

Modeling Higher Level Processing Functions Inherent to the Human Brain

Iren Valova
Department of Precision Machinery Systems
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku,
Yokohama 226-8502, JAPAN
iren@pms.titech.ac.jp

Yukio Kosugi
Frontier Collaborative Research Center
Tokyo Institute of Technology
4259 Nagatsuta, Midori-ku
Yokohama 226-8503
kosugi@pms.titech.ac.jp

Abstract

The most significant feature of the information processing in the brain might be the autonomy based on the motivation and self reward (MSR) to form a processor sequence intending to find out the solution to a problem we are facing. In this paper, we show some preliminary ideas to incorporate the concept of MSR in designing brain-like information processing means, based on physiological and engineering points of view. We propose a hybrid neural network model as an extension of Hebb's rule, to be hypothesized for the function of association areas in the cerebral cortex. The generated neural network model is tested on the problem of segmentation of brain magnetic resonance images (MRI).

formation science is to formulate brain-like information processing paradigms, which are applicable to real world problems. There are many unknown processing mechanisms within the brain, especially those related to the limbic emotional system and higher order processing systems in the neo-cortex. We try to develop a new paradigm, based primarily on neural network models to realize a practical means of taking into account firstly correspondences to physiological findings, and secondly their feasibility in real-world problems.

Unlike the Von Neuman type computer, cortical processors in our brain have to be chained together to form a meaningful process sequence by themselves. This process of learning involves predictive, memorizing and self-rewarding mechanisms. Such a claim is supported by the physiological evidence in investigating the effects of memory, the predictive rewarding function of dopamine neurons, and the effects of lesions in the human brain. From physiological point of view, the reward is given as a part

Introduction

One of the major challenges to the contemporary in-

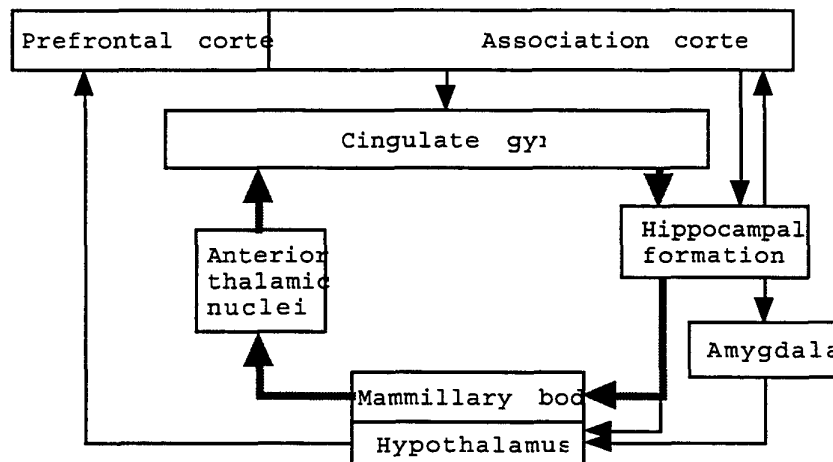


Figure 1 Proposed neural circuit for emotion: J. Papez, 1937 (thick lines) enhanced by P. MacLean (thin lines)

of the learning process. As the task is memorized, mechanisms of internal prediction and self-reward will perform whenever there is a coincidence between behavioral reaction and external stimuli. In a psychological interpretation, we may give some self-reward to those processes which made us feel comfortable. At the same time, from the biological point of view, the self-rewarding criteria have to be very simple to be easily realized in biochemical form.

The circuit in Fig.1 is a simplified diagram of the limbic system, which includes sensory inputs, processing mechanism for the input information, self-rewarding mechanism based on psychological evaluation, and output for controlling human actions. It shows the extended emotion circuit, designed by MacLean and based on the earlier studies of Papez [1]. There are extensive and direct connections between the neocortical areas and the hippocampal formation and the amygdala. Some of the most compelling biological evidence of higher level functions has come from studies of the association areas of the cerebral cortex. In this study the main interest lays with the limbic association cortex and the related limbic system. The limbic area of the cortex is responsible for memory and emotional and motivational aspects of behavior [2]. Biological evidence shows that the amygdala is implicated in the process of learning (particularly tasks that require coordination of information from different sensory modalities) and social judgment as part of the emotion-analyzing mechanism [3, 4].

Neural network model

Our interpretation of the introduced biological neural system is shown in Fig.2. The basic idea is to model the control function of the limbic system and the related association cortex with the "seek a near-optimal solution

control", which is responsible for choosing the correct processor or sequence of processors for the solution of a given problem. The decision of this unit is based on the output of the "inference and induction unit". The latter is bound to consider preliminary knowledge on the topic and perform computational rules in order to "infer" the appropriate information for the "control" unit. Upon achieving some solution, the result is forwarded to the "evaluation of found solution" unit. The self-rewarding mechanisms will provide the feedback for the "inference" unit and therefore "close" the learning cycle.

