

# Modelling Higher Cognitive Functions with Hebbian Cell Assemblies

**Marcin Chady**

The University of Birmingham  
Edgbaston, Birmingham B15 2TT  
United Kingdom  
M.Chady@cs.bham.ac.uk

## Problem Outline

The objective of this work is to develop a model of higher cognitive behaviour using the connectionist paradigm. Examples of such work within the field of artificial neural networks are exceptionally scarce, mainly due to the difficulty of implementing symbol processing in neural hardware. Problems like compositionality and variable binding, as well as the selection of a suitable representation scheme, are a major obstacle to achieving the kind of intelligent behaviour which would extend beyond simple pattern recognition.

This is demonstrated by the following inference example: *if X loves Y, and Y loves Z, and Z is different than X, then X is jealous of Z.* A connectionist system, in order to perform such inference, needs to be able to: i) bind the symbol representations of instances of X, Y and Z to their roles, either by moving them around or by assigning tags; ii) maintain that binding between the stages of stating the proposition and producing the result.

This is inherently hard for connectionist systems due to the locality of their processing. A pattern presented to one set of synapses will be interpreted differently than when it is presented to another set. Although there are connectionist systems which are capable of advanced inferencing (cf. Shastri and Ajjanagadde 1993, or Barnden 1991), they always compromise their flexibility and, most importantly, their capability to learn. All connections in such systems are fixed, having been carefully prearranged by the designer.

What we want to achieve is a self-organising system which is able to learn from a continuous stream of input data and make generalising predictions based on its previous experience.

## Our Approach

The approach suggested here is based on the properties of Hebbian cell assemblies, particularly the associative memory and attractor dynamics. Hebbian cell assemblies are clusters of neurons with reciprocal connections between each other. A very simple and biologically plausible local

learning rule governs the updating of synaptic efficacies:  $\Delta w_{ij} = \eta \xi_i \xi_j$ , where  $\xi_i \in \{0, 1\}$  represents the state of neuron  $i$  (*off* or *on*) and  $\eta$  is a learning rate factor. Such a network tends to prefer states which persistently occurred during its previous history. If just a portion of the original pattern of activity is given, the network will complete it. A preferred state of the network is called an *attractor*.

Inspired by these properties, this study investigates connectionist systems consisting of multiple cell assemblies. Assemblies can be either isolated, overlapping, or connected by long-distance axons. It is expected that activity induced by the input signal in one part of the network will propagate to other parts and, after a while, the network will settle in a global attractor (i.e. the preferred state of the system) representing a balance between all the local attractors. This global attractor may represent the current state of the "working memory", e.g. the proposition "John loves Mary". The next input, say "Mary loves Mark", could then affect the current attractor in such a way that the resulting state will be the one corresponding to "John is jealous of Mark".

Obviously, this general scenario does not answer the question how the variables are bound and transferred from one proposition to another. Some mechanism for "buffering" the variables is required. This can be accommodated e.g. by varying the learning rate. Assemblies with a high learning rate would accept, and sustain for a short time, a wide range of patterns, which makes them a good candidate for a "variable buffer". However, there are many other issues involved. Hopefully they will be addressed in the course of the study.

## References

- Shastri, L., and Ajjanagadde, V. 1993. From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings using temporal synchrony, *Behavioral and Brain Sciences* 16:417-494.
- Barnden, J. A. 1991. Encoding complex symbolic data structures with some unusual connectionist techniques. In Barnden, J. A., Pollack, J. B. eds. 1991. *Advances in Connectionist and Neural Computation Theory, Vol. 1: High-level Connectionist Models*. Norwood, New Jersey.: Ablex Publishing Corporation.