Learning Spatial Navigation Using Chaotic Neural Network Model

Robert Kozma & Prashant Ankaraju
Division of Computer Science, 373 Dunn Hall
The University of Memphis, Memphis, TN 38152
Email: rkozma@memphis.edu

Abstract – **In this work the KIV model is used for the description of the interaction between the sensory and cortical systems, the hippocampus, the amygdala, and the septum. Neural activity patterns in KIV determine the emergence of global spatial encoding to implement the orientation function of a simulated animal. Our results embody the mechanisms, which we believe support the generation of cognitive maps in the hippocampus, based on the sensory input-based destabilization of cortical spatio-temporal patterns. We illustrate learning results using the example of simulated navigation in a 2D environment.**

I. INTRODUCTION

The KIII model is a working example of the implementation of chaotic principles in a computer software environment. KIII exhibits several of the experimentally observed behaviors of brains, like robust pattern recognition and classification of input stimuli, and fast transitions between brain states [1], [2], [3].

KIII consists of various sub-units; i.e., the KO, KI, and KII sets. The KO set is a basic processing unit and its dynamics is described by a second order ordinary differential equation feeding into an asymmetric sigmoid function. By coupling a number of excitatory and inhibitory KO sets, KI_e (excitatory) and KI_i (inhibitory) sets are formed. Interaction of interconnected KI_e and KI_i sets forms the KII unit. Example of KI sets is the dentate gyrus. Examples of KII sets are the olfactory bulb, and prepyriform cortex. In the hippocampus we have CA1, CA2, and CA3 as KII sets. By coupling KII sets with feed-forward and delayed feedback connections, one arrives at the KIII system. KIII shows excellent performance in learning new classes of training input data and it can generalize efficiently the classification of new test data.

The operation of the KIII model is described as follows. In the absence of stimuli the system is in a high dimensional state of spatially coherent basal activity, which is governed by an aperiodic, nonconvergent global attractor. In response to an external stimulus, the system activates a landscape of multiple attractors. It is kicked out of the basal state into a local basin of attraction, which is a memory wing. This wing is usually of much lower dimension than the basal state. It shows coherent and spatially patterned amplitude-modulated (AM) fluctuations. The system resides in the localized wing for the duration of the stimulus then it returns to the basal state. This is a temporal burst process that lasts for about a hundred milliseconds [4]. A memory pattern is defined therefore as a spatio-temporal process represented by the sequence of spatial AM patterns during a burst. KIII-based modeling of the olfactory system is used to classify linearly non-separable patterns. Its performance is compared with those of statistical classification methods and multi-layer feed-forward neural network-based classifications. KIII compares favorably with these methods regarding robustness and noise-tolerance of the pattern recognition, especially for classification of objects that are not linearly separable by any set of features [3].

The next highest level of the K sets is the KIV model. As in the case of all other K sets, the architecture and functionality of KIV is biologically motivated [5]. We extend multiple KIII sets into a KIV set that models the interactions in the cortical-hippocampal system. KIV is intended to have the functionality of planning and selection of action, in addition to classification and pattern recognition represented by single KIII units. KIV consists of three KIII sets, which model the cortical and hippocampal areas. All 3 are involved with learning and memory. The hippocampus is strongly involved in the cognitive processes of spatial and temporal orientation, like cognitive mapping and short-term memory [6], [7], [8].

In the KIII and KIV models several types of learning rules are used simultaneously, including habituation, Hebbian reinforcement learning, supervised learning, and global stability control through normalization. All these learning methods exist in a subtle balance and their relative importance changes at various stages of the memory process. Information is encoded in the KIII and KIV sets in the form of dynamical oscillations of spatially distributed activity patterns. It is hypothesized that the sequence of such activity patterns during the theta cycle belong to the encoding of spatial clues in the form of cognitive maps.

In this paper, we start with the description of the internal organization of the Hippocampal Formation (HF) model, its parts and interconnections. This is followed by the functional description of the HF and its interaction with the sensory cortex and limbic system. Next, details of learning processes in the HF and cortex model are given. We propose to demonstrate the operation of the system, using example of navigation conducted by the mobile agent EMMA in a simple 2-dimensional environment. Our results show that the EMMA learns certain aspects of the environment and uses it for goal-oriented navigation.

