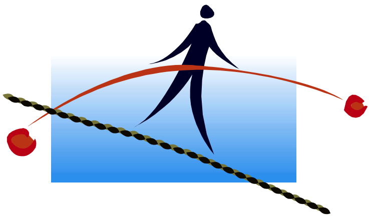
The Robust Organization



Highly optimized tolerance

Max Stewart

Max Stewart

The Robust Organization

Highly optimized tolerance

Published by Decomplexity Associates Ltd
First published 2003

Set in Times New Roman

ISBN 0-9540062-4-0 (European Acrobat edition)

Also available in US Acrobat edition:
ISBN 0-9540062-5-9

and in paperback from booksellers:
ISBN 0-9540062-3-2 (European edition – bound with colour illustrations)

This book is a supplement to the author's *The Coevolving Organization – poised between order and chaos* that is available in several forms:

through booksellers:
ISBN 0-9540062-0-8 (European edition – bound, with full-colour plates)

copyable free from www.decomplexity.com, with different line illustrations and without colour plates:
ISBN 0-9540062-1-6 (European Acrobat edition)
ISBN 0-9540062-2-4 (US Acrobat edition)

AUTHOR

Max Stewart was educated at the Universities of Wales and Cambridge. He wrote the first and widely praised non-specialist account of the application of relational database principles to systems design – something that later became better known as Data Analysis. He was at one time IT Director for the Scottish operations of Leyland Vehicles and later spent many years with Mars, Incorporated. He is a Principal with Decomplexity Associates and lives in Rutland, England's smallest county.

COPYRIGHT AND TRADEMARKS

CONTENTS

Figures

PREFACE

This short book is a supplement to the author's *The Coevolving Organization – poised between order and chaos* published in 2001. This earlier book tried to answer one fundamental business question – how decentralized should an organization be? – using developments in physics and theoretical biology which emerged during 1988-1995. It described how businesses could be positioned, poised and reactive, on the boundary between stodgy stability and decentralized anarchy – using the concepts of 'edge of chaos' (EOC) and 'self-organized criticality' (SOC). However, over the last five years, something new and related has appeared on the horizon: *highly optimized tolerance* (HOT). HOT does not supersede EOC and SOC. Instead, it allows us to exploit the idea of 'decoupling' parts of an organization (divisions, departments, even individuals) such that the decoupled parts can be even more responsive than with EOC/SOC. HOT also highlights the role of deliberate *design* – the antithesis of self-organization. Such self-organization or, alternatively, restructuring using a simple and limited amount of management intervention, can be attempted following the EOC/SOC principles outlined in *The Coevolving Organization*. But if a business is decoupled further using HOT principles, it is possible for the decoupled parts to be even more responsive than would be possible with the EOC/SOC ideas alone. It implies minimising how the decoupled parts can affect one another and having a good understanding of the likely business risks to which each part is subject. In Stu Kauffman's NKCS terminology, lowering the C-coupling between parts allows us to lower the K-complexity of the parts. And the latest HOT findings indicate that a business consisting of many freewheeling (very-low-K) parts can retain coherence and overall stability if the parts connect not to each other but via 'anchor' parts that are more stable (have more high-K 'treacle') and live just below the order – chaos boundary.

Investment banking 'quants' and the more analytical fund managers may have the right mathematical background to read the source texts upon which this book and *The Coevolving Organization* are based. Most business managers do not. This is unfortunate as the concepts are powerful in their own right and provide a way – a 'language' – for managers to discuss and analyze critically the structures of their businesses. The role of the present book is to bridge the gap and show the applicability of HOT to how businesses behave. A guided tour of the relevant academic papers is also included for those such as MBA students who wish to take things further.

The author is indebted to His Grace the 1st Duke of Wellington whose victories in the Iberian Peninsular War have been elaborated to describe the differences between the normal and PLR formulations of HOT.

Max Stewart
Rutland, UK
September 2003

CHAPTER 1

INTRODUCTION

Background

This book takes over from where *The Coevolving Organization* left off. The latter was based upon some developments in condensed matter physics, theoretical biology and economics up to 1998. Things have moved on since then, and a new and related concept – highly optimized tolerance (HOT) – has been introduced. The use of HOT as an extension of the principles outlined in *The Coevolving Organization* should allow organizations to decentralize decision making further than was possible using the 'edge of chaos' and 'self-organized criticality' ideas alone.

Hierarchies and self-similarity

All but the smallest businesses are organized in hierarchies. A parent (holding) business is, perhaps, composed of many operating businesses. Each of these in turn is composed of facilities such as manufacturing sites (with their own employees) plus cross-site staff functions such as Finance and HR. Sales and Marketing teams are probably country-based but with some global or regional marketing staff. Each such group, whether based on geography or function, is in turn built up from divisions, departments, sections and so on down to the lowest-level employee or contractor. Any outsourced operations (such as co-manufacture or IT Service Delivery) have their own parallel hierarchies. Each such hierarchy is what a physicist would call self-similar: a small part such as an asset-accounting team in a small factory is structured and behaves in somewhat the same way as its parent global Finance organization headed by the CFO (the *entire* Finance organization – not just the CFO and his or her direct reports). Within limits, the largest organization unit behaves in roughly the same way as the smallest and lowest within it and every level in between. From our point of view, what matters is the *behaviour* of these groups – especially how they communicate and make decisions; their family tree structure on paper is of lesser importance. But, as described in *The Coevolving Organization,* this similarity of behaviour is closest at the order – chaos boundary. Differences between the behaviour of organization units become blurred at this critical point, and it could be expected that at least some of the behaviour predicted by physicists for potentially critical systems such as forest fires, traffic jams and the now famous 'sand pile' would apply to hierarchical organizations also.

EOC and SOC – recap

The Edge of Chaos (EOC) is Stu Kauffman's phrase for what *The Coevolving Organization* called the 'order – chaos boundary' and is otherwise called the 'critical point' or 'criticality'.
'Self-organization' simply means 'self-adaptive organization' and a self-adaptive system is one that modifies its behaviour without external help. It usually has some goal to seek and some means of feedback to monitor how far away it is from the goal.
'Self-organized criticality' (SOC) is self-adaptation where the goal is the order – chaos boundary.

And even if the coevolving parts of a business, or of competing businesses evolving with each other, did not display self-similarity in their behaviour, the idea of achieving and maintaining an 'active' business poised at its critical point – the boundary between order and chaos – is intuitively attractive: the business is then, by definition, as reactive as possible without degenerating into the chaos which would result if each department (or, in the extreme, each individual) operated without reference to any of the others.

Design

There is, however, a different way to look at an organization. Its hierarchy may look self-similar and may even behave in a self-similar way. But it is also *designed*. Its organization is not random but is structured and restructured by its managers. The business processes too are man-made, and it is these that specify the levels of decision-making: who can decide what and with whose agreement. Although the organization may look both internally self-similar and similar to other (perhaps competing) organizations, it is actually the product of explicit design and has lots of tuning knobs which management can tweak. This alternative – and complementary – perspective is a cornerstone of what follows.

How far can we decentralize?

When *The Coevolving Organization* was written, one salient question remained unanswered:

> Could we deliberately design an organization where each coevolving 'object' (a department, a sales team or whatever) had its decision making decentralized *even further* than its critical point (to the 'chaos' side of the order – chaos boundary) while at the same time any negative side-effects on the rest of the organization resulting from its new and excessive freedom were mitigated?

To use some analogies, can we design a forest such that the trees are planted as densely as possible (thus maximising the yield from commercial forestry) while minimising the impact of a forest fire which could totally destroy a dense forest but could leave a sparsely planted one more or less unscathed?

Public Health Authorities responsible for containing epidemics face a similar problem, and IT practitioners build defences against electronic versions of virus epidemics. If incoming virus-infected email managed to breach a business's outer electronic defences and is opened by the addressee, a single shared central email system containing a mailbox for everyone in the business will spread the virus quickly and the latter will be difficult to contain. Such a system is very effective for the business: mail moves around quickly and with no delay. But it is this very effectiveness that is highjacked by viruses for their malevolent purposes. Say, on the other hand, all internal email were held temporarily in a 'pending transmission' queue for half-an-hour or so after the user had clicked the Send button, and only then were forwarded. The IT people then have a chance to quarantine infected email by freezing the 'pending' queues before the infection spreads too far. Deliberately inserting delays into email transmission would, however, spoil the responsiveness of the business since such delays are enforced in times of health as well as infection. Managing risk thus has a cost, as does creating firebreaks in forests. If, moreover, infected email could cause the central email computer itself to fail, there would be a positive advantage in using several smaller central email computers, with each computer providing mailboxes for a group of users who communicate with each other a lot but with other groups less often. The only pending-queue delay then needed is for email being sent *between* email computers. If a virus hits one user on one central email computer, it can easily spread to the other 'closely coupled' users on the same computer; these other users are, by definition, the users with which the infected user communicates most. But the infection can then probably be isolated to that computer in the way a forest fire is isolated to an area between firebreaks. In this instance, during healthy times we have reduced the potential loss of effectiveness for the business as whole because the only delays are in email between users who communicate infrequently. And during times of infection, we have, with luck, contained the spread of infection to one group of users. We have, in other words, *balanced the likely impact of an infection with the cost of containing it*. This, as we shall see, is a key principle underlying HOT. (Note that this email example is simplistic because real "email" functions are usually split between a central shared computer and the user's own PC. Nevertheless, the process described above can actually be implemented.)

