BIOLOGICAL TEMPORAL SEQUENCE PROCESSING AND ITS APPLICATION IN ROBOT CONTROL

M J Denham and S L McCabe

University of Plymouth, UK

ABSTRACT

We describe an approach to the control of autonomous mobile robots which uses learnt sequences over time of pairs of sensory stimuli and motor actions, so-called sensory-action sequences (SAS), to build an internal model of the world in which the robot is navigating. The hypothesis presented is that the behaviour of an autonomous robot, as with that of an animal, in the learning and performance of a goal-directed, sensory-motor task is determined by the ability of the robot, or animal, to learn and recall spatio-temporal sequences of sensory experiences and associated motor actions In this paper, we propose a model of the underlying learning and recall process which is biologically-inspired, and is based upon a proposed model of the interaction between certain areas of the animal brain, in particular the prefrontal cortex and certain regions of the limbic system.

INTRODUCTION

We consider the problem of controlling an autonomous mobile robot which is able to move in a constrained way, eg in the presence of obstacles, in a dynamically changing world which contains a rich set of sensory cues. The robot is able to acquire these cues as it moves in the world and use them, together with internal knowledge of the actions it has taken, to build an internal model in memory of the world and to use this model to control its future movements. Crucial to this behaviour is the existence at all times of either a specific goal, such as that of re-experiencing a particular previously experienced and remembered sensory stimulus, or the need of the robot to explore its world in order to improve its knowledge of that world.

Our objective is to devise a control system for the robot which will enable it learn such a goal-related internal model of its world and to achieve a specific goal by using its current internal model to make a appropriate choice of future actions.

Fundamental to the solution of this problem are the hypotheses that: (i) the world is perceived not as isolated, individual sensory "snapshots" but as a sequence of sensory stimuli which are associated over time; and (ii) that every sensory sequence is determined by an associated sequence of robot actions, each action in the sequence determining the sensory stimuli which will be received next. We call such a sequence of pairs of associated sensory stimuli and motor actions, a sensory-action sequence (SAS).

There have been a number of attempts to make use of the important concept of temporal sequences of sensory-action pairs, sometimes referred to as perception-action sequences, in the control of mobile robots, eg Scholkopf & Mallot (1) , and approaches developed which are goal-directed, eg Tani & Fukumara (2). The approach presented here tries to combine these two fundamental concepts into a single approach, ie goal-directed learning of temporal sensory-action sequences.

THE CHUNKING PROBLEM

Clearly, the robot's experience of moving in its world will result in the creation of one long, continuously expanding SAS, within which many identical subsequences may be contained. The development of an internal model of the robot's world requires that this one long SAS is segmented, or "chunked", according to some intrinsically determined mechanism, and in such a way that the individual subsequences represent meaningful components of the world model. Associations must also be created between these "chunks" which allow their combination in flexible, dynamic ways, which are appropriate to the robot's current goals. In this sense, a "chunked" subsequence represents a particular piece of memory about some part of the world from which the robot can make predictions or expectations about the results of its actions, on a short term basis, ie whilst it remains within that particular part of the world. The links which have been created between such pieces of memory allow the robot to determine which new piece of memory is appropriate to use for further action selection, ie in a different part of the world, according to the current goal.

In our model the chunking and linking mechanisms are triggered by either the *detection of novelty* or the

UKACC International Conference on CONTROL '96, 2-5 September 1996, Conference Publication No. 427 © IEE 1996

achievement of a current specific goal, respectively. A goal, which is here defined as *the expectation of a particular sensory stimulus or sequence of stimuli*, is achieved when the internal sensory expectation matches, in some sense, the external sensory reality.

