

# Memory, Far from Equilibrium Basins of Attraction of Random Boolean Networks

Andrew Wuensche

Santa Fe Institute  
and University of Sussex  
School of Cognitive and Computer Science  
wuensch@santafe.edu  
andywu@susx.ac.uk

contact address:  
48 Esmond Road, London W4 1JQ  
tel 081 995 8893 fax 081 742 2178  
100020.2727@compuserve.com

### Abstract

Recent work in unravelling the global dynamics of discrete dynamical systems such as cellular automata (Wuensche and Lesser, 1992), and more generally, of random Boolean networks (Wuensche, 1992), allow their basins of attraction to be explicitly portrayed. These are diagrams that connect up the network's global states according to their transitions - typically, the topology is branching trees rooted on attractor cycles. The diagrams are efficiently constructed with an algorithm that directly computes a state's set of pre-images (if any).

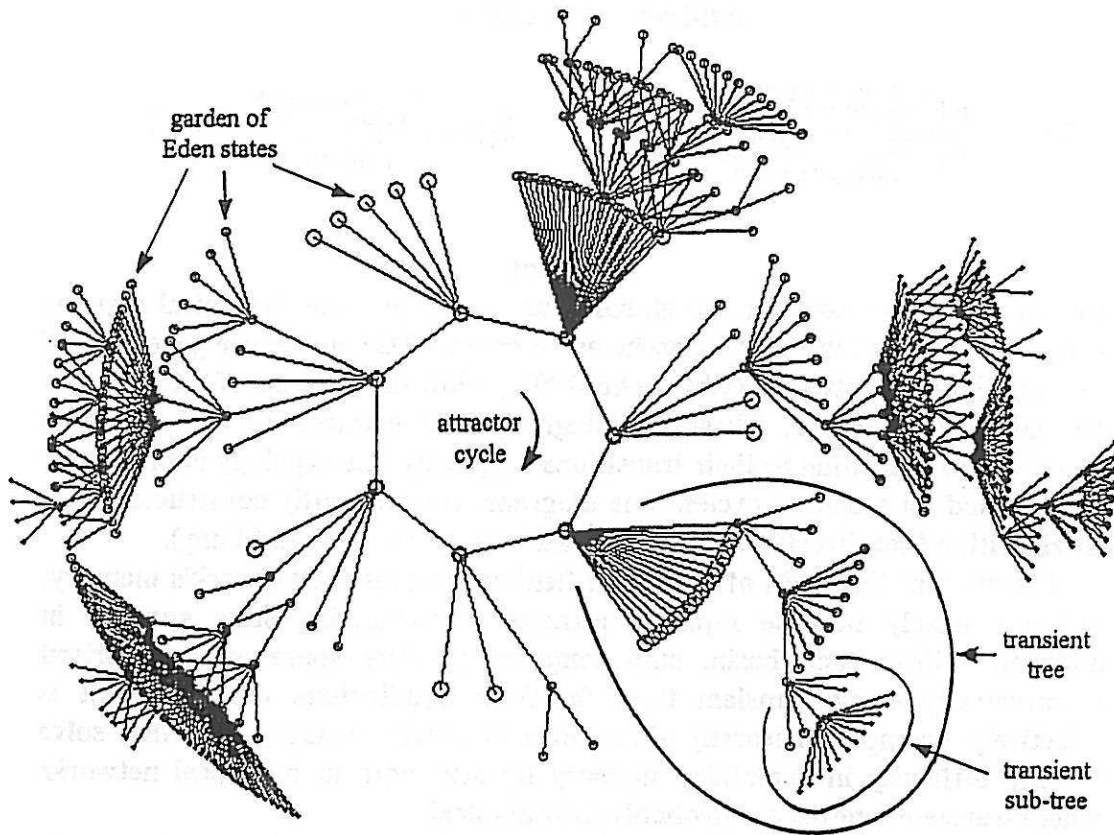
I argue that the basin of attraction field constitutes the network's memory; but not simply because separate attractors categorising state space - in addition, within each basin, sub-categories of state space are categorised hierarchically along transient trees far from equilibrium, creating what is effectively a complex hierarchy of content addressable memory. This may solve a basic difficulty in explaining memory by attractors in biological networks where transient lengths are probably astronomical.

