Basins of Attraction in Disordered Networks

Andrew Wuensche
Santa Fe Institute

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

Abstract
An explicit portrait of *basins of attraction* in *disordered cellular automata networks* are accessible for the first time. Such networks are discrete generalisations of connectionist models. Basin of attraction (fields) may serve as the basis of a mind model, and *sculpting* basin fields may offer a new approach for brain-like computation.

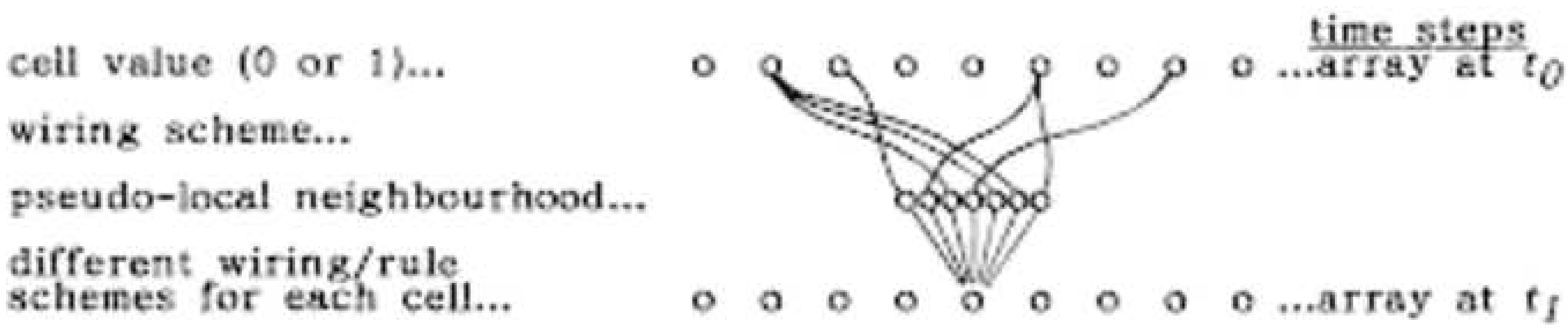
1. Introduction

The basin of attraction fields of *local*, 1-D cellular automata (CA) were presented in *The Global Dynamics of Cellular Automata* [1]. The investigation has been extended to *disordered CA networks*, in *The Ghost in the Machine* [2], of which this paper is a selective summery. Variations on such systems have been studied by Kauffman [3], Walker [4] and Li [5], amongst others.

Different *wiring/rule* schemes in such disordered CA networks result in different basin field structure, suggesting a paradigm for emergent brain-like information processing based on categorisation of input; the basin of attraction field (the ghost in the machine) categorises input hierarchically at many levels, and may serve as a dynamical mind model for understanding cognitive processes such as memory and learning.

A learning algorithm is described which enables the network to learn any number of new *pre-images* to any given global state (and to forget old ones) by small adjustments to its wiring/rule scheme. This opens up the possibility of *sculpting* a basin of attraction field to achieve any desired structure. Disordered CA networks may thus offer a new approach to neural network models for brain-like computation.

2. Disordered CA architecture



A disordered CA network generalises *local* CA architecture [1,2], but disrupts any notion of *space*. It is represented above. The cells are shown as a 1-D array for convenience only; their locations are arbitrary.

A global state of the array is the pattern resulting from the individual values assigned to each array cell, from a finite range of permitted values, generally taken to be binary. At each discrete time step, each cell at t_1 synchronously updates its value as a function of the values of a randomly located set of n reference cells at t_0. The system continues to evolve by the iteration of this global updating procedure, the *wiring/rule scheme*.

In order to define precisely the network's particular wiring/rule scheme, the target cell is wired to the reference cells from a *pseudo-local neighbourhood* of size n, to which a conventional CA rule will apply. The wiring scheme and rule allocated to each target cell may be different, but are fixed over time. The total number of distinct global states (the system's *state space*) = k^L, where L is the array size and k the value range. There are k^n permutations of values in a neighbourhood of size n. A rule table (look up table) with k^n entries will specify the output of all neighbourhood permutations. The total number of distinct rule tables, the size of rule space = ${}_k k^n$. This paper generally limits the value range to binary, where $k=2$. The total number of alternative wiring/rule schemes for an array of size L with connectivity n, equals...

the total wiring schemes x the total rule schemes = $(L^n)^L \times ({}_k k^n)^L$

Thus in a binary CA with an array of 16 cells, with 5 wires per cell, the size of behaviour space equals...

$$(16^5)^{16} \times (2^{32})^{16} = 2^{320} \times 2^{512} = 2^{832}$$

Even for small systems, the behaviour space is vast, and increases at a multiple exponential rate with increasing L, n and k.