THE CONCEPTUAL MODEL

Figure 1 depicts the architecture of the proposed conceptual model. We propose that a robot learns sequences of associated sensory-action pairs, associated over time. The sequences are learnt by a set of one step ahead predictors each of which holds sufficient past history of the sensory-action behaviour of the organism to uniquely predict the next sensory-action pair in the learnt sequence. The current set of predictors compete with each other to determine which holds the closest matching history and therefore which should be used to make the prediction. Sequences are "chunked" based on novelty detection. Exploration is initially used to learn sequences based on randomly selected actions, resulting in short chunks of sensory-action knowledge. Learnt chunks are assembled under a top-down, goal-related influence to produce a given behaviour. Goal-directed learning then is used to form temporal associations between chunks which lead to a given goal; such sequences of chunks can subsequently be recalled by the organism in order to achieve that goal again. We propose that the presence of a goal influences the selection of those chunks which are used to create the behaviour leading to that goal, and that successful achievement of the goal will result in a strengthening of the associations between the chunks which lead to the goal and also between the associated chunks and the goal itself. In the presence of multiple goals and an additional exploratory drive, we propose that the goal-related motivational state of the organism exercises additional control over the learning/recall switching process, over and above that provided by novelty detection. Thus we postulate that strong motivation in relation to a particular goal will favour the recall of learnt chunks previously associated with that goal. On the other hand, weak motivation in the presence of multiple goals will permit the recall of any of a number of chunks associated with each goal, with no strong preference, resulting in the behaviour of being unable to "home in" on the achievement of any one particular goal.

THE BIOLOGICAL BASIS FOR TEMPORAL SEQUENCE LEARNING AND RECALL

The biological basis for the model is derived from a large number of studies over many years by many researchers of the role of various brain regions in controlling human and animal behaviour. Based on thses studies we propose the following model of the interactions between a number of key brain regions in determining behaviour in the presence of both uncertainty, resulting in the need for exploration, and the desire to achieve certain goals.

We propose that the prefrontal cortex plays a primary role in the biological model in providing goal and motivational state related control over the prediction and learning/recall process; that the hippocampal system plays a primary role in the prediction, novelty detection and action selection process; and that the medial septum plays a primary role in the novelty-related learning/recall switching process and, through its input from the prefrontal cortex, in the modulation of this process in relation to motivational state.

The prefrontal cortex receives inputs from both cortical and subcortical areas of the brain, as described in Fuster (3) (Figure 2). Outputs from the prefrontal cortex go to most of the areas from which it receives inputs. Our hypothesis is that two main sub-regions of the prefrontal cortex, the orbital area and the dorsolateral area, carry out two distinct but interrelated functions, each having an influence on the limbic system, in particular the hippocampus, in the formation and retrieval of sensory-action sequences. Orbital prefrontal cortex projects to septum, and therefore influences hippocampal activity via modulation of the action of septum on the hippocampus. The dorsolateral part of the prefrontal cortex is primarily connected to lateral thalamus, anterodorsal caudate, hippocampus and neocortex. In this case, however, it is connected to hippocampus via the cingulate gyrus and more directly via the entorhinal area of the parahippocampal gyrus.

Thus both parts of the prefrontal cortex influence hippocampal activity, but, we suggest, in quite different ways. We propose that the orbital part is primarily concerned with monitoring the motivational state and the association of reward with sensory input in the behaving animal through its connections with hypothalamus, amygdala and medial thalamus, and influences the hippocampus by providing a higher level control of the modulating effect of septum on hippocampal learning and recall. The dorsolateral part, we propose, is primarily concerned with the formation and storage of goals, and influences the learning and recall mechanisms of hippocampus by applying a "top-down priming" action directly through its projections to cingulate gyrus and entorhinal cortex. In this way it influences the selection of which chunks of learnt behaviour should be used to create the prediction of the next sensory-action event, this selection being determined primarily by previously learnt associations between specific chunks and the current goal. Thus the

dorsolateral area is primarily concerned with the goal-directed flexible assembly of appropriate chunks of sensory-action sequences.