0-7803-7898-9/03/$17.00 ©2003 IEEE

II. MAP BUILDING AND NAVIGATION USING DYNAMICAL PRINCIPLES

A. Behavioral patterns

We use a simplified model of the internal motivational system with several internal senses. One is the battery level of the robot, modeling the hunger of animals. Another two are the state of the drive and the tuning motors. The other variable is the exploration/curiosity drive that is expressed in a state variable that promotes forward motion with a random turning component. We define two basic behavioral modes:

- wall following and
- object avoidance.

In the wall following mode, the robot tries to stay close to the wall, once it detected one. This is a more opportunistic strategy. Object avoidance is implemented by turning away from objects upon contact, to avoid collision in the near future. Object avoidance is an exploratory strategy. We combine the above two strategies with changing relative weights depending on the internal state. We use predominantly object avoidance when the robot has plenty of resources, and convert to more and more conservation, as the internal resources are becoming depleted.

In a very simple approach, we define a few basic behaviors and use those for the demonstration of the orientation/navigation capabilities. These can be considered as some reflexive behaviors that are hardwired in the motor system and actuators and will be able to solve the certain relatively simple tasks. Of course, there are the basic behaviors we mentioned above, i.e., wall following or object avoidance. Those will assure, that the robot is in constant movement without harming itself, while exploring the environment. Here we define one additional behavior: backup. Backup behavior is invoked if the robot is stuck or cannot execute a chosen action. A wide range of problems can be solved with these simple behaviors, as it is demonstrated in this study.

B. The role of theta gating

Learning and adaptation is a key component of our model, and it will be discussed in the next chapter. Here we discuss the time periodicity determined by the theta rhythm.

The HF and cortex complete their functions by sampling the environment at a theta rate. To achieve this periodicity, KIV relies on the septum to generate the theta frame rate as a gating function. Temporal framing is done in all sensory systems. Examples of this sampling are the saccadic movement in visual system, sniffing in olfaction, perhaps something similar in the cochlea etc. The present model is simplified by having a single gate generator for all environmental samplings, which is located in the septum.

The theta rhythm will be introduced in the numerical experiments by providing the various KIII units with sensory stimuli periodically, at rates corresponding to the theta frequency. We can simulate the theta sampling in computer experiments by designing a learning cycle as follows. Initial implementation of the model used here has been conducted in [9]. We show pattern A to the system for a duration, say, 100 ms, which corresponds to the drive period in the animal experiments. This is followed by a period of 100 ms without input pattern, corresponding to a resting part of the cycle. Afterward, a new pattern is shown, etc. This will generate a period of 5 Hz to approximate theta cycle. We will have this 5 Hz oscillatory behavior through the simulations. Various learning algorithms take place continuously or during the drive period, as it is described in the next section.

III. RESULTS WITH REINFORCEMENT LEARNING

A. Description of the experiments

Incremental habituation and reinforcement learning are initiated by reinforcement signal during the theta gating. As the testbed, we use a simple 2D environment with several obstacles. In this environment, the movement can take place along a grid, as the one shown in Fig.1. Consequently, at any instance, the robot can chose the next move from one of the 8 direct neighbors of the given grid point.

The orientation signals are the distances and directions with respect to the landmarks, measured from the actual location of the robot Sensory signals to the cortex can be the 6 short-range infrared signals as used in the case of Khepera robot [10]. For the sensory signals, we consider the past several time steps as inputs, in addition to the present time frame.

Both habituation and Hebbian learning take place during the 100 ms window defined by the theta rhythm. Habituation is implemented as a continual degradation of the sensory channels as they process sensory information. The degradation level is proportional to the signal magnitude. The proportionality constant is chosen in such a way that habituation diminishes the weights within several theta period. Habituation should not be overly dominant and should not prevent Hebbian learning through reinforced channels at the same time period.

Both habituation and reinforcement learning are calculated based on the root mean square (RMS) intensity of the gamma-filtered signals at each node, using filter band of 20Hz to 60Hz. Both habituation and Hebbian learning constants are experimentally tuned to have optimum learning performance in the cortical and hippocampal KIII sets.

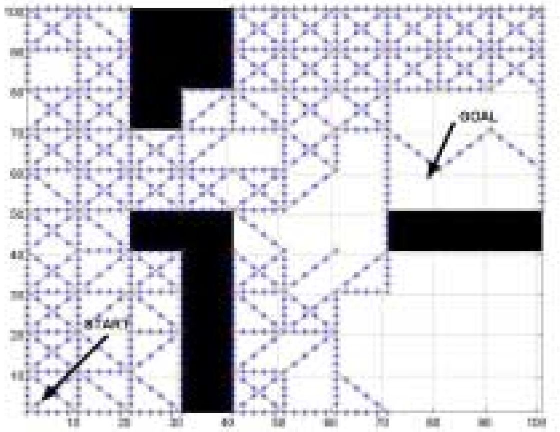

Figure 1: Random exploration of the environment without reinforcement learning. The ‘Start’ and ‘Goal’ locations are indicated by arrows.