We base our artificial neural network model on modified predictive Hebbian learning rules. In other words, we incorporate the idea that the changes in synaptic strengths represent the associations between inputs. However, the Hebb rule alone will not permit predictive relationships formed in conditional tasking [5]. It is known that sensory stimuli through learning can come to act as predictors of reward or punishment. In our model, we view the prediction as a computational goal of the system.

Hebbian learning rules are correlational in the sense that the changes in synaptic strengths represents the association between inputs. There are other relationships that are important to learn about events, such as the temporal order of the inputs. Conditioning experiments have shown that through learning, sensory stimuli can come to act as predictors of reward, punishment, and other salient stimuli. Prediction can be viewed as a computational goal of a system that must operate in an uncertain and variable environment.

The neural network simulating the general model is shown in Fig.3. It consists of feature extractor, marked AIW (Answer-in-Weights neural network), three subnets, and winner-take-all layer. Each subnet is trained to recognize one class only.

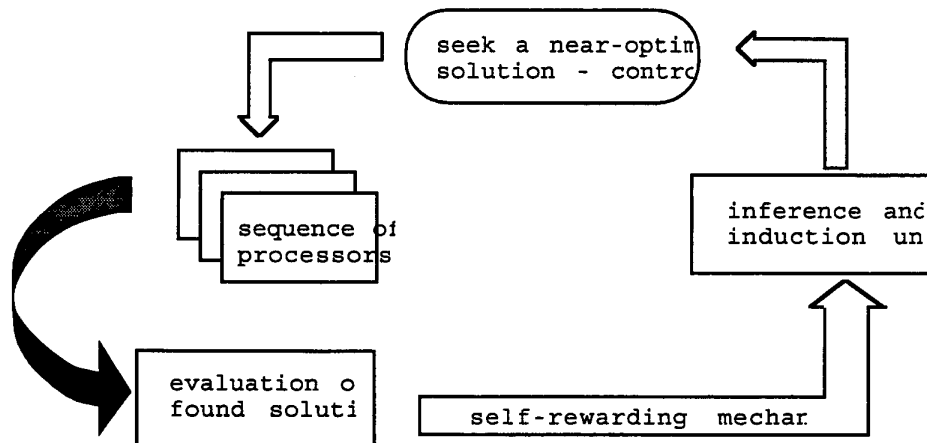


Figure 2 Higher - level processing model based on self-reward

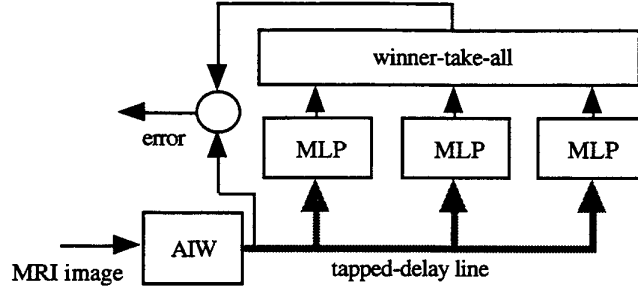


Figure 3 Proposed neural network system

The unit denoted AIW decomposes the image into Hadamard basis functions and provides the weight coefficients as input to the subnets. Details about AIW can be found in [6].

We conjecture that this type of architecture is very flexible in terms of:

- a) choosing the subnet structure and activation function. The latter can be linear, Gaussian or elliptic;
- b) choosing discriminant function of the subnet;
- c) choosing the type of subnets included in the structure.

The subnets in this case are multilayered perceptrons (MLP) with sigmoid activation functions, trained with backpropagation rule and the discriminant function of the subnets is

$$\phi(x, w_i) = \sum_{k_i=1}^{K_i} c_{k_i} \psi_i(x, w_{k_i}),$$

where c_{k_i} are the coefficients in the upper layer and w_i is the weight vector. The idea is to present the inputs in a tapped-delay line and each input vector to produce a prediction for its immediate future sample. During the training phase the object is to minimize the squared prediction error of the training patterns x from one class:

$$E = \sum_{m=0}^2 \sum_{j=0}^{62} \left[\left(\frac{1}{1 + e^{-x_j}} \right) - x_{j+1} \right]^2$$

During the test phase a pattern is input into all subnets. The subnet with the smallest prediction error is the assigned class for the presented pattern. The inputs are 63 dimensional vectors, given serially to the subnets via AIW.

Simulations and discussion

For the simulation we have chosen the problem of segmentation of brain magnetic-resonance images (MRI) into white and gray matter and cerebro-spinal fluid (CSF). The contemporary medical imaging technology provides complementary diagnostic tools such as x-ray computed

tomography, magnetic resonance imaging or positron emission tomography. In routine works, these images are interpreted visually and qualitatively by radiologists. However, the need for quantitative information, i.e. precise detection of structure boundaries or characterization of tissues is becoming increasingly important in the clinical environment. The routine quantitative analysis of these data is infeasible by manual methods, therefore the full potential of modern medical technology will require at least partial automation of the analysis process.