The hippocampal formation is composed of several interconnected regions: the hippocampus, the dentate gyrus and the subiculum. Inputs to the hippocampal formation come primarily from sensory associational areas of neocortex and from the dorsolateral prefrontal cortex, via the cingulate gyrus. All these inputs reach the hippocampus itself via the entorhinal cortex, whereas the medial septum has direct inputs to the hippocampus. Outputs from the hippocampus proper go back to the deep layers of entorhinal cortex, which in turn project back to association cortex. There is also output from the hippocampus, via the subiculum, to the ventral striatum and hence to the motor areas.

The hippocampus (Figure 3) can be further divided into three main areas: CA1, CA3, and the dentate gyrus (DG). The main inputs to the hippocampus come from the entorhinal cortex (EC) and medial septum (MS). The outputs go primarily to the subiculum, the entorhinal cortex and the medial septum.

The inputs to dentate gyrus from EC densely innervate the outer 70% of the dendritic fields of DG granule cells, whereas the septal inputs project primarily to the infragranular layers of hilus, which contain collections of interneurons and polymorph cells. This implies that the EC inputs cause principal (granule) cells of DG to fire, whilst septal inputs determine whether the positive feedback system (granule -> polymorph -> granule) will be activated. In our model, we propose that the selective potentiation of the mossy fibre synapses from DG granule cells to the area CA3 is the means by which a prediction of the next pair in a sensory-action sequence is learnt, based on association between the pre-synaptic activation of DG granule cells and the post-synaptic activation of CA3 cells as a result of inputs from EC.

The high density and extensive ramification of CA3 cell projections, both within CA3 itself and to CA1 implies the existence of self-sustained, recurrent activity in CA3. In our model, we propose that this recurrent activity is the mechanism by which the predicted next pair in a sensory-action sequence is held until the action is taken and the actual resulting sensory input is experienced.

A possible role for CA1, and the one we propose in our model, is to carry out a matching/comparison of the input from CA3, ie the predicted next sensory experience, with the input to CA1 from EC, ie the actual sensory input resulting from taking the predicted next action. This is consistent with the observations made by Hasselmo (4) that the activity of CA1 cells appears to be dependent on the simultaneously activity of previously associated matching inputs on CA3 and EC afferents.

CA1 cells project to the medial septum and appear to directly influence the level of activity of acetylcholine (ACh) cells in MS, which is an important mechanism in controlling the learning/recall process in hippocampus (4). It has been observed experimentally that activity of ACh cells in CA1 causes a decrease in the activity of cells in medial septum (4). Thus a match between expected and predicted sensory input will cause an increase in the activation of CA1 cells and a resultant depression of ACh cells in MS. This will result in sustained activity of the recurrent connections in DG and CA3, and of the CA3 to CA1 projections. Thus the current context will be available for continued recall of components of the presently active, previously learnt sensory-action sequence.

On the other hand, detection in CA1 of a mismatch between predicted sensory input and actual sensory input will cause CA1 activity to remain low and ACh modulation from medial septum to be high. This will have the effect of causing a strong suppression of activity in the recurrent connections in DG, the recurrent connections in CA3 and in the CA3 to CA1 projections. Thus a mismatch will result in the suppression of historical, context-based activity in the DG recurrent connections, restricting further recall based in this context and permitting the subsequent establishment of a new context. Further, it will cause the currently held prediction on the output of CA3 to be suppressed, which is reasonable since the mismatch has indicated that the prediction is not valid. Finally, since it has also been observed that high levels of ACh enhance the potentiation of mossy fibre synapses (4), mismatch will cause enhanced learning of the new association between the current context in DG and the novel sensory input reaching CA3 via its afferents from EC.

In addition, the activity of CA1 will be projected back to EC, and hence to the neocortex, both directly and via the subiculum. Thus we would, in such circumstances, expect to see cell activity in EC, in the neocortex, and in the subiculum, related to anticipated motor actions. Furthermore, the activity in the subiculum, through its projections to the ventral striatum (or nucleus accumbens), could be used to control the motor activity in the animal, perhaps in respect of switching between the present motor action sequence and an exploratory behaviour when mismatch occurs, as has been suggested by Gray (5).

