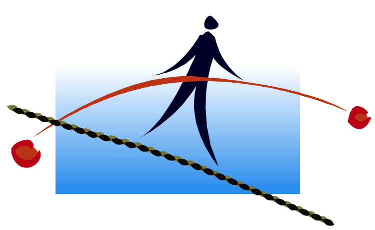
Max Stewart

The Coevolving Organization



POISED BETWEEN ORDER AND CHAOS

Published by Decomplexity Associates Ltd
First published 2001

This publication may be reproduced, stored for later retrieval or transmitted provided that the original Adobe® Acrobat® format is retained and authorship acknowledged. Conversion to editable form or editing in any way is a breach of copyright.

Copyright © 2001 by Max Stewart

The right of Max Stewart to be identified as the author of this work has been asserted by him in accordance with the Copyright, Designs and Patents Act 1988

Set in Times New Roman

ISBN 0-9540062-2-4

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 has latterly spent eighteen years with Mars, Incorporated. He lives in Rutland, England's smallest county, where foxhunting, stout foxes and an agile rabbit population coevolved for a very long time.

COPYRIGHT AND TRADEMARKS

Copyright © Max Stewart 2001

Chaordic is a trademark of the Chaordic Alliance. R/3 is a registered trademark of SAP AG. UNIX is a registered trademark of The Open Group. Windows 98, Windows NT and Windows 2000 are registered trademarks and NetMeeting and Outlook are trademarks of Microsoft Corporation. Lotus Notes and 1-2-3 are registered trademarks and Notes, Domino and SameTime are trademarks of Lotus Development Corporation. Adobe and Acrobat are trademarks of Adobe Systems Inc. ViewStation is a trademark of Polycom Incorporated. SwiftSite is a trademark of PictureTel Corporation. LINUX is a registered trademark of Linus Torvalds. Simula is a registered trademark of the Norwegian Computing Centre. IBM is a registered trademark of International Business Machines Corporation. Mars is a registered trademark of Mars Incorporated. VW is a registered trademark of Volkswagen AG. Yellow Pages is a registered trademark that is held variously (in the UK for example, by British Telecommunications PLC). Other trademarks and trading names are acknowledged.

Wisdom is good with an inheritance:
and by it there is profit...
For wisdom is a defence, and money is a defence:
but the excellency of knowledge is,
that wisdom giveth life to them that have it.

Ecclesiastes (King James version)

CONTENTS

PREFACE

No one really understands how complex businesses compete. For at least two hundred years, governments have grappled with the unpredictable behaviour of economies but have never managed to subdue them. Currencies, stock markets and employment levels continue to lurch from boom to bust and back with a will of their own. Sometimes economies get stuck as happened in the early 1930s when much of the developed world became firmly wedged in the trough of a business cycle. English economist Maynard Keynes was the first to show that the natural state of the economy was not necessarily that of full employment and bullish sales. Being bogged down in a slough of unemployment and stagnation was just as probable. One equilibrium state of the economy was as likely as any other. He suggested that an enlightened government might be able to drag its sick economy out of the mire by stimulating both consumer demand and industrial investment through intelligent manipulation of public spending and interest rates and perhaps even by acting as guarantor for debt in the private sector. Perhaps the economy could then be anchored elastically near the peak of a business cycle. Any jolts would be handled resiliently with the economy soon bouncing back. And perhaps such goals for economies - peak performance, resilience, flexibility and long-term stability - could in some way be achieved also by individual businesses engaged in headlong competition. Those were the hopes. But Keynes, with his unrivalled insight into how markets actually worked, was some sixty years too soon. The theoretical principles needed for a glimmer of insight into the underlying dynamics of how complex things such as brands, businesses and living things interact and compete only became available in the last decade. And the computers essential to model complex behaviour were not around in Keynes' day.

There are now indications that within a group of competing complex things, *each* might be able to pursue its own selfish ends - for a business this might be more growth or higher profit - while also bringing the *group as a whole* to a better state. This may sound paradoxical - competition *apparently* leading to co-operation, but it is simply the inevitable result of the things gravitating as a collection to a state some way between fettered regimentation and wild instability. No altruism is needed. Keynes' 18th Century predecessor Adam Smith was the first to publish such wild ideas. His novel proposal was that "every individual ... intends only his own gain, and he is ... led by an invisible hand to promote [the interest] ... of the society more effectually than when he really intends to promote it". In other words, shortsighted self-interest is better for the common good than trying to be altruistic. But Smith's concern was for the smaller owner-managed business. Global corporations were well in the future. And unlike all too many of the current crop of theoretical economists whose ideas seem divorced from reality, Smith was well aware of the need to validate his recommendations with hard data. Neither was he a total free marketeer. He believed passionately in the need for the State to manage the infrastructure within which a free market could operate successfully without debasing the humanity of individuals. Selfish businesses whose ethics went no further than treating employees as machines certainly needed government restraint, but the State's main role was to change business culture such that everyone, employees as well as employers, behaved 'reasonably' when participating in an otherwise free market. For this intervention to be accepted by business and the populace alike, the State needed to be seen to be promoting their long-term interest and not its own ends.

Selfish or otherwise, businesses grow, change or die. The size needed to retain market share in a stable market becomes a liability when the market or competition changes direction. The larger a business gets, the more inertia it accumulates. Adapting rapidly to the needs of local markets or embarking upon something radically new becomes harder and harder. Businesses competing for market share or trying to open up new markets fight each other like competing predators. Those with large market shares defending their positions against newcomers act like predator and prey. The behaviour of real predators and prey, such as fox versus rabbit, has been honed by evolution and natural selection over a very long time to an apparent equilibrium. But this does not prevent occasional huge upsurges in population or species becoming extinct. Since the early 1990s, the theoreticians of complex self-adaptive systems have shown that these things are to be expected. They are natural consequences of what happens when complex things such as foxes and rabbits interact and compete.

Businesses are themselves complex systems that compete to survive. The weakest go to the wall. Sometimes they collapse individually. Sometimes, when whole economies stumble, they are brought down in cascades like falling dominoes. The skein of theoretical principles that determines the behaviour of complex systems is still being unravelled but appears to apply to businesses also. Answers, albeit incomplete and tenuous as yet, are starting to emerge to a number of questions:

- what is the right level of organization complexity?
- how big should a business unit or division be?
- how can resilience to disturbances such as attacks by competitors be improved while the business still remains poised and reactive?
- why do things which upset the status quo, such as a reorganization or emergence of a new competitor, often have useful and sometimes unexpected side effects and how can a business take advantage of them?
- how can sales and marketing managers get growth in a static market?
- why does de-stocking a supply pipeline or trying to optimize the production schedule of a manufacturing plant sometimes create huge and unwanted fluctuations in material demand?
- are there some businesses already using such ideas successfully?

Such answers could markedly affect how businesses are organized, but they are filtering through slowly. This is partly because the approach is novel and partly because the background material is hard going to those without a scientific background. The source texts demand a knowledge, and in some cases a quite up-to-date knowledge, of mathematical statistics, systems engineering, theoretical physics and biology. Furthermore, translation into how *businesses* behave is almost non-existent - hence this book.

A few popular writers, frustrated poets perhaps, have grasped the bare bones of the subject and written lyrical and sometimes patronizing appreciations of the cleverness of Nature to operate successfully in complex and elegant ways - something Nature clearly sorted out for itself a very long time ago without our assistance. This is not one of those books.

Lastly, an apology: English has not yet evolved gender-inclusive and unstilted variants of 'he' and 'his'. Perhaps communication and natural selection, which together mould languages and ideas as well as species, will in time create something suitably simple and direct. Until then, an author, like Nature, has to work with what tools he has.

Max Stewart
Rutland, UK
January 2001

ACKNOWLEDGEMENTS

I am grateful to the BBC for permission to reproduce part of a transcript of their In Business programme, Methuen and Company for permission to quote from Adam Smith's 'The Wealth of Nations' and CEO Emeritus of VISA Dee Hock for his comments and permission to use his quotations. The description in Chapter 6 of the process by which complex organs are developed is from "The Extended Phenotype: the gene as the unit of selection" © 1982 by Richard Dawkins and used with permission by W.H. Freeman and Company. The prefatory extract from the Authorized Version of the Bible (The King James Bible), the rights in which are vested in the Crown, is reproduced by permission of the Crown's Patentee, Cambridge University Press.

I thank:

- Stuart Kauffman of the Santa Fe Institute and Bios Group LP for clarifying his species invasion mechanism which can induce NKC models on high-K landscapes to self-organize
- Gene Stanley and Luís Nunes Amaral of Boston University who explained some of their mathematical shortcuts
- Ian Stewart of Warwick University for pointing out to me in 1991 the similarity of the dynamics of logistics pipelines and the self-organized criticality of Per Bak's sandpile models
- Michael Rothschild of the Bionomics Institute and Maxager Technology for his views on cost control in capital-intensive batch processing industries
- Christoph Bangerter and Ueli Hagnauer for a wealth of information on Swiss healthcare management.

I am indebted finally to the late A.A. Milne for Winnie-the-Pooh whose game of Poohsticks has been elaborated to demonstrate the branching process of the Reggeon Field Theory universality class.

Errors, particularly those arising from interpretation of the physics of evolution into business behaviour, are quite definitely mine alone.

STRUCTURE OF THIS BOOK

This book is aimed at managers of large multinational businesses that are finding long-term profitable growth harder and harder to get. The size of these businesses creates complexity and complexity smothers responsiveness. In searching for efficiency *and* growth, businesses lurch from centralization to decentralization and back again. Managers tinker with organization models. Perhaps decentralization will make the business more responsive. Perhaps centralization will cut overheads and make plans more consistent. There has been no obvious right answer, perhaps because there is currently little science behind organization design. Business operates against an unstable backdrop. National economies are themselves unstable and unpredictable. Avalanches of change ricochet between economies, and automated financial trading has arguably made things worse. This instability has a knock-on effect on how businesses trade and how they reorganize to react to such changes. Businesses want, as always, stability as well as profitable growth, but it is getting increasingly elusive for most. This book is believed to be the first to offer some *specific* solutions based on evolution plus the science of complex self-adaptive systems. It tries to describe the complex theoretical background in business terms using readily understood analogies and without using mathematics. This inevitably leads to some imprecision. In addition, a blind eye has had to be turned to some obscurities from theoretical physics that have no real existence outside mathematics. Such simplification will be amply justified if a target readership of corporate-level senior managers, business and financial planners, franchise and brand managers and human resources managers find it intelligible and useful or at least stimulating. Even with this simplification, the material is unavoidably novel and covers a wide range of scientific and information technology disciplines: evolution, genetics and other facets of theoretical biology, complexity theory, theoretical physics, knowledge management, telecommunications and so on. This is in addition to the statistical analyses of business performance and similar subject areas more normally featured in an MBA syllabus.

A route-map is given below to guide the reader through this morass. It tries to outline how all these apparently diverse things are linked together into one coherent approach to issues that are at the heart of business performance.

Chapter 1 - Order and Chaos

The creation of decentralized mutual businesses and the relationship to the theory of self-adaptive systems

The challenge facing businesses is how to be large *and* quick to change. One solution might be to reorganize them to sit on the boundary between order and chaos. In summary, this means simplifying and decentralizing into areas that can make decisions autonomously. Such areas may help each other but may also compete. This continual change in each area that results from both co-operation and (particularly) conflict is *coevolution*. Dee Hock employed these ideas intuitively to build VISA into an internally competing self-adaptive business. The evolutionary ideas he used first became well known through the later publication of Michael Rothschild's *Bionomics.* Stuart Kauffman pioneered the related theoretical background on the boundary between order and chaos.

Chapter 2 - Organisation
Coevolution and the use of Kauffman's models
This chapter describes the basic concepts behind Kauffman's models of competing and evolving entities. These concepts - of 'internal complexity', 'external coupling' and 'landscapes' - are more far-reaching than the models themselves and pop up repeatedly throughout this book, particularly the implications of 'rugged' and 'smooth' landscapes. Natural selection and risk taking on a rugged landscape are outlined and the situations for taking risks described. Sometimes it is better to play safe; at other times taking a risk is a better bet than becoming stuck in a dead-end of sub-standard performance and then being ravaged by the competition.

