A constructionist cognitive model of virtual presence

David Nunez

CS04-07-00

18 August 2004

Collaborative Visual Computing Laboratory
Department of Computer Science
University of Cape Town
Private Bag, RONDEBOSCH
7701, South Africa
Email: dave@cs.uct.ac.za

Abstract

This paper proposes a cognitive model of presence which emphasizes the interaction of top-down and bottom-up processes, and the subsequent creation of temporary representations of an environment in working memory. These temporary representations (which we term constructions) are formations of meaning about the environment, which allow a subject to make inferences and interact in that environment. We argue that this constructionist position follows from the existence of hypothesis selection models of presence such as proposed by Slater & Steed. Finally, we contrast our model to that of Wirth et al and discuss some of the implications of this model for presence in unfamiliar or unrealistic environments

.

1. Cognitive models of presence

With the recent move towards a comprehensive theory of presence, cognitive science has surfaced as an important basis for understanding presence. For example [1], in discussing breaks in presence, propose a filtering, information-processing model in which one of two information streams (real environment/virtual environment) is selected for processing. Many other examples of the application of cognitive theory in presence exist - for instance [2] and [3]. Although it is generally accepted that presence is a complex, multi-factor construct, we believe that an understanding of the cognition of virtual environments (and indeed real environments) can lead to invaluable insights into fundamental aspects of presence. Therefore, although this paper presents an exclusively cognitive model of presence, it should be understood as an attempt to understand the role of cognition in presence, rather than an argument that a complete understanding of presence can come from examining cognition alone.

2. Constructionism in cognition

Constructionism involves the interaction of top-down and bottom-up processes which create a mental construct reflecting some external or imagined situation in such a way that inferences and decisions can be made in relation to that situation [4] [5]. Furthermore, constructionism is generally associated with a dynamic state of cognition in which bottom-up processing activates particular top-down processes which in turn lead to a bias in bottom-up processing [6]. The 'rabbit/duck' switching illusion discussed in [1] is an example of two constructions arising from the same bottom-up data, but mediated by a different top-down bias. Constructions are capable of adapting to changes in the external world, but only to a degree; if the products of bottom-up processes lead to a significant mismatch with the construction, a reconstruction will occur so as to better represent the stimulus situation [7]. That constructionism is useful in presence research has already been argued by [8], and indeed, constructionism already exists in the literature in various forms. [9] for instance, argues for the interaction of internal and external factors as a determinant of presence. More recently, [10] describes a concept (perception and action in 3d space) in which a continual re-consideration of the environment in terms of the user's cognition, perception and emotions occurs. Also, [11] proposes a model in which presence is an active process in which sensory and conceptual data are used to construct mental models of the environment (see section 7 below for a discussion of this model).

3. The need for constructionism: The virtual stimuli problem

In our view, constructionism is a useful tool in modeling presence because it solves an important problem in current thinking: namely, the problem of virtual stimuli. Often in the literature a distinction is made between stimuli which originate in the virtual environment, and stimuli which originate from other sources (for example, [1] and [12]). We find this distinction useful because it allows one to ask questions in terms of relevant information and distracters. However, this distinction is also somewhat problematic. From a physiological point of view, there is only one stream of input for a user of a virtual environment – all stimuli are real. All sound, regardless of its origin, is received by the same receptors in the ear. Similarly, all visual stimulation, regardless of its point of origin, is received by the same receptors in the retina. Once these receptors are stimulated, the original stimulus ends its role in perception; the receptors generate neural signals from which cognition proceeds. Therefore, once the stimulated receptors have fired their impulses, all information about the origin of the percept must be inferred via a series of cognitive processes - its position in space, its semantic relationship to other percepts, what object that stimulus represents, and where that stimulus originated.

4. Cognitive constructions

The 'two streams' of information between which a user selects [1] therefore exist not as external streams of information one of which is selected, but rather as two alternative interpretations or constructions of a complex set of perceptions. In one of these constructions (the 'virtual construction') the user interprets their mental state (the perceptions and the inferences drawn from these) as a coherent virtual environment. In the other construction (the 'real construction') the user will interpret their mental state as being in the experimental venue, viewing a display. In each case, information which matches the user's basic idea ('this is the virtual environment') will be added to the construction, while information which does not match the construction will be attended out, or act as a distraction [11]. The construction thus becomes not only a basis of inference for perceptions, but also works to filter out irrelevant stimuli and bias perceptions to create a coherent model of the environment. In the constructionist view, there need not only be two possible constructions – depending on the mental state of the subject, any number of possible constructions are possible, many of which could be considered as evidence of presence. For example, placing the subject inside a cave system displaying the inside of a hospital could lead to, depending on the emotional or semantic aspects of the active construction, the experience of a frightening hospital, a comforting hospital, a friendly VR laboratory or a stressful VR laboratory. It should be noted that the choice of which stimuli are used to form the construction depends not on the actual origin of the stimulus (the real or virtual environment), but rather on the current state of the construction, and the biases on attention and interpretation which it exerts. Thus, it is possible for stimuli arising outside the virtual environment (such as a phone ringing in the office next door) to be incorporated into the construction of the virtual environment.