Thus, as we have described above, the dynamics of learning and recall in the hippocampus appear to be determined by feedback control from CA1 via the

medial septum back into hippocampal areas. We propose that the inputs to the septum from the orbital prefrontal cortex modulates this feedback control system in relation to the motivational state of the animal. Low motivation in respect of a goal or set of goals appears to cause a propensity to be easily distracted towards actions leading to novel experiences or new goals. Thus a low motivational input to the septum from orbital prefrontal cortex might be associated with a lowering of the tendency of heightened CA1 activity to supress ACh cell activity in medial septum. This would result in ACh modulation remaining high, tending to prevent the recall of previously learnt sensory-action sequences which would lead to the current goal. This hypothesis may also be consistent with Damasio's observation (6) that patients with bilateral orbital (ventromedial) prefrontal damage show impairments in their ability to maintain a current course of action and an excessive attraction to novelty.

IMPLEMENTATION OF THE PROPOSED CONTROL SYSTEM

For our initial computational experiments with such a temporal sequence based control system, we have used a modified version of the network introduced by Wang & Arbib (7), which we will describe here in outline only.

In this network, a set of units representing local neuronal populations are allocated to each of the expected sensory-action pairs. The ith unit receives an external input E_i of 1, as long as the sensory stimulus is present and the associated action is active, and 0 otherwise. The internal state of the ith unit, s_i , is determined by the external input and its one step time delayed value, so that a state change signals onset of the particular sensory-action pair. At onset, the excitation level xi of the unit is set to some high value, and remains at this value until another sensory-action pair is active. In this case the excitation level is reduced by inhibition from the unit associated with this pair. By this mechanism, each unit will have a level of excitation which is determined by the time history of the sensory-action pairs, which clearly encodes the sensory-action sequence.

This encoding is transfered into a set of short-term memory connection weights between the set of sensory-action pair units and a detector unit, which is allocated to the current sequence.This learning event only occurs when the network receives a learning signal, which corresponds to the learning trigger described in the previous section, and which corresponds to the detection of novelty. To detect novelty, the network must be able to create an expectation of the next sensory input to be received if a particular action is taken. To do this, the network uses its sequence retrieval capabilities, by which it is able to predict the next sensory-action pair in the current sequence. This is implemented by a set of reverse or feedback connections from the detector units to the sensory-action pair input units. When a sensory-action pair is active on an input unit, a competition amonst the detector units determines the most likely current SAS, ie effectively the part of the world the robot is most likely to be in, based on its current internal model. The backward projections from this unit then specify the most likely next sensory-action pair. The robot then takes the specified action and a comparison is made between the sensory expectation and that which is actually received. Mismatch is regarded as novelty, as described above.

This network is used to implement the first level of the robot's internal world model. Mechanisms are available in the network to handle complex SAS in which sensory-action pairs are repeated. The second level is implemented in an almost identical way. However, in this case the detector units become the input units for the second level, with the state of each unit determined by the onset of selection of the SAS encoded by that unit. Their excitation level at any time encodes the sequence of SAS which form the past history of robot behaviour. Second level detector units are then used to encode in their connections to the first level detector units a learnt sequence of SAS. Learning is triggered in this level however, by the achievement of a specific goal, and the second level detector unit created when the goal is achieved is associated with this goal.

In operation, the activity of the first level detector units is used to stimulate a competition between the second level detector units, as to which SAS should be used next to achieve the specified goal. Only those units which are associated with this particular goal take part in this competition, and thus form a kind of set of chunks of world knowledge which can be flexibly assembled through a competitive mechanism in order to derive the sequence of actions the robot should take in order to achieve its goal.

CONCLUSIONS

An approach to autonomous mobile robot control has been proposed which combines the two fundamental ideas of novelty-based learning and the use of temporal sequences of sensory-action pairs to build an internal model of the robot world. The approach also incorporates the concept that decision making in terms of action selection is carried out in response to a

competion between a set of chunks of sensory-action sequences, that are themselves activated in response to the existence at a given point in time of one or more goals.