Chapter 3 - Self-organization
Self-organization and the use of avalanches as a self-organizing mechanism
It may not be easy to restructure an organization to the boundary between order and chaos. The alternative is letting it *reorganize itself* to the boundary. Simple self-organization often fails, however, and there is a need for an additional 'homing' mechanism. Avalanches of change may occur as a self-organizing business nears the boundary between order and chaos and are a useful telltale that the boundary is being approached. The underlying avalanche mechanism is described using several analogies, all but one of which appear here for the first time.

Chapter 4 - Coevolving business organization
How to structure and manage a coevolving business
The natural world appears to be far removed from the business world but there are strong underlying similarities. In particular, there are business equivalents of 'species' and 'gene'. Some of these genes are special in that they represent potential decisions that can change the way the business works. These are called 'decision genes'. The remainder represent other changeable things in the business that we could vary but choose not to. This chapter describes how to reorganize a business - which generally implies simplifying it. Simplification may give rise to internal competition. In VISA's case this was head-on competition between financial organizations. But internal competition can occur within a single business also. Factories may be encouraged to bid for making a particular product for a specified market. Competition between two or more of a business's products in the same market sector may be actively promoted. When a business is deliberately split into competing areas, the optimum size of an organization unit or division needs to be established. This leads to the idea that competing areas can in turn be composed of other

competing areas like Russian dolls - businesses within businesses. Each competing area can then be viewed as a black box that communicates in a formal way with other black boxes but otherwise keeps its innermost workings secret.

Chapter 5 - Competition
Evolution, business planning and brand management
Some businesses take a scattergun approach to product development and marketing. The more successful products survive and the unsuccessful ones are killed off through something akin to natural selection. This test marketing is described in terms of genetic evolution. It is shown to be limited by a 'fitness catastrophe' that results from injecting more and more test products into the market faster than the losers can be weeded out. Brand managers and business planners have a surfeit of information to guide them but it is rarely consistent and there are benefits in ignoring some of the information.
Business planners use computer models to predict business performance. These models are in competition with models from competing businesses because each model includes (or should include) a representation of the way in which each major competing business will react to moves by its own business. This connection between models is usually via product marketing or pricing assumptions. The models themselves thus evolve with each other in the way that the parent businesses coevolve. Brand management is analyzed in terms of coevolution. Brand managers need to develop each established brand faster and better than the competition while avoiding killing it through too much change. Competing brands can end up in a state of armed truce through jointly adopting 'evolutionarily stable strategies' while managing advertising, trade and consumer promotions.

Chapter 6 - Stability, cohesion and growth
Nudging a business out of a rut while keeping its resilience and adaptability
Competition between businesses *and* between areas of any one business is a never-ending series of 'disturbances'. These disturbances can throw a business or business area off course but can also be exploited. This chapter examines the benefits from the controlled injection of disturbances. Ideally, businesses are resilient to disturbances while at the same time adaptive to desirable change and responsive enough to adapt quickly. Business process re-engineering has recently been in vogue, and some businesses even realize that each bout of re-engineering may not be the last. But most are not aware that, on evolutionary grounds, the very processes by which business process re-engineering is managed themselves need to evolve. Richard Dawkins was the first to highlight this point in the context of the evolution of complex organs such as the eye.

Business decisions are the route through which business coevolution is managed, but decisions are often not 'yes' or 'no' but 'how much'. For example, a decision about capital investment for a project will be a budget figure. The decision to proceed with the project itself was probably an earlier yes/no decision.

The effects of business organization complexity on business growth using published data from a basket of US public companies has been analyzed by a group based at Boston University and MIT under the leadership of Gene Stanley. The relationship of this work to coevolutionary models of businesses is described.

Large decentralized businesses need something to make them hang together as one business. This can be achieved through common Vision and Mission statements or a statement of ethics. At a lower level, the benefit of linking efficiency and effectiveness objectives for individuals or departments is described. Balanced Scorecard objectives are shown to be a good intermediate stage between a high-level corporate Vision and individual objectives. Activity Based Costing is demonstrated to be a good cost management framework for an internally coevolving business.

Finally, management by objectives and centralized control can be overdone, and three real-life cautionary examples are briefly described.

Chapter 7 - Communication

Common computer systems and communication between business areas or between businesses affect the impact of coevolutionary changes

Businesses that coevolve internally need good communication between the coevolving areas. Poor internal communication can negate the benefits of restructuring the organization. It stifles the operation of coevolution and can even fight it. Communication can be via standard company-wide data processing systems using a reservoir of company data. It is also via person-to-person communication. This is shown to have close parallels with the protocols used behind the scenes in data communication. Data processing systems need the flexibility to be customized to individual users' needs without ruining the underlying benefits of using standard transaction processes and data. Using data and processes which are common across coevolving areas does not violate the black-box nature of each coevolving area: what matters is whether *decisions* are local, not necessarily *data*. Person-to-person communication is examined in more detail, and some examples are given of current technology that can be used to assist. This includes the Internet. The data communication industry has already encountered many of the problems inherent in the management of large organizations. Designing a large data (or voice) communication network is a balance between the cost efficiencies, end-to-end billing and fault management capability of using a global network and the resilience and manageability of several smaller networks connected together. These lessons are described and then summarized in terms of coevolution.

Chapter 8 - Knowledge management

Why preserving and accessing corporate knowledge matters and how to do it

Each coevolving area is driven by data, information and knowledge. These are both internal to it and, via business systems and person-to-person communication, external to it. The widespread fallacy that, like information and data, knowledge can be stored in overt form is highlighted, and the implications of this error described. A correct understanding of the nature of business knowledge leads to a renewed emphasis on the importance of human experts. This is in contrast to passive technology such as knowledge management systems. The impact of inadvertently creating 'islands of knowledge' - where experts keep what they know to a confined area - is described. Some business cultures such as consultancies even encourage their creation. Internally coevolving organizations would appear also to encourage islands of knowledge through the black-box nature of each area, but this impression is the result of confusing local data, information and knowledge with the local *decisions* that are the essence of coevolving areas.

Chapter 9 - The free marketeers

Building internal and overt markets. Why some work and some fail

This chapter looks at some real businesses and utilities that have undergone market liberalization. This liberalization was either to an internal market that is hidden from the final customer or to an overt market where the customer decides where to place his business. Two are examined in some detail. One is the UK's state-funded National Health Service that was radically revamped by the Thatcher government to create an internal market where family doctors became budget holders and could place hospital and other referrals where they wished. The other is IGOMED - a much smaller Swiss healthcare management collective founded from the outset on Michael Rothschild's Bionomic self-adaptive and free market organization principles. Some UK businesses such as mobile phone service providers reacted well to market liberalization. There was less success in other areas. The reasons for relative success and failure are described in terms of coevolution.

Chapter 10 - Implementation

How these ideas can be implemented

This final chapter outlines how to go about implementing coevolutionary restructuring of a real business. It describes some real businesses that have made coevolution work for them and gives some views of the future, including those of Dee Hock.

Annex - Theoretical background

Several strands of theoretical work coalesced in the 1990s to give some theoretical backing to the approach outlined in this book. They are evolutionary modelling, thermodynamics, chaos, avalanche dynamics and universality, biology, information technology, the analyses of real businesses at Boston University and economics. This annex and bibliography are included for those who want to take the whole subject further and have the scientific understanding and mathematical ability to do so. Since much of the material is collected here for the first time, this annex may well be of interest to students of the physical sciences and engineering and to anyone studying for an MBA.

CHAPTER 1 - ORDER AND CHAOS

Introduction

US businessman Dee W. Hock has a mission to convince complex businesses that the way many of them operate is fundamentally wrong. To all too many, Hock's ideas sound unworldly and infeasible but his career belies it. He built the VISA organization from a faltering conflict-ridden mess into the world's largest financial services business. He ditched conventional business wisdom and went back to basics to find a novel answer in the principles of evolution and natural selection. And his approach is now starting to be legitimized by theoreticians of complex self-adaptive systems.

The challenge

Business is faced with ever-changing market diversity and widespread instability. The current survivors are large, multinational *and* nimble - a very difficult balancing act. They emerge, collaborate, compete and vanish in complex and ever-evolving combinations. A symbiotic web of suppliers, manufacturers and customers feeding off each other is developing remorselessly. Fleetness of foot, cunning and speed of response dominate. Quality is still writ large but is subservient to meeting customer demand through creative proposals and their accurate delivery. Steady and predictable evolution will be illusory: going back regularly to the drawing board with novel business propositions, following Nature's example of sexual reproduction, will be an accepted way to dislodge entrenched attitudes and fossilized processes. Finally, the successful business combinations will hover *on the edge* of chaos: not in disorder, but rather poised in a state of heightened awareness with hair-trigger reflexes, like a runner on the starting blocks. Progress to that state will not be attained by any business plan, but will be an inexorable path for the survivors. The others will go the wall, and if Dee Hock is right these will be the bureaucratic ones who clamp down on individual creativity.

Dee Hock and the rise of VISA

The forerunners of credit cards as we know them were first issued by New York's Franklin Bank in 1951. In the late 1950s, credit cards hit California when Bank of America launched BankAmericard in San Francisco. In 1966, in response to competition from rival MasterCharge (later renamed MasterCard), Bank of America franchised BankAmericard to other banks across the country. Competition followed suit. The result was a mad scramble for business that brought the credit card industry into disrepute. In 1968, Hock, then president of one of the BankAmericard franchise holders, accepted the job of sorting out the mess. This he achieved by designing a mutual organization not owned by any one financial institution. Thus in 1970 Bank of America ceded ownership to a federation of financial institutions called National BankAmericard and which was renamed VISA International in 1974. VISA was, in effect, a franchiser that was entirely owned by its franchise holders. Each member bank was free to take its own initiatives and poach each other's customers. In addition, merchants who took VISA had to accept cards issued by any participant. Agreement was limited to using a common card format and a way of settling transactions similar to the ones used by banks for clearing cheques. Hock had the prescience to realize that moving money in the form of cash or cheques was a dying concept and one that was slow and expensive to operate. What was needed was a truly global business that transferred 'value' around completely electronically. The result was an organization that grew rapidly worldwide, coped on the fly with national quirks without the drag imposed by a conventional business hierarchy and branched out into other forms of financial transactions. Debit cards and electronic banking are big in Europe, for example, while the credit card still reigns supreme in the US. VISA International, as distinct from its member banks, limited itself to setting technical standards and undertaking some broadbrush market development.

Franchise brand owners such as McDonalds and Holiday Inns do the same. The holders of the individual franchises often have some territory to call their own but they otherwise compete with each other as well as with Burger King and Ramada Inns. Franchises do not, however, always lead to genuine competition with resulting prices that benefit the consumer. Merely having some exclusive territory distorts an otherwise free market. In the UK, retail car prices have for many years been around one-third higher than in other comparable countries. The more obvious differences such as the need to provide right-hand drive (which is not unique to the UK) and arguably a somewhat higher specification cannot account for the disparity. The lack of effective competition is due to the unhealthily close tie between manufacturers and their dealers, an EU exemption from the requirement to supply other retailers and which is designed to protect exclusive dealerships, the power the manufacturers wield to influence product supply and the disproportionate discounts given to fleet buyers at the expense of the individual small dealer and retail buyer. The latter discounts were outlawed in August 2000 but manufacturers were already sidestepping legislation by offering fleet buyers vehicles with different specifications to those that appeared in the dealers' showrooms. The publication of recommended retail prices, reductions in basic dealer margins and a parallel increase in discretionary bonuses to dealers have all limited their freedom to drop prices. But by this time a growing number of retail buyers had had enough of such price rigging and bought their new cars via one of the many specialist Internet-based car importers that were set up by manufacturer-independent organizations such as Consumers' Association, Virgin and Direct Line. Manufacturers such as Mercedes-Benz and Land

Rover accepted the inevitable and cut prices; others - the die-hards - refused. Whatever price-capping legislation is imposed, it is likely to be the retail buyers who will force the manufacturers' hands. They, unlike the dealer network, are free to buy where they want within the EU. All they needed was some easy way to do so.

