Department of Computer Science, University of Exeter, U.K. Technical Report No. 389

Effecting affective communication in virtual environments

Colin G. Johnson[1] and Gareth J. F. Jones[2]

[1] Computing Laboratory,
University of Kent,
Canterbury, Kent, CT2 7NF.
C.G.Johnson@ukc.ac.uk

[2] Department of Computer Science,
University of Exeter,
Exeter, EX4 4PT.
G.J.F.Jones@ex.ac.uk

October 1999

Abstract

Studies of communication between entities in virtual environments have tended to focus on the relevant technical issues and its social impact impact. An important component of human communication is the conveying of affective information via voice, facial expression and gestures and other body language. Virtual environments may be populated by representations of human or virtual agent participants. Communications may be between person-person, agent-agent or person-agent. This paper explores the possible use of the affective communication in virtual environments. The desirability of affective communication is examined and some research ideas for developing affective communication in virtual environments are proposed.

1 Introduction.

The concepts of virtual environments, and of communication within these environments ("virtual communications"), have become widely used in recent years. There are currently no clear definitions associated with these expressions, the only common theme seems be of communication between entities that are physically remote from each other, either because of a spatial or temporal dislocation, or because one or more of the participants is an artificial agent of some form.

Associated with virtual environments there is also a lack of clarity of what is actually expected of communication. Is there an expectation of recreating the rich interaction of physical communication between humans, or is some other alternative scenario, possibly richer or poorer, envisaged?

In this paper we explore the range of virtual environments and the possible participant groups. In this discussion we place an emphasis on communication between virtual entities which make use of other data apart from the factual information being communicated. In particular we would like to explore the capacity of virtual environments to communicate *affective* states, i.e. those aspects of communication concerned with emotional response, and further to explore how agents within virtual environments can model and use affective states to enhance the realism of the virtual environment.

2 Virtual Communication in Virtual Environments.

A core function of virtual environments is to facilitate communication between entities in that environment, be those entities human users or agents. It is often assumed that the ultimate objective of communication within virtual environments is to model communication between humans in the physical world. In order to achieve this objective communication capabilities within the virtual world must not be limited to the simple exchange of information. Everyday human communication involves a level of *affective* communication (communication involving emotional states) that is absent from many virtual environments.

If virtual environments are to be truly representative of the real world they aim to model, they must both (1) facilitate the communication of affect, and (2) agents situated in the environment must react in a way that respects the affective context in which they find themselves. An agent that ignores these aspects of the environment will jar with the realism of the environment as much as one that ignores the laws of physics.

The emergence of *Affective Computing* as an area of practical scientific study is very recent [Pic97], although this builds on experimental and theoretical work established over the last 30 years (see e.g. [SC81, OCC88]). It is argued by Damasio [Dam94] and Picard [Pic97] that affective state is a key component of the human ability to reason :

> *"Evidence like this leads Dr. Damasio to the counter-intuitive position that feelings are typically* indispensable *for rational decisions ; they point us in the proper direction, where dry logic can then be of best use. ...*
>
> *The emotions, then, matter for rationality. In the dance of feeling and thought the emotional faculty guides our moment-to-moment decisions, working hand-in-hand with the rational mind, enabling—or disabling—thought it-*

self. Likewise, the thinking brain plays an executive role in our emotions—except in those moments when emotions surge out of control and the emotional brain runs rampant." [Gol96]

Support for this hypothesis is given by Damasio's studies on humans whose emotional capacity has been limited due to brain injury [Dam94]. The assumptions of this argument have, however, been challenged by Sloman [Slo99] :

> *"Damasio, Picard and others have misinterpreted the evidence about brain damage in Damasio's book as implying that emotions are essential to intelligence. This is a simple non-sequitur.*
>
> *Certain sorts of frontal lobe damage produce two effects : (1) the patients lose the ability to have certain kinds of (secondary) emotional reactions . . . and, (2) the patients become less creative and decisive, and less able to take strategic decisions. . . .*
>
> *It is fallacious to infer from this that (1) is the cause of (2)."*

Resolving this argument is an important question from a cognitive science perspective. However from the perspective of humans and agents interacting within a virtual environment the affective component of the communication is important to the users, whether or not it is a key element of *intelligence* per se.