5. A constructionist cognitive model of presence

We propose the *constructionist cognitive model*, which makes use of a basic information processing architecture (such as used by [13]). In this architecture (depicted in Figure 1), stimuli are selected for further processing by an attention filter. The selected inputs are then processed and transformed in working memory. These transformations are performed using rules and data contained in long-term memory. Thus, the contents of working memory are constructed by means of bottom-up input (stimuli from the attention filter), as well as from top-down input (contained in knowledge structures and transformation rules) from long-term memory. The products of this processing in working memory represent temporary structures such as mental models, themes, cognitive maps and other constructs which allow the subject to make inferences about the world and thus interact with and navigate through it.

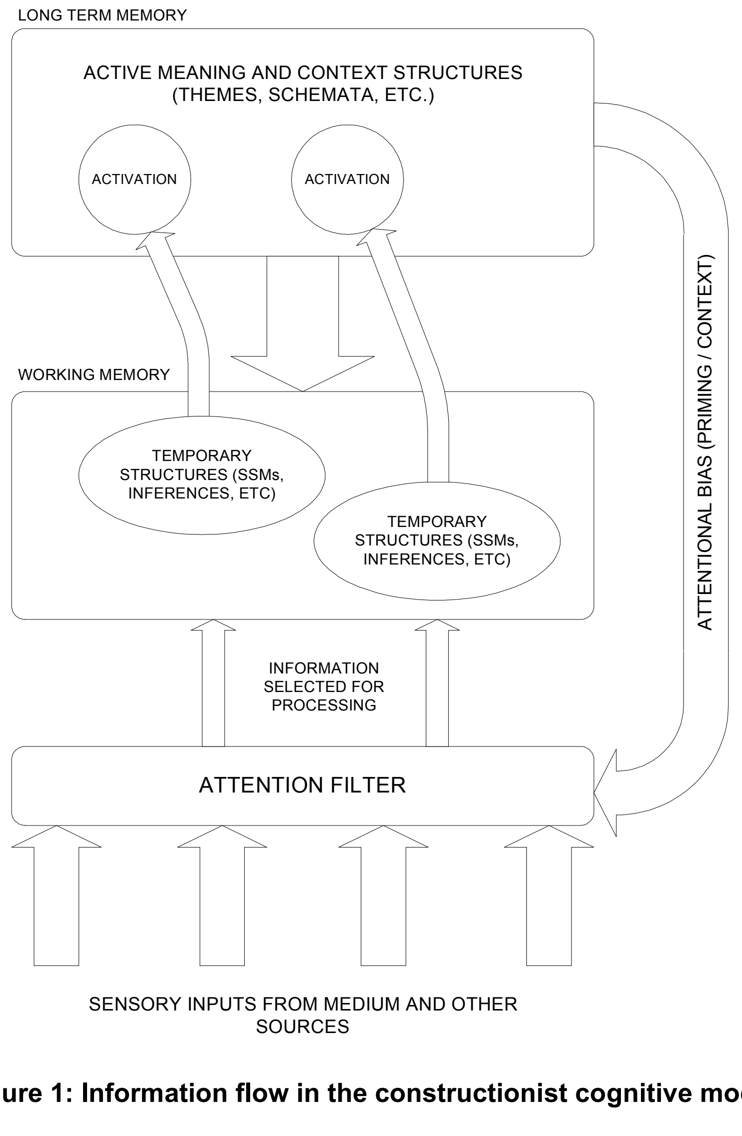

Figure 1: Information flow in the constructionist cognitive model

The model represents a continual upward and downward flow of information and, by consequence, the continual construction and reconstruction of temporary structures in working memory. Information taken in by the senses is first partly processed by the attention filter, which itself contains a bias towards permitting or blocking certain types of information on the basis of relevance [14]. This information is then organized in working memory into temporary constructions such as semantic chunks or mental models [13]. The rules by which this organization occurs are encoded as activated knowledge structures (e.g. schemata or scripts) in long term memory [15]. These structures not only provide rules by which to organize information, but also fill in missing data with that learned in previous experiences [15]. As this occurs, the temporary structures in working memory are adapted to create a state which accommodates the perceptual information provided by the attention filter as well as the active structures in long-term memory. At the same time, the active structures in long term memory are affecting the attention filter by creating a bias towards filtering out unrelated information.