But protectionism extends further. Both Japan and the UK use right-hand drive cars but attempts by Japanese dealers to sell 'gray imports' direct to UK supermarkets and the like have been strangled by the ridiculously small quotas imposed by UK legislation. These quotas are *in addition* to an EU regulation which limits total Japanese car imports to a 12% market share. Achieving true competition is seldom straightforward.

Intervention by government to break supply monopolies and extinguish price cartels can, however, make things worse. Suffocating layers of legislation can stifle the business growth and pricing flexibility that they were intended to promote. A recent study on productivity for the UK Government by management consultants McKinsey highlighted one reason above all others for the relatively high prices in the UK: interference with the free market by the EU and by the Government itself.

Governments in general, but particularly those of Western Europe, act as if they believe they can control employment rates or economic growth. Their record of actually doing so is laughable. Economists who claim to understand the workings of economies at the high ('macro') level are no better. Free marketeers rubbed their hands for joy when the CIS tried to deregulate the former state-managed - or rather state-strangled - consumer goods production sector. But in their euphoria they disregarded Adam Smith's warning that a free market can only work when the national culture shifts to accept it and that it is still desirable that the State intervenes to manage people's expectations and to stop them being exploited. The result of ignoring this warning is all too painfully obvious to those who live in modern Russia. It was not just national attitudes that were in need of change. Communism's bureaucracy was wedded to planning and controlling every facet of production and distribution and was itself largely cushioned against the harsh realities of food shortages and dismal accommodation. It had most to lose by any shift to an open market. Communism has its limits: eventually the innate drives for personal gain or avoiding personal loss win through. But Communism could not give way successfully to a free market without the thorough dismantling and discarding of the hierarchy set up to perpetuate it.

Hock's vision of competition could not be more different from that of the EU or the CIS. He pictured businesses that were decentralized and evolving and where there were just a few mandatory rules about how the participants must work together. The rest was free and fierce competition with resulting successes and failures. It was an environment that was self-adaptive and very reactive without degenerating into disorder. Hock coined for this the name *Chaordic* - something on the boundary between chaos and order. But Hock also recognized that for this to succeed, it needed employees who were treated responsibly and given responsibility. Organization structures, ways of working and goals needed to be readily understandable by all participants. In other words, the freedom within a framework went hand in hand with care for the individual - Adam Smith again.

With its market dominance and federal structure, VISA might seem unassailable, even by rival MasterCard International that has a somewhat similar structure and has a larger share of some international markets than VISA. The expense of setting up a new competitor and building the necessary transaction processing network from scratch would be daunting. But VISA and the other conventional credit and debit card businesses are now threatened by the growth in transaction processing over the one even larger chaordic enterprise - the Internet.

Hock left VISA in 1984 but continued to study the growing literature on genetic processes, complex systems and chaos that was catching up with his vision. In 1993, he visited New Mexico's Santa Fe Institute, then as now the hotbed of complex systems thinking. The theoretical underpinning of Hock's vision lay there with the work of Santa Fe's Stuart Kauffman on *self-organisation.*

The notion of self-organization has been around a long time in both ecology and economics. The natural world lives in an uneasy balance between prey and predator, parasite and host and in competition for limited resources. If predators kill all their prey or if parasites kill their hosts before they or their hosts reproduce, they themselves go out of business. Living things do not *want* competition. It wastes energy and can be fatal. So they seek out niches in which to thrive and where resources are plentiful and competition is weak. This is a natural response to competition. No master plan is needed. Insects such as ants that live in colonies organize their own divisions of labour. The relative numbers of foraging specialists, guards and so on are nicely calculated to optimize the reproductive success of the colony. The queen may control their basic behaviour by broadcasting chemical signals but she is not a cold calculating Lucrezia Borgia. She is merely part of a self-regulating system that has only one aim: to perpetuate the genetic lineage of the colony that, because of its incestuous nature, is mostly hers. Stock markets work in a similar way. There are some overriding and often-unnecessary legal restraints but the actions of lots of buyers and sellers set prices and regulate the market. There is again no master plan. The resulting stock prices are unpredictable, even in theory. The close underlying connection between ecology and economics, and the difficulties of predicting the behaviour of either of them, came to prominence with the wider availability of cheap computing power. This enabled researchers to model ecological and economic activities - and then to stand back bemused at the unexpected and complex behaviour that unfolded. Michael Rothschild called this self-regulating relationship between organisms and their environment *Bionomics* (in his book of the same name). And he pinpointed the essential common feature: that whereas living systems continue to thrive through the evolution of their genetic information, economic systems progress through evolving technical information. Both evolve in parallel but very similar ways.

Ecosystems have a free unregulated market except where we meddle with it with misguided attempts to repair the outcome of our equally misguided spoilage of the natural world. Unregulated economic markets are also effective except when government interferes with their free working in order to satisfy political factions or lobbyists from particular industrial sectors who want ‘protection’. For example, until 1995, through a long-standing anomaly in the law, the retail prices of books in the UK were controlled by the publishers. When some booksellers attempted to secede from this cosy agreement, the remainder lobbied hard to keep protectionism in place. Small bookshops would be forced out of business. Shops would prune their lists down to popular titles only. Education would suffer. And so on. “We thought the world as we knew it would end” said Tim Godfray, chief executive of the Booksellers’ Association trade body. The result of the abolition of this ‘net book agreement’ stunned the critics. True - some inefficient and unattractive bookshops were badly hit, but book sales actually increased overall in every year since 1997 after a temporary blip the previous year when the trade was sorting itself out. Larger bookshops stocking a wider range of titles flourished. Big US retailers such as Barnes and Noble - hitherto discouraged by endemic price fixing - entered the newly liberated market, and this prompted UK retailers to enter into purchasing alliances or simply to merge. But smaller specialist bookshops also thrived in the residual niches and prospered through the renewed interest by the book-buying public who were willing to pay for the value added by the specialist. The same pattern is repeated wherever monopolies are broken or legal protectionism repealed. But it is impossible to predict the outcome accurately. The free market works and it works efficiently if unpredictably. Central control of something as complex as an economy fails to work because there are so many competing interests. Communism in the former USSR drifted on through the 1950s and was certainly effective enough in scientific areas to be a serious military threat to the West in the early 1960s. But it was horribly inefficient. When workers have no incentive to improve the job they do or how they do it, corroding inefficiency and bloated bureaucracy become the norm. The economy falls further and further behind those of countries where individual initiative is rewarded. And unless their homeland can, like pre-war Japan, be kept isolated from the rest of the world - which in today’s interwoven global economy is impossible, something eventually gives way.

Organization is done for a reason. Soviet apparatchiks organized farmers’ collectives with the aim of producing food. Businesses reorganise to make more profit. But what is the aim of *self*-organisation? Is self-organization an all-or-nothing phenomenon? If a disorganized system becomes self-organized, how does it know it has finally got there? And does it stop at that point? Or does a self-organized system live in a world of never-ending change? Finally, is there some underpinning bedrock of theory that can be used to make forecasts of how the system will evolve, or at least to put limits, however broad, on the way it will evolve?

Until around ten years ago there were no satisfactory answers to these questions. Dee Hock arrived at his novel solution for VISA intuitively. Michael Rothschild described the benefits of self-organization and the perils of restraining a free market. But Stuart Kauffman and his colleagues went the next stage and elucidated exactly what happens when self-organization occurs, and described some of the very strange things that happen as a system approaches self-organization or is disturbed when there already.

Order and chaos at Santa Fe

Stuart Kauffman is a theoretical biologist who has spent most of his professional life trying to work out how chemical and biological systems organize and evolve without outside intervention. Kauffman's innovation of a technique to model the processes by which species in competition evolve together finally gave a way to study the dynamics of Hock's organizational ideas. The model uses simple genetics and will be described in some detail because the underlying concepts of *internal complexity* and *external coupling* will be used repeatedly throughout this book.

Take for example a rabbit. Let its characteristics be determined by the values of its genes. Any rabbit gene can have one of two values - a gene for blue eyes or brown eyes for example. Genes work in concert: the physical appearance and behaviour of our rabbit is rarely determined by individual genes working in isolation but by how they work in complex combinations. Let each of its N genes be associated with K others such that the effect of any gene in determining appearance and behaviour is governed both by its own value and by the values of the K other genes to which it is coupled.

Now imagine two *different* types of animal competing for existence. Foxes compete with other foxes to catch rabbits. The successful and well-fed foxes have more offspring. The less successful ones die out over time. The successful rabbits - the ones that 'breed like rabbits' - are the ones that evade predatory foxes through speed and agility and who commandeer the more nutritious patches of grass. Other rabbits become dinner or die of starvation - natural selection in action. Over time, foxes evolve to become better rabbit catchers and rabbits to become better fox evaders. There is an uneasy equilibrium and the populations of both fox and rabbit rise and fall. Foxhunts breed hounds to catch foxes. However, like racehorse owners, huntsmen try to assist evolution by selecting for breeding just those animals who they think have the most desirable characteristics, in this case a combination of speed, stamina, scenting ability and so on. If a *business* could be 'bred' to compete in this way, four questions need answering:

- can it behave in a rational way and evolve when not conventionally 'directed from above' by autocratic senior managers?

- can such evolution nevertheless be *assisted* by structuring the business in some way to speed up the process that brings the business to a reactive state poised on the boundary between order and chaos?
- can this boundary be pinpointed? In which direction does it lie and how do we know when we have got there?
- can a business *find its own way* to the boundary and stay there, reacting to disturbances like a boxer yielding to punches and then coming forward again?

Dee Hock proved that the first is possible. It now appears that the remainder are possible also.

CHAPTER 2 - ORGANIZATION

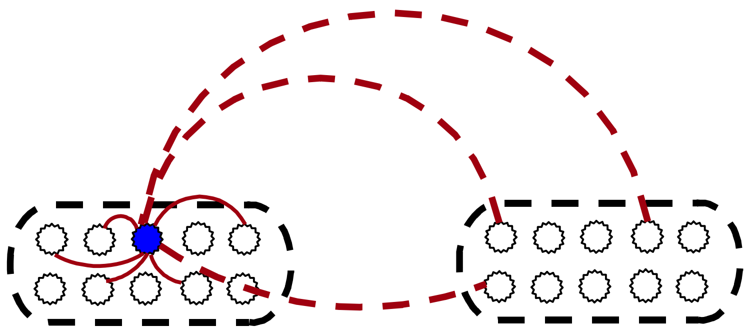
Coevolution

Kauffman's NKC technique for modelling evolution is named after three of the main factors that determine the behaviour of species interacting and evolving with each other. Picture again our hungry fox with N genes, each associated with K others. The effect of its genes working together in this way is to give it some rabbit-catching capability. The rabbit is similarly endowed with fox-evasion skills. Foxes and rabbits slowly evolve through reproduction, subsequent natural selection and some random gene changes (mutations). A combination of gene values that makes a fox a great rabbit catcher will be matched over time by rabbits acquiring combinations of gene values which make them great fox evaders. If not, all rabbits get eaten and ... no more rabbits unless foxes die of rabbit-starvation first - Nature is unsentimental in these matters. Thus gene values in competing species evolve *together* - they *'coevolve'*.

To represent this in Kauffman's model, assume that each of the N genes of our animal is coupled not just with K other genes of its own but, via competition to survive and reproduce, with C genes from the competitor. The rabbit's fox evasion skills are determined not just by (some of) its N genes each acting in concert with K others but also by the values of C genes of the fox. As rabbits get better at evasion through favourable selections of evasion genes, so these selections become less favourable again for the rabbit because the C-coupled genes in the fox are also changing to fight back and be better at rabbit-catching.

This concurrent and linked evolution of the two species can be represented by two 'fitness landscapes', one for foxes and another for rabbits. Fitness landscapes are a concept first introduced many years ago by biologist Sewall Wright. Each landscape has hills and valleys. Each point - each grid reference on the map - represents a particular combination of values of the N genes. Through reproduction and selection, rabbit gene-values change in an attempt to climb the nearest hill. The higher up a hill a rabbit gets, the better - fitter - it is at fox-evasion. Foxes evolve on similar landscapes of their own. Fitness for them is rabbit catching ability. *But the landscapes themselves are not static*. As a rabbit's fitness climbs upwards and the rabbit becomes better at evasion, the landscape beneath changes - the hill gets smaller - because the fox is simultaneously climbing *its* local hill to be better at catching rabbits. Like two children playing on an inflatable ('bouncy') castle or water bed, when one moves, the ground beneath the other also shifts. As the two species evolve, their landscapes continually deform each other through the C-coupling of their respective genes.