I describe a single step learning algorithm for re-assigning pre-images in random Boolean networks. This allows the sculpting of their basin of attraction fields to approach any desired configuration. The process of learning and its side effects are made visible. In the context of many semi-autonomous weakly coupled networks, the basin field/network relationship may provide a fruitful metaphor for the mind/brain.

### Introduction

Hopfield described a neural network model (Hopfield, 1982) with emergent collective properties that spontaneously give rise to a *categorizer* or *content-addressable* delocalised memory. The model is a discrete dynamical system; memories are simply separate regions of state space flowing to stable states or energy minima. He also found unwelcome *simple cycles* and *chaotic wanderings* in small regions. Changes to memory are made by modifying system parameters to alter regions or create new minima. In continuous deterministic dynamical systems this is analogous to the vector field - the field of flow imposed on phase space by the systems dynamics described by its phase portrait. A set of attractors, be they fixed point, limit cycles or chaotic, attract various regions of phase space in the *basin of attraction field*.

Analogous concepts apply to discrete dynamical systems such as cellular automata, and the more general case, random Boolean networks, which update synchronously and are deterministic in the absence of noise or changes to parameters. An important difference, however, is that transients can merge onto one successor state far from equilibrium in these discrete systems, whereas in continuous systems they cannot. Neither does Hopfield's model support deterministically merging transients because



