Dynamic Knowledge Representation in Connectionist Systems

J.M.Bishop, S.J.Nasuto and K. De Meyer

Department of Cybernetics, University of Reading, Whiteknights,
PO Box 225, Reading, RG6 6AY, United Kingdom
J.M.Bishop@Reading.ac.uk

Abstract. One of the most pervading concepts underlying computational models of information processing in the brain is linear input integration of rate coded uni-variate information by neurons. After a suitable learning process this results in neuronal structures that statically represent knowledge as a vector of real valued synaptic weights. Although this general framework has contributed to the many successes of connectionism, in this paper we argue that for all but the most basic of cognitive processes, a more complex, multi-variate dynamic neural coding mechanism is required - knowledge should not be spacially bound to a particular neuron or group of neurons. We conclude the paper with discussion of a simple experiment that illustrates dynamic knowledge representation in a spiking neuron connectionist system.

1 Introduction

"Nothing seems more certain to me than that people someday will come to the definite opinion that there is no copy in the ... nervous system which corresponds to a particular thought, or a particular idea, or memory", (Wittgenstein, 1948).

Over the hundred years since the publication of James' Psychology [8], neuroscientists have attempted to define the fundamental features of the brain and its information processing capabilities in terms of mean firing rates at points in the brain cortex (neurons) and computations. After Hubel and Wiesel [7], the function of the neuron as a specialized feature detector was treated as established doctrine. From this followed the functional specialization paradigm, mapping different areas of the brain to specific cognitive function, reincarnating an era of modern phrenology.

Connectionism mapped well onto the above assumptions. Its emergence is based on the belief that neurons can be treated as simple computational devices. The initial boolean McCulloch-Pitts model neuron [10] was quickly extended to allow for analogue computations. Further, the assumption that information is encoded in the mean firing rate of neurons was a central premise of all the sciences related to brain modelling.

Over the last half century such 'classical' connectionist networks have attracted significant interest. They are routinely applied to engineering problems

[19], and as metaphors of concepts drawn from neuroscience, have also been offered as models of both high [16] and low level [2] [21] cognition. However the classical connectionist models of high level cognition have also been strongly criticized [5] and the situation at the domain of low level neural modelling is little better [1].

More recently spiking neuron, pulsed neural networks have begun to attract attention [9]. Such models no longer aggregate individual action potentials as a mean firing rate but act on temporal sequences of spikes. Like the classical connectionist models spiking neuron neural networks have been studied both for their computational/engineering properties and as models of neurological processes [20].

Although temporal coding of spike trains lends itself more readily to multivariate information encoding than rate encoding, both are typically discussed in a uni-variate framework - an observation that also applies to classicial connectionist frameworks. However uni-variate knowledge representation is limited to the representation of arity zero predicates and, following Dinsmore [5] and Fodor [6], we suggest that this is too strong a restriction for representing the complexity of the real world.

2 Types, Tokens and Arity Zero Predicates

An arity zero predicate is one without an argument, eg. representation of the class of Morgan cars by the arity zero predicate, MORGAN (). Such predicates easily express different 'types' of entities. eg. A car producer may produce half a dozen 'types' of cars in a year, (here 'types' equates to the different models marketed such as the Morgan 4/4 car), but manufacture many thousand individual cars for sale ('tokens'). Knowledge of individual 'tokens', in this case individual cars, is more clumsily expressed in a predicate of arity zero. eg. To represent a particular Morgan car (eg. registration YUY405W), the arity zero predicate MORGANYUY405W () is necessary.

Although a conventional connectionist network can represent 'type knowledge, of the form 'Morgan 4/4', by the activation of a single processing node or group of nodes, because it processes uni-variate information it can only easily instantiate tokens in a similar manner (eg. by an activation on a particular node or group of nodes). A more elegant method of representing the specific member of a class (a token) is by the use of the arity one predicate CLASS (INDIVIDUAL). However, in general this requires the use of bi-variate information to identify both CLASS and INDIVIDUAL.