Finally, in Kauffman's NKC model, there are not just two but S species, and a more refined version has each of the N genes also coupled to W genes of an external world. In our case this external world would, for example, represent the availability of grass for rabbits. There is an assumption here that however much grass rabbits eat, they cannot significantly affect the remaining supply of grass, whereas a shortage of grass may decimate the rabbit population. Thus, unlike other couplings, the W-coupling works in one direction only and is a useful mechanism to disturb a coevolving system from outside in order to find out what happens. To summarize: each coevolving species has a number of genes (N) coupled *within* the species (K), coupled *between* species (C), to a number of species (S) and optionally to W genes of an external world.



A gene in one species ('object') coupled to K=5 others internally and C=3 others in another species (object)

Assigning fitness

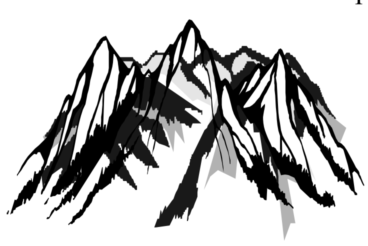
Computer modelling then simply requires selection of particular values for N, K, C and S. The W coupling to an external world will be ignored for the time being. For example, assume that a species has N=10 genes, each of which are coupled to K=5 others of its own and C=3 of another species. Assume also that there are S=2 species. Before starting the model off, one of two fitness values (say 0.3 or 0.4 where 1 is ultra-fit and 0 is lame) needs defining for both of the two possible values of each of the N genes in each species. These values are affected by the values of the K-coupled genes within the species and the C-coupled genes of the competing species. This means that for each of the two values of each gene there can be very many possible values of fitness. For example, if in one of the species gene 6 (of the ten) has the value A (rather than B) and the five (K=5) other genes within the species to which it is coupled have values A, A, B, A, B, and the three genes in the competing species to which it is coupled have values A, B, B, then this particular combination has fitness 0.3 say. Fitness values can be assigned randomly or using some prior knowledge of how genes act in concert.

Rules for survival

To run the model, two rules are needed. The first is a rule for the sequence in which we select each species in order to induce it to evolve and improve its fitness. A simple rule for this could be to select the species' landscapes one after the other on a cyclical (round-robin) basis. There are many other possible rules including one to select the least fit species. This latter rule is the principle of natural selection when applied to individual living things. The second of the two rules we need determines what we do when we have earmarked a species, and hence a landscape, for evolution. One rule could be to pick a point - a particular combination of the N genes - at random and change one gene value from B to A (or A to B, depending upon what value it currently has). If the average fitness of the N genes improves with the new combination, select it and hence move on the landscape and up a hill. Otherwise ignore it and move to the landscape of the next species. A more complex 'greedy' rule could involve trial changes of all the N genes of a species, one after the other, and selection of the one change that gave the best improvement. Even more complex hillclimbing methods exist where several genes are changed at the same time.

Landscape portrait

When this process is simulated on a computer, unexpected things happen. Some combinations of species evolve to coexist. This co-existence is not friendly co-operation but a state of guarded and watchful inactivity. If any one of the species were to evolve further, it would become less fit so it has no incentive to do so. Some species drive others to a state of very low fitness verging on extinction. Yet other combinations of species appear never to stop mutually evolving, each striving for fitness supremacy but being foiled by concurrent evolution of the others. These, however, almost always settle down after a lengthy bout of weaving and dodging. One general characteristic of the models is that those with low K - with genes more or less independent of others in the same species - have *smooth* landscapes with a few very high gradually sloping fitness mountains. Those with high K - with genes linked to many others within the same species in a complex web - have *rugged* landscapes with lots of smaller hills with steep sides.



A rugged landscape which has many peaks. By contrast, a smooth landscape has perhaps just one high peak - like Mount Fuji

Natural Selection

It is worth at this point defining exactly what we mean by evolution through natural selection. Natural selection is colloquially described as the 'survival of the fittest', but since the fittest are the ones who survive, the statement does not say much. Evolution through natural selection firstly means defining fitness as the ability to pass on (or to assist others to pass on) your genes. The more offspring you have the fitter you are. Offspring are just a convenient vehicle to carry genes. Nature implements natural selection in many ways, the most obvious one of which is for predators or the weather to kill off those that are physically less strong or robust. The stronger creatures survive to reproduce, and some offspring at least will inherit their additional strength. They will go on to reproduce further. Traits such physical strength, better reactions, better eyesight and for those creatures which reproduce sexually, attractiveness to the opposite sex will be honed as generation succeeds generation. But for natural selection to work, three conditions are needed. There has to be an *abundance of variants* (in our case different combinations of the values of N genes) to select from, *different fitnesses* for the different variants to allow the fitter to be selected, and *reproduction* to allow this select elite to propagate.

This may sound to a marketing manager like brand development using test marketing and, as with brand development, there are constraints. An animal may build muscle to fight off a predator but muscles cannot keep getting larger. There will come a point when the muscle bulk interferes with agility or needs too much food to sustain it. Excessive bulk may also become too much for the architecture of the skeleton and it will buckle. The levers of brand management - levels of advertising, promotional spend, price protection and relative usage of the more expensive raw materials for example - have to be in line with their benefits. Huge long-term investment in advertising by a small company could conceivably have a good return except that the company would be forced out of business first. There are always limits.

It has been shown that where absolute constraints on development exist, whether limits to an animal's muscle bulk or limits to how much a business can spend on advertising, the constrained traits lead to a stable state of coexistence whereas (relatively) unconstrained traits lead to constant coevolution.

Altruism

There is as little real altruism in Nature as in business. Each living thing has a primordial urge to perpetuate its own genes above all else. If not, its species would have succumbed very early on. It may perpetuate its genes in several ways of which sexual reproduction is only one. Creatures such as bees that live in extended families sometimes behave in an *apparently* altruistic way. A bee may, for example, attack a predatory wasp and die 'for the good of the hive'. But their apparent altruism stems from their close genetic relationship.

Each sister worker-bee is more closely related than any offspring would be. For a worker-bee, dying to protect her numerous sisters is *better* than having offspring: she is more closely related to them (has more genes in common with them) than she would be with any offspring. So it is no surprise that she is sterile. But bee colonies are somewhat unusual genetically in that they have an army of full sisters, a queen and a few reproductive male drones. For animals, including humans, who reproduce sexually and have a more normal mix of sexes, this is rarely a good proposition since having offspring is at least as good as having siblings to carry on a genetic lineage. A mother is certain that her children carry lots of her genes. She can never be *quite* as sure about her brothers and sisters. She only has her own mother's word for it. Was their father really her father? Are they adopted? Is *she* adopted?

The reasons for the existence of altruism in the animal world have probably caused more heated debate among theoretical biologists than any other topic this century. Evolutionary biologist Bill Hamilton who died in March 2000 first proposed what came to be called the 'selfish gene' view of evolution or 'socio-biology'- the principle that individuals are simply vehicles for competing genes that are waging war for supremacy. This stirred up sufficient trouble for the two-part paper concerned to be cited probably more than any other scientific work in history, both by those such as Richard Dawkins and John Maynard Smith who agreed and by those for whom the ideas were abhorrent. Hamilton developed his ideas with political scientist Robert Axelrod to show that what appeared to be altruism could be explained as the by-product of competing genes fighting for their own survival.

The underlying point at issue is whether altruism is simply a side effect of evolution through natural selection at the gene level or whether altruism primarily benefits the *group* to which I belong. Does my altruism benefit my genes because people are altruistic to me in return or is it a result of natural selection at the group rather than the gene level? Charitable institutions apart, business altruism appears to operate as if it were the result of selection at the gene level. If you are a supplier to me, I want to pay you as little as I can get away with. But if I intended to buy from you repeatedly rather than shop around each time, I would allow you to make just enough money to stay in business and to undertake just sufficient R&D so that you remained a competitive supplier. The notion of group selection - the preferential selection of whole groups of individuals (supplier and customer combinations, for example) as opposed to single individuals or single genes, has come back into fashion in the last twenty years in an attempt to explain some of the outstanding oddities of animal behaviour. Behaviour which *looks like* altruism clearly occurs, but whether it is the result of selection at the gene or group level, or whether group selection exists at all, no-one yet knows.

Risk assessment

Random mutations apart (see Chapter 5), a species which gets stuck on the summit of a low hill in a rugged (high-K) landscape might be stuck forever unless a competing species deforms the hill (through C-coupling) and gives the first species a chance to evolve further. Evolution through natural selection is blind: there may be a higher fitness hill nearby but a species stuck on the summit of a lower hill cannot 'see' it. It moves from the summit at its peril. Such a move involves going via an intervening valley of lower fitness that may result in its extinction. But living things *do* have such vision. This vision is innate in most animals, but in humans (at least) is a conscious one as well. We can, for example, invest now - become less fit - in order to reap a payback later. We mentally simulate the likely outcome of such a decision before we take it. We assess the risk before we invest. Other animals may do this instinctively, as we ourselves do when dodging a thrown stone or grabbing for a handhold when falling. Drivers are probably aware of doing both: of reacting instinctively when hitting a patch of ice but also of quickly analyzing a complex and potentially dangerous traffic situation looming up, together with the risks of each course of evasive action. A tiring rabbit fleeing from a fox might take the risk of seeking sanctuary in a farmyard - somewhere it would normally avoid. It *might* become rabbit stew or dog food but the chance of the fox being deterred might make the risk worth taking. It is likely but unproven that evolution has honed a rudimentary risk assessment mechanism within the rabbit's brain. Over the years, rabbits with better risk assessment survive and breed. The others die earlier and leave less offspring.

It may seem contradictory to say that evolution through natural selection is blind to risk assessment but that living things are not. The answer to this conundrum lies in the mechanism of natural selection. Natural selection itself merely weeds out the unfit continually. If a fit creature becomes less fit through taking a risk, such as the last-ditch attempt by our rabbit to evade a pursuing fox by running into the farmyard and unfortunately getting caught by the farm dog instead, it is harvested along with the halt, sick and unlucky. Assume that many rabbits try this and that more survive than become dog food. Then a gene combination which says in effect 'if sorely pressed then an incursion to somewhere normally dangerous is a marginally better bet than trying to outrun a predator which is close on your heels' will spread throughout the rabbit population.

Order and chaos

A high-K landscape is an orderly environment because each gene is enmeshed with many others. Like a fly in a web pulled from all sides, making progress is like wading in treacle. Species on smooth landscapes may be condemned to climb (almost) forever towards a distant high peak, but the peak will be continually on the move as the landscape is deformed by competing species - a non-stop chaotic world. Kauffman found that there were combinations of K and C which *just* enabled each species to surmount the peak of a mountain: each evolved *just* fast enough to keep up with their mountain as it moved.

This combination denotes the boundary between order and chaos.

CHAPTER 3 - SELF-ORGANIZATION

Homing instinct

K and C have until now been treated as if they were knobs for manually tuning the system to the boundary between order and chaos. However, Kauffman asked a more fundamental question: what could a species do to *evolve itself* to this utopian boundary? The answer lay in allowing evolution within each species not just to select gene combinations that lay along uphill paths *but to alter the value of K also*. As before, when attempting to improve the fitness of any one of the landscapes, the model was told to make single gene changes where this improved fitness. But if no improvement in fitness were possible in this way because the species had climbed a small fitness hill and become marooned on the summit, or if changing K were a fitter option, the model was allowed to change K instead. The result is an evolving mix of species that eventually settles down to a life of stability punctuated by occasional bouts of warfare and temporary insanity. At this point the species have jointly homed in on to the boundary between order and chaos. And each species stays there forever, adapting to the evolutionary twists and turns of others, stable but by no means static, with low values of K and C and with K roughly the product of C and S.

Which way is home?

