Constructing a Physiologically Realistic Machine Model of Consciousness

L. Andrew Coward
Department of Computer Science
Australian National University

Abstract

There are a set of theoretical bounds on any system which must learn to perform a complex combination of features with limited information handling resources. A cognitive architecture implemented within these bounds has the capability to generate human-like "conscious" behaviours with mechanisms that are physiologically plausible.

The Characteristics of a Physiologically Realistic Model of Consciousness

As discussed in [1], the objective of a physiologically realistic model of consciousness is not to "understand" subjective individual experience in a philosophical sense. Rather, the objective is to model observable phenomena labelled "conscious" in terms of physiological structures, in such a way that observed cognitive inputs (perceptions etc.) result in physiological states, and these physiological states cause other physiological states that generate the observed high level cognitive outputs.

Take, for example, the phenomenon of dichotic listening [2]. This is the phenomenon that occurs when different texts are presented to the left and right ears of human subjects and the subjects are told to echo (repeat) the text delivered to just one ear. If the texts are switched between ears during the shadowing, subjects switch ears to the meaningful continuation without being aware of having switched. Subjects have no memory of the unshadowed text. Thus although there is no memory of the text presented to the unshadowed ear, such text must be able to influence behaviour because it is able to trigger a switch. A physiological model needs to describe the physiological states that result from the presentation of the two texts and how those states generated other states, and demonstrate that those later states result in the ability to switch from one ear to the other when the texts switched between ears, but create future memories of only one text. It is of course possible that physiological models of this type may give insight into the issues of philosophical understanding.

Many different phenomena have been labelled "conscious", and there is no a priori guarantee that these phenomena are supported by exactly the same physiological mechanisms. It could also be that it is the occurrence of a sequence of physiological processes in a particular order that results in "conscious" phenomena, where all of the same processes in a different order might not result in a "conscious" process. In this context, searches for "neural correlates of consciousness", or detectable physiological states that occur consistently at the same time as a wide range of conscious phenomena, may be somewhat simplistic.

Phenomena labelled conscious have been classified into a number of types including access consciousness and phenomenal consciousness [3]. Access conscious is defined as the ability to report and act upon an experience and requires the existence of some "representation" of the experience in the brain, the content of which is available for verbal report and for high level processes such as conscious judgments, reasoning, and the planning and guidance of action. Phenomenal consciousness refers to the qualitative nature of experience, for example why the experience of the colour red feels as it does and not like something else.

As a starting point for creation of a physiologically realistic model, it is essential to have more specific examples of these phenomena and other phenomena that are not "conscious" for comparison purposes. In the case of access consciousness, one such phenomenon is dichotic listening as described above. To provide phenomena for different types of consciousness, consider the following scenario (similar to that discussed in [1]). In the scenario, a person is out walking with a companion, and encounters a tree partially blocking the path. One behaviour is simple avoidance: stepping around the tree. A second behaviour is to make the comment "That tree is small but looks old. Up in the mountains I saw a whole area covered with trees like it". A third behaviour is to focus attention on the tree and "become aware" of the tree as a tree. The cognitive processes suggested by these scenarios could be visual input from a tree, unconscious activation, conscious activation, avoidance behaviour, and higher cognitive behaviour (including associative thinking, and verbal report of associative thinking).

The first behaviour can generally be unconscious. Sensory input from the tree and the path etc. generates some activation state internal to the brain which leads to avoidance behavior but does not lead to higher cognitive functions, verbal report, or in many cases even later memory. The second behaviour falls within the definition of access consciousness. The internal brain activation state in response to the tree generates both cognitive processing and a complex verbal report.

The third behaviour appears to have some relationship with what is often called phenomenal consciousness. Becoming aware of the tree as a tree is subjectively a richer experience of the tree, the experience is difficult to express in speech and may be very individual specific. The cognitive phenomena resulting from this third behaviour are therefore verbalization of the presence of a richer experience which cannot be described verbally in detail, but may be very idiosyncratic.

The objective of this paper is to demonstrate how a physiologically realistic machine model could be given sensory inputs, and from the physiological states generated by those inputs would in turn generate physiological states corresponding with observed "conscious" cognitive outputs.

