. Although this loop is a great simplification with respect to the real system, it provides a useful initial basis for differentiating between different meanings behind the term.

3.1 Perception

Here, perception refers to the use of the term engagement as it relates to the decoding of basic cues from an other interactor, by a person or by a machine, for example by using computer vision techniques. Of general importance to our sense of engagement with others is our perception of their attention [7], which can be altered by factors such as the effect of distance between interactors on the salience of visual cues and the context of the situation. Important non-verbal cues can be obtained based on head direction and gaze [1], blinking, eyebrow movement, posture and posture shifts [10], smiles [4], feedback and body movement [9], and engagement gestures [17]. These low-level signals can in some cases be interpreted as direct measures relating to engagement and interest e.g. the user is regarded as being engaged/interested when looking at the screen. More sophisticated measures, relying on the integration of multiple sources of information, are categorised here in cognition, Section 3.2.

3.2 Cognition

In relation to the cognition category, engagement here refers to the entities actual state, motivation, goal or tendency towards investing or being concerned with a stimulus. In this sense, interest is often used to describe the motivation or goal towards opening or maintaining engagement. Although an entity may have an interest in something, this is to be differentiated from the action of showing interest in it (Section 3.4). The use of BCI and physiological devices, which may provide direct indicators of attention or affective factors from the body and brain, are relevant here. See for example [15]. Such devices may be particularly effective in conjunction with theories of user state based on inference from their behaviour [13]. Here, engagement details relate to higher-level inferencing based on the multimodal integration of perceived signals for theorising about engagement, that is, making inferences about the complex mental states of the other. For example, interest and concentration may be interpreted at this level based on multiple low level cues [6]. This information helps to keep track of the overall state of engagement from the perspective of the entity, and may help set a context by which ones own behavioural decisions and appraisals of the others behaviour can be judged.

3.3 Experience

In addition to the general perception-cognition-action loop, we found it necessary to add an extra category, called experience, in order to cover many studies of engagement that relate to the subjective experiences felt by individuals who are engaged. For example, during engaging situations when playing computer games, experiences relating to engagement have been reported as feelings of losing oneself in the world of the game, not noticing things happening outside of the screen, or losing all track of the passage of time [5]. These experiences are often related to the concept of immersion, and engagement has been described as the first in three levels of immersion, where the user is required to “invest time, effort and attention in learning how to play the game and get to grips with the controls" [3].

3.4 Action

Here, generation of cues and signals related to displays of engagement are considered. For example, during face-to-face interactions, important information is usually present in the face, so it may be expected that we naturally pay attention to the face if we are engaged with that person, and may also display feedback such as nods to display our interest and/or show empathy by conducting appropriate facial expressions. In this respect, they might signal engagement, for example, by attending to the other and showing interest in what they say. An important distinction here, based on goals and motivations from the cognitive category above, is whether such signals are based on a genuine interest or are superficial displays the have the explicit purpose of communicating to the other that one is engaged.

One may choose to display signals of interest for a variety of superficial reasons, related to the accomplishment of high-

level or abstract goals. Sometimes the display of interest is more important than the actual motivation. For example, while working on the documentary Thin Blue Line, interviewer Errol Morris suggested that in some cases, success in persuading interviewees to disclose facts entailed limiting actual interest and engrossment in the story, so as to focus on showing interest in it [8].

4. CONCLUSION

We have presented a cross-domain consideration of the meanings behind engagement, framing them in terms of the perception-cognition-action loop. We also added considered related concepts which are sometimes used interchangeably with the term and investigated fundamentals that always seem to apply. Relating engagement, similar concepts such as interest, and the use of the term in literature is not an easy task due to the interconnectedness of the terms and the mechanisms they relate to. However, while the categories of perception, cognition, experience and action do not provide a very concise categorisation, we hope that they can provide a start towards elucidating the different circumstances in which the terms may be used.

