Representation and Purposeful Autonomous Agents

Article (Unspecified)

Representation and Purposeful Autonomous Agents

Sharon Wood

Centre for Research in Cognitive Science
University of Sussex,
Brighton, BN1 9QH, UK
sharonw@cogs.susx.ac.uk

Abstract

Although many researchers feel that an autonomous system, capable of behaving appropriately in an uncertain environment, must have an internal representation (world model) of entities, events and situations it perceives in the world, research into active vision, inattentional amnesia (Rensink, 2000b; Wolfe, 1999) and change blindness (Rensink, 2000a; Hayhoe, 2003; Tatler, Gilchrist and Rusted, 2003) has implications for our views on the content of represented knowledge and raises issues concerning coupling knowledge held in the longer term with dynamically perceived sense data. This has implications for the type of formalisms we employ and implications for ontology. Importantly, in the case of the latter, evidence for the 'micro-structure' of natural vision (Hayhoe, 2003) indicates that ontological description should perhaps be (task-related) feature oriented, rather than object-oriented. These issues are discussed in the context of existing work in developing autonomous agents for a simulated driving world (Wood, 1993; 1995a; 1995b; 1998; 2003). The view is presented that the reliability of represented knowledge guides information seeking and perhaps explains why some things get ignored.

Introduction

Traditionally, knowledge representation has been viewed as a prerequisite to informed action. Representations have often been assumed to comprise complete descriptions of the problems solver's environment. Correspondingly, the means by which such representations are constructed and obtain their content are assumed to be comprehensive. For example, early approaches to vision proposed the construction of complete, viewer independent, scene descriptions (cf. Marr, 1982).

The development of autonomous agents has changed that view to one which varies from the extreme of denying the existence of representation underlying activity (Brooks, 1986) to one in which 'representations' of a kind are dynamically generated on a just-in-time (JIT) basis to support interaction with an environment as needed (Ballard, Hayhoe and Pelz, 1995). In this view, the purpose of vision is to actively seek out information pertinent to the agent's current task, rather than passively absorb information to form a complete 'picture' of the world to be held in memory and interrogated at will in determining appropriate courses of action.

Various psychological evidence appears to support this 'active vision' view. The phenomenon of *inattentional blindness* (Rensink, 2000b; Mack and Rock, 1998; Hayhoe 2003) or *inattentional amnesia* (Wolfe, 1999, Rensink, 2000b) demonstrates the selective nature of vision. Even though entities are clearly within view, if they are not central to the task in hand, they frequently remain unseen (Rensink, 2000b). By visually pursuing the selection of information about those entities central to the current task, other entities in the visual scene are actively ignored, no matter how conspicuous they may seem to the non-task-oriented viewer (Simons and Chabris, 1999).

Guiding this process is the rapid orientation of the viewer to the nature or 'gist' of the situation in which they find themselves and rapid feature selection for task-relevant entities. Gist and spatial layout can be rapidly extracted from visual scenes (Intraub, 1980; 1981; Biederman, 1981; Tatler, Gilchrist and Rusted, 2003), encoded and retained (Friedman, 1979; Pezdek, Whetstone, Reynold, Askari and Dougherty, 1989), with retention of object identity, recognition, absolute spatial layout, shape, colour, and relative distance occurring gradually over an interval of between 1-4 seconds.

A related phenomenon of *change* blindness further demonstrates aspects of natural vision which result in failure to notice changes to entities in the visual scene when these take place during a saccadic eye movement (Grimes, 1996; Rensink, 2000a). It appears that change can only be detected when the changing object is fixated (Rensink, 2000a; Rensink, O'Regan and Clark, 2000).

This phenomenon has also been demonstrated during activities in which the changed feature is central to the task in hand (Hayhoe, 2003). Participants asked to pick up blocks, which might be either pink or blue, and to place these in a particular location according to colour, failed to notice when the selected object changed colour between initial selection and final placement. Most often the object was placed in the location appropriate for its colour during initial selection, rather than for the colour to which it had changed. Hayhoe (2003) argues that this demonstrates the 'micro-structure' of vision: that fixation of an object is not sufficient for apprehension of *all* the visual information associated with it. It would appear that during initial selection of the object, participants pay attention to colour, whilst during subsequent fixations they appear to be concerned with location in guiding the object to its resting place (Ballard, Hayhoe and Pelz (1995).

As a consequence of evidence of this kind, 'active vision' proposes a task-related basis for the apprehension of visual information. Several models developed using this approach demonstrate that with sufficient sensory input during performance of a task, recourse to internal representation can be avoided (Brooks, 1986; Bajcsy, 1985; Ballard, 1991). Furthermore, there is evidence that in natural vision, performance reflects the apprehension of visual information just prior to its use (Ballard, Hayhoe and Pelz, 1995).