Architectural Constraints on Complex Learning Systems

It has been demonstrated [4] that any system which must learn to perform a complex combination of behaviours with limited information handling resources will be forced within some very specific architectural bounds by a number of practical considerations. The practical considerations are (1) the need to perform a large number of behavioural features with relatively limited physical resources for information recording, information processing and internal information communication; (2) the need to add and modify features without side effects on other features; (3) the need to protect the many different meanings of information generated by one part of the system and utilized for different purposes by each of a number of other parts of the system; (4) the need to maintain the association between results obtained by different parts of the system from a set of system inputs arriving at the same time (i.e. maintain synchronicity); (5) the need to limit the volume of information required to specify the system construction process; (6) the need to limit the complexity of the construction process; and (7) the need to recover from construction errors and subsequent physical failures or damage. All of these considerations apply in some form to biological brains. For example, a brain which can learn to perform a given set of behavioural features with fewer resources will have a natural selection advantage over a brain requiring more resources for the same behaviours. A brain needs to learn without excessive interference with prior learning and so on.

The architectural bounds have been labelled the recommendation architecture. The bounds include a number of specific requirements. One requirement is for a hierarchy of modules (called clustering) that defines and detects conditions within the information available to the system. This information includes current raw sensory inputs, information indirectly activated by conditions detected within those inputs, and information about the current state of the system itself. One module defines and detects a group of similar conditions. Modules do not correspond with cognitive circumstances like categories, and the groups of modules that detect conditions in response to instances of different categories may overlap. However, these groups are sufficiently different that it is possible to discriminate between instances of different categories.

A second requirement is for a separate hierarchy of components (called competition) that receives a subset of the current condition detections from clustering, interprets each condition detection as a range of recommendations in favour of different behaviours, each with a different weight, and selects the most strongly recommended behaviour. Consequence feedback can change the recommendation weights of recently performed behaviours but not the condition definitions.

A third requirement is that devices within clustering can record conditions, but with a strong tendency for any subsequent repetition of the exact condition at any point in the future to be detected. A fourth requirement is for careful management of when and where additional conditions will be recorded, based upon the degree of condition detection in different parts of the modular hierarchy.

A fifth requirement is for devices in clustering to be arranged in layers to maintain an adequate level of synchronization between conditions within one input state detected by different modules. A modular hierarchy of columns, arrays and areas must be overlaid on the layered structure.

A sixth requirement is for some mechanism to allow simultaneous detection of conditions derived from multiple input states within the same physical group of modules, in such a way that the condition detections are kept separate until it is appropriate to detect combination conditions. For example, if conditions recommending behaviours in response to a group of objects are needed, conditions must first be detected within each object separately, then conditions detected that are combinations of individual object conditions. Resource limitations will tend to require use of the same modules for all the object condition detections.

A seventh requirement is for the components in competition to correspond with specific behaviours, types of behaviours or sequences of behaviours. The competition process is that current condition detections from clustering are first provided to components corresponding with different general types of behaviour. The component corresponding with the type most strongly recommended by these condition detections releases the subset of condition detection instantiating the accepted recommendation (perhaps with some intermediate processing by clustering) to a group of components corresponding with different types of behaviour within its general type. Competition between these components results in selection of a yet more specific behaviour or type until an end point at which a small subset of condition detections are released from clustering to drive behaviour.

The reasons these architectural constraints result from the practical considerations are described in detail in [4]. One critical observation is that if resource limitations are a significant factor, there will be very little direct resemblance between "user manual" type models that describe how cognitive features work in terms easily understood by an outside observer, and "system architecture" type models that describe how physiological resources are organized to provide the processes required to support those cognitive features. The differences arise because the physical modules in a system architecture will be defined in such a way that each module is customized to efficiently perform a group of similar system processes. Many cognitive features could require processes performed by one module, and one module will therefore support many such features. The relationship between "system architecture" modules and "user manual" features will therefore be very complex, a commonplace situation for example in commercial electronic systems. Direct implementation of a user manual would be extremely resource intensive because it would imply separate resources for each user manual defined feature. The implication is that "user manual" type descriptions of consciousness with "modules" corresponding with clearly identified cognitive features or processes (e.g. [5]) may be valuable for understanding at the cognitive level but the mapping between such models and physiology will be very complex, and attempts to implement machine versions of such models will be very resource intensive and physiologically unrealistic.