Note that a sparsely connected neural network with weighted connections may be re-interpreted as a randomly wired CA network by discretizing its weights, and replacing them with multiple discrete connections to a large pseudo-local neighbourhood [2]. The same threshold rule is then applied to all cells. As we have seen, however, in a disordered CA network any other arbitrary mixture of rules from the ${}_k k^n$ in rule spaces is possible.

3. Basins of Attraction

Both local CA and disordered CA networks are discrete 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. The path inevitably leads to a state cycle (the attractor cycle, or *attractor*). The set of all possible paths leading to the same attractor, including the attractor itself, make 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 attractor cycles.

Basins of attraction are portrayed as computer diagrams (*state transition graphs, networks of attraction* [1]), in the same graphic format as presented in [1]. 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* ("in degree"). Fig.1 illustrates a typical basin of attraction, (it is part of the basin field shown in fig.2).

Separate basins of attraction typically exist within state space. A CA transition function will, in a sense, crystallise state space into a set of basins of attraction, known as the *basin of attraction field*. The basin of attraction field is a mathematical object which constitutes the dynamical flow imposed on state space by the transition function. If represented as a graph the field is an explicit global portrait of the systems entire repertoire of behaviour.

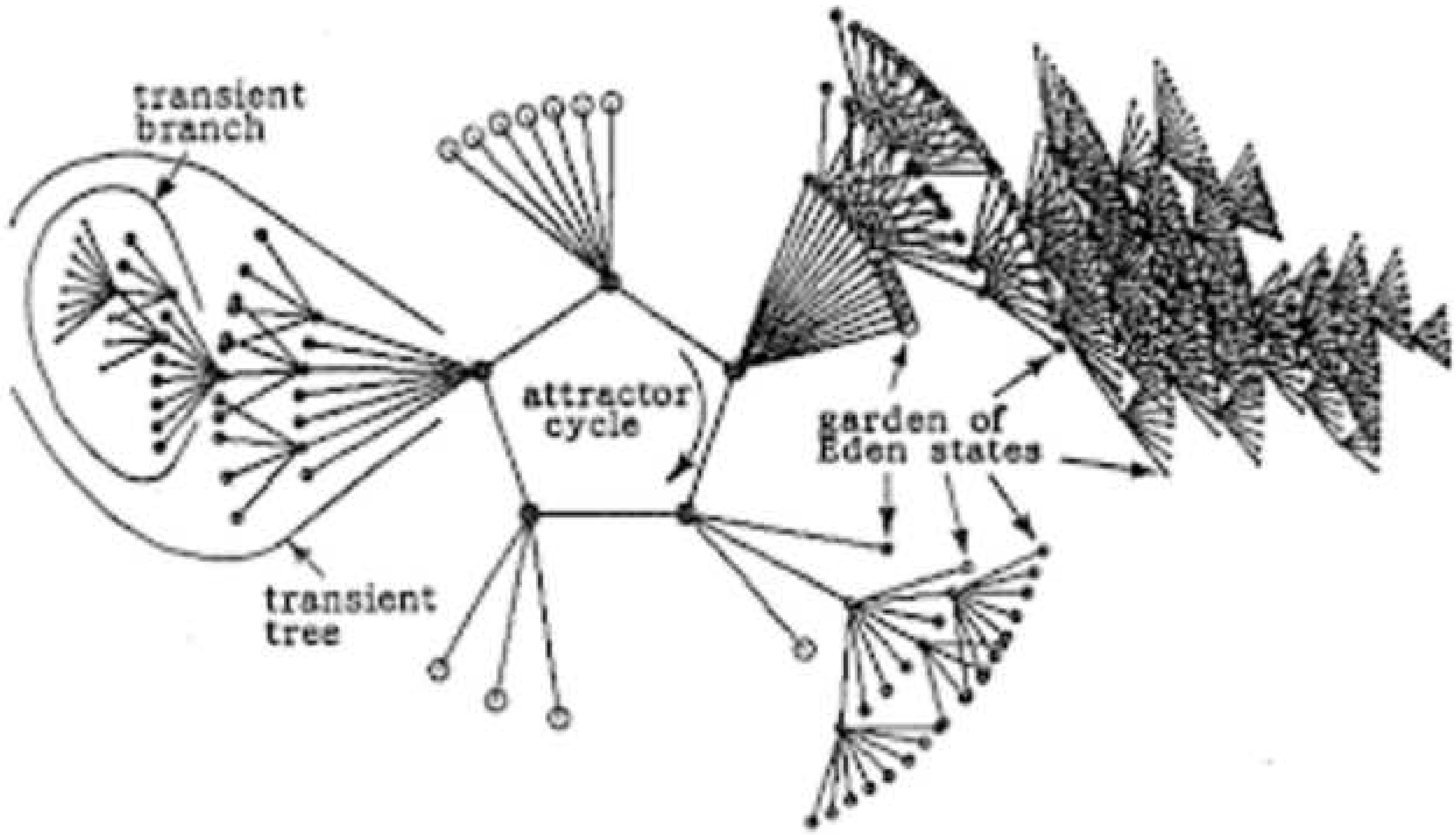
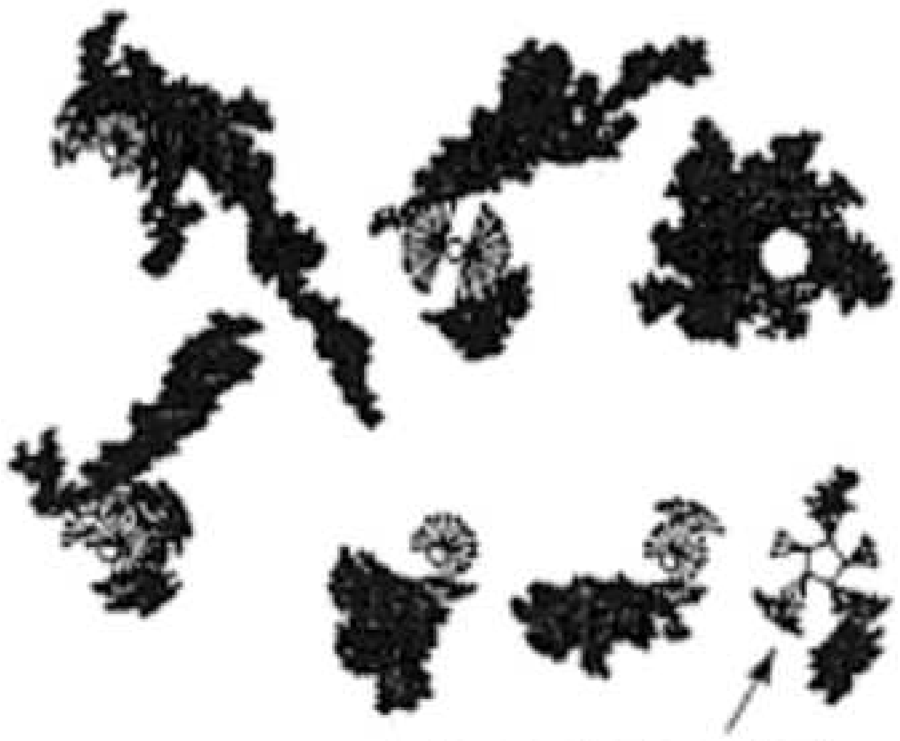


Fig.1. A basin of attraction (state transition graph) with 736 nodes. Evolution proceeds inward from garden of Eden states, then clockwise. the basin is indicated within the basin of attraction field in fig.2.



Cell	Wiring		
1.	15,	15,	1
2.	15,	10,	1
3.	12,	3,	12
4.	6,	10,	9
5.	13,	13,	15
6.	2,	12,	4
7.	1,	7,	6
8.	5,	14,	15
9.	15,	4,	12
10.	14,	10,	12
11.	3,	12,	12
12.	2,	6,	15
13.	13,	3,	7
14.	7,	5,	2
15.	14,	2,	9

Fig.2

The basin of attraction field. The basin in fig.1. is indicated. Randomly wired, single rule, 3-neighbour rule 108, L=15. The field consists of 7 basins of attraction. The total number of states in each basin is as follows: 9100, 8136, 3788, 5520, 3220, 2268, 736. The wiring scheme to the pseudo-local neighbourhood is shown on the right.