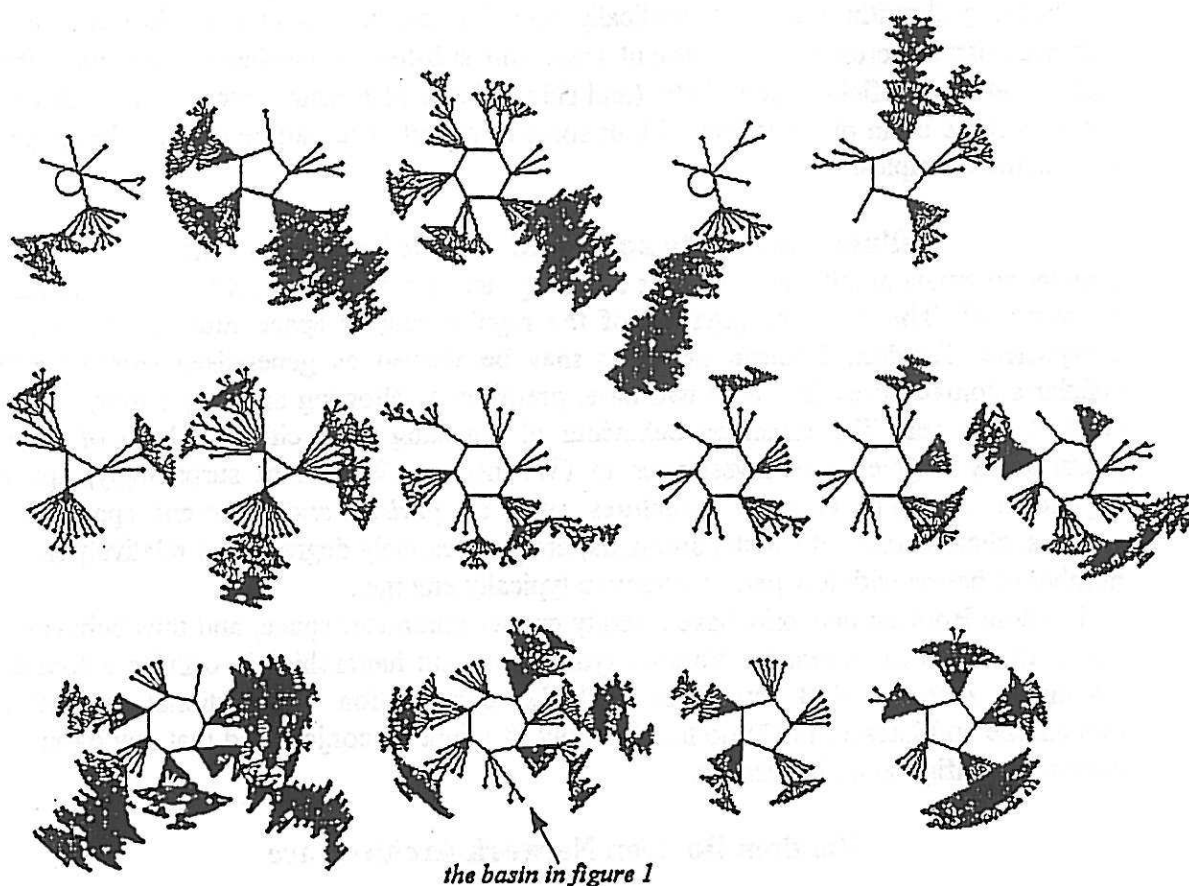
**Figure 1**

A basin of attraction of a random Boolean network ( $N=13, K=3$ ). The basin links 604 states, of which 523 are garden of Eden states. The attractor has period 7. The direction of time is inwards from garden of Eden states to the attractor, then clock-wise. The basin is one of 15, and is indicated in the basin of attraction field in figure 2. Table 1 gives the random Boolean network wiring/rule parameters. Wiring and rules were assigned at random, except that the neighbourhood  $000 \rightarrow 0$ .

Table 1				
cell	wiring	rule	-table	
1	3,12,6	86,	01010110	
2	7,11,4	4,	00000100	
3	3,3,1	196,	11000100	
4	11,3,9	52,	00110100	
5	8,7,5	234,	11101010	
6	1,8,1	100,	01100100	
7	12,4,13	6,	00000110	
8	8,6,8	100	01100100	
9	9,2,6	6,	00000110	
10	5,1,1	94,	01011110	
11	2,7,1	74,	01001010	
12	7,8,4	214,	11010110	
13	1,4,7	188,	10111100	

the updating is randomly asynchronous. It is open to debate whether synchronous or asynchronous updating in a local network is more or less biologically plausible. However, synchronous random Boolean networks may have greater potential as content-addressable memory systems because not only attractors categorise state space. State space is also categorised by a reliable time-sequence of unique states along each transient tree far from equilibrium, creating what is effectively a complex hierarchy of content addressable memory.

The range of topologies of basins of attraction, and the potential for emergent complex categorisation of network states, suggests that the basin of attraction field, a mathematical object in space-time, is the network's cognitive substrate - *the ghost in the machine* (Wuensche, 1992). A basic difficulty in explaining memory by attractors in biological networks has been the probably astronomical transient lengths needed to



**Figure 2.**

The basin of attraction field a random Boolean network ( $N=13$ ,  $K=3$ ). The  $2^{13}=8192$  states in state space are organised into 15 basins, with attractor period ranging from 1 to 7. The number of states in each basin is: 68, 984, 784, 1300, 264, 76, 316, 120, 64, 120, 256, 2724, 604, 84, 428. The arrowed basin is shown in more detail in figure 1, and the network's wiring/rule scheme in table 1.

reach an attractor in large networks, whereas reaction times in biology are extremely fast. The answer may lie in the notion of *memory far from equilibrium* along merging transients.

Basins of attraction of random Boolean networks have been investigated for a considerable time, notably by Stuart Kauffman in theoretical biology and complex systems (Kauffman, 1984, 1989). These studies have built statistical data on basin fields from many separate *forward* simulations. Recently, explicit portraits of basin of attraction fields for cellular automata have been made accessible in (Wuensche and Lesser, 1992), and for random Boolean networks in (Wuensche, 1992). This is achieved by *reverse* algorithms, running the system backwards to directly compute a given state's set of *pre-images* (if any). The diagrams of basins of attraction make content-addressable (*auto-associative*) memory visible. This may help to clarify the process of memory and learning in a variety of neural network architectures, as well as random Boolean networks. Two levels of learning may be distinguished. Where a network's architecture remains unchanged, its dynamics may be shifted from one basin of attraction to another by interacting with other networks or external inputs. More radically, learning (and forgetting) involves adjusting the network's parameters (its wiring/rule scheme and size/conductivity) to achieve a more appropriate basin of attraction field.

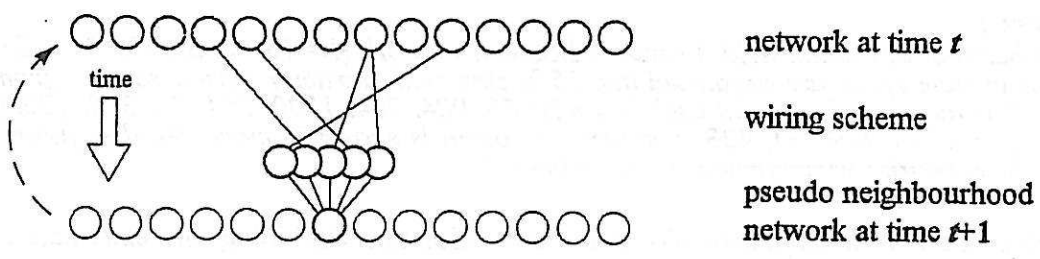
In this paper I describe random Boolean network architecture, basins of attraction, and learning algorithms that automatically re-assign pre-images in a single step. New attractors can be created and transient trees and sub-trees transplanted, *sculpting* the basin of attraction field. The effects, (and side effects) of learning become immediately apparent if the basin of attraction field, or some fragment of it, can be drawn. This paper gives some examples.