However, we do not consider representations of arity zero predicates as sufficient for representation of many complex relationships. Such limitations make it difficult to interpret and analyze the network in terms of causal relationships. In particular, (cf. classical symbolic/connectionist divide), it is difficult to imagine how such a system could develop symbolic representations and quantified logical inference [17]. Such deficiencies in the representation of complex knowledge by classical neural networks have long been recognized [18] [6] [3] [15].

3 Spiking Neurons

Taking into account the above considerations we propose to investigate a spiking neuron connectionist architecture whose constituent neurons inherently operate on rich (bi-variate) information encoded in spike trains, rather than as a simple mean firing rate. NESTER, a network of such neurons, was first proposed in [13] and is further investigated herein. The task of NESTER is to locate an object (memory) projected onto an artificial retina.

The NEural STochastic diffusion search nEtwoRk (NESTER) consists of an artificial retina, a layer of fully connected matching neurons and retinotopically organized memory neurons. The bi-variate information output from retina/memory cells is encoded as a spike train consisting of two qualitatively different parts: a tag determined by its relative position on the retina/memory and a tag encoding the feature signalled by the cell. This information is processed by the matching neurons which act as spatiotemporal coincidence detectors.

It is important to note that matching neurons obtain input from both retina and memory and thus their operation is influenced by both bottom-up and top-down information. As Mumford notices [11], systems which depend on interaction between feedforward and feedback loops are quite distinct from models based on Marr's feedforward theory of vision.

Thus matching neurons are fully connected to both retina and memory neurons and accept for processing new information, contingent on their internal state (defined by the previously accepted spike train).

Each matching neuron maintains an internal representation (a hypothesis) defining a potential location of the memory on the retina and in operation simply conjoins the positional tags of the incoming spike trains from the retina/memory, (corresponding to their retinotropic positions), with its own hypothesis and, dependent upon the result, distributes its successful or unsuccessful hypothesis to other matching neurons.

Effectively NESTER is a connectionist implementation of Stochastic Diffusion Search, (SDS) [4], a simple matching algorithm whose operation depends on co-operation and competition in a population of agents which are realised in NESTER as the matching neurons. Therefore, in the next section we will describe the network operation in terms of the simpler underlying generic mechanism of SDS.

4 Stochastic Diffusion Search

In SDS a group of independent agents processes information from the search space in order to find the best-fit to a specified target pattern. Each agent searches for a micro-feature of the target and once found competes to attract other agents to evaluate this position. In this way the SDS explores the whole search space. Due to the emergent co-operation of agents pointing to the same solution, interesting areas in the search space (those that share many micro-features with the target) are more thoroughly exploited than background areas.

Agents are divided into two classes: active and inactive. An active agent has successfully found a micro-feature from the target in the search space; an inactive agent has not. Thus, the activity label identifies agents more likely to point to an instantiation of the target than to the background. Inactive agents utilise this activity information when deciding whether to communicate with a randomly selected agent in a subsequent phase of processing. Communication only occurs if the selected agent is active and results in the flow of 'where' information from the active agent to the inactive one. Conversely, if the selected agent is also inactive, then there is no information flow between agents; instead, a new random 'where' position is adopted. In this way active agents attract more resources to examining promising regions of the search space.

5 Experiment using NESTER

NESTER was configured with 100 matching neurons, 6 memory neurons and 90 retina neurons. The content of the target memory is defined by 6 symbols from the ASCII character set and the retina by 90 symbols.

The following experiment illustrates NESTER utilizing a dynamic assembly encoding as it locates the best fit of the target memory on the retina. Finding this on the retina causes the onset of time locked activity in an assembly of matching neurons, resulting in a characteristic frequency spectrum of their spike trains.