5.1 Coherence, construction and re-construction

Constructions are formed partly on the basis of previous experience and learning, and as such represent attempts to fit a current stimulus set into a meaningful whole (encoded by existing knowledge structures in long-term memory). In essence, forming a construction is a process of inferring meaning. Therefore, it follows that not every stimulus will be compatible with the construction that is active (for instance, a helicopter would not ordinarily be compatible with a construction of prehistoric Earth). Stimuli which are not able to fit into the construction are selected out by attention, which is biased (by the construction) to only allow in relevant stimuli - in this way, constructions are kept coherent. However, it is also possible for a stimulus, related or not, to demand attention to itself and thus force its way into consciousness [7]. This usually occurs in the case of stimuli which represent large, sudden differences between the outside world and the perceiver's expectations. Thus a sudden noise or movement, an expected tug of a cable or similar stimulus will force its way into consciousness where, if a large disparity between the perception and active construction exists, will force a re-evaluation of the situation. When this occurs, a reconstruction takes place which incorporates the new stimuli. However, when this reconstruction occurs, the re-assignment of cognitive resources is felt – first as an odd sense of mismatch or confusion (as the construction fails to match the situation), and then as a subsiding of that sensation and with a sense of realization of the new situation as the reconstruction occurs [7].

5.2 Meaning activation & feedback

In this model, the meaning which the subject constructs about the environment is crucial for their behaviour, and thus affects all levels of the model. Although meaning originates in the top-most level (knowledge structures in long term memory), it can flow downwards to working memory, as well as to the attention filter in the form of a bias for selection. The subject's current construction also affects the temporary objects being formed in working memory. Thus, knowledge structures contribute to the construction of temporary structures in a 'broad brush' approach (due to the activation of categories, concepts, etc. [15]), although the specific details of the construction are determined by the stimuli. In this way, the meaning inferred by the subject will allow for interpretation and prediction of the state of the environment, while the perceptual stimuli allowed in by the attention filter will ensure successful behaviour which matches the true state of the environment.

5.4 Presence in the model

In this model, presence exists as the degree to which the construction is about the content of the virtual environment rather than about the display system or the experimental situation. This definition is in line with the concept of cognitive presence proposed by [3] as well as with the concept of 'pretence' proposed by Slater, in which presence is taken as when the subject is acting and thinking 'as if' in the virtual environment [17]. In order for this to occur, the construction will need to have a high degree of match to the stimuli (as would occur in a high-fidelity, highly immersive display); if not, the subject will experience a re-construction, which will disturb cognition about the virtual environment and thus reduce presence. Although the subject's own biases and previous experience will contribute to their experience (which one might call the 'willing suspension of disbelief') large disturbances in the stimuli will demand attention and are more likely to cause reconstructions (as occurs in a break in presence).

6. Construction flexibility and unrealistic environments

Although constructions tend to limit the information in consciousness to create a sense of coherence in the world, they do allow a great deal of flexibility in interpreting stimuli. A construction operates simply by means of creating expectations, and attempting to interpret meaning in terms of those expectations. In short, constructions operate on the level of concepts rather than specific objects. Thus, if we have constructed a restaurant from the available perceptions, then we will have an expectation for an interaction with a waiter at some point. We will also have general expectations about that waiter (he will take a food order, will display a particular degree of politeness and so on), but not specific expectations such as whether the waiter will be a man or woman, young or old, dressed in a particular way or not [15]. What exactly those expectations will be will depend on cultural conditioning and personal experience [4], but they will always be general enough to allow flexibility enough to allow successful interaction in the world.

Because constructions rely so heavily on the expectations and data provided by top-down processes, an interesting question arises about constructions of unrealistic or highly unfamiliar environments. In such a situation, our model predicts that the closest fitting available knowledge will be activated, but due to the poor fit with bottom-up data, constant reconstructions will occur until enough information has been gathered to begin forming new top-level structures about the environment. Anti-presence [16] could be taken as an example of this process. Describing anti-presence [16] notes that people witnessing traumatic events often feel as if they were experiencing a film or other mediated environment rather than reality. We would argue that the closest fitting top-down knowledge people tend to have about traumatic events are all from films and other media, and thus it is these structures which become active during the constructions. Gradually however, with increased experience, top-level structures to deal with the new type of situation will form, which will allow more useful constructions to be formed.