The price of risk management

In the foregoing, we have also implicitly assessed the likelihood of an infection or forest fire happening at all. If forest fires occur in a particular locality once every million years, one could simply ignore the risk. This 'likelihood' would probably be specified as a probability 'distribution': 'most likely' may indeed be one in a million years but it would also be possible but much less likely for a fire to break out once every thousand years or once every five million years. Protecting against loss, as we have seen, has a price. In a forest, it is the cost of firebreaks (the loss of valuable 'tree space' plus the cost of ensuring that the firebreaks remain clear). For email virus protection it may be the cost of using more email mailbox computers plus the cost in lost business effectiveness of any artificial delays which are imposed on email transmission in order to provide an opportunity to isolate any email infection. Such loss prevention or containment adds something not explicitly discussed in *The Coevolving Organization*: **design.**

Tuning knobs

Self-organization to the boundary between order and chaos, as described in *The Coevolving Organization,* relied on the automatic (self-adaptive) adjustment of one 'tuning knob': the internal K-complexity of each coevolving object (a department or whatever). And, in both theory and practice, even this is insufficient to drive most collections of coevolving objects to the order – chaos boundary (see Directed Organization versus Self Organization in Chapter 3 of *The Coevolving Organization).* Having only one thing to adjust for each coevolving object does not give us the freedom to hone the collection of objects – to refine their own internal (low K) fitness, their (C-coupling) interaction with other objects and the number (S) of other objects with which they interact – such that we can contain the impact of a disaster (i.e. untoward behaviour) in one affecting the rest.

The Coevolving Organization also described the use of *many* tuning knobs to move an organization to the order – chaos boundary. These could be business drivers or explicit redesign of the organization and processes to decentralize decision-making. But in neither case did *The Coevolving Organization* consider the use of tuning knobs to create organizational firebreaks. It was believed at the time that the order – chaos boundary was optimal in the sense that moving beyond it to an even more decentralized organization would, almost by definition, result in a less effective organization. This was because the latter would have gone beyond the point at which poise and responsiveness were balanced by a stability *just* sufficient to maintain coherence of the business itself (its high level financial goals and ethical principles, for example). Unknown at the time was that:

- driving a complex system such as an organization *beyond* the order – chaos boundary into the 'chaotic' area, thus allowing the business to take advantage of further improved responsiveness from even more decentralized decision-making

plus

- designing in sufficient 'treacle' (delays; lack of impact or responsiveness) *between* the coevolving objects

could be even better.

In NKCS language…

Deliberately reducing the C-coupling between coevolving objects further (to the 'chaos' side of the order – chaos boundary) and thus further decentralizing (decoupling) parts of the business enabled us to reduce the K-complexity of the participating objects. This gave a solution of even greater fitness: greater business effectiveness because of better responsiveness. Parts of the organization were thus allowed to go off more at a tangent and make their own decisions

NKCS – recap

An 'object' is anything that evolves of its own accord or in response to some external influence or both. For our purposes it may, for example, be a department that has cost drivers or other objectives to meet, or it may be competing with another department. Real departments probably have both incentives to respond to.

Each coevolving object has a number of 'genes' (N) coupled to each other *within* the object (K), coupled *between* objects (C), to a number of other objects (S) and optionally to W tuning knobs in the external world. The values of individual genes (which can be Yes/No or a numeric quantity) represent decisions taken.

while limiting the effect of any adverse decisions on the rest of the business. But this presupposes that the likely adverse behaviours are identified in advance and the C-coupling and K-complexity of each coevolving part of the organization are engineered to cope with them.

There remains the impact of *unanticipated* risks, for example the effect of the wholesale defection of a sales force to a rival. The impact of such risks is likely to be greater than if the organization stopped decentralizing when it hit the order – chaos boundary. In other words, *greater effectiveness* as a result of further decentralization brings with it *greater fragility to the unexpected*. This is a defining characteristic of HOT.

In summary, developments in theoretical physics from 1998 have indicated that if we have:

- several organizational tuning knobs to tweak
  and
- a good understanding of the likely risks and their impact

an organization can be pushed beyond the order – chaos boundary and be even more effective.

The remainder of this book describes how these new developments can be utilized. And following *The Coevolving Organization*'s example, material is appended summarising the background academic papers – but without using mathematics to do so.

CHAPTER 2

THE FOREST FIRE

Percolation

The forest fire is the most widely used example of 'percolation' – an event moving from neighbour to neighbour, like the Newton's Cradle executive toy where one swinging ball cannons into another which cannons into the next one, and so on. We will make use of a skeleton forest in which a fire moves from tree to neighbouring tree until it peters out or has obliterated the whole forest. The forest fire example differs from Newton's Cradle in that it is two-dimensional: trees have neighbours on all sides. For ease of illustration, a grid is used for siting trees: trees can thus have neighbours in the North, South, East or West but not at any intermediate position. Note that a fire can only spread from a tree to another tree which is its *immediate* neighbour: jumping gaps or setting alight a neighbour and *its* neighbour in one action is not allowed; it is up to the neighbour to set alight *its* own neighbours.

The grid below in Figure 1 is 50 x 50 and thus has sites for 2500 trees to be planted. The dark squares represent sites with trees and the light squares sites where trees could be grown but are currently vacant.

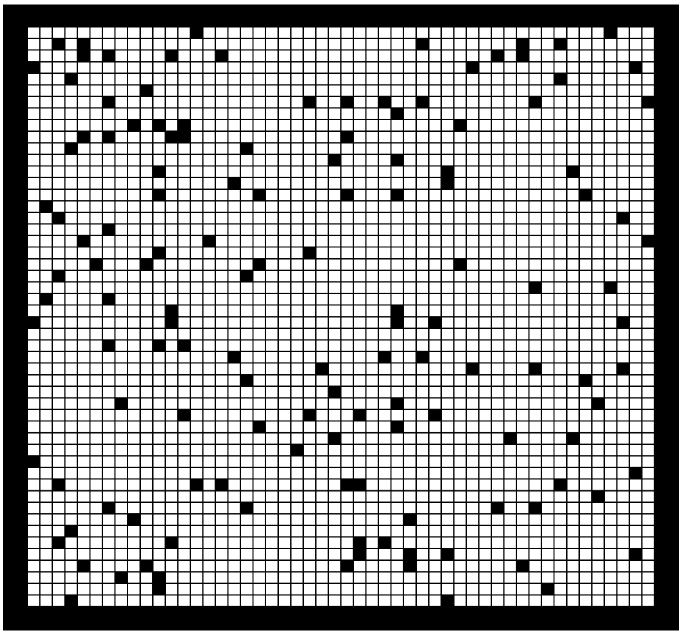


Figure 1 – The simple forest

Two salient measurements which characterise the forest and determine the way in which a fire will spread are:

- the density at which trees are planted – are they closely packed or, as in Figure 1, fairly sparse. Trees may be planted at random as in Figure 1. Or, alternatively, a commercial forest can be *designed* with trees planted in compact clusters with firebreaks (vacant sites) between the clusters.

- the likelihood (probability) of an external spark hitting a particular site. If the site contains a tree, it will then catch fire. The probability may be such that each site has the same likelihood of being sparked as any other. But there may also be parts of the forest that are more likely to be ignited than others – those near picnic sites for example.

These two factors:

- the planting density and, for a designed forest, the pattern in which the trees are planted
- the probability of a site receiving a spark (which, if the site contains a tree will cause it to ignite) and, if some areas are more likely to be sparked than others, the probability ‘distribution’ (which is a way to specify the tendency for sparking some areas more than others)

are at the heart of how percolation works for both self-organized criticality and HOT. Note the difference between sparking and ignition: a sparked site will only ignite if it contains a tree. This is an important difference in a sparsely planted forest.

(In NKCS terms, each tree has zero K-complexity and is C-coupled to each of S neighbouring trees, where S can lie between zero and four).

Optimization

The aim of our simplistic approach to forestry is to plant trees as densely as possible to maximize the ***yield*** from commercial forestry (the average number of trees left standing after a series of fires). But, as we shall see, and as readers of *The Coevolving Organization* may well remember, as we try to optimize, a point of decreasing returns is reached where the system fights back. In the case of our forest, dense planting makes a forest fire spread more easily and increases the potential loss (i.e. reduces the yield) because more trees are burned.

We will now give a series of examples of increasing complexity which show:

- how a ‘simple’ forest which has not been designed in any way demonstrates simple percolation or self-organized criticality

  and

- as design is introduced and the tree planting structured so as to reduce exposure to catastrophic fires by creating firebreaks (i.e. by decoupling clusters of trees from each other), the forest behaves differently and the density of trees increases from that obtainable with simple percolation or self-organized criticality

CHAPTER 3

SELF-ORGANIZED CRITICALITY

Controlled percolation and self-organized criticality

Using the somewhat idealized forest 'grid' introduced in the last chapter, assume that:

- the placement of trees is random
- the likelihood of a spark landing on any one square is the same as that for any other square, i.e. is random also

We shall call this *random percolation*. And the forest may be developed in one of two ways:

- through manual planting by a forester so as to achieve a desired density of trees. We shall call this ***controlled percolation***
- where there is no forester and the trees attain their own balance between burning and new self-seeded growth. This is ***self-organization***.

Controlled percolation

At the outset, let the forest be planted sparsely with just a few trees.

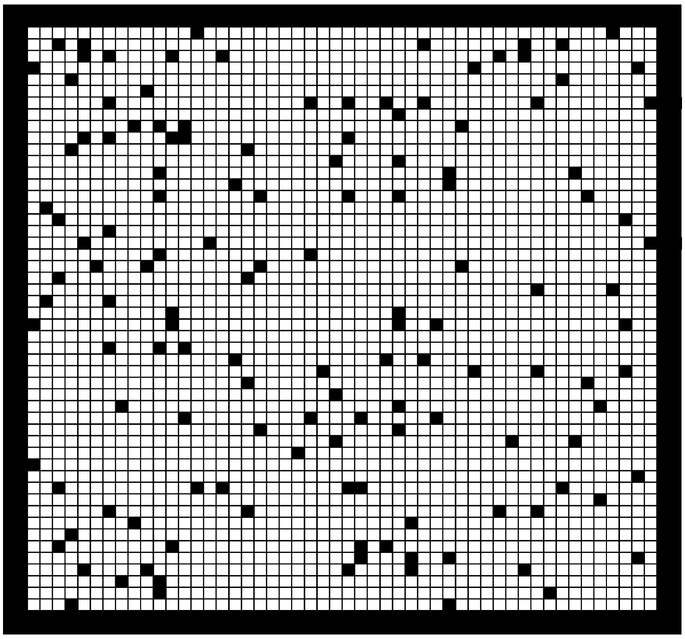


Figure 2 – Initial state

In this example, there are only nine clusters of two trees and one cluster of three. All other trees have space all around them.