The Recommendation Architecture Cognitive Model

A physiologically realistic cognitive model within these architectural bounds has been described [4]. In this model, the cortex corresponds with the recommendation architecture clustering subsystem. Cortex devices are organized into layers, columns and areas as required by the recommendation architecture bounds. The thalamus, basal ganglia and cerebellum correspond with competition subsystems making behaviour selections determining, respectively, the sensory information which will be allowed to influence behaviour, the general type of behaviour and the specific behaviour. Nuclei within the thalamus and basal ganglia correspond with different more specific behaviour types. The hippocampus corresponds with the recommendation architecture subsystem that manages the assignment of clustering (cortex) resources. Other brain structures correspond with other recommendation architecture subsystems [4].

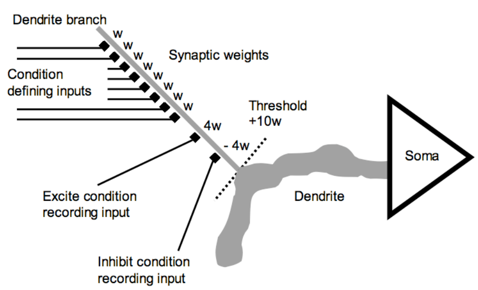

Figure 1. Pyramidal neuron as a condition recording device.

Pyramidal neurons in the cortex correspond with devices that define and detect conditions. The physiological mechanisms involved can be understood by consideration of figure 1. In the figure, there are initially a group of synaptic inputs to a pyramidal neuron dendrite branch. This branch integrates the post synaptic potentials from these synapses, and injects potential into the dendrite if a threshold is exceeded. Further integration occurs within the dendrite as a whole, and then integration within the neuron body (soma) determines the generation of action potentials into the output axon. In information terms, such an action potential indicates the detection of a significant number of the conditions programmed on the dendrites. There are several types of synaptic inputs to the dendrite branch. A number of inputs are from pyramidal neurons in the preceding cortex layer. These inputs in information terms have the capability to define a condition. They are excitatory and for simplicity in the diagram all have the same weight w. Initially these inputs do not have enough total synaptic strength to activate their branch. There are two other types of input. One type is excitatory and in information terms excites condition recording. The other type is inhibitory and in information terms inhibits condition recording. Both these types of input indicate the overall level of activity in some specific module of pyramidal neurons, in general via some neuron (spiny stellate for excitatory, interneuron for inhibitory) located in that module and receiving inputs from the pyramidal neurons in that module. The condition defining inputs alone cannot activate the branch. However, if a high proportion of those inputs are active, and at the same time the excite condition recording input is active and the inhibit input is inactive, there may be enough postsynaptic potential to activate the branch. If this activation is followed a short while later by the soma producing an action potential, the strengths of the recently active synapses is increased, and the absolute strengths of the condition definition management inputs is decreased. These changes are similar to the LTP and LTD effects observed in pyramidal neurons [e.g. 6]. The effect is that the recently active synapses acquire enough postsynaptic strength to activate the branch independent of the state of the management inputs. In information terms the branch has recorded a condition.

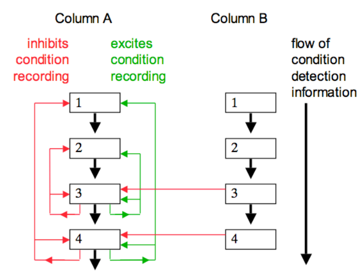

Figure 2. Condition recording change management connectivity within one column and from other columns.

Pyramidal neurons are organized into layers, with columns of neurons penetrating several layers. These columns are organized into arrays in which the same input space is available to all the columns in the array. Columns and arrays manage the circumstances under which conditions will be recorded, as illustrated in figure 2 for a four level column. Note that the recommendation architecture bounds do not specify the number of layers in a column, only the different functions which must be performed. The four layer model is based upon current physiological knowledge of inter and intra layer connectivity in the Macaque cortex (for further details see [4]). Not all connectivity is illustrated, but within column A there is connectivity from layer 3 to layers 2 and 3 which excites condition recording, and similar connectivity that inhibits condition recording. There is analogous connectivity from layer 4 to layer 4 and layer 1. If connectivity from other columns is ignored, the effect is that there can only be a low level of activity in layers 3 and 4, or a high level. An intermediate level would result in condition recording to push activity to a high level. For a more detailed discussion, see [4]. The effect of intercolumn connectivity is that if there is a moderate level of activity in layer 3 of a number of columns, in some columns it will increase to high, in others it will reduce to low.

