

A system-level brain model of spatial working memory and its impairment

Alan H. Bond, email:alan.bond@exso.com,
National Institute of Standards and Technology,
MS 8263, Gaithersburg, Maryland 20899
and Semel Institute of Neuroscience, Geffen School of Medicine,
University of California at Los Angeles, Los Angeles, California 90095

Abstract. A system-level model of spatial working memory is described, using the author's computer science and logical modeling approach. The remembered mental image is located in a lateral parietal module, and plans are executed and sequenced in a dorsal prefrontal module. Mental images are saved and maintained by explicit messages from the currently executed plan and can be reinstated for use in motor control. By attenuating the goal, the plan fell below threshold, matching experimental results of Conklin et al. where schizophrenic and schizotypal subjects could perform a spatial working memory task with a 0.5, but not a 7, second delay.

Keywords: spatial working memory, system-level model, schizophrenia, cognitive model.

Introduction. In this study, we attempted to analyze and understand the underlying mechanisms involved in spatial working memory. This has been shown to be clearly impaired in schizophrenia. In order to understand it, we developed a system-level model of spatial working memory. We used a modeling approach of a modular distributed computational architecture and an abstract logical description of data and control [3], for which we have also analyzed the correspondence to the cortex [4].

Modeling spatial working memory involved developing mechanisms for a frontal area containing a maintenance process, a posterior area containing mental images, and an integration mechanism involving a simple model of episodic memory, corresponding to the hippocampal complex.

We implemented the model as a computer system and studied normal behavior and then abnormal behavior by introducing different types of component deficit. We specifically obtained a match to the work of Conklin et al [9] on schizophrenic and schizotypal subjects. For short delays such as 0.5 seconds they performed normally, but for long delays such as 7 seconds they exhibited a clear deficit.

Spatial working memory. We can perhaps define spatial working memory by describing the basic experiment, which has four steps: (1) fixate - there is a cross at the origin and the subject has to fixate it, duration 2000 milliseconds. (2) note image - there is now also an asterix at a peripheral location, duration 200 milliseconds. (3) delay, maintain image - back to just the cross, duration either 500 milliseconds or 7000 milliseconds. (4) cue, move hand to where the asterix was, duration 5000 milliseconds.

Our brain modeling approach. In the last few years, we have conducted a series of studies and models concerning problem solving [2], episodic memory [6], natural language processing [5], routinization [7], and social relationships [1]. For this project, we have begun integrating all of these mechanisms into a single system which we call our dynamic model.

A system-level brain model is a set of parallel modules with fixed interconnectivity similar to the cortex, and where each module corresponds to a brain area and processes only certain types of data specific to that module. We view all data streams and storage as made up of discrete data packets or chunks and we represent each as a logical literal which indicates the meaning of the information contained in the chunk. An example data packet or chunk is `position(adam,300,200,0)` which might mean that the perceived position of a given other agent, identified by the name “adam”, is given by (x,y,z) coordinates (300,200,0). Thus each module only processes certain types of chunks. In order to allow for ramping up and attenuation effects, we give every chunk an associated strength, which is a real number. Stored chunks are ramped up by incoming identical or related chunks, and they also attenuate with time, at rates characteristic of the module. We represent the processing within a module by a set of left-to-right logical rules which are executed in parallel. A rule matches to incoming transmitted chunks and to locally stored chunks, and generates results which are chunks which may be stored locally or transmitted. Rule patterns also have weights, and the strength of a rule instance is the product of the matching chunk weights and the rule weights, multiplied by an overall rule weight. A rule may do some computation which we represent by arithmetic. This should not be more complex than can be expected of a neural net. The results are then filtered competitively depending on the data type. Typically, only the one strongest rule instance is allowed to “express itself”, by sending its constructed chunks to other modules and/or to be stored locally. In some cases however all the computed data is allowed through. One cycle of the model corresponds to about 20 milliseconds. The uniform process of the cortex is then the mechanism for storage and transmission of data and the mechanism for execution of rules.