In this sense, just in time representation, unlike comprehensive and systematic approaches to representation, appears to refer to currently available sense data, pre-processed to some extent into primitive features, enabling the visual apprehension of task relevant information.

As a consequence of acquiring task relevant information, the information itself (such as the colour of a block) may be retained, but the visual context for that information not – that is, just in time representation is transitory and merely sufficient for the selection of sought information. There is no enduring representation of the entire scene. It persists only as long as the scene is viewed and is not open to manipulation or reorganization.

Karn, Moller and Hayhoe (1997), on the other hand, describe the ability to search for or reach toward an object no longer visible which is crucial to many perceptual and motor tasks, and point to the representation of multiple mutually supportive frames of reference for object location. The representation of a viewer independent frame of reference for spatial layout must be built up over time to support planned activity (Hayhoe, Shrivastava, Mruczek and Pelz, 2003).

This implies that the rapidly acquired gist and spatial layout of a scene (Tatler, Gilchrist and Rusted, 2003) is used to support subsequent visual interrogation of a situation. Rather than constructing a description of all that is in a scene, this finding implies that information retained on visual layout, enables indexing to the scene to further acquire information as needed. It seems likely that indexing to a situation will be both viewer-centred and perspective independent (Hayhoe, Shrivastava, Mruczek and Pelz, 2003), but will not describe the visual information available to the viewer, rather it will support the acquisition of further information as and when required.

Rensink (2000c) proposes a theory of attention that fits in well with this view. He presents an account of the phenomenon that we experience a rich and detailed account of our visual world, based on being able to actively index visually to the world around us, even though, at any given time, we have access to just that limited set of visual data within our focus of attention. The phenomenon arises from the moment by moment construction of JIT scene representations that give us detailed information about our visual world whenever we want it. The feeling of having pre-stored in some way, the detailed information of all we have ever seen, arises from our ability to relocate our focus of attention to any part of the visual scene, as and when we wish, to 're-discover' all that visual detail in its entirety.

The Role of Representation

In the context of this view, what then are the implications for the view that an autonomous system must have an internal representation of the events and situations it perceives in the world?

In a dynamic environment which changes outside the individual actions of an agent, that agent is incapable of omniscience with regard to the 'state' of the world; an agent requires access to sense data to support interactions with its environment (Wood, 1993). Appropriate indexing to just-in-time representations would appear to support this need as such; however, many appropriate interactions, or responses, to the environment require access to information not available *at the time of response* through sense data alone (Wood, 1995a). (This is a different problem from needing to access information momentarily out of view.)

It seems unlikely, therefore, that JIT representation alone, can provide sufficient basis for all purposeful activity. In particular, some interactions appear to require anticipation on the part of the viewer, (Wood, 1993; Wood, 1995b). For example, some kinds of activity would seem to depend upon the ability to model events involving complex interactions between entities and invoking the application of typical scenario-based knowledge (Wood 1993; 1995b).

It is unclear to what extent actively indexing to just in time scene representations can satisfy certain kinds of anticipatory behaviour. Although there is some evidence that very simple anticipation does not require representation *per se* (Schlesinger and Casey 2003; Schlesinger and Barto, 1999; Schlesinger and Parisi, 2001), this has only been demonstrated for cases where the visual information required for generating anticipations (through evocation of learned associations) remains available to the agent's senses throughout. It is unclear whether an agent could sustain those anticipations during change in its focus of attention, building up expectations based on the totality of its observations.

The view presented here then is that there is a case for both views; the focus of concern then becomes a matter of integrating JIT scene representation with more enduring representations of the world and, fundamentally, identifying when it is appropriate to prefer one rather than the other.

Background

The work presented here rests on the view that dynamically constructed situational models might usefully inform the interactions of autonomous agents in rapidly changing multi-agent domains. The context for this view has largely arisen through the development of a computer based agent which carries out the actions of a driver in a simulated driving world (Wood, 1993). The task domain provides a testbed for investigating goal-directed activity, such as following a route to a particular destination, and the integration of this with on the fly responses to situational changes brought about by events, such as changing traffic signals and other agents who may be slowing down for a red light.

A domain such as this imposes demands upon the agent to respond to events in a timely matter. This is achieved through mechanisms for anticipating the outcomes to events. This requires the agent to go beyond immediate sense data in anticipating how events will proceed; but also seems to require the ability to modify expectations in the context of knowledge about the domain (and how agents typically behave within it) combined with the evidence of the behaviour of other agents, in a given context. The resulting situational model characterizes the dynamically evolving sequence of events which provide the context of an autonomous agent's activity.

