Neural networks and the brain : associative learning and/or self-organisation ?

Dedeurwaerdere Tom, Université Catholique de Louvain and National Foundation for Scientific Research, Chaussée de Wavre 434, 1370 Lathuy

Keywords : self-organisation, dynamical attractors, neural networks

1. Associative learning : from individuals to populations

Experimental evidence suggests that modification of synaptic strength in the brain does not depend on co-activation of two connected neurons, as is assumed in most theoretical work since the proposals of D. Hebb. Instead, through independent post- and presynaptic rules multiple modifications occur simultaneously at various sites in the nervous system.
To account for this data, various researchers (G. M. Edelman, J. M. Fuster, ...) propose an extension of the self-organising PDP approach to populational thinking.

1.1. Neural networks, some general principles

Within the classical theory of neural networks two main strategies for the modification of the synaptic weights have been proposed :

1. Supervised learning (error correcting feedback) : We present a list of examples $y = f(x)$ to the network, if the output s of the network is the desired solution y, or sufficiently close to this solution, connection strength increases.

2. Unsupervised learning (self-organisation) : We present a list of data x to the network and if they are correlated in a certain manner connection strengths are increased.

These learning strategies are both inspired by the original propositions of D. Hebb. He proposed a general learning strategy by means of association of neurons in a network. The association is created by modification of synaptic weights : if two neurons are frequently activated simultaneously their connection strength increases, if not it decreases.
We find back this general rule in both cases of supervised and unsupervised learning. In the first case we associate a list of input and output data x and y. In the second we associate input data to each other, following a certain rule of ressemblance.

1.2. Comparison with actual brain research

Since the 1970's much experimental work has been done to look for support for Hebb's rule. As it turns out the experimental evidence suggests that coactivation of two connected neurons is neither a necessary, neither a sufficient condition for modification of synaptic strength. Instead, through independent post- and presynaptic rules multiple modifications occur at different sites in the nervous system.

Presynaptic efficacy ε_j is defined by the amount of transmitter released by cell j for a given depolarization. Postsynaptical efficacy η_{ij} is the local depolarization produced at postsynaptic processes for a given amount of released transmitter. The two rules can be stated as follows :

1. Presynaptic rule : If the long-term average (over times of the order of 1 sec.) of the instantaneous presynaptic efficacy as determined by transmitter release exceeds a threshold, baseline presynaptic efficacy is modified :

$$\langle \varepsilon \rangle_{\sim 1sec} \geq \theta \Rightarrow \varepsilon_{ref}\, modified$$

The long-term average of ε_j is a function of a large population of neurons connected with j.

2. Postsynaptic rule : Modification of the postsynaptic efficacy η_{ij} is a function of the stimulation to other synapses on the same neuron (heterosynaptic modification), coactivated heterosynaptical inputs to a neuron will alter η_{ij} :

$$\Delta \eta_{ij} = f(\varepsilon_{k1},...,\varepsilon_{kn})$$

1.3. Generalisation of Parallel Distributed Processing

The pre- and postsynaptical rule, together eventually with a Hebb-like rule or even other mechanisms, operate in a cooperative manner within the brain. If we combine both rules, we realise that synaptic alterations of a neuron i are not governed by correlated firing with one single neuron, but are depenent on the activity of a large population of other neurons.

As a consequence a slight modification of the presynaptic efficacy in a certain group of neurons will cause a hierarchy of subsequent short-term modifications among various groups. We can say that multiple synaptic modifications occur simultaneously at various sites in the network. These multiple modifications are caused by one single synaptic modification, for example operating on the presynaptical level. This is the fundamental difference with classical parallel distributed processing, where the multiple modifications are caused by multiple parallel operating coactivations.
The degeneracy in the synaptical modifications of the network is of course transitory, if not the brain would give different answers at once to the same stimuli. In the theory of G. Edelman, where the brain dynamics follows the principles of natural selection, this degeneracy fulfills the need for a continual source of variation. After repeated interaction with the environment the most apt answer will be "reinforced" and selected on behalf of the others. In this the extension of the PDP approach that Edelman proposes is a refinement of the classical scheme of operant conditionning for animals, combined with a theory of perceptual categorisation.

2. From unsupervised learning to dynamical attractors

The unsupervised associative learning networks of Kohonen, Fuster and Edelman all account for an evolution towards point attractors. Hebb's rule or the extension to the

pre- and postsynaptical association rules are linear principles (no feedback or feedforward loops). However, experimental evidence on visual perception seem to indicate that even low level visual processes can converge to more than one attractor, to limit cycles or even to chaotical attractors.

A first case is the case of binocular rivalry. Consider that by some optical trick, your right eye is shown something quite different from your left eye. What happens is that after a brief period the percepts start alternating at regular intervals, changing every few seconds. The brain allows you to perceive only one of them at a time. This is called binocular rivalry.
With binocular rivalry, we have a clear example of a perceptual change without any change in the stimulus. Other examples with a similar temporal dynamics of rivalry are obtained when viewing ambiguous figures, such as the Necker cube and other depth reversals.

An example of a limit cycle in visual processing is given by Christopher Zeeman using a sequence of 8 gradual changing pictures causing suddenly a change in perception (Zeeman, 1988). Here the stimulus is changing, first without any change in perception (in the experiment we observe a man's face), and then showing a sudden change to another percept (producing a kneeling woman), the bifurcation occuring when showing the 7th picture. When the series of pictures is showed again in the other sense, from picture n° 8 to picture n° 1, one observes the same phenomenon, but the change in perception occurs not at the same picture. Instead it occurs halfway at picture n° 4.

Thus, experimental evidence suggests to extend the linear associationist account of visual perception to non-linear dynamics (ex. through feedback and feedforward loops) in order to include the multiplicity of forms created by the autonomous brain dynamics. Indeed, as we have seen, perceptual change can occur in an autonomous manner without any change in the stimulus and even without being able to find a unique optimal solution, as in the case of binocular rivalry. To decide between the ambiguous perceptions an active intervention of a higher cognitive level is needed, which suggests a closer interaction between associationist and concept guided recognition processes.