Our approach differs from neural network approaches in using an abstract method of description, so that information is represented by abstract chunks and processing by rules which describe the processing of chunks. Thus we do not model individual neurons, but instead we model an abstraction of a large set of neurons constituting a neural area. It also differs from the abstract neural models of Cohen and Braver [8] in using complex data items and matching, and in using complex computation and control within each module. Our approach is complementary to neural network approaches, and it should be possible, for a given abstract model, to construct corresponding neural network models.

Our approach to the design of a spatial working memory model. Our model is implemented as a set of intercommunicating brain modules that run in parallel. One cycle

corresponded to 20 milliseconds. The design for our model involves:

- (i) generalizing our planning module to use learned plans represented as chunks. so that rules competitively reconstruct chunks in response to their current situation,
- (ii) adding a mental imagery module which stores and represents mental images in terms of image elements and their spatial relationships, and
- (iii) adding a module corresponding to the hippocampal complex, and which receives data from cortical areas and constructs a representation of the current mental event.

Figure 1 diagrams the design of the basic spatial working memory model.

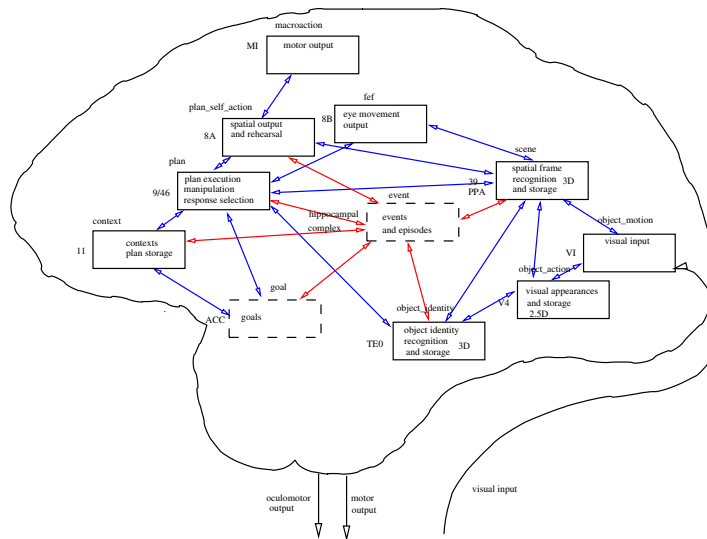


Figure 1: Outline diagram of an initial model

Noting, maintaining and using mental images. In our model, plan steps send messages to the scene module to perform certain operations on mental images stored there. Basically, we need to make a note of an image of the scene to be remembered during the note phase, then we need to maintain this image and stop it from attenuating away during the maintain phase and then we need to reinstate the remembered image during the action upon cue phase.

Given this approach, we were able to precisely define noting, maintaining and reinstating of images:

- (i) Noting the current mental image creates a new scene which is labeled by a name sent from the planning module, instead of a name derived from the current episode key.
- (ii) Maintaining a named mental image consists of simply executing a rule which recognizes and reconstructs its components.
- (iii) Instating a named mental image. When we come to act using the stored mental image, we

instate the noted image to become the current mental image and its spatial properties are then used by the planning and motor hierarchy to execute the desired motor action. This temporarily deemphasizes the percept which is continuously being refreshed from the input visual stream.

The spatial working memory model as implemented. All of the above mechanisms have been designed and implemented and the system will successfully carry out the basic spatial working memory experiment. The system was programmed in Sicstus Prolog and the BAD language. The BAD language and manual can be found at <http://www.exso.com/bad.html>.

Predicted imaging files. We developed a measure of energy consumption based on the number and types of different kinds of information processing operation being performed by the model, and we used this measure to generate predicted brain activation maps. We show afni images generated by our system for the 126th cycle, in the maintain noted image phase, in Figure 2, where we have chosen samples corresponding to sagittal, axial and coronal views for each time, and for two different geometric positions, we use Talairach coordinates:

- (i) sagittal 30, axial 10 and coronal -43, which is the plan module or Brodmann 10 and 46
- (ii) sagittal 44, axial 36 and coronal 41, which is the scene module or Brodmann 40.