4. 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 impractical in terms of computer time as the array size increases beyond modest values.

A *reverse algorithm* that directly computes pre-images for *local* CA was presented in [1], and a *general* direct reverse algorithm for disordered CA networks in [2]. Providing that $n<L$, 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.

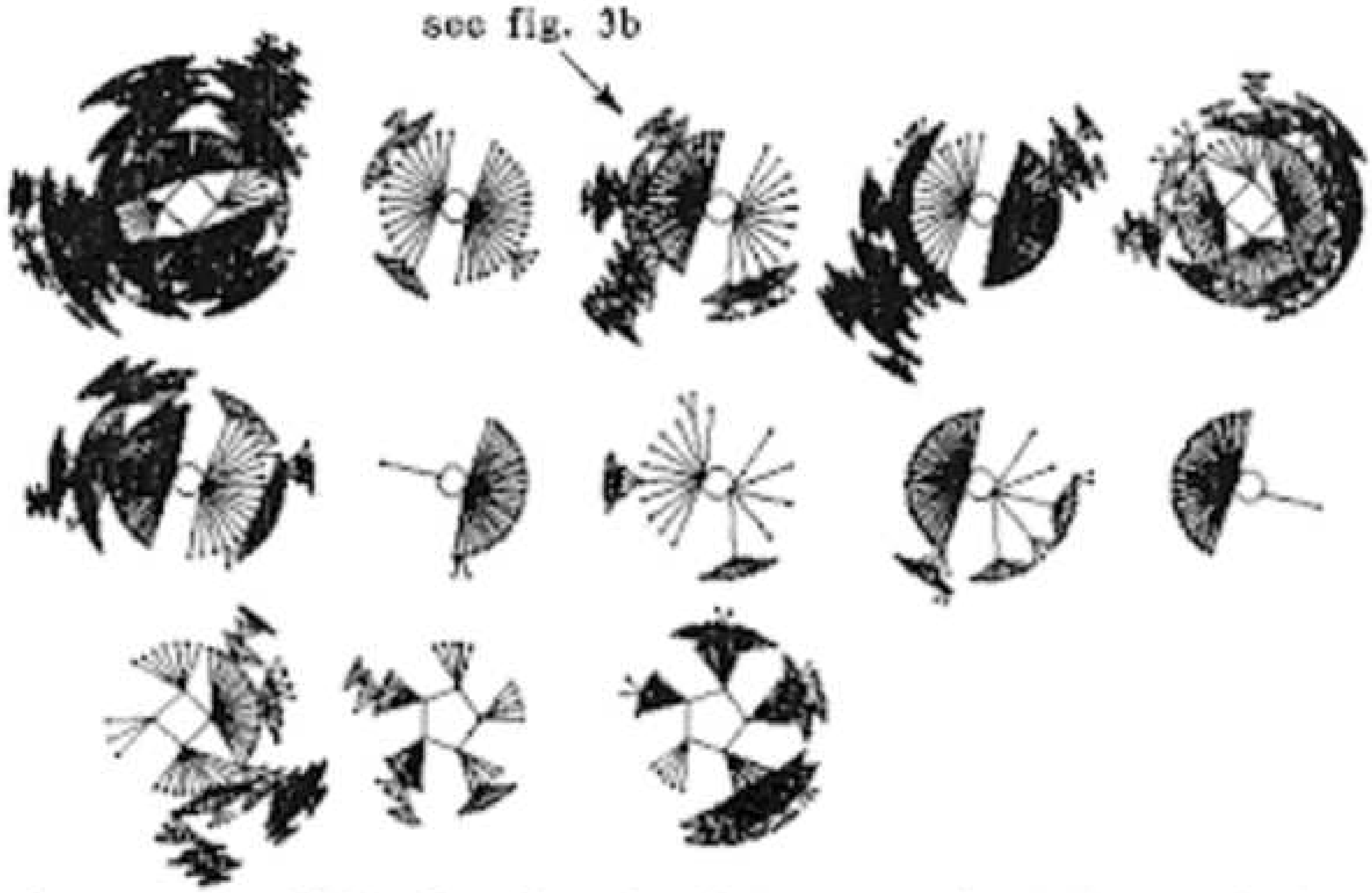
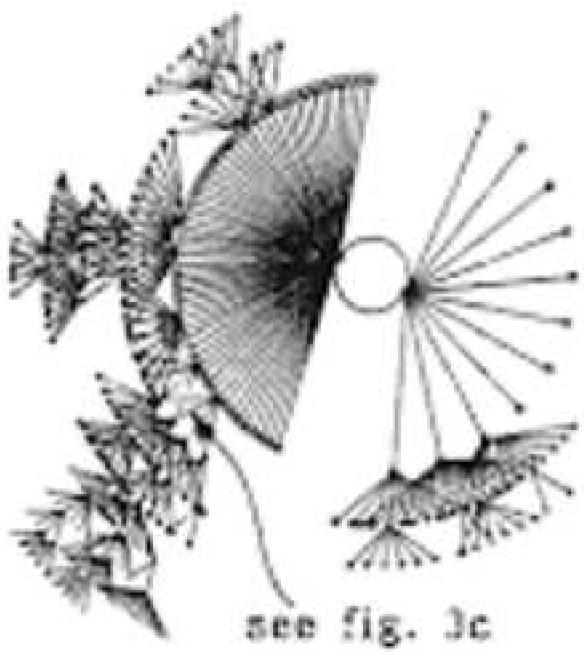