Integrity and Completeness of the Situational Model

Situational modeling, then, provides an informed context for analyzing current sense data, and a platform for realizing its implications for future scenarios (Wood, 1993; 1995a; 1995b; 1998). Of course, the advantages gained through situational modeling depend upon the quality of the knowledge held in the model. As a model of current events based in previously sensed data, its validity diminishes over time. Uncertainty in the sensed data and its implications contributes to this effect. Therefore, objects and agents remaining within view must be repeatedly observed if the integrity of the situational model is to be maintained.

Typically the viewer is limited in the rate at which it can apprehend new information. Consequently the model cannot provide a 'complete' description of the agent environment, rather it is selective in the information described. The incompleteness of the model upon which an agent relies in order to interact effectively with its environment is not inherently problematic, provided the model is good enough (Simon, 1981). Indeed, the findings on inattentional amnesia and change blindness clearly support this view. It is the inherent 'goodness' or integrity of the model which can guide us here: rather than seeking 'information about the world' an agent might seek 'information to maintain the integrity of its situational model', as it relies upon this to inform action. Identification of lapses in the integrity of knowledge held in the model can be used to guide sensing priorities, informing focus of attention and selective perception, with the aim of maintaining the quality of knowledge held (Wood, 1998).

To Represent or Not to Represent is *not* the Question

Viewing the situational model as not only a means of informing action but also of informing viewing strategy would also appear to offer a basis for understanding perhaps why some things get ignored.

We have already considered the phenomenon of 'change blindness' when a change in a central aspect of a scene remains undetected (Rensink, 2000a). What is intriguing is that a similar phenomenon is observed even though the changed feature is central to the agent's activity, as in Mary Hayhoe's (2003) experiments. Participants behave in accordance with having apprehended the colour of the objects to be manipulated, so this information would appear to be held in memory; why *then* failure to observe subsequent changes to this feature?

Hayhoe (*in conversation*) made a further observation that on occasions when participants *did* notice the new colour of the held object, they would frequently conclude they had mistakenly picked up another object to that intended, rather than that the object's colour had changed (even though this possibility had been mentioned).

Both phenomena described: (i) failure to interrogate the visual scene for object colour following its initial designation, and, (ii) when the object *was* re-interrogated for colour, the assumption that colour change was a result of mis-perception rather than actual change, could be explained through a particular characteristic of visual information seeking. It would seems that knowledge about our visual world, learned through experience, tells us that certain aspects of a visual scene are more enduring than others. Objects rarely change colour; following initial identification, therefore, there is little reason for checking an object's colour again. Consequently, a change in colour is more attributable to an error its initial perception or, more likely still in a cluttered scene, the failure to direct an action towards the correct object. Experience is a powerful determinant of visual experience (Lotto and Purves, 2002; Purves, Lotto and Nundy, 2002) and it would seem to be at least possible that experience might also play this role in guiding viewing strategy (cf. Rensink, 2002). Expectations regarding persistence effects would point to it being safe to ignore some aspects of our world over others, once the crucial information required has been initially apprehended.

The observation, if correct, that visual behaviour is consistent with experience, might also guide us in determining where and when an autonomous agent would be best advised to situationally model, and when not.

Summary

Purposeful autonomous agents would seen to require the ability to model the world around them if they are to be able to interact effectively within that world. On the other hand, evidence from human studies suggests that much activity can be supported through just in time scene representation. This poses the new problem of how to integrate in a coherent and useful way, knowledge reflecting situational understanding stored in memory, with transitory information from our visual surroundings.

The problem of identifying where and when it is preferable to construct more enduring representations, and the problem of identifying when to update them, appear to be related. It is suggested that evidence from human studies on the relationship between learned expectations and the likelihood that change will be ignored, could guide further investigation of these difficult problems.

References

Bacjsy, R (1985) Active perception vs passive perception. In *Proceedings of the Workshop on Computer Vision*, 55-59.

Ballard, D (1991) Animate vision: an evolutionary step in computational vision. Journal of the Institute of Electronics, Information, and Communication Engineers, **74**, 343-348.

Ballard, D, Hayhoe, M and Pelz, J (1995) Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, **7**, 66-80.

Biederman, I (1981) On the semantics of a glance at a scene. In M Kubovy and JR Pomerantz (Eds), *Perceptual Organisation*. Hilldale, NJ: Lawrence Erlbaum Associates.

Brooks, R (1986) A robust layered control system for a mobile robot. *IEEE Journal of Robotics and Automation*, **2**, 14-22.

Ferguson, I. (1992) TouringMachines: An architecture for dynamic, rational, mobile agents. TR Report No. 273, Computer Laboratory, University of Cambridge.

Friedman, A (1979) Framing pictures: the role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology – General*, **108**, 316-355.