That is the theory. In practice, species often fail to home in to the boundary between order and chaos without some external help. Particular values of C might need to be manually selected, for example. Kauffman's model was actually a little more complex and allowed a species to become extinct if another fitter species could invade and take over its niche while still staying fit. In a note to the present author, Kauffman explained that these take-overs are the subtle underlying mechanism needed to drive the homing-in process. A fitter species which takes over from a less fit (and thus now extinct) species brings with it a new value of K. If this species remains fitter than the competition, it may invade yet another species' niche and make that species extinct also. So better values of K spread throughout the population faster than the mere individual adjustments to K could manage. One mechanism reinforces the other. Kauffman's model also had a larger number of species than S. In other words each species was directly C-coupled to only *some* of the other species. But this still permits knock-on effects (chain reactions) to occur in which one species affects a second which in turn affects a third and so on. This is common in Nature: when one item in a food chain is removed the repercussions can be far reaching. For example, in our simplistic foxes and rabbits model, no grass leads to no rabbits which in turn leads to no foxes. However, since an abundance of foxes is only linked to the grass supply indirectly via a supply of rabbits, any grass shortage does not immediately affect the fox population. Foxes will perhaps switch their attentions to nearby hencoops until local farmers get wise to their activities. Only when the hungry rabbits die of starvation in large numbers do foxes get seriously malnourished. Foxhunts are even further removed from the availability of grass and are well buffered for some time. But finally, when the rabbit population crashes, the knock-on effect sweeps through the system as a series of avalanches affecting both foxes and foxhunting.

Similar indirect links occur in many other spheres. For example, the difference in behaviour between the market performance of a unit trust and of a similar investment trust perplexes many amateur investors. The market price of a unit trust's shares directly reflects the current market prices of its underlying portfolio. The share price of an investment trust - typically a publicly quoted company in its own right - with an identical portfolio reflects the market's view of its performance. It asset value is only *indirectly* related to the value of the underlying portfolio and is usually less.

Avalanches

Avalanche behaviour occurs in many different areas of physics and biology as well as on snowy hillsides. It is best exhibited by a simple experiment with a pile of sand. Trickle dry sand by hand on to a table. As trickling continues, the sand will build up into a cone-shaped pile until it reaches a critical height. Further trickling will cause avalanches down the side of the pile. An avalanche may be tiny - just a few grains - or take half the pile with it. It is impossible to predict what will happen in any one instance, merely that small avalanches will be much more common than large ones. The sandpile 'wants' to stay at its critical height and sloughs off attempts to pile it higher. It lives in a state of what physicist Per Bak called 'self organized criticality'. Bak and colleague Kim Sneppen at the University of Copenhagen's Niels Bohr Institute have studied at length some simpler variants of Kauffman's model. Their aim was to find ways to make the boundary between order and chaos sharper and thus a stronger attraction for species to home in to - like moths being more strongly attracted by a brighter light. Instead of selecting for evolutionary development a species at random or on a cyclical (round-robin) basis, Sneppen decided to select for development or extinction the *least fit* each time and, like a falling 'roped' climber, for this species to drag down its immediate neighbours with it. The result of this continuing process is to raise rapidly the fitness of all the species to similar levels, like reducing the allowable sizes of all intermediate stock-holding points in a logistics pipeline to similar small values. As species approach the same fitness, avalanches of species' extinctions flood the system as the extinction of one species has a knock-on effect on the next. Bak's self-organized criticality in such an ensemble of species is what Kauffman's self-organization to the boundary between order and chaos is in one species.

How will we know we have got there?

A.A. Milne's Winnie-the-Pooh found a real Pole on his epic 'expotition' to the frozen North. Pooh was lucky: the boundary between order and chaos is not so clear-cut. Imagine standing beside a frozen lake on a warm day. If we make a small hole in the ice by the edge and drop a stone in, the ripples will be blanketed by the surrounding ice and peter out quickly. As the day gets hotter, heat percolates downward and the ice starts to melt. Tendrils of water appear between still-frozen sheets of ice. Ripples from the stone might now travel further across the lake via the zigzag of tiny rivulets between ice sheets. If we keep trying, the magic moment will come when the ripples manage to cross the lake to the other side, albeit by a roundabout route skirting sheets of ice. If the ice is melting uniformly, ripples from stones dropped in at other places on the edge may well also manage to cross the pond. The lake is still a patchwork of areas of frozen stability, but things that happen in one part of the lake can now affect unfrozen parts at any distance. This, in the lake's terms, is the boundary between frozen order and liquid chaos.

Avalanches of fir-cones

However, we are not dealing with something continuous like water but with a collection of discrete species, so it seems appropriate to borrow from Pooh's game of Poohsticks and elaborate the model of the lake. Assume that it was densely covered with fir-cones, each cone representing one species. Assume also that the cones were contentedly bobbing up and down until they became trapped in the ice when the lake froze over. As the ice melts again, meandering lines of fir-cones are freed up across the pond between the swathes of other cones still held in the sheet ice. A sharp nudge to a free cone at one side of the pond will cause that cone to bump the next one and so on. The series of collisions of one cone into the next might peter out or might go right across the pond. The bumping between cones may take many routes, some of which may converge again after traversing either side of an ice floe.

With luck, however, something a little more interesting may occur. A gentle impact from one of the newly freed cones may liberate a cluster of cones in an almost-melted sheet of ice. These cones may exercise their newly found freedom by colliding with all their neighbours and causing waves of collisions in all directions. Further cones in adjacent almost-melted sheets of ice may also be freed and bump into yet others. This chain reaction eventually settles down. The effect of the initial nudge that started it all dies away and the cones stop moving. As they do so many will freeze back into the remaining semi-frozen ice sheets.

It is worth summarizing salient points from this admittedly imperfect analogy:

- ❖ at the point when the lake was just melting, most of it was still stable (frozen) but collisions could find their way in any direction across the lake between the frozen ice sheets
- ❖ a small initial jolt gave rise to an avalanche of collisions whose size was unrelated to how hard the jolt was
- ❖ without avalanches caused by the release of cones from their frozen state, any small nudge would peter out quickly and would probably not get to the other side of the lake
- ❖ with avalanches causing other avalanches, any or all of the cones on the lake could be affected
- ❖ if most parts of the lake are frozen quite hard, perhaps shadowed from the sun, the effect of avalanches is contained
- ❖ if the cones are widely spaced such that sideways movement by one has little effect on many others (i.e. the cones have low C-coupling), each cone moves very predictably away from the nudge and is unaffected by rebounds from others. Cones tightly packed (lots of C-coupling) bang into each other repeatedly in unpredictable ways as a result of a nudge to one of them.
- ❖ those cones which are in water which has fully melted are free to move in any direction: they are moving on a low-K landscape. Cones that are in still partly frozen water are constrained like the fly in the high-K web.

The biggest difference between this analogy and the structure of an NKC model lies in what K represents. In the analogy, K is represented by the *external* environment - freezing water. K in an NKC model is, however, *internal* to the species.

Survival of the fittest is elimination of the weakest

Although the lake and fir-cone analogy demonstrates C-coupling (cone interaction), K-complexity (the state of the water - frozen or not) and the transition from frozen order to liquid chaos, it ignores *fitness.* A cone is either free to move or frozen in place; one cone is otherwise the same as any other. A different analogy is needed to illustrate fitness. This is worth pursuing since fitness is an essential input to at least some of the rules we could use when selecting which species in an NKC model should evolve next. The analogy must combine four concepts: fitness, selection, modification of what has been selected, and the resulting interaction between species via C-coupling.

Imagine an earth embankment holding back river water. If the rainfall is unusually heavy and the river is unusually swollen and rising very gradually, a point may come when the embankment will start to leak at its weakest points. Some leaks will be pinpricks; others will be more serious. Let each length (a metre say) of the embankment represent a species and let the strength of the embankment at any point be a measure of fitness of that species. Assume that each length has one leak. Some leaks will be small (the embankment there is relatively fit) and others will be spouting torrents. In an attempt to shore up the embankment, we need to block the leaks and thus need some rule for which leak to seal first. We could choose them at random, or work from one end. Or we could choose the smallest or largest leak first and then the next smallest or largest and so on. Assume that all leaks can be seen from wherever we stand, and that when we have sealed one leak we can go directly to the next one without traipsing up and down the length of the embankment. If the fitness of the embankment as a whole is measured by the total amount of water being lost, it is commonsense to seal the largest leak first. This seems to be the fastest way to raise the embankment's fitness but also the fastest way to put more pressure on the remaining leaks: with the first large hole plugged, the river will now rise faster as less water is being lost. As soon as we have sealed a leak and have moved on to the next smaller one, the length of embankment just sealed will start to leak again. But we cannot predict whether the new leak will be a large one or a small one.

However, if we change the rules slightly such that when one leak is plugged we also plug the two immediately adjacent ones on the left- and right-hand sides, *however big or small they are*, something remarkable happens. At the start, our quest for the largest leak to plug takes us backwards and forwards along the embankment at random. But, after a while, we find that the next largest leak to plug seems closer at hand. We encounter large leaks in clusters. We plug one large leak and the next large one pops up nearby. Large leaks have *not*, however, clustered in particular areas *at any point in time*.

An aerial snapshot would show that our remedial work on the embankment is being only partly successful. The embankment looks in better shape but still leaks. With a few exceptions, the leaks seem to be concentrating into a band of sizes ranging from a pinprick to one-third of the largest size we started with. Within that band, all sizes are equally likely. But - strangely - however hard we try, we make less and less headway. It seems to be impossible to reduce all leaks below the 'magic one-third' barrier that impedes further progress.

It is the very act of plugging a leak that causes another large one to appear nearby, most but not all of the time. Plugging a large leak can thus trigger off an avalanche of large leaks. Sometimes we find we have plugged an entire cluster of large leaks - one 'avalanche' - and the next large one is some distance away. As plugging progresses, this separation - the distance we have to walk between largest leaks - acquires a characteristic mathematical signature. There are many small walks but much fewer long ones. Note the difference between:

- a snapshot which shows no marked clustering of leaks

and

- the places we encounter largest leaks as the plugging process continues

The first is at a point in time. The second is the result of the process's evolution in time.

We now need to look a bit more closely at the barrier we encountered. The size of leak we select for plugging, still the largest one each time, hits a size barrier which prevents further progress. This leak is around one-third as big as the large leak we started with. But however long we carry on plugging, we never get around to plugging the really small leaks by *directly* selecting them although we may plug some small 'neighbours'. What happens is that when the leak selected for plugging becomes as small as the barrier size, it sets off avalanches of change to the leak sizes up and down the embankment. This creates waves of change leading to more leaks of all sizes and means that we have further larger ones to plug again.

Where do these avalanches come from and why did they not appear when we merely plugged individual leaks and not the two adjacent ones also? The answer is simple but, surprisingly, the detailed mechanism has only been elucidated since 1990 and is still being studied. When, as well as our own 'largest' leak, we plug the two neighbouring leaks irrespective of size, these neighbours (which may have had small leaks) are now given the chance to create leaks of any size. The ones that are now larger than the 'largest' one we have just plugged immediately become candidates for plugging in their own right. If one of these is plugged next time round it causes *its* neighbours to change size. Clusters of consecutive leaks change size in this way and groups of changed clusters appear up and down the embankment. A cluster that has just changed may be forced to change again when two clusters widen, become adjacent and bump into each other. Waves of change of leak size move backwards and forwards along the embankment as we move up and down still plugging the largest leaks (and the two adjacent ones) each time. We are on a treadmill: the more successful we are in plugging large leaks, the more we hit the barrier in leak size and set off avalanches which create both large and small leaks again. *The system fights back to stop the smallest value of the largest leak sizes we plug ever becoming less than the barrier value*. It fights back by creating avalanches, and a spate of avalanches only peters out when there are some more large leaks to plug again.

The system has, remarkably, organized itself to the critical barrier: it is the boundary between order and chaos. It does not mean that there are no leaks of a size smaller than the barrier value, but it does mean that we will never select them as the largest leak for plugging. But they are still vulnerable because they have neighbours of larger leak size who *do* get selected and which take their fitter (smaller leak) neighbours with them into a pell-mell of avalanches.

The behaviour of this model closely resembles that of Bak's sandpile. The sandpile reaches a critical height and slope; our embankment reaches a critical fitness and the water level a critical height. When more sand is added to the sandpile, the pile partly collapses so as not to exceed its critical height and slope; when more leaks are sealed in the embankment, other leaks spring up to keep the water pressure (and hence level) more or less constant.