In this experiment two patterns with 5/6 correct symbols (ie. 16.6% noise) were projected at different locations onto the retina. Matching neurons first located one pattern and formed its representation - a dynamic assembly of neurons with time locked activity. Figure 1 shows matching neuron activity against time while locked to the first pattern. The activity is periodic, as indicated by the dominant frequency in the power spectrum shown alongside. After some time a

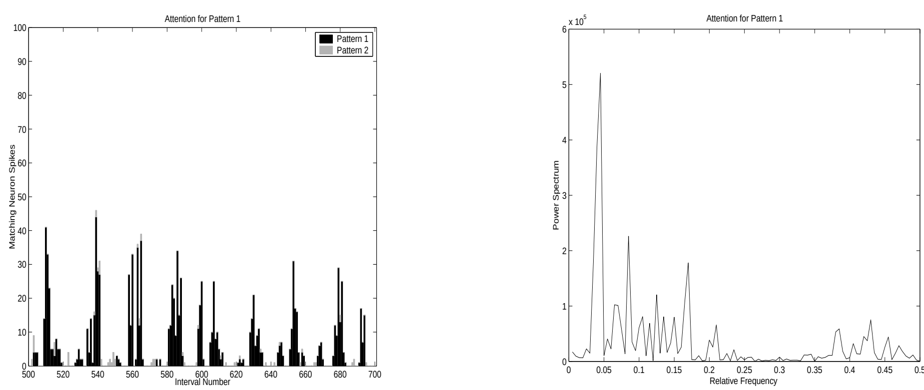

Fig. 1. Network activity and its frequency spectrum while locked to the first pattern

second assembly of matching neurons emerges reflecting the presence of the second pattern at a different location on the retina. The activity of a newly formed assembly corresponding to the second pattern and the resulting power spectrum are shown in Figure 2. It is clear that this spectrum is very different from that shown in Figure 1 although the patterns constitute equal instantiations of the memory.

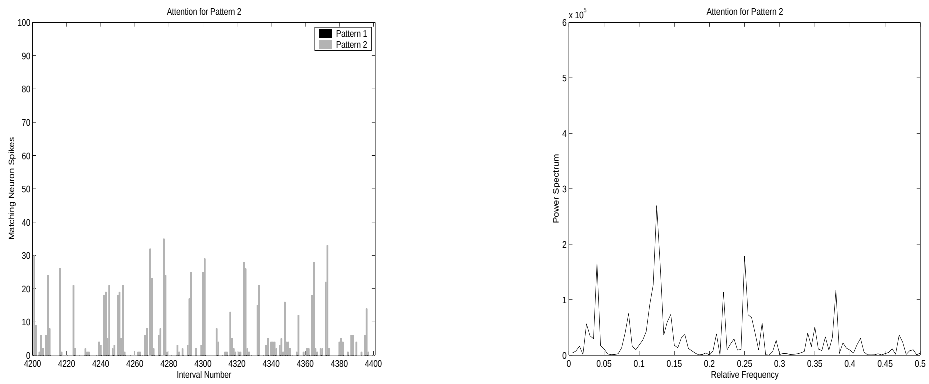

Fig. 2. Network activity and its frequency spectrum while locked to the second pattern

6 Conclusions

NESTER uses dynamic assembly knowledge encoding. The experiment illustrates that locating the target memory on the retina results in the onset of time locked activity in an assembly of matching neurons, as indicated by the characteristic frequency spectrum of matching neuron spiking. Dependent on the quality of the target instantiation (i.e. how many micro-features it has in common with the target), matching neurons will spend different periods of time maintaining particular hypotheses (retinal locations). On average, those matching neurons maintaining the best location hypothesis will spend a longer period examining the same retinal location than matching neurons with a poor location hypothesis (i.e. pointing to areas of the retina with few symbols in common with the target). Hence, such neurons will have more possibilities to communicate their hypothesis to others, and in this way a population of neurons will rapidly converge onto the current best instantiation of the target on the retina.