### Cellular Automata and Random Boolean Networks

Cellular automata architecture requires a homogeneous neighbourhood template and rule at every cell. The "laws of physics" of the regular cellular space must be the same everywhere. Random Boolean networks may be viewed as generalised (*disordered*) cellular automata, breaking these two basic premises by allowing arbitrary wiring and/or rules at each cell. The effect on behaviour of deviating from either or both of these premises by degrees was investigated in (Wuensche, 1992). Not surprisingly, space dependent emergent complex structures such as *gliders* and coherent space-time patterns, characteristic of cellular automata, are progressively degraded. A relatively small number of basins with low period attractors typically emerge.

Random Boolean networks have a vastly greater parameter space, and thus behaviour space, than cellular automata. Various symmetries and hierarchies in cellular automata dynamics, such as shift invariance and the conservation of rotational symmetry (Wuensche and Lesser, 1992) no longer apply. It might be conjectured that any basin of attraction configuration is possible.

### Random Boolean Network Architecture



**Figure 3.** Each cell in the network synchronously updates its value according to the values in a pseudo neighbourhood, set by single wire couplings to arbitrarily located cells at the previous time-step. Each cell may have a different wiring/rule scheme. The system is iterated.

Random Boolean networks, as the name implies, allow network elements to take on one of two possible attributes (the cell's value, say 0 or 1), but in principle the arguments in this paper could equally apply to a network where cells have more than two values. A global state of a network of  $N$  cells is the pattern of values assigned to each array cell, from a finite range of values  $V$  (usually  $V=2$ ). Each cell synchronously updates its value in discrete time steps. The value of a cell at time  $t+1$  depends on its particular cellular automata rule applied to a notional or *pseudo* neighbourhood, size  $K$ . Values in the neighbourhood are set according to single wire couplings to arbitrarily located cells in the network at time  $t$ . The system is iterated. The system's parameters are set by specifying the cellular automata rule and pseudo neighbourhood wiring, which may be different for each cell. Once set, the system's parameters are fixed over time.

### The Network's Behaviour Space

There are  $V^K$  permutations of values in a neighbourhood of size  $K$ . A rule table (look up table) with  $V^K$  entries will specify the output of all neighbourhood permutations. By convention (Wolfram, 1983) the rule table is arranged in descending order of the values of neighbourhoods. For example, the rule table for rule 30 ( $V=2, K=3$ ) is,

	111	110	101	100	011	010	001	000	...	neighbourhoods
rule table...	0	0	0	1	1	1	1	0	...	outputs (0 or 1)

The total number of distinct rule tables, the size of rule space =  $V^{V^K}$ . The number of alternative wiring schemes for one cell =  $N^K$ . The number of alternative wiring/rule schemes,  $S$ , that can be assigned to a given network turns out to be vast even for small networks, and is given by

$$S = (N^K)^N \times (V^{V^K})^N$$

for example, a network where  $V=2, N=16, K=5, S = (2)^{832}$

Random Boolean network architecture is in many ways similar to weightless neural networks (Alexander *et al*, 1984), where standard memories (RAMs) hold each cell's look-up table. Classical neural network architecture uses weighted connection and threshold functions. A random Boolean network may be regarded as a discrete generalisation of a sparsely connected classical neural network. Connections with higher weights may simply be replaced by multiple couplings, and the threshold function applied. However, a threshold function is a tiny sub-class of the  $V^{V^K}$  possible cellular automata rules.

Intermediate architectures between cellular automata and fully random Boolean networks may be appropriate for particular applications. For instance, notions of space and boundary conditions may be retained if random wiring is confined within a local zone.

### Basins of attraction

Cellular automata and random Boolean networks are examples of discrete deterministic dynamical systems. They evolve along a deterministic trajectory consisting of a succession of global states that represents one particular path within a basin of attraction, familiar from continuous dynamical systems. As the number of global states,  $V^N$ , is finite, any path inevitably leads to a state cycle (the attractor). The set of all possible paths leading to the same attractor, including the attractor itself, makes up the basin of attraction. This is composed of merging trajectories linked according to their dynamical relationships, and will typically have a topology of branching trees rooted on the attractor.