Bak has described another analogy that has behaviour similar to that of his sandpile. To improve the performance (fitness) of an organization, a draconian CEO might simply pick the worst performer in the business and lay him off plus the two workers immediately adjacent in the organization chart. (He may use the specious excuse that the latter two have been tainted by contact with the worst performer.) He then hires three replacements but without testing their competence before hiring them: they may be better or worse than the workers they supersede. He then picks for laying off the new worst worker in the business and two adjacent workers ... and so on. The organization soon hits a barrier of competence. The more it dismisses people and hires replacements, the more it creates avalanches of change. Incompetent worker A is dismissed along with colleagues B and C whose competence is irrelevant (it is not a fair world). Workers A, B and C are replaced by D, E and F respectively. F (C's replacement) is found to be the new most incompetent worker and is dismissed along with his colleague D (who has only just joined) and another colleague G. Colleague G's new replacement is the least competent and is replaced taking with him F's replacement plus another colleague. And so on. As with the embankment model, small avalanches of change in each department which occur as we move around the business laying off the most incompetent workers can impinge upon each other. This occurs when a common colleague, perhaps the manager of both departments or a liaison person between two departments, becomes part of an avalanche. We then get avalanches setting off other avalanches.

For embankments, sand piles and the laying off of incompetent workers, not much happens initially when there are many large leaks, the sandpile is low and there are many incompetent workers. As the process of plugging the largest leaks, adding sand to the pile (and thus putting strain on grains of sand lower down), and raising the standard of competence continues, there comes a point when the fitnesses in each model hit the barrier - they can get no higher. The leaks selected for plugging (not the neighbours) never get smaller than a certain size, the sandpile reaches a critical size and slope, and the worker with lowest competence selected for dismissal is never more competent than the barrier level of competence. All these systems have hit self-organized criticality and maintain themselves at the boundary between order and chaos without explicit intervention.

Traffic jams and forest fires

We will now leave analogies aside for something more familiar. The behaviour of fir-cones causing avalanches of other cones and the behaviour of the embankment as the larger leaks are plugged may sound familiar to any driver: it is how phantom traffic jams (and some real ones) occur. Phantom traffic jams are the ones that appear to have no cause. Lines of moving cars stretching into the distance in front of a driver appear to slow down and speed up markedly for no apparent reason. If, when cruising along a trunk road, I slow down slightly then speed up again, as all drivers and even cars with cruise control do, this will have no effect on the car behind providing it is sufficiently far away for the driver not to have to react. As the traffic builds up - perhaps the rush hour starts - the spacing between cars drops. Reducing my speed slightly now causes the car behind to do likewise. If the cars behind this latter one are some distance away, we are a tiny traffic jam of two cars that quickly sorts itself out. As the traffic builds up further and the road starts to saturate, our mini-jam starts avalanching rearwards from car to car. This avalanche will only stop when it hits sufficient clear space between cars. Worse is to come though. My little jam could still be in the process of sorting itself out when I become involved in a much larger one from cars ahead of me. This in turn percolates via my car to cars behind me. A small temporary reduction in speed by one car on a congested road can reverberate backwards a long way when the cars behind are already at a critically small separation. When a procession of cars on a congested road slows down *just a little* as they pass an accident on the other carriageway, in order to rubberneck or to avoid scattered debris, the knock-on effect can be felt a long way behind.

Forest fires spread in a similar way to traffic jams. Each burning tree can only set alight those trees that are its immediate neighbours. These trees in turn set alight other trees that are *their* neighbours. Firebreaks stop the conflagration from spreading in the same way that gaps in otherwise busy traffic prevent sporadic jams from percolating backwards.

Directed organization versus self-organization

The cone-filled lake had a little external assistance. We assumed that the day was just hot enough for much of the ice to be on the point of melting; that the temperature had been adjusted to the right level. If it were too hot, all the frozen islands would melt. If it were too cold, everything would be solid ice and any impacts between cones could not travel at all. Adjusting the temperature is the equivalent of adjusting K in a manually tuned NKC model.

It is not always obvious whether to increase or decrease K to bring the different species within the system to the boundary between order and chaos (for the system to ‘go critical’). There is no absolute value of K that denotes a melting point in the way 0°C does for ice. When the landscape is smooth (low K), it is possible to hill-climb towards the boundary between order and chaos successfully by increasing and then decreasing K by a small amount and selecting whichever K is fitter. But on a high-K rugged landscape, this might take a species up a small local hill but away from a much better mountain of fitness that lies in the other direction. It has even been proved theoretically that, on landscapes that are very rugged, the system simply cannot converge to the boundary between order and chaos if it merely self-adjusts K in this way.

But life would be simpler if we could find different ways to let the system change K indirectly such that it *did* converge. This is what happens when (following Kauffman) we select a species of low fitness and replace it by a clone of another species of higher fitness. This clone brings with it its own 'better' value of K. It is also what happens if (following Bak and Sneppen) we replace the species of lowest fitness (plus some others to which it is directly connected) by species of random fitness.

Kauffman's method allows values of K that are held by species of higher fitness to reproduce and spread throughout the population. On the assumption that higher fitness is related to how near species are to the boundary between order and chaos, an assumption that remains unproven, self-adjustments of K will continue until the system is at the point of self-organized criticality.

Bak and Sneppen's method ignores K and simply pumps the system up to the point of self-organized criticality by removing the least fit species (and replacing it by a species of random fitness), then removing the new least fit species, and so on. This technique has the advantage of being closely related mathematically to what happens in some other more established areas of physics, and its convergence to the boundary between order and chaos has been proven for some special instances but not in general.

How does this apply to organizations?

Creating an organization that organizes itself to the right degree of complexity (K and C) has two advantages over attempting this manually through deliberate restructuring:

❖ the business does not know at the outset what the right degree of complexity is and hence whether to make the organization more or less complex. It is difficult to know exactly when something as complex and non-uniform as a business arrives at the boundary between order and chaos. In principle, we could let a self-organizing business just get on with it. In practice, this may take too long, so intelligent observation of which direction self-organization is taking the business can allow us to speed things up manually through intelligent tinkering with K and C. For example, we may reduce the number of authorization signatures needed for capital project approval. From an evolutionary standpoint, this resembles the *managed* breeding of fox-hounds or race-horses. Promising bitches or mares are mated with proven stallions in the expectation that some of the resulting offspring will share the fitness of the parents and maybe even surpass it.

- once the organization arrives at the critical point, it needs to be stable to disturbances. For example, an added complexity such as an additional imposed authorization level will need compensating for somehow. It needs also to be able to keep up with the slower change of K and C that denote a drift of the order/chaos boundary itself. Self-organization automates both of these.

The notions of K-complexity and C-coupling can apply to businesses competing with *other* businesses, where C is the degree of coupling of one business to another. This C quite likely represents their brands stealing sales from one other. But C can also apply to the constituent parts of *one* business.

Many large traditional businesses are like the well-frozen lake: the various divisions and departments are so internally constrained (high K) by the need to get lots of internal approvals to do anything that they end up doing nothing. They are tightly coupled (high C) upward and downward in the hierarchy. Instructions coming downward from on high and business performance results coming upwards do cause some movement, but the high K offsets the effect of high C. With a low C, total lack of direction from above for example, they probably would not move at all...

Smaller newer companies behave like the well-thawed lake: K is lower and the various parts of the business are decomplexed such that decisions can be made quickly and with the minimum of people involved. There is still direction from above and this direction can be acted upon quickly.

Some - very few - businesses have little direction from above (low C), apart from an overall business strategy, some common ethics and some common methods of financial reporting and communication (a single electronic mail system for example). They are not continually being pushed in new directions. A high C-coupling to top management is not necessarily a bad thing, but high C-coupling to several other divisions sideways or as a result of matrix management can cause a low-K division to jump this way and that in a fruitless quest to satisfy everyone. In this case a higher K is a useful damper. If this low-C (but not no-C) business also has low K (decomplexed) divisions, then we have the recipe for a reactive enterprise that can quickly take initiatives as they arise. Hierarchical C-coupling in this instance is easier to pinpoint than the somewhat nebulous K-complexity within a division. One drastic way to sort out those divisions who are in roughly the same line of business and which have low K from those which have high K is to let them compete in the same market. The ones with sufficiently low K and quick reactions will drive the ones with high K to the wall. Ones with K much lower than C (or ones with zero K - the typical one-man band in a new territory) will be too reactive. The survivors will be the ones whose K and C are in balance and whose K is sufficiently low to enable them to seize opportunities in a controlled rather than totally opportunistic way.

A coevolving business is not necessarily one that is poised exactly on the boundary between order and chaos, whether brought there by manual tuning or through self-organization. The idea of coevolution, whether within a business or between businesses, is a useful concept on its own. It gives a framework for us to study both internal complexity and the coupling between businesses or between business areas. A business can be improved as a result of the fresh insight gained. But a coevolving business that is on the boundary between order and chaos has special characteristics. We will call any business that has attained this state of grace a *critical* coevolving business.

This was Dee Hock's intuitive vision of VISA.

CHAPTER 4 - COEVOLVING BUSINESS ORGANIZATION

From genes to business

Dee Hock built a coevolving business before there was any real theoretical underpinning - he left VISA the same year that the Santa Fe Institute was set up. But things have moved on since then. For us to use some of the new theoretical ideas to create a business whose divisions behave like coevolving species, we need first to define the business equivalents of ‘species’ and ‘gene’. We need also to identify business meanings of K- and C-coupling of genes. These concepts apply equally to whole businesses, to parts of businesses such as major manufacturing plants, and even to competition with other conventional businesses that are not internally coevolving.

In business terms, a *species* could be any convenient major subdivision of a business or plant (division, department, factory, sales area...). To generalize, call each such species simply an *object*.

A business’s *genes* are the levers with which the performance of the business can be intentionally changed. Unlike genes in Nature, a business’s genes are not frozen from the time of its launch forever afterward - until perhaps it spawns another business. They are more akin to the controls of a machine. A car, for example, has a steering wheel, accelerator pedal, footbrake and perhaps a clutch pedal and gear lever. The positions of all of these are varied repeatedly in complex combinations in order to get the car from A to B. A car with automatic transmission has options, however. It can be actively driven using the D, 1 and 2 positions of the transmission selector and perhaps with an economy/sport switch or overdrive switch as well. Or it could be left in D with the transmission managing as best it can. We will divide business genes into two groups: those whose values we want to change intentionally and the rest. We will call the former *decision genes*. The remainder are like the 1 and 2 positions on the transmission selector when we choose to let the transmission manage everything in D: available but not used. The decision genes are the N genes of the NKC model. (The analogy should not be pushed too far: positions 1 and 2 are mutually exclusive and it then becomes a moot point whether each represents a gene or gene value).

The values of a business’s genes - plus uncontrolled or uncontrollable events that are external or internal to the business - give rise to the state of the business at any one time. This state is represented by the combination of knowledge, information and data that describes the business. At any one time these will have particular *values*. Some of these values are the values of all the measurable things in a business. They might be the dispatches of a given product to a particular branch of a retailer, or the current throughput of a widget die-stamping press or anything else measurable. Other values represent the innumerable things in a business whose values change and which we cannot measure, and which are connected in equally innumerable ways.

No management accountant knows everything about his business or how changing one thing affects others. He uses his best judgment (or follows corporate policy) to select things to measure which are:

- significant from the point of view of overall business performance
- can be controlled in some way (for example, a shop's product sales are influenced by the manufacturer's promotional expenditure or product quality)
- measurable!

Business performance measurements also need to be understandable by non-specialists. I cannot meet my objectives if I do not understand them or how they are measured or how I can influence them. We will return to this subject in Chapter 6.

A decision is an intentional change in the value of one or more decision genes that affects the values of others. The ways in which it affects the values of others may be simple. They are literally low K. A decision to accept a contract for a lower unit cost of electricity supply to a manufacturing plant has a direct and immediate effect on manufacturing overhead costs, provided that usage does not change. The effect may on the other hand be complex, ill defined, risky and delayed, such as a decision to introduce a new product.