5. REFERENCES

[1] S. Asteriadis, K. Karpouzis, and S. Kollias. Feature extraction and selection for inferring user engagement in an HCI environment. HCI International, San Diego, CA, 19-24 July 2009, 2009.

[2] C. Breazeal and R. Brooks. Robot emotion: A functional perspective. In J.-M. Fellous and M. A. Arbib, editors, *Who Needs Emotions: The Brain Meets the Robot*, Series in Affective Science, chapter 10. Oxford University Press, 2004.

[3] E. Brown and P. Cairns. A grounded investigation of game immersion. In *CHI '04: CHI '04 extended abstracts on Human factors in computing systems*, pages 1297 - 1300, New York, NY, USA, 2004. ACM.

[4] G. Castellano, A. Pereira, I. Leite, A. Paiva, and P. W. McOwan. Detecting user engagement with a robot companion using task and social interaction-based features. In *International Conference on Multimodal Interfaces*, Cambridge, MA, 2009. ACM.

[5] C. Jennett, A. Cox, P. Cairns, S. Dhoparee, A. Epps, T. Tijs, and A. Walton. Measuring and defining the experience of immersion in games. *Int. J. Hum.-Comput. Stud.*, 66(9):641{661, 2008. [6] R. E. Kaliouby and P. Robinson. Real-time inference of complex mental states from facial expressions and head gestures. In *CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 10*, page 154, Washington, DC, USA, 2004. IEEE Computer Society.

[7] S. Langton, R. Watt, and V. Bruce. Do the eyes have it? cues to the direction of social attention. *Trends in Cognitive Sciences*, 4:50 - 59, 2000.

[8] R. Levine. *The Power of Persuasion*, chapter Whom do we trust? Experts, Honesty and Likability. Wiley, 2006.

[9] M. Michalowski, S. Sabanovic, and R. Simmons. A spatial model of engagement for a social robot. In *Proceedings of the 9th International Workshop on Advanced Motion Control (AMC 2006)*, pages 762 - 767. IEEE, March 2006.

[10] S. Mota and R. Picard. Automated posture analysis for detecting learner's interest level. In *Computer Vision and Pattern Recognition Workshop, (CVPRW '03)*, volume 5, page 49, 2003.

[11] H. O'Brien and E. Toms. What is user engagement? A conceptual framework for defining user engagement with technology. *J. Am. Soc. Inf. Sci. Technol.*, 59(6):938 - 955, 2008.

[12] C. Peters, S. Asteriadis, K. Karpouzis, and E. de Sevin. Towards a real-time gaze-based shared attention for a virtual agent. International Conference on Multimodal Interfaces (ICMI), Workshop on Affective Interaction in Natural Environments (AFFINE), Chania, Crete, October 2008.

[13] C. Peters, S. Asteriadis, and G. Rebolledo-Mendez. Modelling user attention for human-agent interaction. In *WIAMIS 2009 - International Workshop on Image Analysis for Multimedia Interactive Services*, London, 6-8 May 2009.

[14] R. W. Picard. *Affective Computing*. The MIT Press, Cambridge, September 1997.

[15] G. Rebolledo-Mendez, I. Dunwell, E. Martinez-Miron, M. Vargas-Cerdian, S. Freitas, F. Liarokapis, and A. Garcia-Gaona. Assessing neurosky's usability to detect attention levels in an assessment exercise. In *Proceedings of the 13th International Conference on Human-Computer Interaction. Part I*, pages 149 - 158, Berlin, Heidelberg, 2009. Springer-Verlag.

[16] I. Roseman and C. Smith. Appraisal theory. In K. Scherer, A. Schorr, and T. Johnstone, editors, *Appraisal Processes in Emotion: Theory, Methods, Research*, Series in A®ective Science. Oxford University Press, 2001.

[17] C. Sidner, C. Lee, C. Kidd, N. Lesh, and C. Rich. Explorations in engagement for humans and robots. *Artif. Intell.*, 166(1-2):140 - 164, 2005.