An extension of these arguments is that communication of cues to affective state is a significant component in human communication. If this is truly the case, rich communication in virtual environments should attempt to model this communication of affective state, either explicitly or more likely by giving and reading cues typical of those observed in human communication. Sloman argues, quite properly, that voluntary or even involuntary emotional responses are generally very broad and we are not able to perceive the underlying condition which led to the affective state. While this is of course true, the context of a communication is likely to provide significant disambiguation of affective states. The context provides a filter through which the many factors that could suggest emotions (e.g. speed of talking, visible physiological changes, "body language" et cetera) are refined to an understanding of the emotional aspects of the communication. There is of course no guarantee of disambiguating such cues correctly, different people behave differently and people may incorrectly "see" cues they wish to see, these mistakes can lead to confusion, but in general our unconscious use of affective cues aids communication.

3 Virtual Environments.

In order to explore the theme of affect in detail it is first important to give clear definitions to some important concepts. A *virtual environment* is any place whereby (people and/or agents) can come together to communicate. Under this definition virtual environments cover a wide range of possibilities. At one end of the spectrum there are video teleconferences between human participants, and at the other virtual environments could be wholly text based, as exemplified by MUDs and Internet Relay Chat. The more popular definition is probably to consider environments which involve techniques drawn from virtual reality, such as a teleconferencing system with humans represented by avatars. An example which falls between the extreme of a pure text based system and a fully immersive virtual environment is the Microsoft *Comic Chat* system [KSS96], which uses static graphics paired with text, which generates dynamic comic strips as participants enter the discussion via text input.

A key taxonomy of communication concepts in virtual environments is to classify the participants in the communications, and to study the capacity of the virtual environment to facilitate communication between those participants. We explore such issues in the remainder of this section.

3.1 Person–Person Communication.

Full person–person communication, as exemplified by video conferencing, enables all natural communication cues to be conveyed between the participants. Thus not only is the basic information of the spoken words conveyed but also a rich stream of visual and aural information which is perceived consciously or unconsciously. Some of this additional information will convey details of the affective states of the individuals which can assist in contextual understanding of the spoken data and further understanding beyond this.

A naive view is that less natural virtual environments should attempt to give the users of an environment the same affective capabilities that they would have in real-world interactions. This view however misses out a number of important features of virtual environments. The *limited* affective capability of certain virtual environments may be a positive feature, rather than a hindrance, to some users. For example, participants in text based chat rooms enjoy their anonymity and their ability to control the interaction precisely by careful use of text [Tur96, Fon93, Fon97] :

> *"The Internet is another element of the computer culture that has contributed to thinking about identity as multiplicity. On it, people are able to build a self by cycling through many selves."* [Tur96]

The nature of such interactions would change dramatically if these environments conveyed information relating to the affective states of participants. The idea of people experimenting with alternative personae would be much disrupted.

This situation may be rather different for participants in a synthetic graphical world where they might be happy to have affective information conveyed by alterations in speaking style, posture or facial expression. However it is not clear what the role of such systems is. For a formal discussion a video conference would perhaps be better, especially as increased bandwidth is increasing the power of video conferencing systems, arguably making simpler representations redundant, and for the chat scenario anonymity is important.

There are two interestingly contrasted ways of communicating affective states within a virtual environment. The first is for the user's computer to interpret information about the user's affective state and to convey that information as some kind of caricature of human expression on the user's avatar. Such a method would involve complex pattern matching techniques, which could draw inspiration from work on face recognition (e.g. [JN95]) and on the work of Ekman and others [Ekm92a, Ekm92b, Ekm93] on classifying basic human emotions via facial expression. There are many other techniques that could be used to extract this same information, for example the use of affective wearables (chapter 8 of [Pic97], [Sta96, SMR$^+$97]). There are disadvantages to this approach, however, as the consequences of the wrong emotion being communicated are potentially dangerous—this touches on the well known issue of *trust* in agent systems [Mae94]. There is also the problems that avatar's expressive capability is likely to be restricted to a small range of physical responses, thus reducing the affective bandwidth of the communication channel.