Because only a few sites are occupied, the number of sparks hitting occupied sites (i.e. trees) and starting fires is small relative to the number of sparks that fall harmlessly on fallow sites.

Secondly, if a spark hits an occupied site and ignites the tree, the likelihood of the fire spreading beyond that tree is small because, as we saw above, most of the trees are isolated. The worst that could happen is that one of the trees in the cluster of three is ignited – we would lose three trees.

Assume that the forester plants more trees at vacant sites within the forest, *and that he or she selects the sites at random*. Figure 3 through Figure 6 show growth from a point where around 20% of sites are occupied to a point where around 60% are occupied.

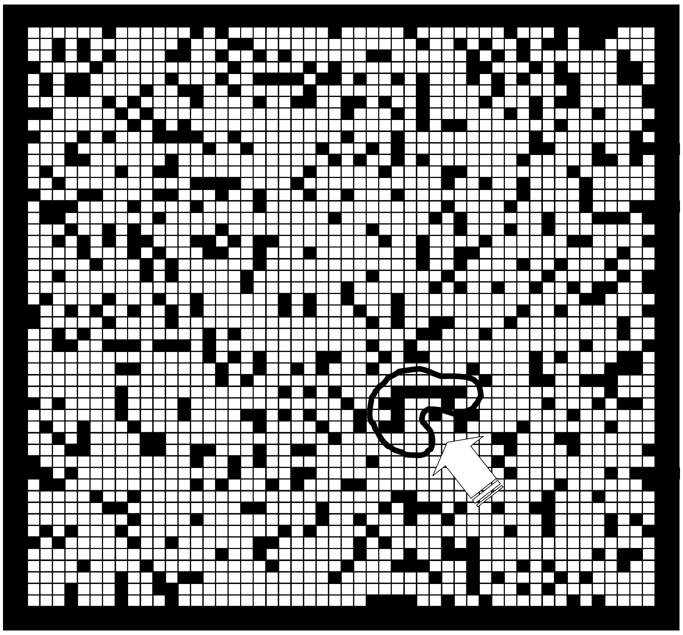


Figure 3 - 20% of sites occupied

In the Figure 3 forest, which is still relatively sparsely populated, most sparks strike vacant sites. The odd spark which strikes an occupied site will still generally do little damage and only the tree at that site will be burned. The worst damage a single spark can do is to ignite a tree in the cluster of twelve trees highlighted.

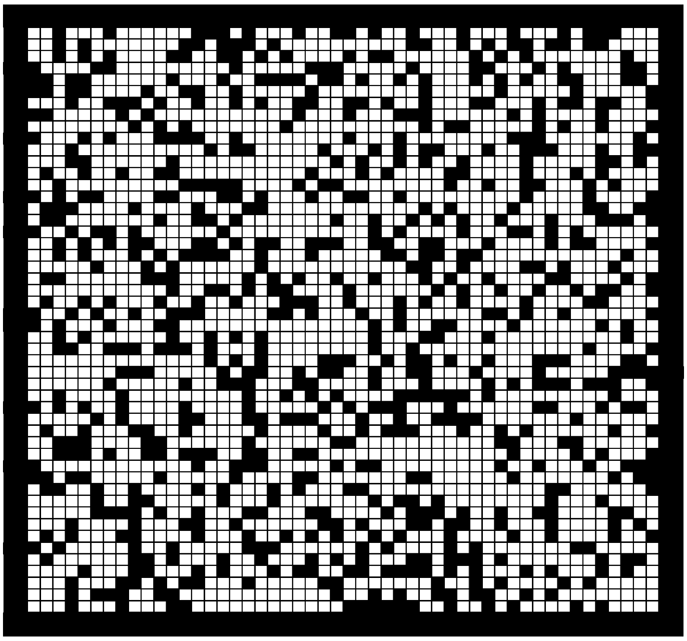


Figure 4 - 30% of sites occupied

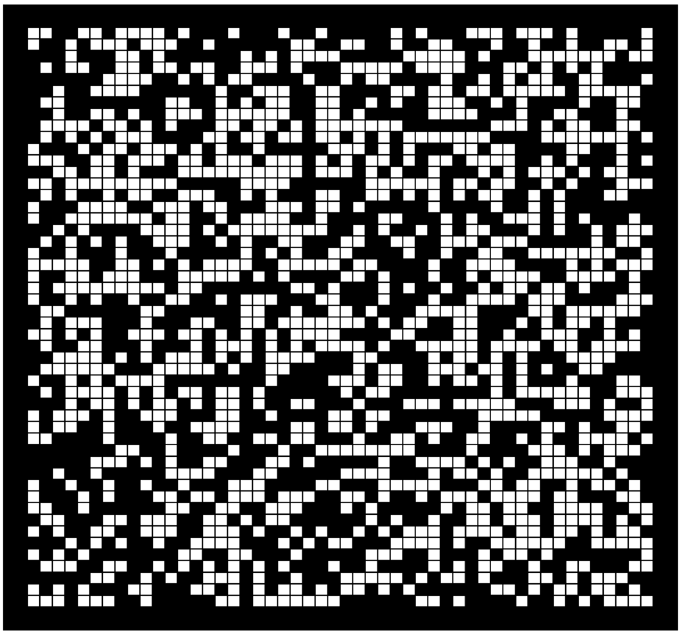


Figure 5 - 50% of sites occupied

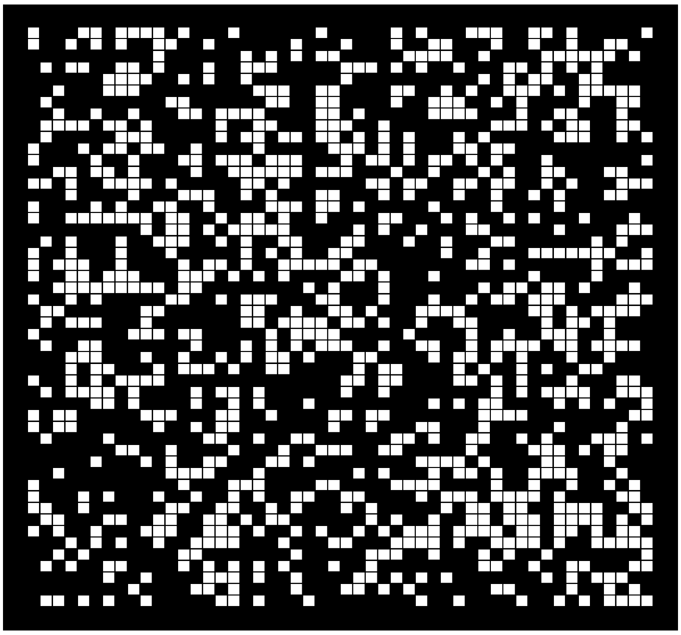


Figure 6 - 60% of sites occupied

Close inspection will show that as we move from the Figure 5 forest (50% occupied) to the Figure 6 forest, the sizes of clusters have increased radically: the density of trees has become such that existing clusters have joined. It would be possible for a squirrel to traverse the entire forest swinging from tree to tree from the bottom left-hand side to the top left-of-centre, or to the mid-right-hand side or, with the exception of a one-site gap, to the top right-hand corner also. A spark hitting a random site in such a forest may hit a vacant site, but is more likely to hit an occupied one. Worse, many sites are now part of large clusters, so hitting one of those sites causes many trees to ignite. But much worse is for a spark to hit a site in the one very large cluster (the '***percolating cluster***') described above which spans the forest in several directions. Since it is a large cluster, one of its sites is reasonably likely to be hit. And, moreover, the impact from doing so can be devastating. Attempts to plant trees randomly at densities greater than this critical density (around 59% of the sites) are doomed to failure since forest fires will then destroy a disproportionate number of trees and reduce the density back to the density at which the percolating cluster first appeared.

Readers of *The Coevolving Organization* should now be on familiar ground: the tree density at which the percolating cluster appears is the order – chaos boundary and the forest has been pushed by the forester to a state of criticality – but *not self-organized* criticality.

Self-organization

In the preceding example, we have assumed that a forester deliberately planted trees at random vacant sites so as to obtain a specific desired density of trees (for example, Figure 5 shows a density of 50% – half the sites occupied). But primeval forests did not have foresters. Instead, the growth rate of new self-seeded trees was roughly in balance with the rare incidence of trees burning: no-one was deliberately managing the planting to achieve any target density of trees. As would be expected, as the forest matures and the trees become denser, a point is reached where the infrequent forest fires prevent the density rising any further. The forest has organized itself to reach the order – chaos boundary without external intervention: it has attained self-organized criticality. The density at this point – at around 40% - is lower than that of similar controlled percolation but the behaviour is otherwise very similar.

But things are very different if we have more tuning knobs – more design parameters – to play with than merely the density of trees (for controlled percolation) or the balance between the rates of tree growth and tree burning (for self-organization).

CHAPTER 4

HIGHLY OPTIMIZED TOLERANCE

From random to designed percolation

We saw in the last chapter how a forester could plant trees in random vacant sites such that the forest eventually became dense enough to hit the order – chaos boundary and could go no further. We also saw how a maturing forest could move itself to the order – chaos boundary and achieve a state of self-organized criticality. The behaviour of the forest in both cases was very similar in spite of their arriving at the order – chaos boundary through separate routes. But both assumed that trees were either planted at random vacant sites or self-seeded randomly, and this raises one obvious question:

If we allow the forester to ***design*** the forest by specifying exactly where he or she plants the trees (as opposed to planting at random sites), or if some form of evolution though natural selection can do likewise, will the forest have a greater yield?

The answer, with some caveats, is Yes, although surprisingly the precise mechanism was not elucidated until 1998.

For our purposes, we can rephrase the question as:

where does the forester place firebreaks to maximize yield?

Firebreaks can range from simple lines of vacant sites stretching vertically and horizontally across the forest to lines of contiguous vacant sites in almost any pattern. The number, placement and shape of firebreaks are tuning knobs available to the forester in addition to tree planting density. And, as we shall see, the firebreaks need not be simple lines of vacant sites: they may be wider than a single site but sparsely occupied by trees.