Basins of attraction are portrayed as computer diagrams in the same graphic format as presented in (Wuensche and Lesser, 1992). Various other names are sometimes used, for example, flow graphs, state transition graphs, networks of attraction. Global states are represented by nodes, or by the state's binary or decimal expression at the node position. Nodes are linked by directed arcs. Each node will have zero or more incoming arcs from nodes at the previous time-step (*pre-images*), but because the system is deterministic, exactly one outgoing arc (*one out degree*). Nodes with no pre-images have no incoming

arcs, and represent so called *garden of Eden* states. The number of incoming arcs is referred to as the *degree of pre-imaging* (or *in degree*). Figure 1 shows a typical basin of attraction of a random Boolean network (it is part of the basin of attraction field shown in figure 2).

### Computing Pre-Images

Construction of a single basin of attraction poses the problem of finding the complete set of pre-images of every global state that is linked together in the basin. The trivial solution, exhaustive testing of the entire state space, becomes intractable in terms of computer time as the network's size increases beyond modest limits. To overcome this problem, methods have been invented for computing pre-images directly, without exhaustive testing. The network's dynamics can, in effect, be run *backwards* in time; backward trajectories will, as a rule, diverge. A *reverse algorithm* that directly computes pre-images for one-dimensional cellular automata was presented in (Wuensche and Lesser, 1992), and a *general direct reverse algorithm* for random Boolean networks (which includes higher dimensional cellular automata) in (Wuensche, 1992). Providing that  $K < N$ , the average computational performance is the many orders of magnitude faster than exhaustive testing, making basin portraits for these systems accessible for the first time.

### A biological model

Memory far from equilibrium along merging transients may answer a basic difficulty in explaining memory by attractors in biological neural networks. A view of the brain as a complex dynamical system made up of many inter-linked specialised neural networks is perhaps the most powerful paradigm currently available. Specialised neural networks may consist of further sub-categories of semi-autonomous networks, and so on, which contribute to re-setting each other's global state. A biological neural sub-network is nevertheless likely to be extremely large; as a dynamical system the time required to reach an attractor from some arbitrary global state will probably be astronomical. Even when an attractor is reached, it may well turn out to be a long cycle or a quasi-infinite chaotic attractor. The notion of memory only as attractors seems to be inadequate to account for the extremely fast reaction time in biology.

A discrete dynamical system with synchronous updating categorises its state space reliably along transient trees, far from equilibrium, as well as at the attractors. A network that has evolved or learnt a particular global dynamics may be able to reach memory categories in a few steps, possibly just one. Moreover, the complex transient tree topology in the basin of attraction field, makes for a much richer substrate for memory than attractors alone, allowing hierarchies of memory sub-categories.

There is evidence that the firing of nearby neurons is strictly related in time. Phase locking of spike discharges between neighbouring cells has been observed, extending up to 7mm across the cortex (Grey et al, 1989, Singer, 1991, Wasserman, 1992). Synchronous firing may be mediated by interneurons, which lack axons (Shepherd, 1990), or mechanisms relying on close physical proximity between neurons (their dendrites, cell bodies and axons). Gap junction effects (physical connections between neurons made by large macro molecules), and ephaptic interactions (the local electrical field) serve to synchronise local neuronal activity (McCormick, 1990).

A random Boolean network may serve as a model of a patch of semi-autonomous biological neurons whose activity is synchronised. A cell's wiring scheme models that sub-set of neurons connected to a given neuron. Applying the cellular automata rule to a cell's pseudo neighbourhood models the non-linear computation that a neuron is said to