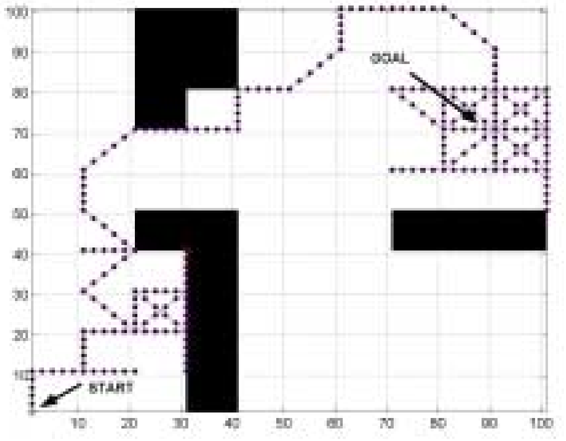

Figure 2: The path of the agent when reinforcement learning takes place. Note the significantly reduced path length after learning, c.f. Fig. 1.

B. Dynamical map generation

At first, the only landmark the animal is given is the ‘Start’ beacon, which is set by the human controller. In an explorative mode, the home acts as a repeller with a monotonic gradient field centered at the home, that drives the animal away from home. At each step, the animal’s action is determined by its present position and the location of home. In other words, it goes straight away from home, until it meets an obstacle. Constrained by the obstacle, it continues its path along the steepest possible gradient. Soon or later it will not be able to move further, it stuck. That is a conflict, which generates a reinforcement signal to learn. Reinforcement learning takes place both in the cortex OB/PC (“What?”) and in the hippocampus CA3/CA1 (“Where?”).

The above learning mechanisms are complemented with the following algorithm to form additional landmarks based on the experience during exploration. When the animal is stuck, the controller is notified about this event and its location. As a result, a new landmark is generated and its position is added to the existing ones. From now on, the animal gets orientation signals from all the beacons, including this new one. The sign of this one will be negative, which will be taken into account as a vector sum of field intensities at each time step when the decision about the next action (step direction) is made. As this new beacon will be repeller, it is less likely that the animal gets close to it another time.

Based on this approach, each learning experience establishes a positive or negative orientation beacon in the environment, or more concretely, at the location of the event in the map of the controller, and also in the cognitive map being constructed and maintained in the HF, as the roving device explores in search of positive goals and avoiding negative sites.

At each episode of being stuck, as behavioral response, we use the ‘back up’ motion. We use back when the stuck state is detected and the reinforcement learning has been initiated. The back up action simply can be a step in a direction the animal came from. This information is available, as the sensory signal at each time instant includes the present and a few (8) recent time frames as well. An example of such exploration is shown in Fig. 1. It took the system about 250 steps to get from ‘Start’ [0, 0] to the ‘Goal’ [80, 60].

It should be noted that the direction of the next step is selected based on the above algorithm, complemented with a small random noise component. We used this additive noise to simulate real life uncertainties, and also to avoid the system to get stuck deterministically in certain repeating situations. In the present experiments, additive noise has been selected at the level of 3%. This means that in 3% of the cases the system select a direction that has not been determined as the optimal one based on the given learning level.

Once the exploration phase has been conducted extensively, we can switch the home beacon to ‘attract’ mode. In order to test the system’s performance, we re-start it from home and give a goal location to go to. If the robot is properly learned the environment, it will navigate efficiently and find a reasonably optimal path to the goal based on the combined use of the internally formed cognitive map, using only the home beacon and its classification landscape learned in the cortical areas. This is illustrated in Fig. 2. After learning, the length of the trajectory from ‘Home’ to ‘Goal’ is reduced to about 40 steps.

IV. DISCUSSION OF THE RESULTS

Results of the previous section clearly demonstrate that our leaning algorithm produces significant learning gains, which are converted into improved navigation through the environment. Now we evaluate in details the nature of the observed gains.