The overall effect of the column connectivity is that at least a minimum number of columns in an array will detect conditions in response to every input state, with conditions present within that input state being recorded if necessary to reach the minimum level. An array of columns defined in this fashion will divide up an input space into statistically relatively independent components (analogous with but significantly different from independent components analysis [7]). Such arrays of columns will discriminate between different types of input states, although no one column will correspond with one type of input space [8].

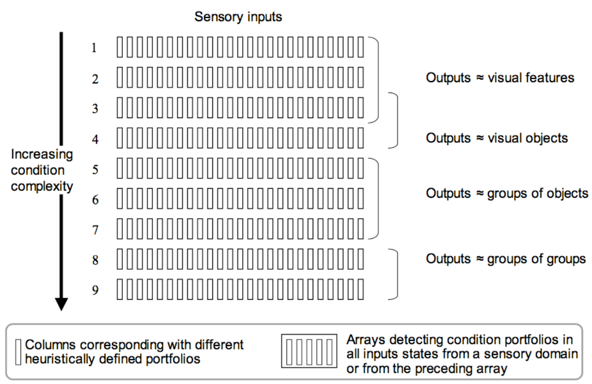

Figure 3. A sequence of arrays of columns detecting conditions on different levels of complexity that can discriminate between different types of cognitive circumstances.

Arrays will be arranged in a sequence as illustrated in figure 3. In the figure, array 1 of columns detects conditions which are combinations of relatively raw sensory inputs. Array 2 detects conditions which are combinations of the conditions detected by array 2 and so on. Successive arrays therefore detect conditions of increasing complexity, where the complexity of a condition is the total number of raw sensory inputs that contribute to it, either directly or indirectly via intermediate conditions. Conditions with different levels of complexity will be effective for discriminating between different types of cognitive situations. Relatively simple conditions will be effective for discriminating between cognitive features, somewhat more complex conditions for discriminating between cognitive objects, yet more complex conditions for discriminating between different groups of objects and so on.

In order to detect conditions within a group of objects, it is necessary for conditions to be detected simultaneously within the individual objects in arrays 1 through 5 in figure 3, then combination conditions detected in the following arrays. However, it is important that the "object" conditions are kept separate. In other words, in arrays 1 through 5 it is important not to detect conditions containing information derived from several objects. As described in detail in [9] the requirement to keep independent populations of condition detections active without interference within the same physical groups of neurons can be achieved by a frequency modulation mechanism.

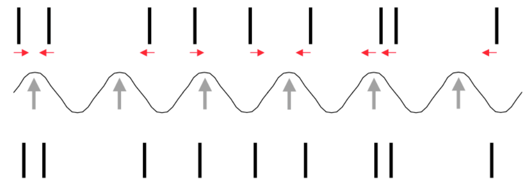

Figure 4. An irregular sequence of output action potentials from a neuron at the top is frequency modulated by an imposed signal in the middle to generate the bottom, modulated spike sequence.

As illustrated in figure 4, if the outputs of neurons are sequences of voltage spikes, frequency modulation means that each spike is shifted slightly towards the nearest peak in the modulation frequency. Neurons can be viewed as integrating their spike inputs over an integration time. For appropriate threshold levels, a modulated group of inputs will generate a much stronger response than the same inputs unmodulated. Hence modulation of the inputs derived, say, from one object in a visual field, will result in condition detection within that object and not with the rest of the field.

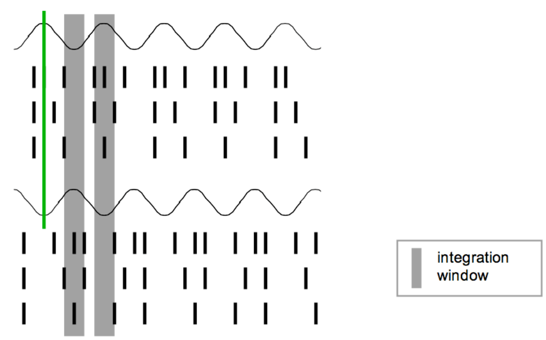

Figure 5. Two groups of neuron outputs, modulated with different phases of the same frequency, tend to have most of their output action potentials within different integration windows.

As illustrated in figure 5, if one group of inputs are frequency modulated with the same phase, and another group with the same frequency but a different phase, then provided the integration window is significantly smaller than the modulation frequency, spikes from the different groups will tend to occur in non-overlapping integration windows. The spikes generated by the neurons targetted by those inputs will tend to have the same frequency modulation as their inputs. Hence in figure 3, by placing different phases of frequency modulation on raw sensory inputs derived from different visual objects, condition detection in the upper arrays can proceed for the different objects with limited interaction between them. If the outputs from (say) layer 4 are brought into the same phase, layer 5 and subsequent layers will detect conditions within the group of objects.