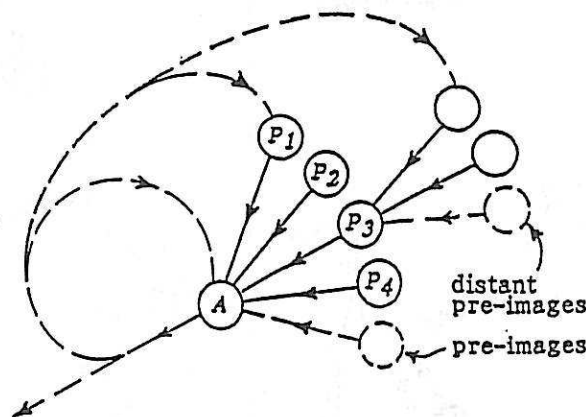
apply to these inputs to decide whether or not to fire at the next time-step. This is far more complex than a threshold function (Shepherd, 1990). The biological computation may be a function of the topology of the dendritic tree, the microcircuitry of synaptic placements, intrinsic membrane properties, or properties of the cytoskeleton, said to act as the neuron's 'internal nervous system' (Hammeroff *et al*, in press). There appears to be no shortage of biological mechanisms that could perform the role of a cellular automata's look-up table.

The model may be elaborated by weakly inter-connecting a number of random Boolean networks (or perhaps 3-state networks) so that they are able to influence each other. Communication between networks may be *asynchronous*, and at a slower frequency than a particular network's internal synchronous clock. Such an assembly of networks will have implicit in its particular pattern of connections at any instant, a vastly more complex but intangible web of interacting basin of attraction fields - *the ghost in the machine* ?