Continuing exploration of the retina by inactive matching neurons ensures that this process will eventually discover, and converge to, the best-possible fit of the target on the retina. This convergence to the global-best solution occurs because NESTER implements Stochastic Diffusion Search. This is formally demonstrated in [14] and the time complexity discussed in Nasuto et al. [12].

References

1. Abbot, L.: Learning in Neural Network Memories, Network: Computation in Neural Systems **1**, (1990) 105-122
2. Ahmad, S.: VISIT: An efficient computational model of human visual attention, PhD Thesis, University of Illinois, USA (1991)
3. Barnden, J., Pollack, J. (eds.): High-Level Connectionist Models, Ablex, Norwood NJ, USA (1990)
4. Bishop, J.M.: Stochastic Searching Networks, Proc. 1st IEE Int. Conf. on Artificial Neural Networks, London, UK (1989) 329-331
5. Dinsmore, J.: Symbolism and Connectionism: a difficult marriage, Technical Report **TR 90-8**, Southern University of Carbondale, USA (1990)
6. Fodor, J.A., Pylyshyn, Z.W.: Connectionism and Cognitive Architecture: a critical analysis, Cognition **28**, (1988) 3-72
7. Hubel, D.H., Wiesel, T.N.: Receptive fields, binocular interaction and functional architecture in the cats visual cortex, Journal of Physiology **160** (1962) 106-154
8. James, W.: Psychology (briefer course), Holt, New York, USA (1891)
9. Maass, Bishop, C. (eds.): Pulsed Neural Networks, The MIT Press, Cambridge MA, USA (1999)
10. McCulloch, W.S., Pitts, W.: A logical calculus immanent in nervous activity, Bulletin of Mathematical Biophysics **5** (1943) 115-133
11. Mumford, D.: Neural Architectures for Pattern-theoretic Problems. In: Koch, Ch., Davies, J.L. (eds.): Large Scale Neuronal Theories of the Brain. The MIT Press, London, England (1994)
12. Nasuto, S.J., Bishop, J.M., Lauria, S.: Time Complexity Analysis of the Stochastic Diffusion Search, Neural Computation 98, Vienna, Austria (1998)
13. Nasuto, S.J., Bishop, J.M.: Neural Stochastic Diffusion Search Network - a theoretical solution to the binding problem, Proc. ASSC2, Bremen, Germany (1988) 19-20
14. Nasuto, S.J., Bishop, J.M.: Convergence of the Stochastic Diffusion Search, Parallel Algorithms **14:2**, (1999) 89-107.
15. Pinker, S., Prince, A.: On Language and Connectionism: Analysis of a Parallel Distributed Processing Model of Language Acquisition. In: Pinker, S., Mahler, J. (eds.): Connections and Symbols, The MIT Press, Cambridge MA, (1988)
16. Rumelhart, D.E., McClelland, J.L.: Parallel Distributed Processing: Explorations in the Microstructure of Cognition, The MIT Press, Cambridge MA, (1986)
17. Russel, B.: On Denoting, Mind, (1905); repr. in Marsh, R.C. (ed.): Bertrand Russel: Logic and Knowledge. Essays 1901-1950. London, UK (1965)
18. Sejnowski, T.J., Rosenberg, C.R.: Parallel networks that learn to pronounce English text. Complex Systems **1** (1987) 145-168
19. Tarassenko, L.: A Guide to Neural Computing Applications, Arnold, London (1998)
20. Van Gelder, T., Port, R.F.: It's About Time: an overview of the dynamical approach to cognition. In Port, R.F., Van Gelder, T. (eds.): Mind as Motion. The MIT Press, Cambridge MA (1995) 1-45
21. Van de Velde, F.: On the use of computation in modelling behaviour, Network: Computation in Neural Systems **8** (1997) R1-R32
22. Wittgenstein, L.: Last Writings on the Philosophy of Pyschology (vol. 1). Blackwell Oxford, UK (1948)