Individual columns will not correspond exactly with specific cognitive circumstances. Rather, each column will have recommendation strengths in competition in favour of a wide range of different behaviours appropriate in response to different cognitive circumstances which may contain conditions recorded within the column. Some of the behaviours which may be recommended by a column are listed in table 1.

Table 1. Different types of behavioural recommendation strengths which may be possessed to different degrees by even a single column

A. Motor Behaviour Management
1. Perform a general sequence of motor behaviours
2. Perform a specific sequence of motor behaviours
3. Perform an individual motor behaviour

B. Speech Behaviour Management
1. Generate a sound
2. Speak a word
3. Say a phrase
4. Express meaning verbally by a sentence

C. Attention management
1. Perform a general sequence of attention and internal activation behaviours
2. Perform a specific sequences of attention and internal activation behaviours
3. Perform an individual attention behaviour

D. Managing Information Availability
1. Prolong the activity of currently active portfolios
2. Shorten the activity of recently activated portfolios
3. Synchronize the activity of several different populations of currently active portfolios in the same array
4. Release outputs of one array to the next array

E. Activation on the Basis of Frequent Past Activity
1. Activate portfolios which have often been active at the same time in the past as the currently active portfolio
2. Activate portfolios which have often been active just after past activity of the currently active portfolio
3. Activate portfolios which have often been active just before past activity of the currently active portfolio

F. Activation on the Basis of Past Condition Recording
1. Activate portfolios containing conditions recorded at the same time in the past as some conditions in the currently active portfolio
2. Activate portfolios containing conditions recorded just after some conditions in the currently active portfolio
3. Activate portfolios containing conditions recorded just before some conditions in the currently active portfolio

G. Activation on the Basis of Recent Activity
1. Reactivate portfolios which have recently been active
2. Reactivate portfolios which recently recorded conditions

Recommendation strengths are instantiated in the connection strengths of the outputs from cortex columns into neurons in different nuclei within subcortical structures such as the thalamus and basal ganglia. Behaviour types A and B ultimately result in physical muscle movements outside the brain. However, the other types result in different information management behaviours within the brain that are important intermediate steps in the course of generating behaviour.

Consider first attention behaviours (type C). Columns detecting conditions correlating with the presence of closed boundaries at different places in the visual field all have recommendation strengths in favour of shifting the attention domain to correspond with their boundary. Acceptance of one such recommendation may require eye movements, but the key result of acceptance is that all the sensory inputs derived from the area within the closed boundary are modulated with the same phase, and conditions within the selected object are therefore preferentially detected. Sequences of attention behaviours exist for learned cognitive processes. For example, in processing an arithmetic equation, attention shifts in a particular sequence between different sub-objects within the equation [4].

Information availability recommendation types (D) are required to manage how long a group of condition detections will be allowed to influence behaviour selection, and also to manage when the outputs from several separate active column populations (at different phases) in one array will be synchronized and released to the next layer.

Indirect activation recommendations (E, F, and G) can expand the population of condition detections available to influence behaviour. The conditions detected within the visual objects currently the focus of attention will have recommendation strengths in favour of externally directed behaviours in response to the object. However, there may be other conditions which are not currently being detected but which may have appropriate recommendation strengths for current circumstances. For example, columns which are currently inactive (i.e. not detecting conditions) but which have often been active in the past at the same time as currently active columns may have relevant recommendation strengths in some circumstances.

Columns therefore have recommendation strengths in favour of activating such other columns. Similarly, if two columns record conditions at the same time, each column acquires recommendation strength in favour of activating the other. Such recommendation strengths will in general decay with time unless reinforced by actual use followed by positive consequence feedback. As discussed in [10], indirect activation on the basis of frequent past simultaneous activity supports semantic memory, and indirect activation on the basis of past simultaneous condition recording supports episodic memory. Such indirect activations must be recommendations that compete with, for example, externally directed behaviours, otherwise the brain could be swamped with irrelevant information in the course of performing any behaviour.

Modelling Conscious Phenomena

This cognitive model makes it possible to model conscious phenomena in terms of physiology.