Changing a decision gene is literally a change in knowledge, information or data that modifies how an object appears to operate, from introducing a new product to cutting capital investment. The current state of the business is changed by decisions and also by external influences, by accident ('random mutations') and by prior events - intentional or otherwise. Decisions may be weighty and strategic or merely operational judgments about which product to schedule down the production line next. Decisions can be made at a formal 'approvals' meeting, by an informal consensus or by an individual decision-maker. They could also be made semi-automatically as happens in arbitrage trading when matched offers and bids are made to exploit fleeting price gaps. We shall call any mechanism by which a decision can be made a *decision point*. If a decision point is a meeting it must, as we shall see shortly, be treatable as a 'black box' that is asked to decide something but whose internal machinations are invisible to the outside world - like the process for selecting a new Pope.

In the following four chapters, examples will be taken mainly from consumer goods manufacturing. This does not mean that there is something special about the applicability of coevolution to manufacturing industry but rather that it has most of the constituents of other businesses including service industries. It has direct suppliers (of raw material, sub-assemblies or parts), indirect suppliers (suppliers of raw materials to sub-assembly manufacturers, for example), direct customers (the wholesale trade) and indirect customers (the consumer or end-user). Because of this, manufacturing industry is more complex. Remedies for over-complexity may well be easier to apply in simpler businesses. To counterbalance this emphasis on consumer goods manufacturing, Chapter 9 will cover a complex service industry - healthcare management.

A conventional business works by directing each lower level area to play its part in meeting the same rather remote targets such as return on assets. Commands cascade down from the top. They get fleshed out and not infrequently distorted on the way down. A coevolving business would, on the other hand, work by setting parochial and perhaps even conflicting business targets for each object. The objects would then be coupled or decoupled from adjacent objects to the point where a controlled low number of significant C-couplings exist. The objects themselves would be internally decomplexed (low K) to improve flexibility and speed of adaptation and to avoid their becoming stuck on the peaks of hills of low fitness. If the changes in K and C are deliberately controlled to drive the business to the boundary between order and chaos, whether through manual tuning or through self-organization or a bit of both, we will have created a *critically coevolving* business. Note, however, that reorganizations of this type are not all or nothing: the process of identifying and adjusting K (and C where possible) may on its own give substantial improvements in responsiveness even when these K and C values are not those of a critically coevolving business.

The choice of targets of individual objects must be such that they are not all freestanding - some must interact with the targets of other objects. Hence the attempts of each object to improve its fitness to meet its own targets deform the landscape of other objects. Deformation *may* be constructive and not competitive and objects may evolve a close and co-operative (symbiotic) relationship.

Some businesses have for years dabbled with the idea of internal competition by allowing their factories to bid for making a particular product for a specified market. Others promote the head-on competition of two or more products in the same market sector; Proctor and Gamble for example started doing this as long ago as 1931. For some it works; for others it encourages a very short-term view of return on capital, stifles new product development and creates a barrier to branching out into new markets. Perhaps the successful ones understand intuitively how best to translate complexity and coupling into organic growth.

There are several questions which need resolving before anyone would ‘bet their business’ on re-engineering it to coevolve critically. In particular, they would need to know whether dividing up a business or factory into objects is resilient to disturbances and whether there is a best way to divide it up.

Bounces back

Resilience to disturbances is increasingly rare in otherwise efficient businesses. Anyone who has used mathematical optimization methods to schedule a factory will know that given lots of computing time and a model which is not too complex, an optimal result may be obtained. 'Optimal' can be almost anything: maximal output, minimal cost of production, highest quality or something else worth having. In practice, trying to optimize one of these factors almost always upsets one of the others. Optimization is then either a compromise (for example, maximize *the sum of* some measure of output and some measure of quality) or subject to constraints (maximize output such that quality never falls so low that it is unacceptable). Compromised or constrained optimization is a result of *unavoidable* K-complexity inherent in the cost structure of the business and will be revisited in the next two chapters.

Placing undue reliance on the outcome of mathematical optimization methods brings new problems, however. Any deviation, however small, caused for example by slightly late arrival of raw materials *could* have a disastrous effect on the way the factory operated. What was optimal might turn out to be very sub optimal indeed. Just in Time must *never* be Slightly too Late. Factories or - worse - whole logistics pipelines from raw material supplier via the factory to the customer become much more susceptible to this as they reduce or eliminate stock and remove stock-holding points. The pipeline is like a guitar string: a 'twang' when it is slack has little effect but, when all the slack is taken up, the twang propagates down the string and lingers. But the effect is at least predictable. It is not easy to predict the effect of a disturbance to a logistics pipeline. It might be trivial and die away quickly or it might be catastrophic. Tautening such a pipeline by removing stock and by scheduling manufacturing operations with no slack time between them *can* result in huge fluctuations in response to minor ones. Jay Forrester of MIT, the father of pipeline dynamics studies, investigated the effect of pipeline disturbances back in the late 1950s but never stumbled upon some of the stranger things that can happen in a de-stocked pipeline. Perhaps he lacked computer power. The underlying behaviour is somewhat similar to that of Bak's sandpile. But the result may not merely be the short-term delays and stock accumulation that would normally happen in a simple distribution pipeline on either the inbound or the outbound side of a factory. If a production *sequence* is disturbed by, for example, needing to bring forward the production of another product because the materials for making the scheduled product are late, the impact of the resulting chaos - for (mathematical) chaos it is - is unpredictable

How big should an object be?

Into what size lumps should a business be split: whole factories and major divisions, a much larger number of small departments or each individual? The work of Stuart Kauffman and his colleague Bill Macready provides some pointers. Remember that an object is selfish - it tries to improve its fitness irrespective of the effect on others. A complex business (not necessarily a big one) with many internal connections benefits from having several smaller objects. A simple business, however big, may be better run as an entity. This is not a matter of how much any one individual in the management hierarchy can physically manage, however talented. Instead it reflects the underlying logical problem that a hierarchy cannot easily manage a mesh (something with cross connections). If a manager cannot break the connections between objects, he can at least allow objects to seek for their own *local* best fitness. We can make decisions for such an object without consulting everyone else in the business or worrying about the effect on them. Even better, objects should be barred from peeking into the internal workings or data of other objects. This principle of 'information hiding' first came to prominence in the early 1970s in the context of what is now called object-orientated programming. The optimum number of objects - the amount we carve the business up - is dependent upon the way in which the objects are managed. A business with 'greedily managed' objects - ones where there are conscious attempts to allow only those decisions which are *the best* for the object (as opposed to allowing all decisions which make *some* improvements) may thrive better with many smaller objects.

When we split up a business into objects, there is a natural tendency to do so by business function, geographical area or some other natural grouping. Business process re-engineering is partly about restructuring the organization into groups which are as self-contained as possible. In other words, the genes *within* an object are likely to be associated with other genes within the object rather than genes within *other* objects. For example, we work mostly with colleagues in the same area.

Boxes within boxes - which is K and which is C?

Objects may be black boxes but, like Russian dolls, can be made up of other smaller objects. A business may be made up of divisions that in turn are composed of departments ... and so on. The lowest level might be an employee. Simulations by the present author confirm that Kauffman's aim of balancing the values of K-complexity and C-coupling to achieve a critical state balanced between order and chaos works also for a hierarchy of objects. And if K is low, the resulting system is reactive but relatively stable. Small disturbances generally result in the hierarchy of objects evolving back in balance. One way to apply the model to a business made up of objects that are composed of other sub-objects is for the sub-objects to take advantage of an existing management hierarchy. Perhaps the object managers are responsible for the decision genes associated with other objects. In other words, the association between two departments at the same level could be via the two senior departmental managers and not between their respective staffs.

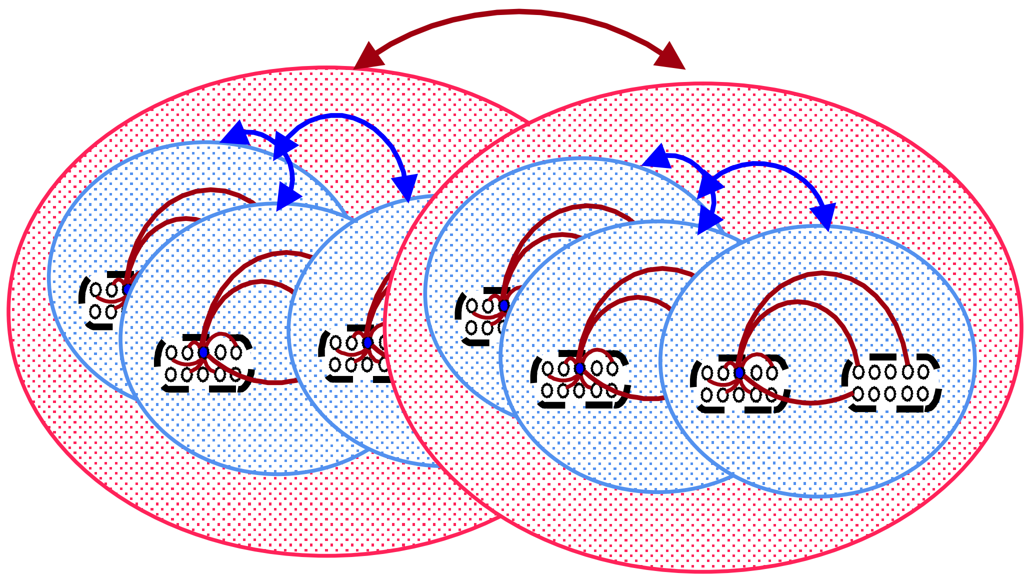
In a hierarchical collection of objects - business units, divisions, departments, teams and so on, it may be unclear which is K and which is C. Such a collection is not limited to individual businesses, however. The concept can be extended to a collection of competing businesses in which each *whole business* is an object. The solution is firstly to decide on the *highest* level of object at which we want to:

- build a notional wall around the object and treat it as a 'black box'. Its internal workings are then invisible to other objects except via any C-couplings

- allow the object to evolve in competition with others

The consequence is that each of these objects has its own fitness objectives that are targets it can pursue without reference to other objects. This is true because other objects are black boxes also. At this point we will have defined which (N) things in an object are the equivalent of decision genes. These will then be given couplings to other genes within the object (to give K) and to genes in the other objects with which it is to coevolve (to give C).

If we now decide to break things down to the next level, we will be treating one or more of the objects as a collection of other smaller objects. These we also want to make coevolve. For example, we could select an entire business and have its several divisions coevolve. We could then select one of the divisions and want the several manufacturing plants within that division to coevolve. Finally, we might want one or more of the larger manufacturing plants to be broken down into functional areas - raw material intake, processing, packaging and distribution for example, and let them coevolve. At the highest level, the divisions will each have decision genes defined. These are, for example, at decision approval points that are K-coupled to other decision points (other authorization points within the division, for example) and C-coupled to points in the other divisions. Each division now evolves on a landscape whose ruggedness is defined by K. When we now break a division down into manufacturing plants, we may well pick different decision genes - ones more appropriate to the performance of plants. If each plant contains one or more of the decision genes that have also been selected at the highest level, this is fine: it gives a tight link between divisional evolution and plant evolution. If *all* the decision genes are the same, then divisional evolution is irrelevant as the plants of each division can all evolve with each other. All the plants of all the divisions would participate in one large coevolution. Divisional coevolution will then happen as a by-product of the coevolution of its plants.



Pairs of coevolving objects that represent other larger objects. These larger objects are coevolving in two groups of three. Each collection of three objects is itself another even larger object and two of these are also co-evolving

When selecting decision genes at any level but the top level, it was mentioned earlier that we might choose to use the fact that hierarchies operate through managers. This means that decision genes at the divisional level may all be C-coupled to the general manager of each plant and not to several decision points in each plant. This reflects the fact that the plant general manager takes orders from above, takes advice from those below him but acts as a funnel for decisions. There is no strict need for hierarchies however. As far as coevolution is concerned the flatter the organization the better. In the extreme, this gives rise to a simple collection of individuals with autonomous decision-making power.