The alternative to this is to directly communicate this measured data (what we could call "affective meta-factors") about the state of the user, without making any attempt to interpret these features in terms of a discrete set of emotions or emotional dimensions. The data communicated could be that that would be available during face-to-face communication, such as skin temperature (indicated by changes in skin colour and texture) or data which are unique to the fact of being in a virtual environment (such as an indication of how hard the user is typing on a keyboard or manipulating a joystick). In this situation there is no need for one communicated factor to correspond to one emotional factor, the user can learn which combinations of factors contribute to which emotional states as they become accustomed to working within the virtual environment. The main disadvantage of this scenario is that the user will have to spend a long time situated in the virtual environment in order to learn how to interpret these cues. In particular there would be a need for standards across virtual environments so that such learning in one environment could be transferred effortlessly to another. Choosing appropriate meta-factors is in itself a challenging task.

3.2 Agent–Agent Communication.

A more open question is the use of affective states in agent–agent communication. These communications may involve just a simple transaction involving information transfer, or it may be much more complex involving brokering of access to resources and bargaining over information and access to other agents or a user. Decisions in this environment could use standard AI methods, expert systems of some sort, searching thought possibilities to find an optimum, et cetera. It is suggested by Damasio's work that in order for such communication to follow the path of a person-person transaction each agent should have a model of affective state and that there should be some communication of affective information between the agents. Sloman disagrees [Slo99] :

> *"Damasio assumed that only emotions can perform this kind of high-level control. However recent work in AI has produced alternative meta-level control mechanisms, including planners which use "anytime algorithms" [BD89], which allows an interrupted planning process to produce some initial partial plans in intelligent machines."*

More human-like reasoning in agent–agent communication has consequences for the kind of algorithms used in agent architectures and multi-agent communication [BD89] and in section 3.3 of this paper.

3.3 Person–Agent Communication.

A final form of virtual environment is that inhabited by a mixture of people and artificial agents [Fon93, Fon97]. In this environment the ability of the agent to respond may be enhanced if the agent is able to hypothesise about the emotional state of the person. In general the agent will not be following its own agenda, like the participants in a chat room, but will be attempting to better serve its human masters.

There is a further question here of how human emotional states should be represented to agents. Should the agent "see" and "hear" emotions from an avatar? Or should the agent have direct access to signals used to produce particular affect driven responses in the avatar, creating its own concepts of affective states by a process akin to *data fusion* [RM88]?

There are advantages to both. In the first scenario the user is able to tailor their avatar to their own personal affective repertoire, making it easier for the agent to understand the affect being communicated, and release only that information which they are willing to give agents access to. In the second scenario the agent would develop a deeper understanding of the connection between the "affective meta-factors" and the desired affective expression. This issue is particularly relevant in the scenario where an agent will only meet a small range of emotional situations — it could learn to interpret only that range of affective states that it encounters, and be more accurate than if it had to deal with large amounts of irrelevant over-interpreted information. The choice of meta-factors is also governed by the acceptability of the communication of these factors to users. It has been noted (Mark Nixon, personal communication) that in biometric person-recognition users are generally much happier with machines attempting to analyse factors that could be recognised by other humans (e.g. face shape, gait) than more hidden factors (e.g. heart rate, skin moisture level, or DNA structure).

In order to maintain the naturalness of the interaction the agent should perhaps respond in an affect driven manner, perhaps indicating frustration or pleasure. A simple, rough representation of emotion in the agent may acknowledge the fact that human decisions may sometimes not follow an entirely rational path, while not attempting to actually model the emotions that give rise to this irrationality.

Also the agent should be designed so that its own decisions are controlled not just by internal criteria of optimality, but by the virtual social context in which it is situated. This has parallels in Damasio's work with brain-damaged patients [Pic97] :

> *"This disorder is exemplified by "Elliot", whose IQ and cognitive abilities are all normal or above average, but who suffered damage to frontal lobe brain tissue as the result of a brain tumor. When confronted with a simple decision as to when to schedule an appointment, Elliot will disappear into an endless rational search of "Well, this time might be good,", or "Maybe I will have to be on that side of town so this time would be better," and on and on. Although a certain amount of indecisiveness is normal, in Elliot it is apparently not accompanied by the usual feelings, such as embarrassment, if someone is staring at you for so long to make up your mind. Instead, Elliot's tendency is to search an astronomically large space of rational possibilities. Moreover, Elliot seems to be unable to learn the links between dangerous choices and bad feelings, so he repeats bad decisions instead of learning otherwise. Elliot's lack of emotions severely handicaps his ability to function rationally and intelligently."*

This has implications for the kind of algorithms used in agent systems. Such algorithms need to be *interruptible*, i.e. they need to give a reasonable approximation to the solution even if they have to stop half way through. Some algorithms, for example those which use a highly distributed, parallel processing model, may not be able to bring the relevant data together to give a response in a sufficiently quick time to cope with the social context within their virtual environment.