7. Cognitive constructions vs. spatial situation models

Wirth *at al* [11] have proposed a model which emphasizes the subject's construction of a mental model of the spatial situation. In this model, a spatial situation model (SSM), which encodes information about a physical space and the observer's relation to it, is constructed from information presented by the encoding medium as well as from internal factors which include previous knowledge and context. In this model, the SSMs are formed and then tested against further input as the subject interacts with the world. The cognitive construction model we propose, although similar in its basic constructionist philosophy, is different from Wirth *et al*'s model in several important respects. Firstly, the model we propose differs in that we see constructions operating not only at the level of spatial or mental models of the environment, but rather at the higher level of meaning about the environment. Furthermore, we propose that the construction of meaning affects lower level cognitive processes (such as attention), which in turn affects the activation of higher level processes, thus making cognitive construction a highly interactive, continual process. Finally, we propose explicitly modeling the effect of the construction on attention, in line with findings from priming studies in attention (for example, [14]). Thus while this model is completely compatible with Wirth *et al*'s (in that spatial situational models can be seen as intermediate level constructed structures), we see construction as a far more fundamental process which involves all levels of cognition in an interactive way.

8. Conclusion

The constructionist cognitive model of presence is, we believe, an extremely useful tool not only for investigating the role of cognition in presence, but also for asking such questions as the importance of previous experience and thematic coherence as factors in presence. Although the model is only in its conceptual stages, we believe it is a useful framework for the theoretical development of the role of cognition in presence.

Acknowledgements

We would like to acknowledge the support of the CAVES (Collaborative African Virtual Environments System) project in providing the means for conducting this research.

References

[1] M. Slater, & A. Steed. 2000. A virtual presence counter. *Presence: Teleoperators and Virtual Environments 9,* 413-434.

[2] K. E. Bystrom, W. Barfield & C. Hendrix. 1999. A conceptual model of the sense of presence in virtual environments. *Presence: Teleoperators and Virtual Environments 8,* 241-244.

[3] D. Nunez & E.H. Blake. Cognitive presence as a unified concept of virtual reality effectiveness. In *Proceedings ACM AFRIGRAPH 2001*. 115-118. November 2001.

[4] J. Bruner. Acts of meaning. Harvard University Press. 1990

[5] L.B. Resnick. Shared Cognition: Thinking as social practice. In L.B. Resnick, J.M. Levine & S.D. Teasley (eds.), *Perspectives on Socially Shared Cognition*, American Psychological Association, 1991, 1-20.

[6] D.E. Rumelhart, G. Hinton, & J.L. McClelland. A general framework for parallel distributed processing. In D. Rumelhart & J.L. McClelland (Eds), *Parallel Distributed Processing, Vol. 1: Explorations in the microstructure of cognition*. MIT Press 1986, 45-76

[7] C. Martindale. Cognitive Psychology: A neural network approach. Brooks/Cole. 1990

[8] C. Fencott. Presence and the content of virtual environments. *Paper presented at the 2nd International Workshop on Presence*, April 1999.

[9] M. Slater & S. Wilbur. 1997. A Framework for Immersive Virtual Environments (FIVE): Speculations on the role of Presence in Virtual Environments, *Presence: Teleoperators and Virtual Environments, 6,* 603-616.

[10] W.A. IJsselsteijn. Elements of a Multi-level Theory of Presence: Phenomenology, mental processing and neural correlates. *In Proceedings of PRESENCE 2002*, 9-11. October 2002.
[11] W. Wirth, P. Vorderer, T. Hartmann, C. Klimmt, H. Schramm, & S. Böcking. Constructing Presence: Towards a two-level model of the formation of Spatial Presence experiences. *Presentation at the 6th Annual International Workshop on Presence*, October 2003.
[12] C. Heeter. 2003. Reflections on real presence by a virtual person. *Presence: Teleoperators and Virtual Environments, 12,* 335-345.
[13] A.D. Baddeley. Working Memory. Clarendon Press. 1986.
[14] J. Deutsch & D. Deutsch. 1963. Attention: Some theoretical considerations. *Psychological Review*, 70, 80-90.
[15] D. E. Rumelhart & A. Ortony. The representation of knowledge in memory. In R.C. Anderson, R.J. Spiro & W.E. Montage (Eds.), *Schooling and the acquisition of knowledge*. Lawrence Erlbaum Associates, 1977.
[16] L. R. Timmins & M. Lombard. When real seems mediated: Anti-presence. *Presentation at the 6th Annual International Workshop on Presence,* October 2003.
[17] M. Slater. Presence or pretence? *Presentation at the 6th Annual International Workshop on Presence*, October 2003.