The hippocampus has been proposed as playing a role in the learning and recall of temporal sequences by Taylor & Reiss (8) and generally in the learning of spatial and temporal mappings by Schmajuk (9), Grossberg & Schmajuk (10) and Grossberg (11). In particular in (9) it is proposed that the hippocampus acts to compute "aggregate predictions" of environmental events that are used to control associative learning, ie predictions of what event is going to occur, when in time and where in space. Such a function would be consistent with the concept of the hippocampus as facilitating the learning and recall of spatial and temporal sequences, as proposed in this paper, since prediction is essential to the detection of novelty, ie the mismatch between learnt expectations and actual experiences, and therefore supports novelty-related learning.

In this paper, we have used the insight gained from a wide range of experimental studies to propose a model of the hippocampal process and of the interactions between prefrontal cortex, sensory association cortex and the limbic system, which would form the basis of a system for the learning and recall of complex, goal-directed behaviour. We have proposed that the hippocampus is involved in the prediction of the upcoming motor action and resultant sensory experience. If we assume that this prediction is intimately involved in the selection and assembly of appropriate chunks in order to meet a current goal, then we would expect to find evidence of the predicted outcomes of a current event to be present in the recurrent connections from the hippocampal formation to the dorsolateral part of the prefrontal cortex, to the posterior parietal cortex and to the inferior temporal cortex. This is indeed the case. Anticipatory neurons have been detected in posterior parietal, inferior temporal and supplementary motor areas of neocortex. In (3) it is reported that one half or more of the neurons detectable by conventional microelectrode methods in dorsolateral and dorsomedial areas of monkey's prefrontal cortex show increased activity in the period of enforced delay between cue and response in delayed response trials in trained animals, but not in untrained animals.

In addition, we have postulated that, in accordance with recent findings in experiments on septal-hippocampal interplay in rats (4), the presence of novelty in a sequence of sensory experiences can control the recall and learning of sequences.

The proposed system is easily implementable in its basic form using a neural network for temporal sequence learning, recognition and reproduction, and simulations of a simple robot world have given initial indicatations of its validity as a control system. The simple two level memory model described has the potential to be developed to incorporate many levels of sub-goals, each sub-goal, as it emerges from the robot's behaviour and its environment, resulting in the dynamically self-organised creation of associations between goal-determined "chunks" of the robot's internal world model. This will permit the robot to work within a dynamically changing environment in an autonomous way, where one of the prime requirements is the ability to flexibly combine chunks of world knowledge in order to respond effectively to a changing world.

ACKNOWLEDGEMENTS

The authors acknowledge many stimulating discussions with colleagues in the Neurodynamics Research Group, in particular Professor Jack Boitano, Visiting Professor in the NRG during 1995, and with Professor John Taylor of Kings College London.

REFERENCES

1. Scholkopf B & Mallot H A,1994, "View-based cognitive mapping and path planning", Technical Report No. 7, Max-Planck-Institut fur biologische Kybernetik, Tubingen, Germany.
2. Tani J & Fukumura N, 1994, "Learning goal-directed sensory-based navigation of a mobile robot", Neural Networks, 7, 553-563.
3. Fuster J M, 1989, "The Prefrontal Cortex", Raven, New York.
4. Hasselmo M E, 1995, "A network model of hippocampus combining self-organization and associative memory function", Proc World Congress on Neural Networks (WCNN'95), Washington DC, July 1995.
5. Gray J A, "The contents of consciousness: a neuropsychological conjecture", Behavioural and Brain Sciences, 18, 659-722
6. Damasio A, 1994, "Decartes' Error", Putnam, New York.
7. Wang D & Arbib M A, 1993, "Timing and chunking in processing temporal order", IEEE Trans. Systems, Man, and Cybernetics, 23, 993-1009.
8. Taylor J G & Reiss M, 1992, "Does the hippocampus store temporal patterns?", Neural Networks World, 3-4, 365-384.
9. Schmajuk N A, 1990, "Role of hippocampus in temporal and spatial navigation: an adaptive neural

network", Behavioural Brain Research, 39, 205-229.
10. Grossberg S & Schmajuk N A, 1989, "Neural dynamics of adaptive timing and temporal discrimination during associative learning", Neural Networks, 2, 79-102.
11. Grossberg S (ed), 1987, "The Adaptive Brain, Volumes I and II", Amsterdam, Elsevier.