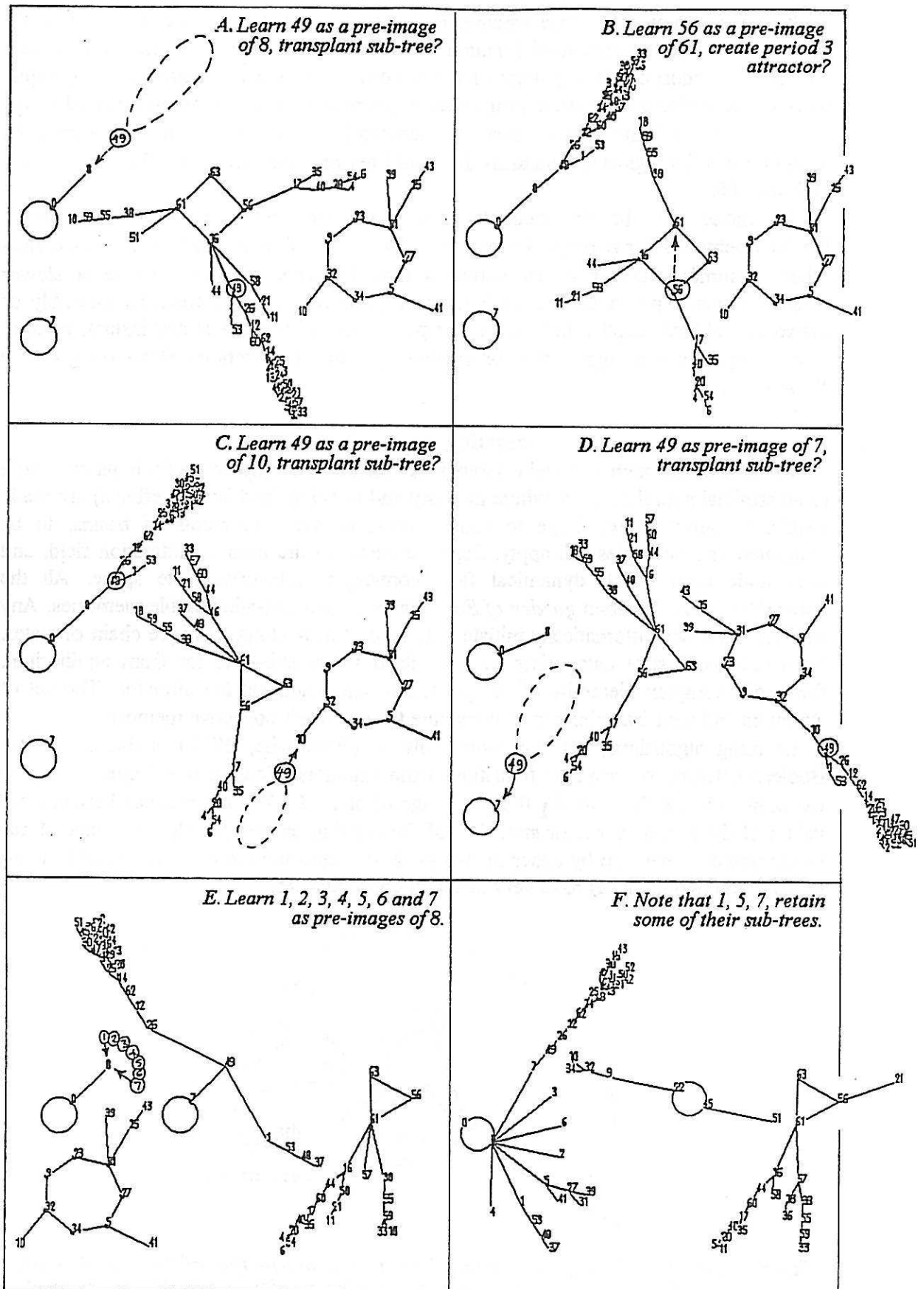
### Learning Algorithms

Whether or not such a model is biologically plausible, it may be useful in its own right as an artificial neural network where memory and learning (and its side effects) are made visible. In networks too large to allow basins, or even fragments of basins, to be computed, the principles still apply. Separate basins in the basin of attraction field, and each node onto which dynamical flow converges, categorise state space. All the network's states other than *garden of Eden* states are context-addressable memories. Any external input will automatically initiate a dynamical flow along a unique chain of states. Each successive state categorises states in its transient sub-tree, far from equilibrium, forming a complex hierarchy of categorisation culminating at the attractor. The set of attractors and their branching trees constitute the network's collective memory.

Learning algorithms, set out more fully in (Wuensche, 1992) enable a random Boolean network to learn new transitions from experience (and also to forget). Suppose we want to make the state  $P_1$  the pre-image of state  $A$ . Any mismatches between cell values of the actual successor state  $B_1$  (of the aspiring pre-image,  $P_1$ ) and state  $A$  can be corrected in one step by either of two methods, adjusting the network's wiring or rule scheme; the two methods have very different consequences.



**Figure 4.** States  $P_1, P_2, P_3, \dots$  etc may be learnt as pre-images of the state  $A$ . Distant pre-images of  $A$  may also be learnt, for instance the pre-images of  $P_3$ . Learning  $A$  as a pre-image of itself creates a point attractor. Learning  $A$  as a distant pre-image of itself creates a cyclic attractor. If  $A$  is learnt as the pre-image of some other state in the basin of attraction field, the states flowing into  $A$ , its transient tree, may be fully or partially transplanted along with  $A$ .



**Figure 5.** A sequence of learning steps, A to F, in a 6 cell network. Note the stability of basin structure as well as the side effects of learning at each step. Bit patterns are represented by decimal numbers.



To correct a mismatch by rewiring, one (or more if necessary) of the cell's wiring couplings is moved to a new position. Any move resulting in a pseudo neighbourhood with opposite output (according to that cell's look up table) will correct the mismatch. This is a stochastic method as there may be many alternative successful options. If another aspiring pre-image,  $P_2$  is to be added to  $A$  there is a chance that  $P_1$  may be forgotten.

On the other hand, to correct a mismatch in a particular cell by adjusting the rule scheme, one specific bit in the cell's rule table is flipped - the output of the pseudo neighbourhood. There is only one option, certain to succeed. Adding another aspiring pre-image,  $P_2$  by the same method cannot cause  $P_1$  to be forgotten. This is because any mismatch between a particular cell in the successor state  $B_2$  (of the *aspiring* pre-image,  $P_2$ ) and  $A$  can not relate to the same rule table entry that was 'looked up' to determine  $P_1$ 's successor. Otherwise there would have been no mismatch. Any change to correct the mismatch must be to a different rule table entry;  $P_1$ 's successor cannot be affected. It turns out that there is no limit to the number of pre-images of a given state that can be learnt by this method, and no risk of forgetting previously learnt pre-images of the state, but of course there will be side effects elsewhere in the basin field.

Re-wiring appears to have a greater effect on basin structure than mutating the rule scheme, but in either case the stability of basin structure is noteworthy. As illustrated in figure 4, using these methods, point attractors, cyclic attractors and transient sub-trees can be created. Transient sub-trees are sometimes transplanted along with the repositioned state. Generalisation is present, because bit patterns in the same pre-image fan are likely to be close in Hamming distance from each other, and so may be learnt by default from examples.

Figure 5 shows an arbitrary example (with no particular aim) of visible learning (and side effects) in a 6 cell network with regular wiring but randomly allocated rules. At each learning step, a state (or set of states) is made the pre-image of a target state. Fig 6 shows the basin of attraction field of a 6 cell network that has been taught to segregate strings with odd and even parity into two separate basins.