In the two contrasting examples that follow, the first is the forest at the order – chaos boundary (arrived at by either controlled percolation or self-organization) and the second is the same forest with firebreaks:

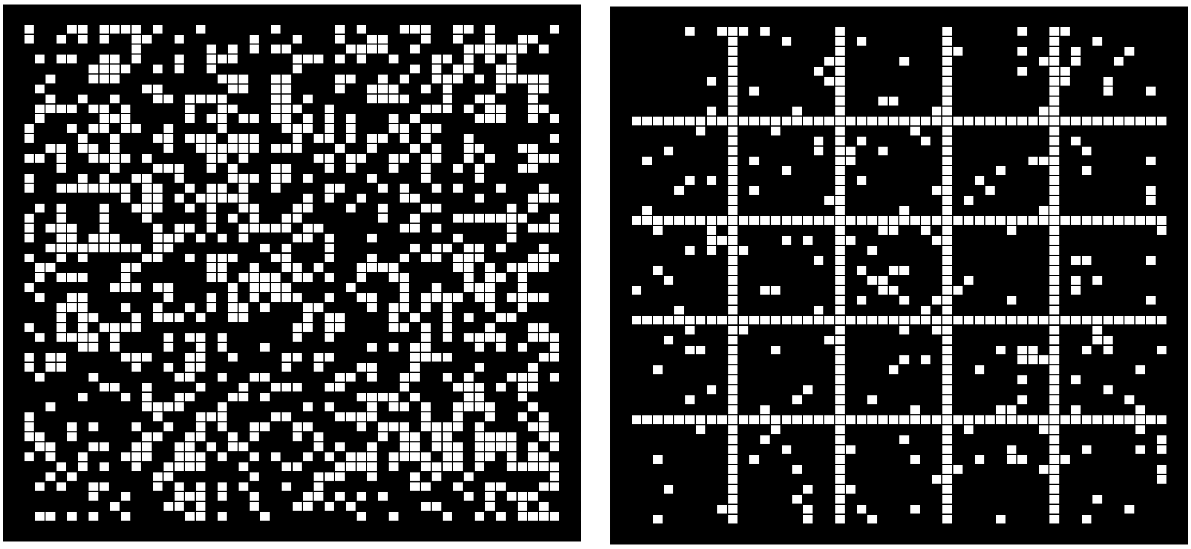


Figure 7 – Self-organized critical forest vs. forest with firebreaks

In the right-hand example in Figure 7, the firebreaks allow the 25 isolated areas of trees to grow to a higher density than would be possible in the self-organized or controlled percolation forest. The firebreaks stop large clusters from developing. Including the vacant sites which constitute the firebreaks, trees occupy around 76% of the forest sites.

This leads to a related question: what is the optimum number and positioning of firebreaks such that the yield of the forest in the face of fires is optimal. Firebreaks cost money, however: at the very least, vacant sites mean lost revenue from logging. So we are torn between:

- creating firebreaks to improve the yield of the forest by reducing the spread of fires
  and
- bearing the cost of the firebreaks themselves

One offsets the other.

If sparks are concentrated in particular areas of the forest (i.e. the distribution of sparks is not random), then it is clearly better value for money to place firebreaks closer together in those areas where fires are more likely to start and to space them widely elsewhere. For example, assume that there is a picnic site at the centre of the forest and that sparks from careless picnickers are thus more likely in the neighbourhood of the centre than elsewhere. The optimum spacing of straight-line firebreaks would look something like that shown below in Figure 8:

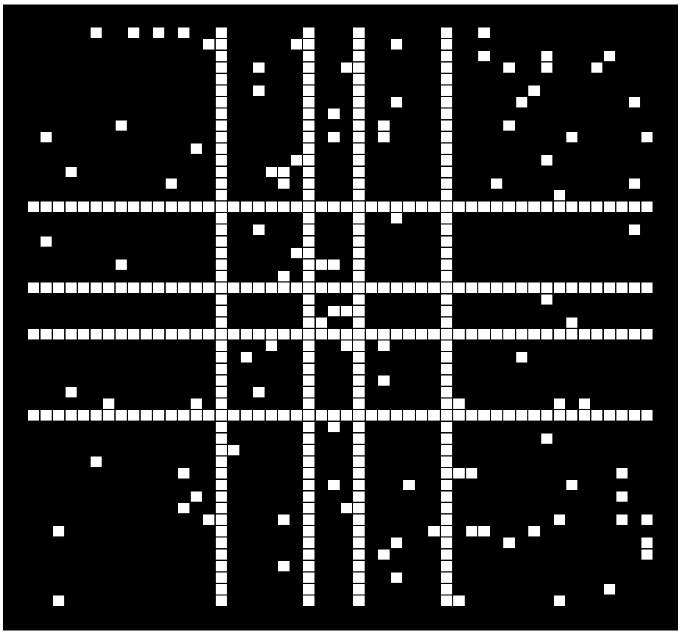


Figure 8 - Firebreaks concentrated near likely sparks

Firebreaks need not be straight lines, and an alternative set of possible firebreaks for this example is shown in Figure 9.

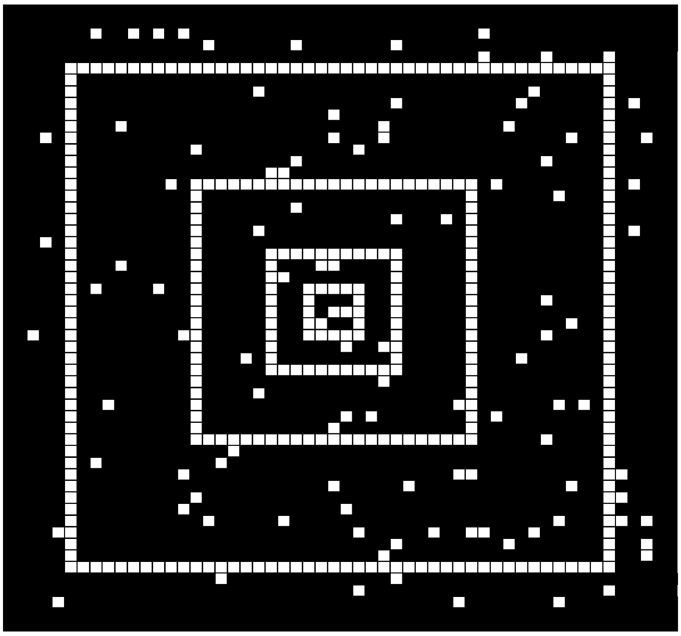


Figure 9 - Firebreaks with other shapes

The distribution of sparks determines the optimum shape and positioning of firebreaks. For some of the simpler spark distributions – a bell-shaped Normal distribution with its peak at the centre of the forest for example – it is possible to calculate exactly where firebreaks should be, given that they cost money (the smaller the total length of firebreak, the more forest space can be devoted to trees).

So what exactly is HOT?

The examples given of highly optimized tolerance have three notable characteristics:

- design is used to apply a resource (firebreak) such that the overall yield is maximized (which is normally the same as minimising losses). The resource is either limited or has a cost associated with it which offsets the value of the yield: applying too much resource can reduce the yield

- the resource reduces the total losses sustained because of some external event (spark). These losses may be caused by a chain reaction of the initial event (an external spark ignites a tree) causing other events (fire spreading to neighbours)

- the external events happen with some known probability distribution (some areas of the forest may be more likely to receive an external spark than others)

One consequence of HOT is that the greater yield (average tree density) renders the forest more vulnerable to *unanticipated* external events (perhaps the firebreaks were concentrated in the neighbourhood of picnic sites, but the forest was instead struck by lightning that prefers to hit the higher more exposed areas). But our HOT forest is also the most robust *for the particular amount of resource deployed*. And 'robustness' here is simply a measure of how stable the yield is in the face of anticipated risks. The firebreaks ensure that any fire in a vulnerable area is small. Fires in less vulnerable areas are larger but occur less frequently, and the damage of these can also be contained (with a small reduction in yield) through the COLD (!) variant of HOT described in Chapter 6.

How much better is HOT?

This depends on how sparks are distributed and the cost of the firebreaks – whether just the cost of lost trees or the additional cost of keeping the vacant land clear. Further, there is the design of the firebreaks themselves: whether they are simple lines of single vacant sites (i.e. one tree wide) or wider.

It would be counterproductive for the firebreaks to be wider if they were entirely vacant; after all, a firebreak one tree wide is sufficient to stop the spread of a fire in our model forest. But it appears that a possible optimum solution is a firebreak several trees wide which is sparsely planted rather than left free of trees entirely. The behaviour of such a forest was only discovered in 2001 and is remarkable and quite unexpected.

The areas of forest isolated from each other by the firebreaks grow to be almost fully occupied (i.e. attain a very high density). The firebreaks must be maintained at a density slightly lower than that needed to hit the order – chaos boundary. We thus end up with a patchwork of areas very densely populated separated by narrower bands of firebreak areas which are, say, 55% occupied with trees but no more. Because the firebreaks are maintained just below the order – chaos boundary, a large ‘percolating cluster’ (see page 14) of firebreak trees cannot form. For a random distribution of sparks, it is even possible to calculate the optimum relative sizes of the dense areas versus the firebreak areas.