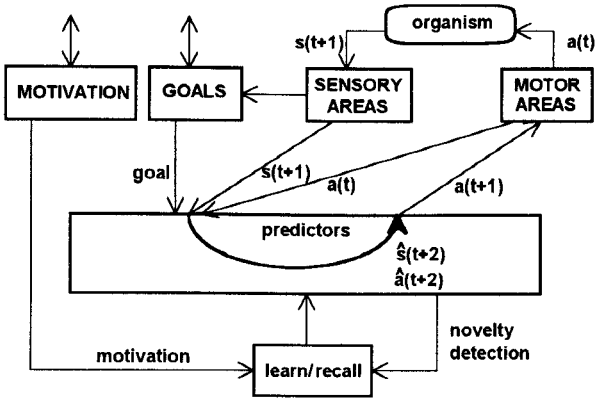

Figure 1. Conceptual model for the learning and recall of sensori-motor sequences.

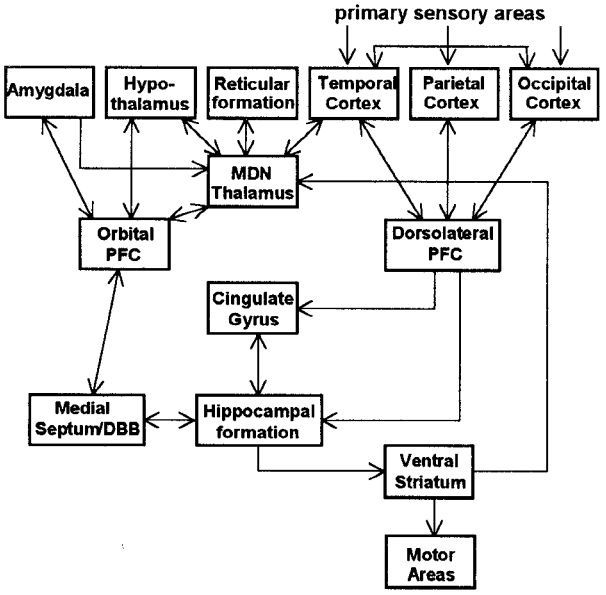

Figure 2. Connectivity of the orbital and dorsolateral areas of the prefrontal cortex (from Fuster, 1989).

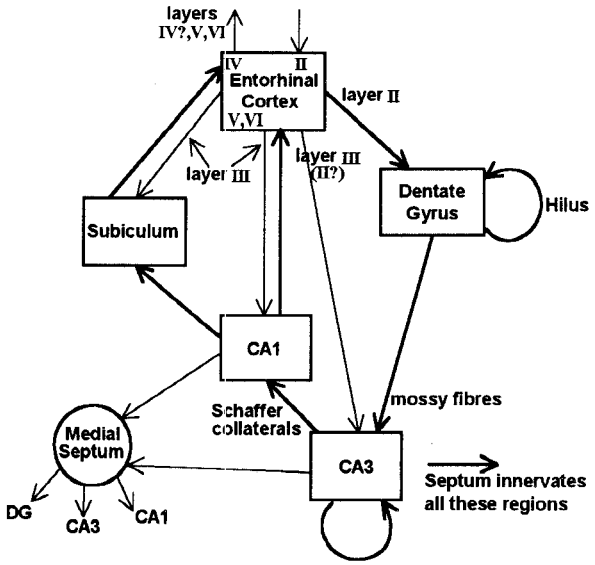

Figure 3. Hippocampal regions, showing septal projections.