Grimes, J (1996) On the failure to detect changes in scenes across saccades. In KA Atkins (Ed.) *Perception*. Vancouver Studies in Cognitive Science, 5. Oxford: OUP.Hayhoe, M. (2003) What guides attentional selection in natural environments? In *Abstract Proceedings of Fifth Workshop on Active Vision*, University of Sussex, September 22, 2003.

Intraub, H (1980) Presentation rate and the representation of briefly glimpsed pictures in memory. *Journal of Experimantal Psychology – Human Learning and Memory*, **6**, 1-12.

Hayhoe, MM, Shrivastava, A, Mruczek, R and Pelz, JB (2003) Visual memory and motor planning in a natural task. Journal of Vision, **3**, 49-63.

Intraub, H (1981) Rapid conceptual identification of sequentially presented pictures. *Journal of Experimental Psychology – Human Perception and Performance*, **7**, 604-610.

Karn, KS, Moller, P and Hayhoe, MM (1997) Reference frames in saccadic targeting. Experimental Brain Research, **115,** 267-282.

Lotto, RB and Purves, D (2002) The empirical basis of colour perception. *Consciousness and Cognition*, **11**, 609-629.

Mack, A and Rock, I (1998) *Inattentional Blindness*. Cambridge, MA: MIT Press.

Marr, D (1982) *Vision*, Oxford: WH Freeman.

Pezdek, K, Whetstone, T, Reynolds, K, Askari N, Dougherty, T (1989) Memory for real-world scenes – the role of consistency with schema expectation. *Journal of Experimental Psychology – Learning, Memory and Cognition,* **15**, 587-595.

Purves, D, Lotto, RB and Nundy, S (2002) From the cover: Why we see what we do. *American Scientist*, **90**, 236-243.

Rensink, RA (2002a) Change detection. *Annual Review of Psychology*, **53**, 245-277.

Rensink, RA (2000b) Seeing, sensing and scrutinizing, *Vision Research*, **40**, 1469-1487.

Rensink, R (2000c) The dynamic representation of scenes. *Visual Cognition,* **7**, 17-42.

Rensink, RA, O'Regan, JK and Clark, JJ (2000) On the failure to detect changes in scenes across brief interruptions. *Visual Cognition*, **7**, 127-145.

Schlesinger, M. and Barto, A. (1999) Optimal control methods for simulating the perception of causality in young infants. In M.Hahn & S.C. Stoness (Eds) *Proceedings 21st Conference of the Cognitive Science Society,* Erlbaum.

Schlesinger, M and Casey, P (2003) Visual expectation in infants: Evaluating the gaze-direction model.
Proceedings of the Third International Workshop on Epigenetic Robotics, Boston College, Boston, MA, August 4-5, 115-122.

Schlesinger, M and Parisi, D (2001) The agent-based approach: A new direction for computational models of development. *Developmental Review*, 21: 121-46.

Simons, D.J., & Chabris, C.F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. Perception, 28, 1059-1074

Tatler, B, Gilchrist, I and Rusted, J (2003) The time course of abstract visual representation. *Perception*, **32**, 579-592.

Wood, S (1983) Dynamic World Simulation for Planning with Multiple Agents. In Proceedings of the Eighth International Joint Conference on Artificial Intelligence (IJCAI-8), Karlsruhe, W.Germany, 8-12 August 1983, 69-71.

Wood, S. (1993) *Planning and Decision-Making in Dynamic Domains*. Chichester: Ellis Horwood.

Wood, S. (1995a) When Being Reactive Just Won't Do. *Proceedings of the AAAI Spring Symposium on Integrated Planning Applications*. Stanford University, Palo Alto, CA, March 27-9 1995, 102-6.

Wood, S. (1995b) The role of plan recognition in reducing situational uncertainty and directing attention in planning. Proceedings of the IJCAI Workshop on The Next Generation of Plan Recognition Systems: Challenges for and Insight from Related Areas of AI. Montreal, Canada, August 20, 1995, 124-9.

Wood, S (1998) A Probabilistic Network Approach to Prioritising Sensing and Reasoning: a step towards satisficing modeling. Proceedings of the AAAI Spring Symposium on Satisficing Models, Stanford University, Palo Alto, CA, March 23-5 1995, 87-90.

Wood, S (2003) Attentional Shift and Enactive Perception in Purposeful Agents. Proceedings of 7th Annual Conference of the Consciousness and Experiential Psychology Section of The British Psychological Society on Enactive Consciousness, Perception, Intersubjectivity and Empathy, St Anne's College, Oxford, June 28-29 2003.

Woolfe, JM (1999) Inattentional amnesia. In V. Coltheart (Ed) *Fleeting Memories*, Cambridge, MA: MIT Press.