A final thought. Does an entity in a virtual environment need to be aware if it is communicating with an agent or person. This of course provokes standard AI arguments about Turing tests etc, but if we have a graphical environment where some communication is by gesture and affect synthesis, say in the face, poor natural language and inability to answer questions properly could be overcome by encouraging the human participant to give more information.

4 Coda.

Affective states are one of the most important factors in communication, and for a virtual environment to be truly realistic it cannot ignore such factors. However simply to replicate existing affective communication channels in a virtual environment is both naive and misses a fascinating opportunity to enhance the value of virtual communication. One of the most interesting features of virtual environments is that it is possible to create a world that operates according to entirely different laws of physics to the natural world. Analogously, to explore alternative realms of affective communication in such environments is an interesting area for the future development of virtual environments.

Note.

This is an extended version of a paper from the Proceedings of the Second Workshop on Intelligent Virtual Agents (Ed. Daniel Ballin), University of Salford, September 1999.

References

[BD89] M. Boddy and T. Dean. Solving time dependent planning problems. In *Proceedings of the 11th International Joint Conference on Artificial Intelligence, Detroit*, 1989.

[Dam94] A.R. Damasio. *Descartes' Error : Emotion, Reason and the Human Brain.* Gosset/Putnam Press, 1994.

[Ekm92a] P. Ekman. Are there basic emotions? *Psychological Review*, 99(3):550–553, 1992.

[Ekm92b] P. Ekman. An argument for basic emotions. *Cognition and Emotion*, 6(3/4):169–200, 1992.

[Ekm93] P. Ekman. Facial expression and emotion. *American Psychologist*, 48(4):384–392, April 1993.

[Fon93] Leonard N. Foner. What is an agent, anyway? a sociological case study. Agents Memo 93-01, MIT Media Lab, 1993.

[Fon97] Leonard N. Foner. Entertaining agents : A sociological case study. In *Autonomous Agents'97 Proceedings*. ACM, 1997.

[Gol96] Daniel Goleman. *Emotional Intelligence*. Bloomsbury, 1996.

[JN95] X. Jia and M. Nixon. Extending the feature vector for automatic face recognition. *IEEE Transactions on Pattern Matching and Machine Intelligence*, 17(12):1167–1176, 1995.

[KSS96] D. Kurlander, David H. Salesin, and T. Skelly. Comic chat. In *Proceedings of SIGGRAPH96*, pages 225–236, 1996.

[Mae94] Pattie Maes. Agents that reduce work and information overload. *Communications of the Association of Computing Machinery*, 37(7), 1994.

[OCC88] A. Ortony, G.L. Clore, and A. Collins. *The Cognitive Structure of Emotions*. Cambridge University Press, 1988.

[Pic97] Rosalind Picard. *Affective Computing*. MIT Press, 1997.

[RM88] J.M. Richardson and K.A. Marsh. Fusion of multisensor data. *International Journal of Robotics Research*, 7(6):78–96, 1988.

[SC81] Aaron Sloman and Monica Croucher. Why robots will have emotions. In *Proceedings of the 1981 International Joint Conference on Artificial Intelligence, Vancouver*, 1981.

[Slo99] Aaron Sloman. Review of "Affective Computing". *AI Magazine*, 1999. In the press.

[SMR+97] T. Starner, S. Mann, B. Rhodes, J. Levine, J. Healey, R.W. Picard, and A. Pentland. Augmented reality through wearable computing. *Presence*, 1997.

[Sta96] T.E. Starner. Human-powered wearable computing. *IBM Systems Journal*, 35(3/4):618–629, 1996.

[Tur96] Sherry Turkle. *Life on the Screen : Identity in the Age of the Internet*. Weidenfeld and Nicholson, 1996.