Dichotic Listening

Consider first the dichotic listening phenomenon discussed earlier. A version of the cognitive architecture simplified for the purposes of explanation is illustrated in figure 6. The sequence of clustering arrays illustrated in figure 3 has been compressed to three arrays in the case of visual processing, and one array for each ear in the case of auditory processing. The columns in the first visual processing array detect conditions which can discriminate between different types of visual object, those in the second array can discriminate between different types of groups of objects and so on. The columns in the array detecting conditions in the auditory information presented to the left ear can discriminate between different words, and similarly for the right ear.

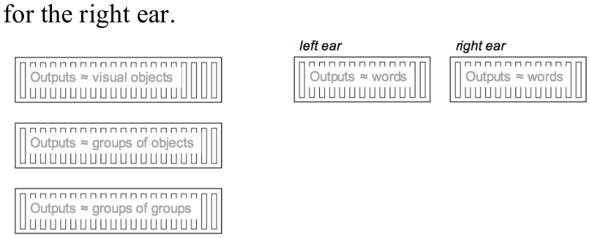

Figure 6. A simplified set of column arrays for the purpose of describing dichotic listening.

The activation state at the point at which the text switch occurs is illustrated in figure 7. When a word is heard in the left ear, columns detect conditions within the auditory information contained in the word. To simplify speech learning considerably (for a somewhat more realistic discussion see [4]), the "auditory" columns activated by hearing a word like "dog" have often been active in the past when visual columns activated by seeing a dog have also been activated, because a teacher has often spoken the word when the learner's attention was directed towards the visual object. The auditory columns have therefore acquired recommendation strengths in favour of activating the visual columns.

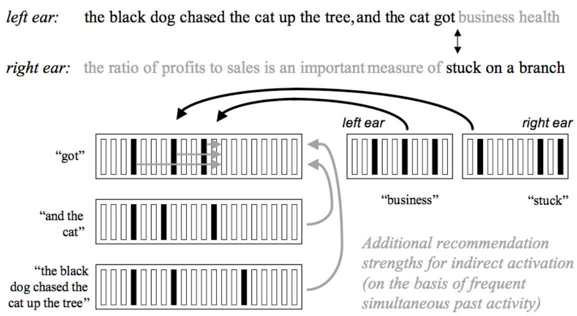

Figure 7. The pattern of column activation just as the switch of texts between ears occurs.

Hence the word "dog" results in activation of a visual image of an average dog, although only in arrays detecting conditions at intermediate levels of complexity (i.e. not at the simpler levels where the result would be a visual hallucination, because there is much less consistency in past activity between the auditory columns and the visual columns at these levels). Hearing a phrase like "the black dog" first indirectly activates columns in the top visual array corresponding with "the" and holds them active, then columns (at a different modulation phase) corresponding with "black", then columns corresponding with "dog". The outputs from the three column populations in the top array are then synchronized and released to the middle array and conditions detected corresponding with a visual image of the phrase, or group of objects. This population is held active.

A similar process then occurs in response to the phrase "chased the cat", resulting in a second activated column population in the middle array. After the phrase "up the tree" there are three separate activated column populations in the middle array, and the outputs from these populations are synchronized and released to the bottom visual array. A population of columns is activated in that bottom array that detects conditions in the group of phrases "the black dog chased the cat up the tree". Hearing the next phrase "and the cat" leads to an activated column population in the middle array, then hearing the word "got" results in a single activated population in the top array. At this point, as illustrated in figure 7, there is one population active in the bottom array, one population in the middle array, and one in the top array.

At each point in this process, there were also recommendation strengths in favour of indirect activation of visual columns corresponding with the words heard in the right ear, but the recommendation strengths of words heard in the left ear were enhanced by the instruction to echo the text heard in that ear.

Consider now the situation when the word "business" is heard in the favoured left ear, and the meaningful continuation "stuck" in the right ear. There are some additional recommendation strengths which must be considered. The visual columns currently active in the three arrays have recommendation strengths in favour of activating portfolios which have often been active at similar times (the same time or shortly after) in the past. Because "stuck" is the meaningful continuation, these strengths reinforce the indirect activation strengths of the auditory portfolios corresponding with the left ear. These strengths are sufficient to shift the predominant recommendation strengths over in favour of that meaningful continuation. Note that these strengths were always present, but until the switch reinforced contiuation in the targetted ear. The indirect activation recommendation strengths of columns activated by the untargetted right ear were also always present, but did not result in pseudovisual activation and therefore (through condition recording) memory. However, their constant presence results in the switch when they are reinforced by the recommendation strengths of the currently active visual portfolios.