Fig. 3a

The basin of attraction field of a disordered CA network, L=13, consisting of 13 separate basins. The total number of states in each basin is as follows: 3072, 80, 432, 1872, 628, 812, 34, 62, 114, 62, 512, 84, 428. Wiring to the pseudo-local neighbourhood (see section 2), and the rule scheme is shown in the table below.



cell	wiring			rule
1.	3,	12,	6	87 - 01010111
2.	7,	11,	4	4 - 00000100
3.	3,	3,	1	194 - 11000100
4.	11,	3,	9	52 - 00110100
5.	8,	7,	5	235 - 11101011
6.	1,	8,	1	101 - 01100101
7.	12,	4,	13	6 - 00000110
8.	8,	6,	8	101 - 01100101
9.	9,	2,	6	6 - 00000110
10.	5,	1,	1	95 - 01011111
11.	2,	7,	1	74 - 01001010
12.	7,	8,	4	215 - 11010111
13.	1,	4,	7	189 - 10111101

Fig. 3b

The third basin in the basin of attraction field above (fig 3a) drawn at a larger scale. The transient branch detailed below is indicated.

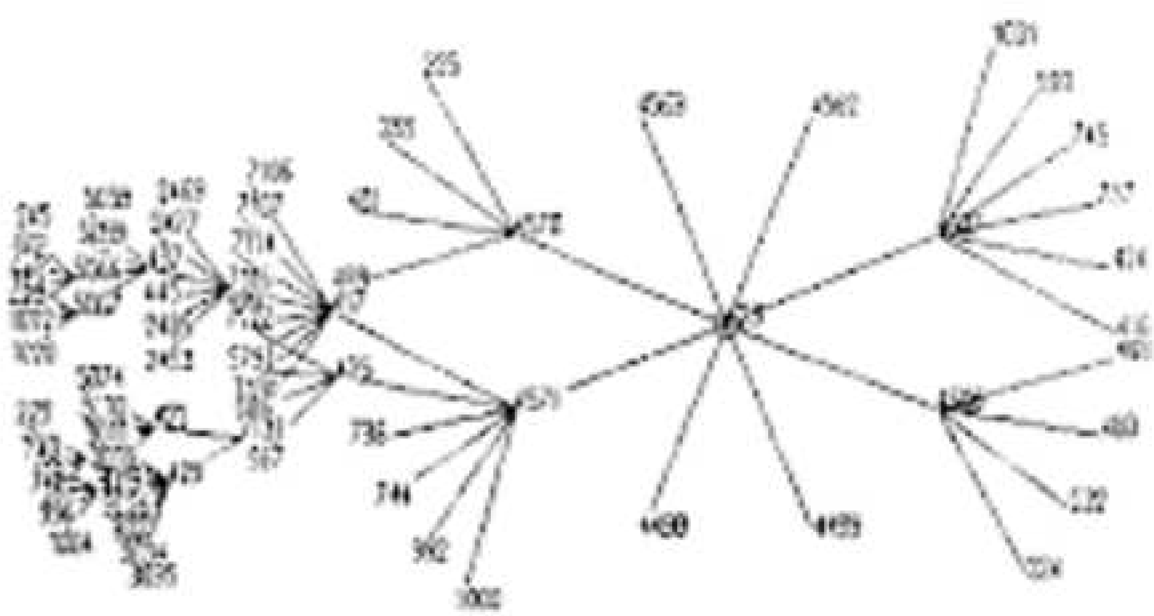


Fig. 3c

The transient branch indicated in the basin of attraction above (fig 3b), showing the decimal equivalents of the binary CA states at each node.

5. Brain–Like Computation

Because of the many possible permutations of wiring/rule schemes, one may conjecture that a wiring/rule scheme for a population of cells exists that would result in a basin of attraction field that will categorise any set of inputs onto an appropriate output; a learning algorithm to achieve this is outlined in [2]. Separate basins in the field, and each node onto which dynamical flow converges, are potential information storage and recognition systems.

Fig.3a shows an example of a basin of attraction field that has categorised state space into 13 basins, the most fundamental order of categorisation. Fig 3b shows the third basin at a larger scale; fig.3c shows a *transient branch* in the basin, showing the decimal equivalents of the binary CA states at each node.

Further categorisation takes place within each basin. Any input will immediately initiate a dynamical flow along a unique chain of states (the transient trajectory). Each successive transient state categorises potential input from states that belong to its transient branch. At each successive state the transient branch may expand, increasing the proportion of states categorised, thus forming a hierarchy of categorisation culminating at the attractor.