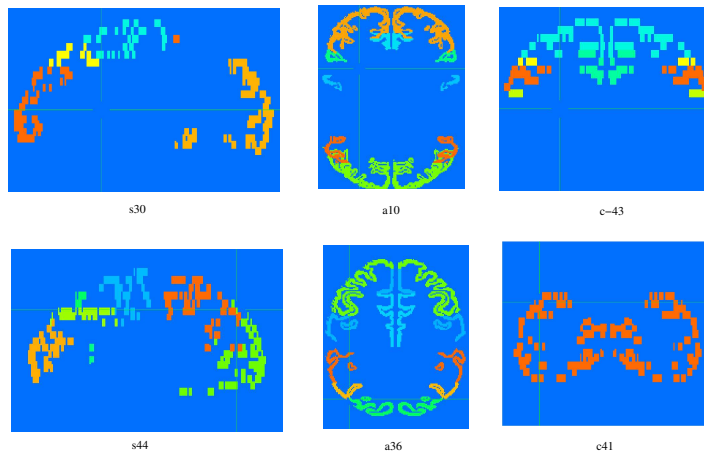


Figure 2: Predicted imaging files for cycle 126, in the maintain noted image phase

Results for the time course of energy consumption. Figure 4 show the time course of energy consumption by the scene module (blue and uppermost), the plan module (red and second), the context module (purple and third) and the goal module (green and lowest). The ordinate is a measure of the instantaneous energy consumption. The abscissa is time and is in units which are cycles, i.e., 20 millisecond increments, the marked divisions being at 50 cycle intervals, from 0 to 900 which is the duration of the longest experiment. We can see the spurts in energy at the main transition points between phases, at cycles 10 (fixate), 111 (note), 126 (maintain) and 471 (move hand). We also see the imaging system staying active as it visually

tracks the movement of the hand, after which the system falls into a rest state.

Modeling the spatial working memory deficit in schizophrenia. We investigated in what ways the system could be compromised. By examining the dependencies of one data type on others, we could systematically determine all the effects of lesioning each component of the model.

Figure 3 outlines the logical structure of the model as implemented, showing the main de-

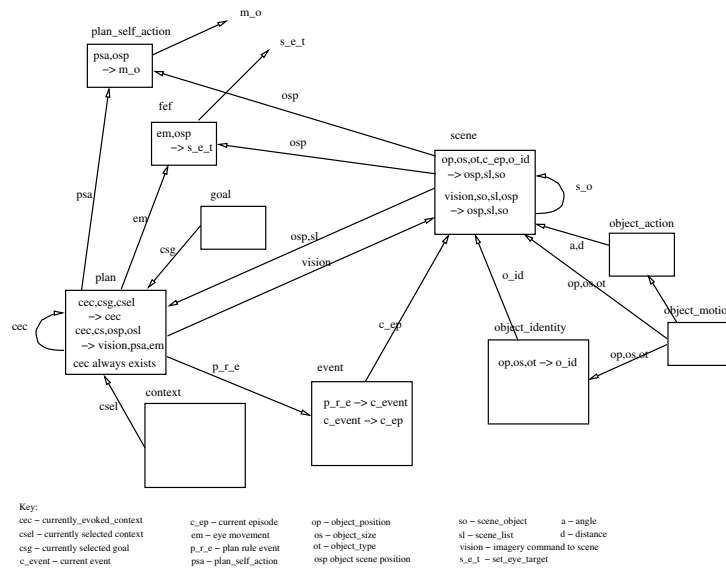


Figure 3: Logical diagram of system as implemented

scriptions and their logical dependencies. This diagram does not show the detailed temporal dynamics of how the model operates, it is a description of all the data types involved and how they are derived from each other.

We were also able to obtain a graded deficit in spatial working memory performance by allowing the goal to attenuate faster in time. We found that attenuating the goal was the best way to obtain the experimentally observed deficit, since attenuating the plan, or the image store, directly made it difficult to get normal behavior in the short delay case.

