Explorations with the Dynamic Wave Model

Thomas P. Rebotier
Department of Cognitive Science
UCSD, 9500 Gilman Dr
LA JOLLA CA 92093-0515
rebotier@cogsci.ucsd.edu

Jeffrey L. Elman
Department of Cognitive Science
UCSD, 9500 Gilman Dr
LA JOLLA CA 92093-0515
elman@cogsci.ucsd.edu

Abstract

Following Shrager and Johnson (1995) we study growth of logical function complexity in a network swept by two overlapping waves: one of pruning, and the other of Hebbian reinforcement of connections. Results indicate a significant spatial gradient in the appearance of both linearly separable and non linearly separable functions of the two inputs of the network; the n.l.s. cells are much sparser and their slope of appearance is sensitive to parameters in a highly non-linear way.

1 INTRODUCTION

Both the complexity of the brain (and concomittant difficulty encoding that complexity through any direct genetic mapping), as well as the apparently high degree of cortical plasticity suggest that a great deal of cortical structure is emergent rather than pre-specified. Several neural models have explored the emergence of complexity. Von der Marlsburg (1973) studied the grouping of orientation selectivity by competitive Hebbian synaptic modification. Linsker (1986.a, 1986.b and 1986.c) showed how spatial selection cells (off-center on-surround), orientation selective cells, and finally orientation columns, emerge in successive layers from random input by simple, Hebbian-like learning rules. Miller (1992, 1994) studied the emergence of orientation selective columns from activity dependant competition between on-center and off-center inputs.

Kerzsberg, Changeux and Dehaene (1992) studied a model with a dual-aspect learning mechanism: Hebbian reinforcement of the connection strengths in case of correlated activity, and gradual pruning of immature connections. Cells in this model were organized on a 2D grid, connected to each other according to a probability exponentially decreasing with distance, and received inputs from two different sources, A and B, which might or might not be correlated. The analysis of the network revealed 17 different kinds of cells: those whose output after several cycles depended on the network's initial state, and the 16 possible logical functions of two inputs. Kerzsberg et al. found that learning and pruning created different patches of cells implementing common logical functions, with strong excitation within the patches and inhibition between patches.

Shrager and Johnson (1995) extended that work by giving the network structure in space (structuring the inputs in intricated stripes) or in time, by having a Hebbian learning occur in a spatiotemporal wave that passed through the network rather than occurring everywhere simultaneously. Their motivation was to see if these learning conditions might create a cascade of increasingly complex functions. The approach was also motivated by developmental findings in humans and monkeys suggesting a move of the peak of maximal plasticity from the primary sensory and motor areas towards parietal and then frontal regions. Shrager and Johnson classified the logical functions into three groups: the constants (order 0), those that depend on one input only (order 1), those that depend on both inputs (order 2). They found that a slow wave favored the growth of order 2 cells, whereas a fast wave favored order 1 cells. However, they only varied the connection reinforcement (the growth Trophic Factor), so that the still diffuse pruning affected the rightmost connections before they could stabilize, resulting in an overall decrease which had to be compensated for in the analysis.

In this work, we followed Shrager and Johnson in their study of the effect of a dynamic wave of learning. We present three novel features. Firstly, both the growth trophic factor (hereafter, TF) and the probability of pruning (by analogy, "death factor", DF) travel in gaussian-shaped waves. Second, we classify the cells in 4, not 3, orders: order 3 is made of the non-linearly separable logical functions, whereas the order 2 is now restricted to linearly separable logical functions of both inputs. Third, we use an overall measure of network performance: the slope of appearance of units of a given order. The density is neglected as a measure not related to the specific effects we are looking for, namely, spatial changes in complexity. Thus, each run of our network can be analyzed using 4 values: the slopes for units of order 0, 1, 2 and 3 (See Table 1.). This extreme summarization of functional information allows us to explore systematically many parameters and to study their influence over how complexity grows in space.

Table 1: Orders of logical complexity

ORDER	FUNCTIONS
0	True False
1	A !A B !B
2	A.B !A.B A.!B !A.!B AvB !AvB Av!B !Av!B
3	A xor B, A==B

2 METHODS

Our basic network consisted of 4 columns of 50 units (one simulation verified the scaling up of results, see section 3.2). Internal connections had a gaussian bandwidth and did not wrap around. All initial connections were of weight 1, so that the connectivity weights given as parameters specified a number of labile connections. Early investigations were made with a set of manually chosen parameters ("MAN-

UAL"). Afterwards, two sets of parameters were determined by a Genetic Algorithm (see Goldberg 1989): the first, "SYM", by maximizing the slope of appearance of order 3 units only, the second, "ASY", byoptimizing jointly the appearance of order 2 and order 3 units ("ASY"). The "SYM" network keeps a symmetrical rate of presentation between inputs A and B. In contrast, the "ASY" net presents input B much more often than input A. Parameters are specified in Table 1 and, are in "natural" units: bandwidths and distances are in "cells apart", trophic factor is homogenous to a weight, pruning is a total probability. Initial values and pruning necessited random number generation. We used a linear congruence generator (see p284 in Press 1988), so that given the same seed, two different machines could produce exactly the same run. All the points of each Figure are means of several (usually 40) runs with different random seeds and share the same series of random seeds.

Table 2: Default parameters

MAN.	SYM.	ASY.	name	description
8.5	6.20	12	Wae	mean ini. weight of A excitatory connections
6.5	5.2	9.7	Wai	mean ini. weight of A inhibitory connections
8.5	8.5	13.4	Wbe	mean ini. weight of B excitatory connections
6.5	6.5	14.1	Wbi	mean ini. weight of B inhibitory connections
5.0	6.5	9.9	Wne	m.ini. density of internal excitatory connections
3.5	1.24	12.4	Wni	m.ini. density of internal inhibitory connections
0.2	0.20	0.28	DW	relative variation in initial weights
7.0	1.26	0.65	Bne	bandwidth of internal excitatory connections
7.0	2.86	0.03	Bni	bandwidth of internal inhibitory connections
0.7	0.68	0.98	Cdw	celerity of dynamic wave
1.5	3.0	-3.2	Ddw	distance between the peaks of both waves
9.87	17.6	16.4	Wtf	base level of TF (=highest available weight)
0.6	0.6	0.6	Btf	bandwidth of TF dynamic wave
3.5	1.87	3.3	Tst	Threshold of stabilisation (pruning stop)
0.6	0.64	0.5	Bdf	bandwidth of DF dynamic wave
0.65	0.62	0.12	Pdf	base level of DF (total proba. of degeneration)
0.5	0.5	0.06	Pa	probability of A alone in the stimulus set
0.5	0.5	0.81	Pb	probability of B alone in the stimulus set
0.00	0.00	0.00	Pab	probability of simultaneous s A and B

3 RESULTS

3.1 RESULTS FORMAT

All Figures have the same format and summarize 40 runs per point unless otherwise specified. The top graph presents the mean slope of appearance of all 4 orders of complexity (see Table 1) on the y axis, as a function of different values of the experimentally manipulated parameter, on the x axis. The bottom left graph shows the mean slope for order 2, surrounded by a gray area one standard deviation below and above. The bottom right graph shows the mean slope for order 3, also with a 1-s.d. surrounding area. The slopes have not been normalized, and come from networks whose columns are 50 units high, so that a slope of 1.0 indicates that the number of such units increase in average by one unit per columns, ie, by 3 units