Combining wiring and rule scheme adjustments, may result in a powerful method of cumulative learning (supervised and unsupervised) in random Boolean networks. In future work I anticipate using genetic algorithms to improve the start parameters of networks prior to learning, to extend the learning methods to include system size and connectivity, and to investigate assemblies of weakly coupled semi-autonomous networks.

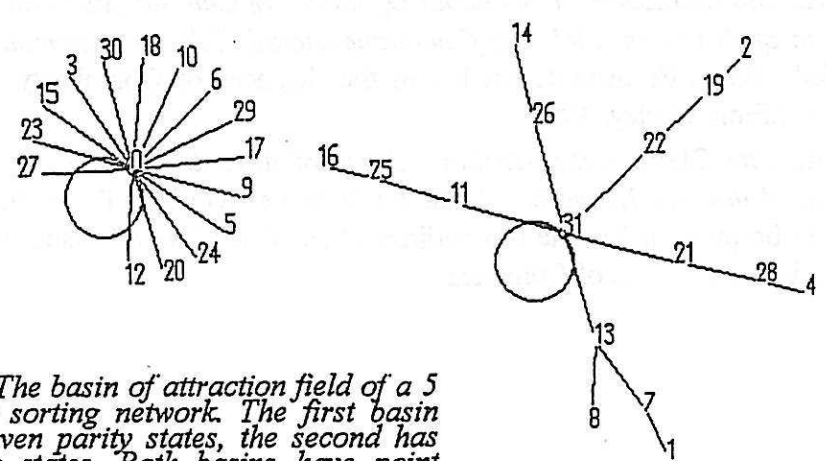


Figure 6. The basin of attraction field of a 5 cell parity sorting network. The first basin has only even parity states, the second has odd parity states. Both basins have point attractors. Bit patterns are represented by decimal numbers.

### Acknowledgements

I am grateful to colleagues at the Santa Fe Institute and the University of Sussex for discussions and comments.

### References

Alexander, I., W. Thomas and P. Bowden, *WISARD, a radical new step forward in image recognition*, *Sensor Review* (1984) 120-4.

Grey, C.M., P. König, A.K. Engel, W. Singer, *Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties*. *Nature* 1989, 338:334-337.

Hameroff, S.R., J.E. Dayhoff, R. Lahoz-Beltra, S. Rasmussen, E.M. Insinna and D. Koruga, *Nanoneurology and the Cytoskeleton: Quantum Signaling and Protein Conformational Dynamics as Cognitive Substrate*, in *Behavioral Neurodynamics*, K. Pribram and H. Szu, eds., Pergamon Press, in press.

Hopfield, J.J., *Neural networks and physical systems with emergent collective computational abilities*, *Proceedings of the National Academy of Sciences* 79 (1982) 2554-2558.

Kauffman, S.A., *Emergent properties in random complex systems*, *Physica D*, vol 10D (1984), 146-156.

Kauffman, S.A., *Requirements of evolvability in complex systems; Orderly dynamics and frozen components*, in *Complexity, Entropy and the physics of information*, ed. W.H. Zurek. Santa Fe Institute Studies in the Sciences of Complexity, Proceedings Vol. VIII, Addison-Wesley, 1989, 151-192.

McCormick, D.A., *Membrane Properties and Neurotransmitter actions*, in *The Synaptic Organization of the Brain*, 3rd edition, Oxford University Press, 1990, 32-66.

Shepherd, G.M., ed, *The Synaptic Organization of the Brain*, 3rd edition, Oxford University Press, 1990.

Singer, W., *Response synchronization of cortical neurons: An epiphenomenon or a solution to the binding problem?* *IBRO news* 1991, 19:6-7.

Wolfram, S., *Statistical Mechanics of cellular automata*, *Review of Modern Physics*, vol 55, no 3 (1983) 601-64.

Wuensche, A., and M.J. Lesser. *The Global Dynamics of Cellular Automata; An Atlas of Basin of Attraction Fields of One-Dimensional Cellular Automata*, (diskette included), Santa Fe Institute Studies in the Sciences of Complexity, Reference Vol. I, Addison-Wesley, 1992.

Wuensche, A., *The Ghost in the Machine; Basin of Attraction Fields of Disordered Cellular Automata Networks*, Santa Fe Institute Working Paper 92-04- 017, 1992, to be published in the proceedings of *Artificial Life III*, Santa Fe Institute Studies in the Sciences of Complexity.