6. A Mind Model

The basin of attraction field, consisting of one or more basins, categorises input at many levels, dependent entirely of the wiring/rule scheme of the CA network. New information may be stored (re-categorised) by adjusting the rule/wiring scheme, analogous to learning. Once learned, the system reacts to input automatically, directly homing in on an appropriate sequence of outputs, without the need to sequentially search memory addresses as in conventional computer architecture. This is the process that operates in auto-associative neural network models, and possibly in the brain.

A disordered CA networks may thus serve as an idealised model of the activity of a semi-autonomous population of inter-connected neurons in the brain. The model is elaborated in [2] by inter-connecting many semi-autonomous networks so that they are able to reset each other's initial state. Networks activate each other asynchronously, and at a slower frequency than a particular network's internal synchronous clock. Such a nested hierarchy of networks 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 capable of categorising and re-categorising information - a mind model.

7. Learning

Whether or not such a model is biologically plausible, it may be useful in its own right as a *transparent* connectionist computational system, where learning is equivalent to *sculpting* the system's basin of attraction field.

A CA network can learn (and also forget), by small adjustments to its wiring or rule scheme, particular transitions between global states. This allows the possibility of finding the appropriate wiring/rule scheme to produce any desired set of pre-images to a given global state. It may be conjectured that given any desired structure of a transient branch, transient tree, basin of attraction, or basin of attraction field, a wiring/rule scheme can be evolved by degrees that will result in progressively closer approximations to that structure.