For normal performance the attenuation rate for goals was set to a very small amount, corresponding to extinction in 1000 cycles, whereas for the pathological case we set it to attenuate faster, corresponding to extinction in about 400 cycles. This resulted in the pathological case being able to carry out the task with a 0.5 second delay but not for a 7 second delay. Figure 4 shows the energy curves in the 7 second delay case. The graph shows the reduction of the goal and context energy, and it also shows that the move hand operation did not occur.

This approach reduces the activation level of the frontal planning area, which agrees with some experimental findings.

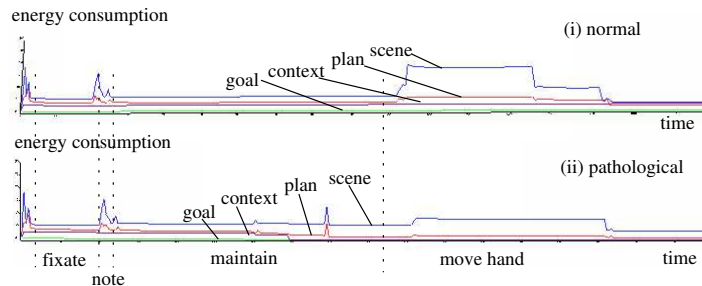


Figure 4: Time course for 7 second delay experiment, (i) normal case and (ii) pathological case

References

- [1] Alan H. Bond. Modeling social relationship: An agent architecture for voluntary mutual control. In Kerstin Dautenhahn, Alan H. Bond, Dolores Canamero, and Bruce Edmonds, editors, *Socially Intelligent Agents: Creating relationships with computers and robots*, pages 29–36. Kluwer Academic Publishers, Norwell, Massachusetts, 2002.
- [2] Alan H. Bond. Problem-solving behavior in a system model of the primate neocortex. *Neurocomputing*, 44-46C:735–742, 2002.
- [3] Alan H. Bond. A Computational Model for the Primate Neocortex based on its Functional Architecture. *Journal of Theoretical Biology*, 227:81–102, 2004.
- [4] Alan H. Bond. An Information-processing Analysis of the Functional Architecture of the Primate Neocortex. *Journal of Theoretical Biology*, 227:51–79, 2004.
- [5] Alan H. Bond. A psycholinguistically and neurolinguistically plausible system-level model of natural-language syntax processing. *Neurocomputing*, 65-66:833–841, 2005.
- [6] Alan H. Bond. Representing episodic memory in a system-level model of the brain. *Neurocomputing*, 65-66:261–273, 2005.
- [7] Alan H. Bond. Brain mechanisms for interleaving routine and creative action. *Neurocomputing*, 69:1348–1353, 2006.
- [8] Todd S. Braver, Deanna M. Barch, and Jonathan D. Cohen. Cognition and control in schizophrenia: a computational model of dopamine and prefrontal function. 46:312–328, 1999.
- [9] Heather M. Conklin, Clayton E. Curtis, Monica E. Calkins, and William G Iacono. Working memory functioning in schizophrenia patients and their first-degree relatives: cognitive functioning shedding light on etiology. *Neuropsychologia*, 43:930–942, 2005.

Biosketch. Alan H. Bond was born in England and received a Ph.D. degree in theoretical physics in 1966 from Imperial College of Science and Technology, University of London. From 1966 to 1969, he was on the faculty of the Computer Science Department at Carnegie-Mellon University, Pittsburgh. During the period 1969 to 1984, he was on the faculty of the Computer Science Department at Queen Mary College, London University, where he founded and directed the Artificial Intelligence and Robotics Laboratory. From 1985 to 1992, he lead research in applied artificial intelligence at the University of California, Los Angeles. He has published research in autonomous robotics, multiagent systems and parallel computer architectures. From 1996 to 2003, he was a Research Scientist and Lecturer at the California Institute of Technology. He is currently a Specialist in Cognitive Neuroscience in the Neuropsychiatric Institute at University of California at Los Angeles and a Visiting Professor at the National Institute of Standards and Technology in Gaithersburg, Maryland. His main research interest concerns the system-level modeling of the human brain.