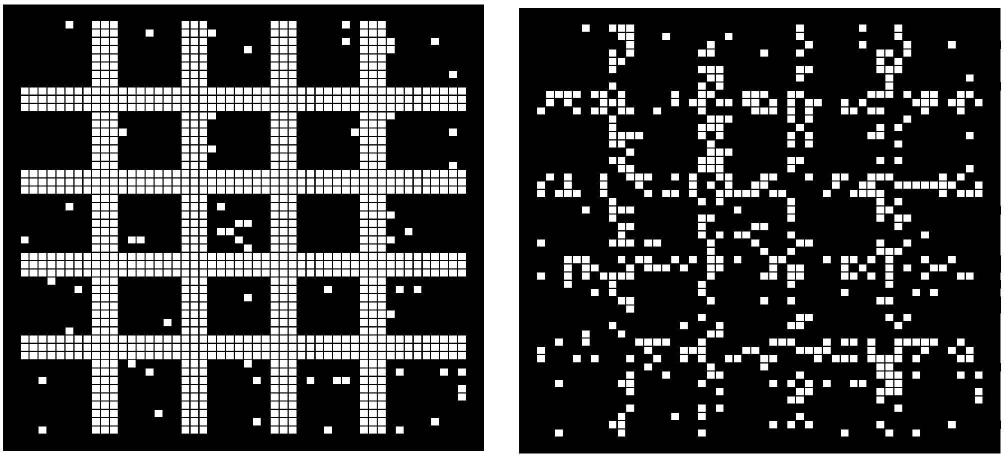


Figure 10 – Clear firebreaks versus firebreaks almost at criticality

Summary

So far we have used a single two-dimensional example to illustrate how highly optimized tolerance differs from controlled percolation and self-organization, but the principle works in many dimensions. Think, for example, of the spreading of fires though a ‘cube’ rather than a two-dimensional forest – a virus infection spreading through the offices in a skyscraper. It also works for many other phenomena that operate at the order – chaos boundary such as the ‘sand pile’ (which gets there through self-organization) and ones that only get there through deliberate tuning.

CHAPTER 5

HOT and BUSINESS ORGANIZATION

Recap

The Coevolving Organization described businesses operating at the order – chaos boundary: how to get them there, how they might behave when they got there and what the ensuing business advantages might be. It described both simple tuning and self-organization using the NKCS model of coevolving objects as a language. Tuning entailed adjusting K and C for each object such as a department. Each object's K-complexity was reduced in line with reducing its C-coupling to each of its S neighbours. Each object then became:

- more responsive (because it was largely free of the K-complexity 'treacle')

  while remaining

- still relatively stable because the external buffeting it received from each of its S (C-coupled) neighbours was also reduced

Chapters 4 and 5 of *The Coevolving Organization* described ways to tune K and C and the advantages and disadvantages of objects living on simple, low complexity (low-K) landscapes. Having a low K was usually but not always best. The peaks and troughs on a high-K landscape were steeper than those on a low-K landscape although the peaks were lower. This meant that if the high-K 'treacle' otherwise permitted a quick response (which it usually does not) to a competitor's attack on market share for a key brand, the business could fight back more quickly because a small movement towards the peak *in distance* on the steep slopes represented a relatively large movement *upwards*. But as a rule of thumb, 'low K is best'. Chapter 4 of *The Coevolving Organization* also described how to split up a business into coevolving objects – whether to take each division, each department or what? And how many objects? Once the split had been determined, K, C and S were also determined in the sense that the connections *within* each object were represented by K while the connections (couplings) *between* objects were represented by C.

The HOT examples described in the preceding Chapter took this one stage further. Assume that removing one or more C-couplings (i.e. further decoupling an object) or reducing their strength has a price in lost business effectiveness. This might be the result of added lack of cohesion, with the newly freed-up object moving in a direction contrary to head office strategy. Internally competing coevolving objects also need C-coupling to spur each other on. For example, the processing part of a manufacturing plant wants raw materials just when it needs them; not too early or late. The logistics department, on the other hand, wants to ship by the truckload or trainload to reduce transport costs. There is thus a compromise between:

- The *cost* of reducing C-coupling: less business cohesion; less incentive for coevolving objects to spur each other on to greater efficiency (for example, lower unit costs) and effectiveness (for example, improving product quality). This is dealt with at some length in Chapter 6 of *The Coevolving Organization.*

- The *benefits* from lower K-complexity (speed of decision making in response to a competitive threat, for example, or speed to market of a country-specific product). This lack of inertia can be dangerous if K-complexity is reduced without concurrently reducing either C-coupling or the number of other objects to which an object is C-coupled or both: an object can then become very unstable and will lurch unpredictably in one direction and then another. For example, a feather in the breeze has little inertia (low K-complexity) and will literally 'blow with the wind'. A kite has more inertia but also has a strong C-coupling to little Johnny at the other end of the string. An aeroplane in flight has large inertia from its mass and thrust of its engines. And frequent flyers will have noticed how susceptible small planes are to gusts of wind and how stable Boeing 777s are.

Note that decoupling objects does not necessarily mean that common computer systems and shared flows of common data cannot be used; decentralizing decisions does *not* imply decentralized data flows and systems.

Implications of HOT

In organization terms, moving an object further from the order – chaos boundary (in the 'chaos' direction) implies:

- reducing C-coupling to zero or the strengths of the couplings to a low value (or a mixture of the two). This must be done for each other object with which it coevolves

- allowing K-complexity to drop commensurately

The latter change gives greater responsiveness and, if business effectiveness is determined directly or indirectly by responsiveness, a better business. But there are several disadvantages:

  - lack of business cohesion: the business becomes a collection of unrelated smaller businesses. This may well be acceptable at the level of a self-contained operating unit, but is not acceptable for functional units such as Finance or HR which must work closely (but not always in co-operation…) with the operating units and other functions

  - lack of any spur to competitive coevolution

An object with very low K-complexity has a smooth landscape to work on. There are fewer small local peaks on which to become marooned (see 'Homing Instinct in Chapter 3 of *The Coevolving Organization*). And if it has very low or zero C-coupling to other objects, there is little or no external disturbance from elsewhere in the business to knock it off course.

When discussing the HOT forest fire model, we noted the importance of understanding risks. But what does this mean to a low-K and very low-C collection of business objects which comprise a coevolving organization?

Risks

In the present context, the meaning of a 'spark' is any event – internally or externally generated – which could cause an object to do something which was deleterious to itself (a snap decision which prices a major product far too low, for example). Because the object has low K-complexity, it will react quickly and will be pushed off course easily. So being more effective through having low K-complexity has the disadvantage that the object may be completely derailed. This is the price of greater effectiveness when things are going well. But the zero or very low C-coupling to the rest of the business should limit the extent of a backlash on the rest of the business. This is only true, however, for known (previously identified) C-couplings. If the wrongly priced product had an *unanticipated* impact on the rest of the business – perhaps stealing market share from other units' products – the effect could be catastrophic for the business as a whole. And since the other parts are also low-K, the entire business may react violently and in an unpredictable way. As with the forest fire (see Figure 9, for example) risky areas of the business such as Sales and Marketing need fencing off with lower C-coupling than areas such as HR which inherently carry less risk.

Anchor points

One organizational compromise to help ameliorate this problem is to create island-of-stability objects which have medium K-complexity and are C-coupled to several more reactive (lower K) objects in a star formation. Each such 'anchor point' acts as a buffer between any pair of more reactive objects instead of allowing them to buffet each other directly. Anchor point objects live just on the 'order' side of the order – chaos boundary but not beyond; they are equivalent to the 'critical' firebreaks in the Figure 10 right-hand forest.

Robustness

HOT may enable us to obtain a more responsive business by decoupling it further and thus decentralize decision-making. The risk is thus that *unanticipated* side-effects allow one object to knock other objects off course far more than if the collection of objects stayed at the order – chaos boundary. If, however, some areas are naturally more risky than others, we could:

- decouple them further (reduce C-coupling)
  without
- reducing their K-complexity

The result would be more robust but effectiveness would be unaltered. An analogy would be ring fencing an awkward entrepreneurial manager in a nominally freestanding business unit without giving him or her any concomitant decision making powers. This would protect the rest of the business while not giving the manager more freedom than previously.

CHAPTER 6

REFERENCE MATERIAL

Introduction

HOT is a mechanism first proposed by Jean Carlson of the University of California at Santa Barbara and John Doyle of Caltech at Pasadena as an alternative to self-organization and which also exhibited power-law[1] behaviour. Their scope was any system that was optimized to provide robust performance in a risky environment. They conjectured that power laws arose from compromises between yield (whatever was to be optimized – like fuel consumption in an aircraft), the cost of making the system robust, the degree of robustness achieved and the risks the system was likely to encounter. They demonstrated that systems designed on HOT principles had high performance and were resilient to risks (disturbances) for which they had been designed but were also very sensitive to disturbances for which they had not been designed. Carlson and Doyle noted that self-organization in complex systems is based on the assumption that, when near the order – chaos boundary[2], such systems have a self-similar internal structure where different parts of whatever size look and behave more or less the same. They pointed out that this is very different from a typical *real* complex and optimized structure such as a car where the different major parts have very different functions, and many of these parts are there simply to provide robustness and are invisible to the driver[3].

The structure of the Internet has been widely used as an example of self-organization since detailed data on its physical structure and performance are available and such data do indeed show all the 'power-law' hallmarks. Carlson and Doyle believe, however, that the freely evolving self-modifying nature of the Internet is an illusion. Although the Internet has no central control and the traffic patterns may appear to adapt automatically to congestion or failure of a link without intervention by the user or even by the communications link supplier, they consider that this is a consequence of the vast amount of ***design*** for both performance and resilience which has gone into the Internet's TCP and IP communications protocols and their physical implementation in routers (see 'Lessons from Telecomms' in Chapter 7 of *The Coevolving Organization*); it is not a natural consequence of the self-evolution of the Internet. The Internet provides other examples: 'good' web pages are designed as a compromise between size (a big page takes a proportionately long time to appear on the user's screen) and usability (users do not like following a long series of URL links from one short page to the next short page). If only the first part of a large page were

[1] where the probability of avalanche and similar unpredictable events is inversely proportional to some power of their size: small events are common and large events are relatively uncommon. Systems where uncommon large events are still nevertheless sufficiently frequent to cause problems – large earthquakes on an earthquake fault-line for example – are said to have a 'fat [or heavy] tail'(from the shape of the right-hand side of a probability / size graph).

[2] which physicists call the 'critical point' or '[second-order] phase transition'

[3] automobiles typically have power-*assistance* for 'heavy' manual operations such as braking and steering. Such systems are usually not duplicated (backed up) because if they fail, the manual element is still there. Contrast this with totally fly-by-wire aeroplanes such as the Boeing 777 where critical systems are (at least) triplicated. Adding robustness thus adds complexity.

frequently read, it is probably advantageous if the remaining parts were separate pages linked to the first page.