We have conducted 50 independent sessions of numerical experiments of navigation through the environment with and without learning. The experiments were conducted until the robot reached the target. However, we have terminated the experiment after 300 steps, for practical reasons. If a session lasted very long, it meant that the robot has got stuck at a particular location, and it took very long time to escape, or it could not get away at all from such a ‘trap.’ Getting trapped in a corner in spite of the learning advances could not be excluded. Even the probabilistic component of the decision making at each step could not completely remedy such potential problems. In order to avoid a bias caused by the deformation of the distribution function of the paths in the experiments, we have excluded paths with length of 300 or more from the present analysis.

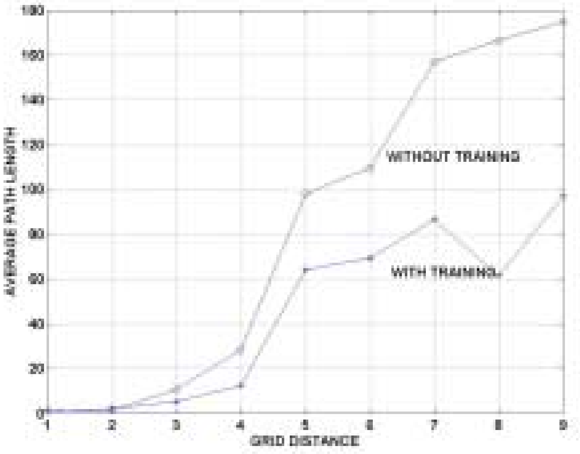

Figure 3: Comparison of the average traveled path without learning and with reinforcement learning as the function of the distance between the goal and the starting position.

In Figure 3, the average length of the path traveled by the robot is shown, as the function of the distance between the start and goal locations. For simplicity, the start has been always at position [0, 0], while the goal has been moved across the diagonal of the map, up to position [100, 100]. For small distances the learning gain is relatively small. One can see that the learning caused a reduction of the travel by about half. Clearly, this result is far from perfect. Still, it indicates the potential of our method.

V. CONCLUSIONS

We have introduced a novel method of learning spatial maps using hippocampal model, as part of the KIV set. We have demonstrated the feasibility of the methodology, and showed that K models are promising dynamic chaos neural networks to address navigation tasks. Previously, the KIII models have shown robust performance as classification and pattern recognition devices. With this new advancement, we have expanded the potential application areas of the K sets from the classification task to a more complex decision making and behavioral generation domains.

ACKNOWLEDGMENT

This research is supported by NASA grant NCC-2-1244 and NSF grant EIA-0130352.

REFERENCES

[1] Chang H.J. & Freeman W.J. (1996) Parameter optimization in models of the olfactory system, *Neural Networks*, Vol. 9, pp. 1-14.

[2] Freeman, W.J. (2000) *Neurodynamics. An exploration of mesoscopic brain dynamics.* London UK. Springer Verlag.

[3] R. Kozma, W.J. Freeman, “Chaotic Resonance – Methods and Applications of Noisy and Variable Patterns,” *Int. J. Bifurcation & Chaos*, 11(6), 2307-2322, 2001.

[4] Barrie J.M., Freeman W.J., Lenhart M.D. (1996) Spatiotemporal analysis of prepyriform, visual, auditory, and somesthetic surface EEGs in trained rabbits, *J. of Neurophysiology*, 76: 520-539.

[5] R. Kozma, W.J. Freeman, P. Erdi , “The KIV Model – Nonlinear spatio-temporal dynamics of the primordaial vertebrate forebrain,” *Neurocomputing*, 2003, in press.

[6] Arleo, A. and Gerstner, W. (2000) Spatial cognition and neuro-mimetic navigation: A model of hippocampal place cell activity. Biological Cybernetics, 83: 287-299.

[7] Bliss, T.V.P., and Lomo, T. (1973) Long-lasting potentiation of synaptic transmission in the denate area of the anaesthetized rabbit following simulation of perforant path. *J. Physiol.* 232: 331-356.

[8] Burgess, N., Recce, M., and O’Keefe, J. (1994) A model of hippocampal function. *Neural Networks*, 7 (6/7): 1065-1081.

[9] Ankaraju, P. “The Hierarchy of K sets – From Pattern Recognition to Navigation,” *Masters’ Thesis*, The University of Memphis, 2002.

[10] Harter, D., Kozma, R., Graesser, A. (2001) Models of Ontogenetic Development for Autonomous Adaptive Systems. *Proc. 23rd Annual Conference of the Cognitive Science Society,* CogSci’02, pp. 405-410.