The behavioural scenario

To understand how physiology could lead to the behavioural scenarios described earlier, it is necessary first to briefly outline the mechanisms supporting episodic memory of events in the recommendation architecture cognitive model (for more details see [10]). In the architecture illustrated figure 3, consider the pattern of column activation generated in response to viewing news of a novel event on television. The pattern of activation will extend through all the layers, and the novelty of the event means that there will be significant condition recording particularly at the group of objects and group of groups levels.

Later, words are spoken with some relationship with the original event. The auditory portfolios directly activated by the words will generate an indirect activation in the "object" levels 3 and 4. This activation will drive column activation in the group of objects levels 5 to 7 and perhaps beyond. Suppose now that the behaviour type *indirect activation of columns on the basis of simultaneous past condition recording* is favoured.

The currently active columns will all have recommendation strengths of this type, derived from all the past events during which they recorded conditions at the same time as other columns. Each such past event will have resulted in some such recommendation strength. However, if the words spoken have been chosen well, the reactivation strengths established by the target event may be relatively small but will be greater than for any other event.

Hence the population of columns activated in response to the words will evolve to a secondary population with a higher proportion of columns that recorded conditions during the target event. This in turn could evolve on the same basis to a tertiary population with an even higher proportion of target columns. The end point is a population of columns in which a major proportion are columns which were active during the original event. This population will have recommendation strengths in favour of, for example, speaking about that event. Hence use of indirect activation on the basis of correlations in past condition recording can support cued episodic memory for events.

In the behavioural scenario discussed earlier, a person is out walking with a companion, and encounters a tree partially blocking the path. One behaviour is simple avoidance: stepping around the tree. A second behaviour is to make the comment "That tree is small but looks old. Up in the mountains I saw a whole area covered with trees like it". A third behaviour is to focus attention on the tree and “become aware” of the tree as a tree.

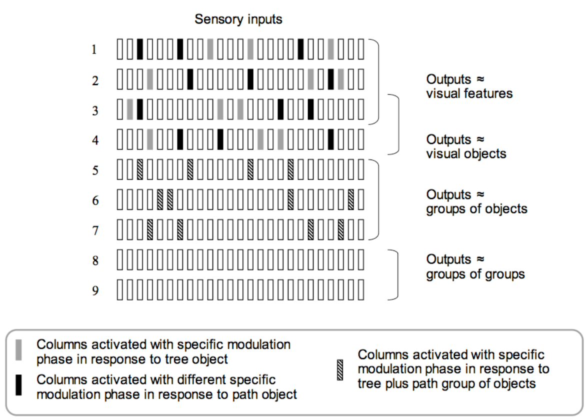

Figure 8. Pattern of column activation supporting motor behaviours avoiding a tree without “conscious” awareness.

To understand the physiological processes supporting this scenario, consider the pattern of activation illustrated in figure 8 for the same architecture as illustrated in figure 3. The figure illustrates the activation pattern when the tree is first perceived. There is an activated population of columns in arrays 1 to 4 in response to visual input from the tree, and another activated population in response to visual input from the path.

These two populations are modulated at the same frequency but different phases to maintain separation. The outputs from the two populations in array 4 are synchronized and released to array 5, and activated populations are generated as a result in arrays 6 and 7. The columns activated in arrays 6 and 7 have recommendation strengths in favour of appropriate walking behaviours. All these column activations are the result of detection of conditions directly within current sensory inputs. If there is little novelty in the tree and path objects, there may be little condition recording. Hence motor behaviours are generated, but there may be little future memory of the event.

Now suppose that indirect activation behaviours are encouraged. This will mean that the activated population will be larger (with higher biological cost), and may have overall motor recommendation strengths somewhat less appropriate for the walking motions required. However, these indirectly activated populations have recommendation strengths in favour of columns that would be directly activated by objects or groups of objects that are not currently present. Such objects or groups would be any that resulted in condition recording at some past time when the currently active columns were also active, such as the mountain area covered with small trees, and have recommendation strengths in favour of speaking about those objects. If there is one object or group of objects with a strong overlap (in terms of simultaneous past condition recording) with the current directly activated column population, such a verbal behaviour is quite probable.