They stress the difference between a 'large' system consisting of many similar and simple parts and a 'complex' system consisting of many dissimilar parts which themselves may be complex (see 'How big should an object be?' in Chapter 4 of *The Coevolving Organization*). In real non-trivial designed systems – cars, aeroplanes and the like – complexity comes about through the need to provide robustness against likely stresses while still providing good performance. But this robustness comes at a price: the higher levels of performance possible using a HOT system can result in catastrophic failure when such a system is hit by a disturbance the possibility of which the designers had ignored. Carlson and Doyle characterize this behaviour as "robust yet fragile".

Background material

The notion of Highly Optimized Tolerance appeared first in Carlson and Doyle (reference **1**). Forest fire percolation is used to demonstrate HOT, but other examples such as road traffic flow, computer networks, electric power networks and biological systems are touched.

Reference **2** by the same authors introduces HOT in a way more abstract than the forest fire model. They consider a 'continuous' version of the forest – one where there is no discrete grid but rather an area where trees could be anywhere, not on grid sites – and a forest floor which has an arbitrary number of dimensions (not just two). Instead of simply assuming that the cost of a firebreak is the value of the trees that could otherwise grow there, they use a more general notion of a firebreak as an abstract resource that would restrict the size of subsequent 'events' (fires). They then limit the total quantity of this resource – equivalent to limiting the aggregate lengths of firebreaks in our forest model. They then assume that the size of an event (fire) at any point in the forest is inversely proportional to the amount of resource deployed at that point; in other words, more firebreaks in any particular part of the forest mean less spreading of a fire in that area. Finally, instead of sparks occurring at random places throughout the forest, they assume any arbitrary probability distribution of sparks. Minimising the expected loss through fires (i.e. maximising the yield) subject to the constraint on the maximum amount of resource (firebreak) available gives power laws of fire size (many small fires and a small number of large fires). They call this particular formulation of HOT '*probability – loss – resource*' ('*PLR')* from its three main facets: the probability (distribution) of an event happening, the loss if it indeed happened, and the resource which could be deployed to prevent it from being worse. They then proceed to examine the two-dimensional forest fire and the sand pile model in more detail (as in reference **1**) *not* overtly using the PLR formulation, and contrast the designed version with self-organized criticality – as we did earlier.

Reference **3** by the same authors is a summary of the foregoing but with some additional material on varying the 'amount' of design used.

Reference **4** by the same authors is an elaboration of the more abstract presentation of HOT (see the commentary above on reference **2**) as a PLR problem. They apply the same constrained-optimization process to three apparently very different problems: the forest fire (as in references **1** and **2**), data compression ('minimize the amount of data transmitted'), and web page length (as described earlier in this chapter).

Carl Robert et al (reference **5**) examine the spread of epidemics (infections = sparks) in a HOT context using a simple extension of the forest fire model which contains up to three 'cells' of population (forests) which are linked. An infection from one cell is allowed if necessary to spread to a neighbouring cell. They examine epidemic containment globally (i.e. across all the cells) versus containing epidemics locally within each cell. The authors then add an extra dimension – time, start a notional clock, and explain what happens when population growth is allowed to occur in cells that have not been infected between one clock tick and the next. A number of familiar patterns emerge including a chaotic one which occurs when the growth rate – and hence density of potential infections within a cell – is high (a similar effect should be observable in a forest if trees re-grew after a fire quickly relative to the incidence of sparks.

David Reynolds et al (reference **5**) is a detailed study of the effect of altering the number of tuning knobs[4] within the HOT forest fire model. Instead of attempting to place vertical and horizontal firebreaks at optimum positions, this paper takes a different tack and superimposes a coarse grid ('design lattice') on the finer grid of the forest. Initially, for the sake of example, the forest is divided into four areas using one vertical and one horizontal division (these do *not* represent firebreaks). The simplification is then made that the density of trees within any design lattice area (i.e. within each area delineated by the divisions) is the same, although it may be different in different design lattice areas. The densities within each area are allowed to vary independently. The authors demonstrate that for a single tuning knob (when the design lattice consists of one area that covers the whole forest), optimal yield is at the order – chaos boundary as would be expected. They then proceed to increase the number of tuning knobs by dividing the forest into four, nine (like tic-tac-toe), sixteen and finally twenty-five areas. Since the density in an area can be varied independently from that in other areas, each area corresponds to a tuning knob. For each of these examples, the optimal yield and tree density in each area is calculated. For few (but greater than two[5]) tuning knobs and at optimal yield, the density of each area tends to alternate across the forest: one area at almost 100% density with a neighbour at critical density[6]. If the distribution of sparks across the forest is random, the optimal solution when the number of tuning knobs is high (i.e. when the forest is broken down into lots of small areas) is merged clusters of small areas at almost 100% tree density with vertical and horizontal bands of single areas just below critical density. These are firebreaks which are not lines of vacant single tree

[4] 'design degrees of freedom'

[5] with two tuning knobs, i.e. with the forest divided in two, the optimal density for each area is the critical density if the distribution of sparks is random. It switches to one area at critical density and the other at 100% if the spark distribution is skewed, i.e. one area is then hit by sparks more than the other

[6] i.e. order – chaos boundary density, the density at which the percolating cluster appears

spaces but rather thin areas; they appear naturally between the areas of different tree density[7].

Tong Zhou and Jean Carlson (reference **7**) explore what happens when a HOT system is perturbed from its optimum by slightly changing the positions of its cuts (firebreaks).

Zhou et al (reference **8**) look at the evolution through natural selection of competing organisms, where *each* organism is represented by a 'forest'. The fitness is the yield after a fire but calculated differently for different organisms because the population of organisms is divided into two groups: in any one generation, there are:

- a group whose members' fitness is the yield after a single spark. In other words, the same spark applied to all members of the group will result in some low-yield ones (where the fire did major damage) and some high yield ones which protected themselves best.

- a group whose members' fitness is an average fitness resulting from *all* the possible sizes of fire which could occur and weighted by the probability of each size occurring

From Aesop's best-known fable, the first group are called 'hares' and the second 'tortoises'. Because a hare's fitness in any one generation is calculated from how it responds to *one* disturbance (spark), the weeding out of unfit organisms takes account only of how the hares have responded to the present disturbance and not to previous ones or possible future ones. The fitness of a tortoise in any one generation, on the other hand, is an average fitness of the tortoise's response to *all* possible disturbances. Hares are thus taking a short-term view and their evolution is very responsive to their environment whereas tortoises are taking a longer-term and more balanced view and their evolution responds more slowly to changes in the environment. As they evolve, hares can become over-specialized to the disturbances they have encountered; for example, following a series of small fires over several generations, hares may quickly evolve a high fitness when faced with small fires. The downside is that they may be decimated when a less common large disturbance occurs. The tortoises, on the other hand, will have evaluated their response to events of *all* sizes (weighted by their likelihood) when evaluating fitness and will thus be more resilient to the less common disturbances. Evolution occurs asexually by breeding two child organisms (child 'forests') from each single parent. Each child has the possibility of a mutation in one site (tree space) – from 'vacant' to 'occupied' or vice versa, and the probability of a mutation happening is set such that, on average, only one of the two children has a mutation. As each new generation is born, the parents are killed off and the population culled down to a fixed limit (Zhou used 1000) of hares and tortoises. Culling is done by removing any organisms whose fitness is less than a preset lower limit and then removing the least fit organisms, irrespective of whether hare or tortoise, in ascending order of fitness, until the population is back within the

[7] such a 'fully designed' HOT forest with lots of firebreaks does not exhibit power-law behaviour if the distribution of sparks is random because all the clusters of trees are the same small size. An infrequent 'large' event (part of a 'fat tail') is thus unable to form

limit. Diversity into hares and tortoises is, however, maintained by creating niches for each, in which the fittest few (Zhou used 50) of each type are ring-fenced from culling. An organism which finds itself within a niche competes only with the others in the niche for that generation. Niches are useful to, among other things, protect the tortoises from total extinction shortly after the start of evolution where the quickly evolving hares may dominate the population. The limit on the total population can be thought of as finite living space.
Organism 'forests' start out with random 'tree placement' but this soon evolves into the familiar clusters of densely occupied sites separated by barriers ('firebreaks') of vacant sites. Using a spark distribution which is not random (i.e. where 'sparks' regularly hit some 'forest areas' more than others), hares evolve higher fitness quicker than tortoises but tend not to develop barriers in the areas of low risk (few sparks). Zhou studied several variants of this basic model – removing niches; replacing tortoises with hares which had a low mutation rate (i.e. evolved slowly) and so on.

Carlson and Doyle (reference **9**) is written more for engineers than mathematicians and physicists and is a summary of much of the then current (2002) work with descriptions of the real world applicability of HOT. The non-descriptive material is roughly the contents of references **1** + **4.**

Mark Newman et al (reference **10**) extend the primary Carlson and Doyle paper (reference **2**) to look in more detail at the 'fat tail' – the impact of unlikely events. HOT protects forest areas which are most likely to receive a spark at the expense of those areas where fires are unlikely. The effect of this is to make the impact of an unlikely event catastrophic. The authors use a continuous model of the forest (roughly, very thin trees separated by very small spaces). They consider both the conventional forest fire model and the probability – loss – resource (PLR) formulation of it (see comments above on reference **2** and the Q&A in Chapter 7), and derive algebraic[8] solutions for both. They then – and this is the point of the paper – study what happens if the 'fat tails' (the effects of unlikely events) were weighted such they became thinner (even less likely). The effect of this is that instead of the firebreaks concentrating around the areas where sparks were most likely, they are more spread out such that areas where fires are unlikely are nearer to a firebreak that they would otherwise be. This drops the yield: the HOT version is the optimal one but has then been tinkered with to reduce exposure to unlikely catastrophic events. But the authors show that for a considerable reduction in the risk of unlikely but catastrophic events, the drop in yield (i.e. the amount the system is sub-optimal) is relatively small. Unsurprisingly, they gave the name *COLD* (constrained optimization with limited deviations) to this 'safer' version of HOT.

