Reducing connectivity by using cortical modular bands

Julien Vitay, Nicolas Rougier, Frédéric Alexandre

Loria laboratory
Campus Scientifique, B.P. 239
54506 Vandœuvre-lès-Nancy Cedex, France

Abstract. The way information is represented and processed in a neural network may have important consequences on its computational power and complexity. Basically, information representation refers to distributed or localist encoding and information processing refers to schemes of connectivity that can be complete or minimal. In the past, theoretical and biologically inspired approaches of neural computation have insisted on complementary views (respectively distributed and complete versus localist and minimal) with complementary arguments (complexity versus expressiveness). In this paper, we report experiments on biologically inspired neural networks performing sensorimotor coordination that indicate that a localist and minimal view may have good performances if some connectivity constraints (also coming from biological inspiration) are respected.

1 Introduction

In a princeps paper in 1986 [2], D. H. Ballard has set the basis of a taxinomy for neuronal coding schemes that has been extensively discussed later. Based on biological evidences, he proposes that parameter encoding may be done in two complementary ways. In the variable encoding scheme, the intensity of activation of a neuron directly represents the parameter while in the value encoding scheme, a neuron responds only to an interval of the parameter and a population of neurons has to be used to code for the whole parameter span. In theoretical connectionism, these encoding schemes respectively correspond to sigmoidal and gaussian activation functions. In computational neurosciences, these schemes have been respectively observed in the peripheral nervous system and in cortical areas.

This work was partly supported by the MirrorBot european project and the Robéa CNRS French initiative.

Even if these representations have been theoretically described as basis functions with potentially universal approximation capacities [5], practical studies have shown complementary properties for networks with reasonable size [4]. To sum up, concerning information representation, the variable encoding scheme allows for better generalization and extrapolation properties and the value encoding scheme is better for complex (e.g. multimodal) information representation. Another important issue is about the connectivity of such networks. As classical sigmoidal multilayer perceptrons indicate, populations (or layers) of variable encoding neurons are generally fully interconnected and one neuron in a hidden layer can combine all the neurons of the inferior and superior layers. The number of neurons in the hidden layer can accordingly be compared to the complexity of the association function that has to be computed between the other two layers. In contrast to this fully distributed connectivity scheme, value encoding neurons generally use localist connectivity, with the associated 'receptive field' principle, also coming from biological inspiration. In an extreme view (receptive field restricted to one neuron), a hidden (or associative) layer is composed with neurons, each of which connecting only one neuron in the superior and inferior layer. To compensate this very low connectivity, associative layers are generally very large (in that extreme view, the size is the product of the sizes of the associated layers if a systematic combination is needed).

In a word, the situation can be described as follows. On the one hand, variable encoding neurons lead to very compact networks, where information is fully distributed and learning is very efficient. This solution, recommended by classical connectionism, has revealed itself very efficient for relatively simple function approximation. But, as the function becomes more complex or as the information has to be interpreted or used in parallel for different purposes, this kind of networks becomes untractable, precisely because of its compactness. On the other hand, value encoding neurons propose larger networks where more complex information can be more easily represented and processed. If the problems of generalization and fuzziness of information can be solved by overlapping and redundant encoding schemes permitted by receptive fields (which lead to further increase in size), the remaining problem is about the size of associative layer which can lead to combinatorial explosion and also untractable networks if no adapted strategies of combination are possible.

The goal of this paper is to propose and experimentally evaluate such a strategy, based on biological considerations. As described below, we propose, for the sake of comparison, two sensory and motor encoding of information made with the value encoding scheme, that we try to associate with two strategies of connectivity (fully and partially connected). The partially connected strategy will be described as modular, by bands and will lead to comparable though inferior performances for perfect inputs, whereas the connectivity is highly reduced and resistance to noise increased.

2 The model

2.1 A not so simple task

To illustrate our proposal we used a simple task where an object has to be centered onto an image perceived via a mobile camera device. The object is roughly localized relatively to the image using normalized coordinates and the instruction is to center this object onto another position using camera motors. The camera can be moved along both a vertical axis (pan) and an horizontal one (tilt) that allow to move the relative position of any perceived object. More generally, any position (x,y) of an object is normalized into space $[-1, +1]^2$ where position $(0, 0)$ is the centre of image and motor commands are controled using normalized displacements $[-1, +1]^2$ where position $(0, 0)$ corresponds to no movement at all. The difficulty of the task is then to convert absolute coordinates of object and target positions into a relative camera displacement.