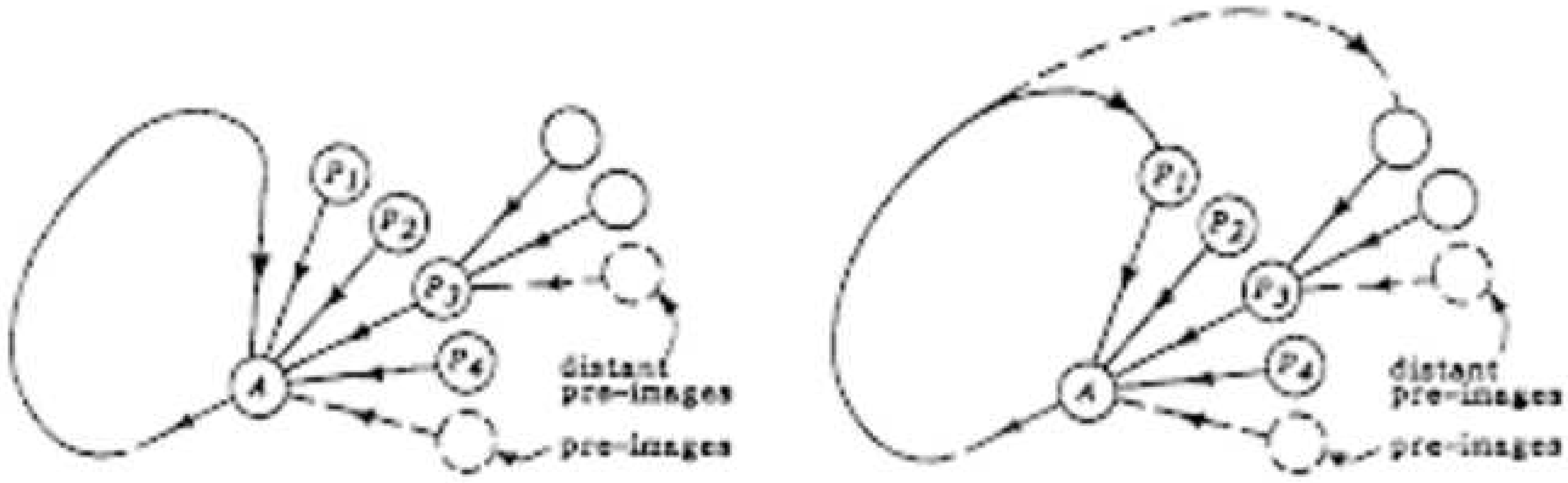


Fig.4 a) Learning the global state *A* as a pre-image of itself, making a point attractor. b) Learning the global state *A* as a distant pre-image of itself, making a cyclic attractor.

The network can learn either by re-wiring or by mutations to the rule scheme. Re-wiring is a more fundamental adjustment, as will be shown below, analogous to changes in a neuron's synaptic connections so that the set of neurons sampled is slightly altered. Mutating the rule scheme is potentially a finer adjustment, analogous to changes in the topology of the dendritic tree, the microcircuitry of synaptic placements and intrinsic membrane properties; a neuron's output is thought to result from a non-linear computation dependent on these factors [6]. There appears to be no shortage of biological mechanisms to permit each neuron to express a rich and flexible repertoire of computation, which is modelled by the rule table.

The learning algorithm is described in detail in [2]. The capacity of the network to learn new pre-images by re-wiring (without forgetting those previously learnt) dependents on a number of factors: the original wiring/rule scheme, the similarity of the new pre-images, the size of the network and the extent of connectivity. However, the network may have additional capacity to learn *distant* pre-images, further upstream in the transient tree as in fig 4. Note that if the network learns the given state itself as its own pre-image, this will result in a point attractor as in fig.4a; the state itself learnt as a distant pre-image will result in a cyclic attractor with a period equal to the distance as in fig 4b.

Surprisingly, in learning by mutating the rule scheme, it turns out that there is no limit to the number of pre-images to a given state that may be learnt by the network. The network can learn more pre-images without any risk of forgetting previously learnt pre-images; in the limit all states in state space (including the given state itself) may be learnt as preimages.

Combining learning by rewiring and learning by mutating the rule scheme may result in a powerful method of categorising input, recognition, and cumulative learning in disordered CA networks.

References

1. Wuensche,A.,and M.J.Lesser. *The Global Dynamics of Cellular Automata*, Santa Fe Institute Studies in the Sciences of Complexity, Reference Vol.I, to be published by Addison-Wesley, June 1992.
2. Wuensche,A. *The Ghost in the Machine, Basin of Attraction Fields in Disordered Cellular Automata Networks*, Santa Fe Institute Working Paper, 1992.
3. Kauffman,S.A., *Emergent properties in random complex automata*, Physica D, vol 10D (1984), 146-156.
4. Walker, C.C., *Attractor dominance patterns in random-structure cellular automata*, (1990), to be published in Physica D.
5. Li,W., *Phenomenology of Non-Local Cellular Automata*, Working Paper 91-01-001, Santa Fe Institute, Santa Fe, NM, Jan. 1991.
6. Shepherd, G.M., (editor) *The Synaptic Organization of the Brain*, third edition, Oxford University Press, 1990.