References **11** and **12** are not HOT articles per se. Stu Kauffman et al (reference **11**) describe the effect of splitting a 'forest' into areas such that each area is allowed to optimize its own yield and ignore its effect on others. It is described in *The Coevolving Organization* Chapter 4 (How big should an object be?). Reference **12** by Maya Paczuski et al is discussed at some length in *The Coevolving Organization* Theoretical background Section 3 (Chaos, avalanche dynamics and universality) and is still the best and most detailed review of

[8] i.e. as opposed to the results of computer simulations

avalanche effects (forest fires; sand piles; ..) in both self-organized and designed systems. It was written midway between the first sand pile analyses by Per Bak[9] and Carlson and Doyle's discovery of HOT.

Because HOT is relatively new, there are no authoritative books on the subject. Reference **13** (edited by Erica Jen) which, at present[10] awaits publication, describes many of the applications of HOT. In particular, there are lengthy articles by Doyle, Carlson et al on the design, evolution, robustness and fragility of the Internet.

Reference **14** by Dietrich Stauffer and Amnon Aharony is the standard textbook on percolation. And unlike most reference works, it is written with a dry sense of humour…

Christopher Alexander (references **15, 16** and **17**) is a practising architect, and the relevance of these publications to HOT or coevolution may seem tenuous. But he is also a Cambridge-educated mathematician, and his approach is to analyse how abstract 'things' – which may be supporting or conflicting – interact, and how misfits between these 'things' and their environment can be minimised. Alexander's work has spawned considerable interest from other areas, notably object-orientated software design (see *The Coevolving Organization* Annex – Information Technology). Appendix 2 of reference **15** contains the proof of a highly relevant theorem: "given a system of binary stochastic variables, some of them pair-wise dependent, which satisfy certain conditions, how should this system be decomposed into a set of subsystems such that the information transfer between the subsystems is a minimum". The significance of this to designing an organization should be readily apparent to readers of *The Coevolving Organization* (see Chapter 4 – How big should an object be?): one design criterion for selecting coevolving objects is that they *naturally* communicate between themselves as little as possible (i.e. communication needed by business processes is primarily *within* objects). If this is not true, the carving up of the business into objects has been done wrongly and there is a better way to do so which concentrates communication *within* objects and reduces it *between* objects. One can (loosely…) apply the formulation of PLR: if we have a fixed maximum number of barriers between business areas, we want to place the barriers such that the communication between areas (i.e. across the barriers) is minimized relative to any other way of placing barriers. Alexander introduced the idea of 'patterns' which can be used at a local (decentralized) level to create structures – which in our case are the internal processes of organization units – each of which has the most appropriate fit for its purpose.

Finally, references **18** and **19** by Duncan Watts describe the recently elucidated Small World effect that could cause a HOT business organization to experience an unplanned catastrophic event. To recapitulate, our model forest fires spread from neighbour to neighbour and do not jump directly from one tree to another tree which is not immediately adjacent. HOT creates clusters of trees separated by firebreaks of vacant sites, with firebreaks concentrated around sites that are most likely to catch fire and with areas less likely to catch fire being less

[9] Per Bak died in October 2002 and was married to Maya Paczuski

[10] September 2003

well protected. A spark in an 'unlikely' area can thus create a fire that can spread widely and do catastrophic damage before it hits a firebreak. *But there are other ways that catastrophic events can happen.* One way is for a small fault to develop in a firebreak (for example, an unplanned tree grows in a firebreak site that should be kept vacant) which enables a fire on one side of the firebreak to spread to the other side. But, as described earlier (Risks in Chapter 5), the same unplanned fault could occur if one coevolving object (a marketing team in one country for example) makes a decision about one of its products which then *unintentionally* steals sales from a product marketed by another coevolving object (team). Coevolving objects in business organizations do not have the same 'nearest neighbour only' spreading of events, i.e. being physically adjacent is fast becoming unimportant in the upper reaches of an organization. Nevertheless, teams – at the lower levels at least – *are* physically clustered into countries, offices, manufacturing plants and the like. Most planned risks ("if we make a decision to launch new product X, do you have the capacity to manufacture it in the volumes which Sales might need if customer take-up is 30% higher than planned?") impact physically adjacent teams. But risks or side effects no one had thought of may impinge anywhere in the organization (which is probably why no one had thought of them…). And the Small World effect can enable or exacerbate this.

The Small World effect is a feature of clusters of objects which are linked to each other locally but which have little communication with remote objects. For example, picture the spread of a contagious disease through a collection of isolated villages. If everyone stays in his or her own village, the disease remains localised. But if just one individual travels, the disease has the chance to spread. If this one individual infects just one person in another village (not necessarily a neighbouring village), it can spread rapidly within that village also. Put more abstractly, imagine a 'network' of objects made up of many clusters with many connections within members of a cluster and few connections between clusters. Then the addition of a small number of connections between members of different clusters selected at random can make the spreading of information (or infections or anything else that can be passed on) disproportionately easy. Without the additional random links, the spreading might need to pass through many individuals consecutively. The few additional links manage to short circuit this lengthy chain to a degree which was first highlighted by Harvard's Stanley Milgram in 1967 but not properly analyzed until 1997 by Watts and colleague Steve Strogatz of Cornell University.

The NKCS formulation is straightforward. Let each village (cluster) be an object. Individuals within a village are then linked (communicate) using K-complexity links. A few individuals may also be linked to one or more individuals in other (probably neighbouring) villages using C-coupling links but this would be rare. Adding just a few C-couplings to individuals in a random selection of other (perhaps remote) villages has the potential to disturb the stability of both sides radically. But what is more significant is that a message from an individual in one village to another in a remote village would, under normal circumstances, take a very roundabout route and pass through intermediaries in a succession of other villages before it was delivered. The few extra random C-couplings make a disproportionately huge difference and the number of intermediaries is cut drastically.

CHAPTER 7

QUESTIONS AND ANSWERS

Q: I think I understand the 'controlled percolation' forest fire formulation of HOT, but cannot see the connection between this and the probability – loss – resource (PLR) version. Does PLR occur in real-life?

A: The Duke of Wellington[11] was outnumbered when defending against the French at Torres Vedras (near Lisbon) during the Iberian Peninsular War. He had two conflicting constraints: *winning* while *minimizing casualties* (loss) and he was, with some restrictions, able to place his troops such that the probability of casualties *overall* was minimized. Some soldiers would be in advanced positions most likely to be attacked but Wellington ensured that these were in small groups heavily protected by gun emplacements, palisades and earthworks that were built at considerable cost by several thousand Portuguese labourers. At the other extreme, he spent little on protecting his reserves that were further from the firing line. Given this strategy:

❖ a 'normal' HOT formulation would be: win whilst minimising the cost of casualties *plus* the cost of flank protection (more small groups = more flanks to protect). The difficulty is that turning either casualties into money or the cost of flank protection into equivalent 'avoided casualties' is subjective.

❖ the PLR HOT formulation would be: win whilst minimising the cost of casualties subject to a *limit* on the cost of flank protection. His tactical problem was this: with a fixed-sized war chest for spending on defences, where should he spend the money on building these defences such that his overall casualties were minimized, *given his assessment on the likely casualties in each area*. In this formulation, there is no need to put a price on casualties.

The 'normal' formulation is thus:

▪ optimize yield where the yield (value) is offset by the cost of flank protection which insulates one area from another

whereas the PLR formulation is:

▪ optimize yield subject to a limit on the total cost of flank protection

The first assumes no overt limit on the cost of flank protection, but assigns a cost such that the minimization process itself puts a brake on the amount of flank protection used. It makes a compromise between the value of the yield and the

[11] 1769-1852; the UK's best field commander since the (1st) Duke of Marlborough. Although at the time not yet a Duke, he was on fast track promotion during the war as progressively Sir Arthur Wellesley; Baron Douro; then Viscount, Earl and lastly Marquis of Wellington.

cost of protection. The second formulation does not assign any cost per unit length of flank protection, but limits the total amount that can be employed. From the above example, Wellington had a fixed army; his latitude was how to deploy them in groups geographically. More small groups limit the *overall* impact of a sudden and successful assault on his troops: some small regiments may be totally wiped out but, since he had deployed his troops such that the ones most at risk were protected by the best defences, Wellington had done his best[12].

Q: You said that, without firebreaks, a controlled percolation forest would have a yield that peaked at a tree density of around 60% at the point where a cluster of trees (the percolating cluster) spanned the forest. When one of the trees in the cluster ignited, the forest would have hit the order – chaos boundary and a large fire would ensue, bringing the density down again. Now, if the forest were split into areas separated by firebreaks, why do these smaller areas not act in the same way? i.e. why can these areas approach 100% tree density while the overall forest cannot? After all, a patch of forest is itself a forest…

A: In a randomly planted forest, the growth rate of new trees is in balance with the rare incidence of trees burning and the almost-as-rare incidence of sparks. New trees are assumed to be planted at random places. Below the critical density, any fires involve a few trees only, and so the density can creep up inexorably to the critical density. When the forest reaches its critical density, the widespread fires that occur because the percolating cluster is large check any further growth, and the density falls again. The density of the forest oscillates around its critical density, with lots of small fires and much fewer large fires preventing the forest from growing much denser.

A designed forest behaves very differently. If the placement of trees is such that there is a large number of firebreaks, it is impossible for a large cluster of trees to form and this limits the size of the largest fires. The firebreaks have a cost that, at its simplest, is the loss of trees that could otherwise grow there. In the 'normal' HOT formulation, this overt loss is more than offset by the increased overall yield that results. But there is an obvious but hidden limit on the total amount of firebreaks: they cannot occupy more than the total number of sites in the forest (i.e. the limit when the forest is all firebreak and no trees). In the PLR formulation, the yield is not overtly reduced by the vacant tree space in the firebreaks, but an upper limit is put on the aggregate length of firebreaks.