2.2 Information encoding

The input and output implementations are defined as two dimensional maps of $n \times n$ neurons, where each neuron has a preferred attribute in the space $[-1, +1]^2$, which is regularly distributed across the map. The response of a neuron to a given stimulus in that space is defined by a gaussian activation depending on the distance between its preferred attribute and the stimulus, with a given standard deviation (which has a great influence on performance). Decoding is then directly done by determining the center-of-mass (COM) of the preferred attributes, ponderated by the activities of the corresponding neurons

This population coding clearly is a value coding scheme. It is quite loose in the sense that any value induces a large bubble of activity within the concerned map (virtually spreading over one quarter of the map).

2.3 Architecture

The architecture of the model is quite constrained by the task which requires at least one input map for actual position of the object, one input map for the instruction (target to be reached) and one output map for the motor command. There is also an associative map that allows to learn the required sensory-motor coordination since the task cannot be solved by a direct association. The key point concerning the associative map is then the pattern of connectivity that is necessary to ensure that virtually any association between perception and action can be learnt. All maps are chosen with a $n \times n$ size.

One naive and quite expensive way of connecting maps is then to use a full connectivity (i.e. like for a multi-layer perceptron) where each associative unit receives connections from the two input maps ($2 \times n^2$ connections) and projects itself onto each output unit (n^2 connections) for a total of $3 \times n^4$ connections. One of the legitimate question that comes to mind is then to know if it is really needed to make any such arbitrary associations. Therefore we propose

to drastically reduce it by virtue of cortical modular bands and distributed representation of information. The biological plausibility of this modular connectivity has been often confirmed and particularly in [3].
Let's name E the map that represents actual perception, C the map that represents the desired perception and M the output motor map for controlling the camera device (cf. Figure 1).

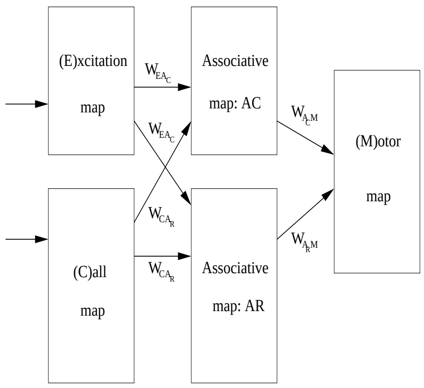

Figure 1: The partially connected model architecture.

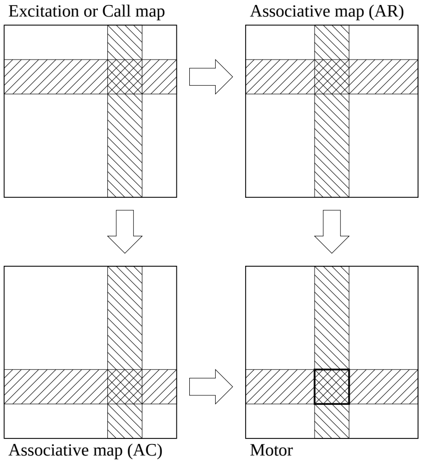

Figure 2: The pattern of connectivity between a perception map and a motor map.

In the proposed strategy, we have splitted the previously unique associative map into two distinct associative maps A_C and A_R that link inputs maps to motor maps using two disctinct patterns of connectivity as illustrated in Figure 2.

Each column of the A_C map (or associative column map) receives projection from corresponding column of input maps while each line of the A_R map (or associative row map) receives projection from corresponding line of input maps. Furthermore, each line of the A_C map projects itself to the corresponding line of the motor map while each column of the A_R map projects itself to the corresponding column of the motor map. More generally:

$$\text{Each unit } A_C(x,y) \text{ is linked to } A(i,y) \text{ and } C(i,y),\ i \in [0,n[$$
$$\text{Each unit } A_R(x,y) \text{ is linked to } A(x,i) \text{ and } C(x,i),\ i \in [0,n[$$
$$\text{Each unit } M(x,y) \text{ is linked to } A_R(i,y) \text{ and } A_C(x,i),\ i \in [0,n[$$

Overall, each unit of any associative maps receives $2 \times n$ connections and each unit of the motor map receives $2 \times n$ from associative maps for a total of $6 \times n^3$ connections, to compare with the $3 \times n^4$ connections of the perceptron.

2.4 Learning

The goal of our experiments was to compare a full connectivity with our restrictive pattern of connectivity combined with a gaussian-type distributed representation of information. We consequently implemented two multilayer perceptrons with similar input and output and respectively the partially connected and the fully connected structure. We tested both networks 10 times to get the average behavior that is quite stable over each run. We also tested the partially connected model using the Continuous Neural Field Theory (CNFT, [1]) that allows to suppress noise in the output layer by asynchronous competition between neurons using “mexican-hat” shaped lateral connections[6].