The experience of “becoming aware” of the tree can be understood in a similar fashion. If sensory input is taken only from the tree, a population of columns is directly activated. Suppose now that recommendation strengths in favour of indirect activation, on the basis of both past correlated activity and past correlated condition recording, are encouraged. The effect is to generate a much larger population of activated columns. Different subset of this population will be subsets of the populations active during the very large number of past experiences in which trees were present. In general, none of these subsets will be large enough to generate speech behaviour appropriate to the corresponding past experience. The effect will therefore be a much richer mental experience which it is not possible to describe verbally, and which will be specific to the past experience of the individual.

Note that the behavioural value of such “conscious” activations is to search a much larger space of possible appropriate behaviours than is available from conditions actually present in current sensory inputs. The costs are higher biological effort, and the risk of interference with the most appropriate behaviour in response to the currently perceived objects. Hence selection of “conscious” activation behaviours will depend upon the value of such behaviours compared with simpler motor behaviours. Thus in the case of verbal behaviour, the social value of communication is greater that the (potentially slight) reduction in the effectiveness of motor behaviour. In the case of the awareness behaviour, the implication is that there is nothing in the current environment that requires an immediate response, and there is value in an almost random search for a possible behaviour in response to one particular object.

Machine Implementations of Required Information Mechanisms

The objective of the discussion has been to outline the architecture for a machine system which could implement conscious like phenomena in a physiologically realistic manner. There have been various electronic implementations of simplified versions of different subsets of the architecture which demonstrate that the various information mechanisms can be implemented. The use of very simple condition recording devices to heuristically define columns that divide up an input space in a way that discriminates between input states that have behaviourally different implications has been demonstrated [11]. The ability of such column arrays to learn with limited interference with prior learning has also been demonstrated [12]. The effectiveness of indirect activation mechanisms in supporting activation of pseudovisual images in response to verbal inputs, and to support activation of pseudovisual images of objects often present in the past at the same time as currently perceived objects has also been shown [13]. The capability of the frequency modulation mechanism to implement attention functions at the physiological level has also been demonstrated [9].

The electronic implementation of the key information mechanisms required by the architectural model of consciousness supports the feasibility of electronic implementation of the full model.

References

[1] Coward, L. A. and Sun, R. (2004). Some Criteria for an Effective Scientific Theory of Consciousness and Examples of Preliminary Attempts at Such a Theory. Consciousness and Cognition 13(2), 268 - 301.

[2] Triesman, A.M. (1960). Contextual cues in selective attention. Quarterly Journal of Experimental Psychology 12, 242–248.

[3] Block, N. (1995). On a confusion about a function of consciousness. *Brain and Behavioral Sciences* 18, 227 - 287.

[4] Coward, L. A. (2005). A System Architecture Approach to the Brain: from Neurons to Consciousness. New York: Nova Scientific Publishers.

[5] Baars, B. (1997). In the Theatre of Consciousness. *Journal of Consciousness Studies*, 4,4, 292 - 309.

[6] Sourdet, V. and Debanne, D. (1999). The role of dendritic filtering in associative long-term synaptic plasticity. Learning and Memory 6, 422 – 447.

[7] Hyvärinen, A., Karhunen, J. & Oja, E. (1999). Independent Component Analysis. New York: Wiley.

[8] Coward, L. A., and Gedeon, T. G. (2005). A model for representation of concepts in the brain. Proceedings of the International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning. Espoo, Finland.

[9] Coward, L. A. (2004). Simulation of a Proposed Binding Model. *Brain Inspired Cognitive Systems 2004*, L. S. Smith, A. Hussain and I. Aleksander, (editors), University of Stirling: Stirling.

[10] Coward, L. A. (2005). Accounting for episodic, semantic and procedural memory in the recommendation architecture cognitive model. Proceedings of the Ninth Neural Computation and Psychology Workshop: Modelling Language, Cognition, and Action.

[11] Coward, L.A. (2000). A Functional Architecture Approach to Neural Systems. International Journal of Systems Research and Information Systems, 9, 69 - 120.

[12] Coward, L. A., Gedeon, T. D. and Ratanayake, U. (2004). Managing Interference between Prior and Later learning. ICONIP 2004, Calcutta.

[13] Coward, L.A. (2001). The Recommendation Architecture: lessons from the design of large scale electronic systems for cognitive science. Journal of Cognitive Systems Research 2(2), 111-156.