But designed forests *can* gain even more over self-organized (or controlled percolation) forests when the probability of sparks is different in different parts of the forest. If the more vulnerable areas were split into many smaller chunks by the firebreaks (at the expense of less vulnerable areas which remain in large chunks), the overall impact on yield of a spark hitting one of these smaller chunks is small – because the chunk is small. So a few small chunks might be totally wiped out but the remaining chunks are free to grow to a high density because one burning chunk cannot set fire to its neighbour.

Self-organized forests behave slightly differently to controlled percolation versions which involve manually tuning the tree density (but not manually tuning

[12] he won…his defence and logistics were so good that the French under Marshal Massena, with very tenuous supply lines, themselves starved and retreated

the tree *placement* – this is still random). The percolating cluster (order – chaos boundary) occurs at a lower density (around 40% versus 60% for random percolation) and the relative incidence of large events is somewhat smaller[13] [14]. But the essential behaviour is the same even if occurs for different reasons.

Q: In a self-organized forest, what happens if the growth rate is relatively much faster or slower than the incidence of sparks?

A: If the growth rate is faster, the density will rise much higher than the critical 40% but the first spark which hits a tree (rather than a vacant area) may then burn most of the forest. So the density will swing wildly (see reference **5**) from very high (approaching 100%) to that of total devastation. If, on the other hand, growth is slow relative to the incidence of sparks, the forest density will approach the critical density very slowly and density oscillations will be small.

Q: The 'forests' which have been used as examples are all two-dimensional. When HOT is applied to organizations, the dimensionality may be higher and it is determined by the way K-complexity and C-coupling links actually connect genes within objects. What is the impact of a higher number of dimensions on self-organized criticality and HOT?

A: In self-organized criticality and controlled percolation models, as the dimensionality increases, the relative incidence of large events (large fires) becomes less[15]. In HOT systems, the relative incidence of large events *increases* although this tendency can be contained using COLD 'fat tail reduction' ideas (see reference **10** and the preceding commentary on it).

Q: HOT seems to be most effective when the spark distribution is skewed (not random) such that some areas are more vulnerable (and need closer firebreaks) than others. Can power law behaviour (the incidences of various sizes of fires, for example) occur when the distribution is random? Does the shape of the spark distribution drive the power law shape?

[13] i.e. the power law curve which shows the (logarithm of) the probability of an event versus the (logarithm of) the size of that event becomes steeper. For controlled percolation forests, for example, such curves can be constructed by simulating the impact of single sparks (drawn from a given probability distribution) on different forests with the desired density and averaging the result.

[14] even our idealized forests do not follow perfect power laws. At criticality, fires occur in compact high-density clusters that are decimated when they burn *and* in straggling low-density clusters which, because of their shape and sparseness, partly survive. The imperfect power law curve reflects this odd mixture

[15] i.e. the power law curve becomes steeper as dimensions increase. Large avalanche events in random systems become more and more unlikely as dimensions rise because the connectivity needed for the percolating cluster to span, say four or even five dimensions, is very unlikely indeed to occur at random

A: Skewed spark distributions lead to skewed HOT firebreak positioning, with more firebreaks in the areas most likely to receive sparks. This is simply because HOT seeks to maximize yield, and the only way to achieve this is to ring fence with firebreaks the areas most likely to be hit and thus limit the size and spread of the resulting fire if a spark lands in one of these vulnerable areas. Having different sizes of area, the most vulnerable being small and the least vulnerable large, gives power-law distributions of fire sizes (lots of small fires and fewer large fires). It is the HOT process's aim to position these firebreaks optimally such that the *overall* loss is minimized. HOT power law behaviour cannot occur when the spark distribution is random and the forest is large and has many firebreaks[16].

Q: You described the HOT tuning knobs ('design degrees of freedom') needed to create and position multiple firebreaks. What is the equivalent of these tuning knobs for self-organized criticality and for controlled percolation?

A: The self-organized criticality equivalent is the ratio of spark frequency to the growth (or random planting) of new trees – the faster the growth relative to the incidence of sparks, the higher the resulting tree density. The controlled percolation model's tuning knob is simply the density. In either case, the forest fights back with repeated fires when the density approaches the critical density

---

Q: HOT seems to depend on a designer placing firebreaks at optimal positions. Can a system (forest) evolve such firebreaks itself through, for example, natural selection? (and if so, isn't this self-organization under another name?)

A: No – at least a normal forest cannot. A *collection* of systems can indeed evolve into HOT states with well-placed firebreaks (see the commentary above on reference **8**), but the evolution mechanism needs to select the best systems of each generation (or cull the worst of each generation) based on their yields. If there is a mechanism (mutation, for example) to generate sufficient variation[17] in each generation (i.e. if there is always sufficient diversity in the offspring in each generation), natural selection will cause firebreaks to appear automatically because systems with them have higher average yields than systems without. As the systems evolve, they are clearly 'organizing themselves' with ordered patterns of firebreaks, but this is not 'self-organization' as it is understood in the context of criticality. A self-organized system starts out as a random system; for example, trees are planted or self-seeded at random places in the forest. Any excess of new growth over burnt trees moves the system inexorably to the order – chaos boundary, at which point the system fights back with avalanches of fires to stop

[16] behaviour in a small forest is influenced by what happens at its edges. And if there are only a few tuning knobs (i.e. only a few firebreaks), the areas delimited by the firebreaks are each a substantial portion of the forest; the behaviour when they catch fire is thus 'lumpy' compared to the much smoother behaviour when there are lots of smaller areas to catch fire

[17] lack of sufficient variation can cause evolution to stop at a local optimum – a small peak in the foothills – and not continue to reach the global summit. See 'The different faces of K' in Chapter 5 of *The Coevolving Organization*

the density of trees becoming any greater. At the order – chaos boundary, there is no pattern to tree placements (i.e. no HOT-like clustering) and behaviour of the system is unpredictable: the next fire may be large or small and there is no way to forecast which. All we could say is that there will be lots of small fires and few large fires. In a self-organized critical forest, trees grow and re-grow at random positions. Like the proverbial monkeys trying to type the works of Shakespeare, a self-organized critical forest *might* conceivably evolve an optimal set of firebreaks, but the probability is vanishingly small. A collection of forests evolving to a HOT state through natural selection will, however, get there in a reasonable time simply *because there is a selection mechanism*. A manually tuned HOT system has used mathematics or rules of thumb to bypass evolution and jump straight to the optimal arrangement of firebreaks.

Self-organized criticality is only optimal in the sense that the order – chaos boundary represents the maximum density achievable[18] for a *random* placement of trees. This density is *not* optimal for *other* 'designed' placements of trees and this is why HOT tree densities are much greater than those for self-organization.

[18] this might be exceeded for a short time if the growth rate of trees is fast relative to the frequency of sparks, but the ensuing big fire following the next one or two successful sparks will rapidly bring the density down again to the critical point

BIBLIOGRAPHY

Papers

1. Carlson J.M. and Doyle J. "*Highly optimized tolerance: Robustness and power laws in complex systems"* (Condensed Matter 9812127 8 Dec 1998)

2. Carlson J.M. and Doyle J. "*Highly optimized tolerance: A mechanism for power laws in designed systems*" (Phys. Rev E 60 1412-1427 1999)

3. Carlson J.M. and Doyle J. "*Highly optimized tolerance: Robustness and design in complex systems*"(Phys. Rev. Lett 84 No 11 2529-2532 13 March 2000)

4. Carlson J.M. and Doyle J. "*Power laws, highly optimized tolerance, and generalised source coding*" (Phys. Rev. Lett. 84 5656 12 June 2000)

5. Robert C., Carlson J.M. and Doyle J. "*Highly optimized tolerance in epidemic models incorporating local optimization and regrowth*" (Phys Rev E 63 056122 1-13 2001)

6. Reynolds D., Carlson J.M. and Doyle J. "*Design degrees of freedom and mechanisms for complexity*" (Phys Rev E 66 art. 016108 2002)

7. Zhou T. and Carlson J.M. "*Dynamics and changing environments in highly optimized tolerance*" (Phys Rev E 62 No 3 September 2000)

8. Zhou T., Carlson J.M. and Doyle J. "*Mutation, specialization and hypersensitivity in highly optimized tolerance*" (Proc Nat Acad Sci Vol 99 No 4 2049-2054 19 Feb 2002)

9. Carlson J.M. and Doyle J. "*Complexity and robustness*" (Proc Nat Acad Sci 99 2538-2545 2002)

10. Newman M.E.J., Girvan M. and Farmer J.D. "*Optimal design, robustness, and risk aversion*" (Phys. Rev. Lett. 89, 028301 2002)

11. Kauffman S., Macready W.G. and Dickinson E. "*Divide to co-ordinate: coevolutionary problem solving*" (Santa Fe Institute 15th October 1994)

12. Paczuski M., Maslov S. and Bak P. "*Avalanche dynamics in evolution, growth and depinning models*" (Phys Rev E Vol 53 No 1 January 1996)

Books

13. Jen E. (ed.) "*Robust design: a repertoire of biology, ecology and engineering case studies*" (Oxford University Press – to be published Dec 2003)

14. Stauffer D. and Aharony A. "*Introduction to percolation theory*" (2nd ed. – Taylor and Francis 1994)

15. Alexander C. "*Notes on the synthesis of form*" (Harvard University Press 1964)

16. Alexander C. "*Centre for Environmental Structure Series*" (Oxford University Press)
    16a. "*The timeless way of building*" (1979)
    16b. "*A pattern language*"[19] (1977)
    16c. "*The Oregon experiment*" (1988)

17. Alexander C. "*Nature of order*" four-volume series (Centre for Environmental Structure 2003)
    17a. "*The phenomenon of life*"
    17b. "*The process of creating life*"
    17c. "*A vision of the living world*"
    17d. "*The luminous ground*"

18. Watts D.J. "*Small Worlds: The dynamics of networks: between order and randomness*" (Princeton University Press 1999)

19. Watts D.J. "*Six degrees: the science of a connected age*" (Heinemann 2003)

20. Stewart M. "*The coevolving organization*" (Decomplexity Associates 2001)

[19] with Sara Ishikawa and Murray Silverstein

INDEX

The Robust Organization

ISBN 0-9540062-4-0