3 Results

As expected and as illustrated on figure 3 the fully connected model benefits from best performances while partially connected model and its variant (using CNFT) get slightly worse performances. The measured performance is the distance from actual camera position to the desired location after one running step and then, from a practical point of view, it means that both models are able to bring camera to the desired location in one or two steps only (one large move followed by a very small correction one). The loss in terms of pure performances is then not really significant for a robotic application. Moreover, if we had uniformly distributed noise on the input maps, we can see on figure 4 that the partially connected model rapidly becomes more robust to noise levels greater than 10% than the fully connected model.

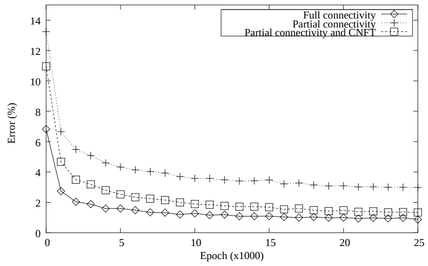

Figure 3: Performance comparison between the fully connected model, the partially connected model and the partially connected model with CNFT correction.

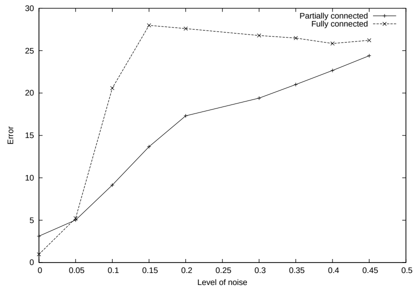

Figure 4: Effect on the performance of the networks of an uniformly distributed noise level on input maps.

It is also interesting to note that the size of the gaussian for information encoding/decoding within input and output map has a direct influence on performances: thin gaussians, as well as too large ones, are unable to produce learning. There is a unique value of the gaussian radius for each network where performance is maximal (cf. Figure 5). This value is higher for the reduced-connectivity network.

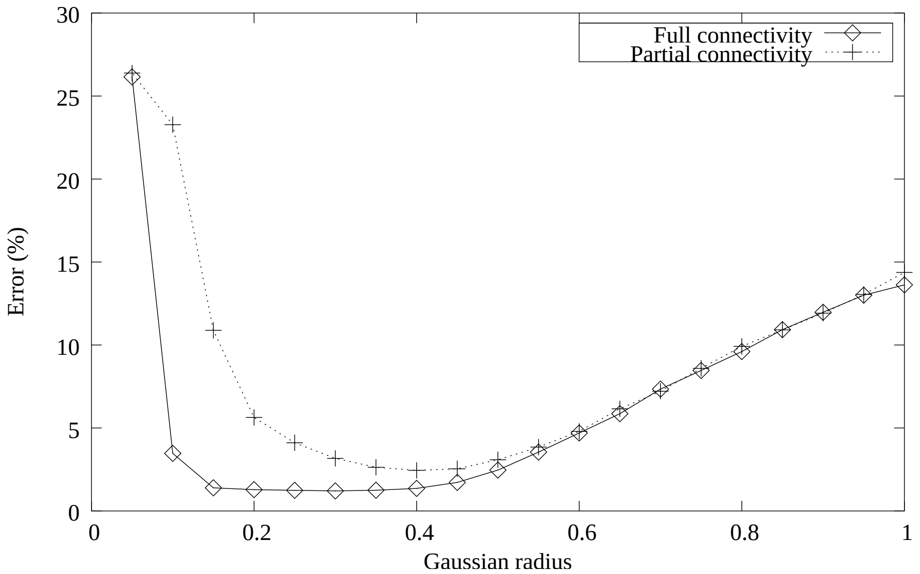

Figure 5: Influence of the standard deviation of the tuning curve of input neurons on network performance.

These results also show that for real-world autonomous robotics, both patterns of connectivity have similar performances, whereas the one we propose is highly cheaper. Using this connectivity and local mechanisms like the CNFT also indicate that the value encoding principle is a very powerful and general principle with interesting learning and representational capacities. Our goal was to show here that such principles can be used with networks of reasonable size if other connectivity constraints from biology are added.

References

[1] S.-I. Amari. Dynamical study of formation of cortical maps. *Biological Cybernetics*, 27:77–87, 1977.

[2] D.H. Ballard. Cortical connections and parallel processing: Structure and function. *Behavioral and Brain Sciences*, 9:67–120, 1986.

[3] Y. Burnod. *An adaptive neural network: the cerebral cortex*. Masson, 1989.

[4] E. Guigon. Computing with populations of monotonically tuned neurons. *Neural computation*, 15:2115–2127, 2003.

[5] Tomaso Poggio and Federico Girosi. A theory of networks for approximation and learning. Technical Report AIM-1140, 1989.

[6] A. Pouget, K. Zhang, S. Deneve, and P.E. Latham. Statistically efficient estimation using population coding. *Neural Computation*, (10):373–401, 1998.