There are many difficulties which prevent the full automation of the MR images analysis, namely: intensity variations due to the choice of time parameters, radio frequency coil, imager imposed inter-patient variations, interslice intensity variations, and problems with assessing the accuracy of the methods due to lack of gold standard.

The multidimensional nature of MR image data is a characteristic to be considered, as it provides information about three tissue dependent parameters: the spin-lattice relaxation time, the spin-spin relaxation time, and the proton density. Moreover, recent developments of techniques that permit the registration of several imaging modalities, such as X-ray CT, PET, and MRI increases the dimensionality of the data thus strengthening the potential of multidimensional classification techniques.

Possible areas of application include the automatic or semiautomatic delineation of areas to be treated prior to radio surgery, the delineation of tumors before and after surgical or radio surgical intervention for response assessment, or the volumetric measurement of the white and gray matter and of (CSF) for the study of the degenerative diseases.

For our simulations we have adopted a T₂-weighted MRI slice. The image is 256x256 pixels, 256 gray levels within 8 bits/pixel presentations. The image to be segmented is divided into blocks of 8x8 pixels; each block is decomposed and the respective coefficients from each pixel presentation are taken as features for this pixel. We conjecture that this type of input vectors provides "normalized", intensity-independent input information and, therefore, unbiased pixel feature vectors, which is essential when seg-

menting MRI. After this preprocessing stage, each pixel's feature vector is fed to all three MLPs in the proposed hybrid architecture (Fig.3). Each MLP is trained to recognize only one class, i.e. gray matter, white matter, or CSF. There are 63 inputs to each MLP. Although the number of coefficients after AIW is 64, the largest-value coefficient is discarded from the feature vector.

The neural network was trained on randomly chosen pixels from the shown slice (Fig.4a) and tested on the whole image. The resulting segmentation is pictured in Fig.4b, where the black portion denotes CSF, the gray is the gray matter and the white zones refer to the white matter. The overall performance for the volume averaged 90% correct segmentation. Though not excellent, this result supports our conjectures about the learning and generalization properties of our model for modeling higher-level functions in the brain.

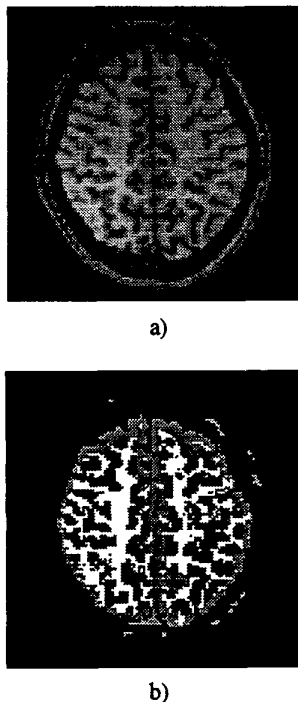


Figure 4 Segmentation simulation: a) original; b) segmented MRI

Conclusion

In this paper we have formulated our preliminary concerns and experimental results in modeling the motivation and self-reward mechanisms when designing brain-like information processing means. The proposed hybrid neural network represents the controlling functions, inherent to the limbic system and associative cortex, which we have chosen as a model. As future research goals, we aim to incorporate more pronounced predictive and reward mechanisms based on matching (or coincidence) between the sensory information and the system response. On a more engineering level, the MLP has many defects, which prevents this type of network to serve as true model of brain functions. Hence, we view the processing units as a hybrid between supervised and unsupervised modes of learning.

The proposed neural network model supports our general concept and theories and the demonstrated results are satisfactory in terms of generalization and learning abilities. The model can be expanded and new features added in accordance with the biophysiological evidence.

References

- [1] Papez J., A proposed mechanism of emotion, Arch. Neurol. Psychiatry, vol.38, pp 725-743, 1937
- [2] Kupfermann I., Hypothalamus and Limbic System: Peptidergic neurons, homeostasis and emotional behavior, Chapter 47 in Principles of Neural Science, (eds. Kandel E., Schwartz J., Jessell T.), Appleton&Lange, 1993
- [3] Whalen P., Rauch S., Etcoff N. at all, Masked presentation of emotional facial expressions modulate amygdala activity without explicit knowledge, The Journal of Neuroscience, vol.18, No.1, pp 411-418, 1998
- [4] Adolphs R., Tranel D., Damasio A., The human amygdala in social judgment, Nature, vol.393, No.4, pp 470-474, 1998
- [5] Montague P., Sejnowski T., The predictive brain: temporal coincidence and temporal order in synaptic learning mechanisms, Learning&Memory 1, pp 1-33, 1994
- [6] Valova I., Kosugi Y., Image Decomposition by Answer-in-Weights Neural Network, IEICE Transactions on Information and Systems, vol.E78-D, No 9, pp 1221-1224, 1995