Where we introduce multiple levels of coevolution, it would be helpful to know if there was a theoretical optimal number of objects at each level that could be used as a guide. Unfortunately there is no general solution to this at present but any answer is likely to depend upon:

- ❖ the length of time for which the organization has existed. Is it still rapidly evolving or has it reached some degree of internal stability
- ❖ the K-complexity of each object. This also helps to determine stability
- ❖ its internal complexity. Is it a pure hierarchy or partial mesh or what?
- ❖ the relative degrees of co-operation versus competitiveness of the C-couplings between objects. These couplings may well determine the K-complexity of the objects at the next level above, but this depends on how the decision genes at this latter level are selected
- ❖ its natural internal rate of change: how often new products or major variants are introduced, new factories opened and new markets entered, for example

- the strength and complexity of C-coupling to competitors which is discussed in the next chapter

This is all internal to a business. We need now to look at how businesses themselves interact. Businesses that sell items or services interact via their brands, where a 'brand' is something that has a market share. There is a market for hotel room occupancy of a particular class; there is another for baked beans. Market size and growth are watched closely, but the real war is fought over market share. Between businesses, market share is an important (perhaps *the* most important) C-coupling. Businesses react to changing sales, but if their shortfall is a competitor's gain, alarm bells ring and, in a low-K business at least, actions are taken.

What follows is a more detailed look at how coevolution can be applied to business competition. Battles between brands are fought with advertising and with other means of promotion such as trade price support. We will investigate more closely how these weapons interact.

CHAPTER 5 - COMPETITION

Adapt fast or die

Disturbances to an object can be the result of landscape movements caused by other objects fighting their own corner. But they can equally well be from other external events such as unusually good or bad weather - for those who sell ice cream or umbrellas. Twenty years ago, Manfred Eigen and Peter Schuster proved that if random changes occur faster than an object can evolve to keep up with them, a previously fit object drifts slowly downhill. To ameliorate this, a business must learn to adapt faster but also must stop reacting to *everything.* The difficulty is knowing what is significant and what can be ignored. Businesses such as the 'old' IBM often took a scattergun approach to product development and marketing. This is how it appeared to its customers at least. Several competing products were marketed in the same market sector and the unsuccessful ones allowed to perish. There is nothing wrong in principle with this approach - letting something akin to natural selection by your customers decide your product strategy - apart from the inevitable customer confusion. But those who practise it have to make sure that the unsuccessful products do not linger in the market. The loss on developing and marketing the unsuccessful ones must always be outweighed by the profit made on the successful ones *during the same time span.* If not, the fitness of the business itself goes downhill temporarily. Even worse, if there are no obvious winners and losers, the business may be tempted to inject more and more products into the market faster than its losers can be weeded out. The result is an Eigen and Schuster fitness 'catastrophe'. The more products a business has in its product line, some fraction of which is injected into the market, the sooner the catastrophe happens.

If market forces are strong enough such that all but the most successful products perish overnight, there are few surviving variants remaining from which to 'breed'. If the marketing environment changes - customers suddenly become more cost-conscious as a recession bites, for example - there may be no suitable products around to satisfy them. This matters more in Nature where reduced biological diversity creates problems for the future; when a species or major variant goes extinct it is gone forever - but it might be just what was needed to survive new harsh winters or thermonuclear pollution, for example.

It would be wrong to give the impression that low-K landscapes are always preferable. Recall now that S is the number of objects to which any one object is C-coupled. We saw earlier that if K is roughly the product of C and S, the incessant jostling which results from being nudged by other objects is somewhat counterbalanced by the damping provided by K. The greater the degree to which objects are C-coupled to others and the greater the number of coupled objects (S), the more beneficial is the 'treacle' conferred by K-complexity. If, while bathing in the sea, I were seized by a shoal of S malevolent octopuses (C in this case cannot exceed 8), my concern would be directly related to the number of tentacles (S x C) ensnaring me.

A low-K landscape can be less than ideal if an object is being buffeted hard by many other C-coupled objects or is evolving slower than random changes are occurring. The reason lies in the structure of the landscape itself. A low-K landscape has few (perhaps only one) high but gradually sloping mountains. Picture an object halfway up one of these gradual slopes. The wind of random change can blow it in any direction but, since the slope is gradual, it will end up each time at a very similar height. Any voluntary hillclimbing upward will similarly progress at a stately pace. The object may get blown about a lot but its fitness does not change very much. A rugged (higher-K) landscape works differently. Peaks are lower than on low-K landscapes but the slopes are steeper. And steep slopes mean that there is a significant difference in height between an object's position as it gets blown about. It is likely that there is a compromise value of K that gives a landscape of sufficient ruggedness - and thus steepness of slope - which allows the object to recover from adverse random changes quickly whilst still having acceptably high peaks for the object to climb. It may sound as if objects on a low-K landscape were the ones most resilient to disturbances but they are also less responsive to attempts to climb upwards and away from where a random disturbance had blown them or a competitor had nudged them by deforming the landscape itself. But were not decomplexed objects - the ones evolving on low-K landscapes - the poised and responsive ones? The answer depends upon how far we need to move on the landscape and how quickly. Do we need, for example, to make a sudden reaction of limited scope and duration or a large strategic move? In order to clarify the link between responsiveness and K, we need to have a deeper look at the meaning of K.

If many variants of a prototype product are being test marketed, the last thing a brand manager wants is for there to be little consumer preference between the best and worst. K needs to be raised to sharpen this difference and K in this context is one measure of how different the products are from each other. Consumers can only preferentially elect for one product over another if there are perceived differences. Exactly *why* K is a measure of product differentiation may not be obvious. After all, hitherto it has been treated as a measure of the internal complexity of an object.

The different faces of K

K was defined earlier as the number of connections between any one gene and others in the same object. C was defined similarly except that it was the number of connections between a gene in one object and genes in *other* objects. Each object lives on a fitness landscape. Its topography could resemble the Alps, the rolling dunes of the Sahara desert, Mount Fuji or (most unlikely) Cape Town's Table Mountain. Its ruggedness is determined by K. In order to clarify *why* it is determined by K we need to look at an object's freedom to improve its fitness. Remember that the value taken by each gene contributes to this fitness. It is possible that an object has genes whose values are irrelevant; the business analogy would be those nominally redundant employees like Japan's 'window gazers' who have been sidelined. We will ignore such window-gazer genes. Consider the extremes. In an object with zero K, each gene can take a new value to improve the object's fitness without affecting the fitness conferred upon the object by the values of its other genes. Each gene is thus working independently, and any improvement in fitness caused by changing the value of one gene does not have the side effect of reducing or enhancing the fitness contributed by other genes. There are no sudden surprises as a result of changing the value of a gene. The landscape is one gradually sloping high mountain. Climbing it is a long haul but straightforward (unless C-coupling moves the mountain itself - which is another matter altogether). By contrast, an object on a high-K landscape has genes coupled to lots of others of its own. Every change in the value of one gene to improve the proportion of the object's fitness conferred by that gene has unpredictable side effects on the fitness contribution of other genes. It is like tuning an old-fashioned radio that has plenty of knobs. On a zero-K landscape, an improvement made by turning one knob is always an overall improvement. There is no adverse effect on the results of the settings of the other knobs. On a high-K landscape, turning one knob upsets the results of the settings of the others. They in turn need adjusting. Tuning becomes very difficult because of these side effects. There are lots of combinations of knob settings that represent the tops of small fitness hills and the slightest adjustment (equivalent to a step down a steep slope) can cause the radio to go off-tune very easily.

This might, however, be the source of an unexpected advantage in one specific instance: when our object is not at the top of a fitness hill *and* we need to react - improve fitness - *very* quickly in response to an attack by a competitor. Speed of response in this instance is more important than for how long the response can be sustained or whether the response is strategic. A small change in gene value in such circumstances can cause the object to move uphill faster - because the slope is steeper - than it would on a low-K landscape. Thus a sudden reaction of limited size and duration *may* happen faster on a high-K landscape if there is scope for it to happen. But there will only normally be scope if the object is not already at a peak. If the peak is a small one it might also be the wrong move strategically. The distance we can climb upward and thus improve fitness over the longer term will be less than would be possible on a less rugged landscape. However, this phenomenon does not contradict the general rule that K should be roughly the product of C x S. In a high-K business, it will almost always be a purely theoretical advantage. The time needed to agree the change with the multiple decision makers in a high-K business will counteract the quicker but limited improvement in fitness that could then result.

In the last chapter, the possibility was raised of unwittingly creating an unstable (brittle) logistics pipeline or manufacturing plant through over-reliance on mathematical optimization techniques. But even if the result were relatively stable, we still cannot optimize different operating criteria such as plant output volume without adverse effects elsewhere. Higher output may need more labour, for example, and manufacturing costs will go up accordingly. If such operating criteria are each managed autonomously, they can be treated as the effects of C-coupled genes of separate objects. If such criteria are lumped together into one compromise fitness super-criterion as described in the preceding chapter, they are being treated as the effects of K-coupled genes of one object. If, for example, increasing plant output increases manufacturing costs in a proportional way and we want a compromise of the two, the result is an apparently low-K smooth landscape. At the point where output rises to the extent that a new labour shift is needed, the landscape suddenly breaks up and a sharp-sided valley of low fitness appears. Plant managers will avoid descent into the valley when only marginal extra output volume is needed.

There are two ways an object can evolve. Until now we have painted a picture of the genes of an object seeking to improve the object's fitness by continually changing values. This is the correct interpretation for business organizations. Each decision gene is associated with either a decision or a particular type of decision, made perhaps at a formal meeting or by an individual who approves or rejects requests for capital investments for example. The object whose fitness is decided by different combinations of gene-values then becomes fitter through making the right decisions. *The object itself changes*. But this is not the interpretation needed for test marketing of different variants of a product. Each product variant has a fixed fitness: customer acceptance. If we ignore the fact that one variant may steal sales from another, we have a picture of all the variants sitting *concurrently* at different fixed positions on the landscape. The variants with low customer take-up are removed from the test market ('become extinct'). At the same time, new variants are added with features more closely resembling those of variants that are currently selling well, like the selective breeding of pedigree dogs for competitive showing. Sufficient differentiation must, however, still be preserved to allow the selection process to continue. The population of variants drifts slowly up the hill. So by a process of trial and error the brand manager homes in on the fitter ones - the ones with highest acceptance and at the highest point on the hill. This is of course a bit idealistic: a test marketing campaign typically cannot afford to launch a continuing and almost endless stream of new variants with the later ones hijacking the best features of those which already have a high customer take-up.

This second interpretation is close to what happens in Nature. But there is one fundamental difference: Nature does not have a brand manager with an overall view. Each variant of each species (or more precisely each gene of each variant) is out to improve its own fitness by reproducing: by producing new variants with fitter combinations of gene values than the parent produces and by producing as many of them as possible. The less fit become extinct. In practice of course, genes do not know in what direction to change value in order to improve the fitness of their variant. They merely produce variants with different fitnesses and natural selection then weeds out the less fit. There is a trade-off between creating new species from scratch and capitalising on the work already undertaken in creating existing species by generating many variants. Creation of variants is relatively quick while a new species take a very long time. But a new species - which in business terms would equate to a radically different product, service or manufacturing process - might be just what is needed to survive in a new environment where mere variants are just not novel enough and go extinct. It is easier for a species variant of low fitness to have offspring of higher fitness than it is for a variant that is already fit. In landscape terms, it is easier to climb upwards when near the bottom; climbing any further when near the top is hard going. The cause of this is not some genetic equivalent of exhaustion or lack of oxygen but the shape of the mountain. Near the bottom, many directions are uphill, so as new variants of a species are generated through reproduction, many will be fitter than the parent. Near the top, most directions lead downhill and most variants will be less fit than the parent. The longer you climb, the harder it becomes to climb further. At the top every direction leads downhill.

One feature of a low-K landscape is that an object (or variant in the case of a reproducing species) does not easily become stuck; it is almost always able to evolve in a fitter direction. This is because the landscape has perhaps just one large mountain with a distant peak way above in the mists. An object on a high-K landscape easily becomes marooned on one of the many peaks of low fitness. Further evolution without risk is impossible unless a C-coupled competitor deforms the landscape.

Less data - more information

Having too much data can be counterproductive, even with sufficient analysts and ample computer power available. Market data for example are rarely consistent and anyone trying to derive information or make a decision based on such data has to judge how much weight to give to each data source and how to make compromises between data that disagree. One possible way out of this impasse is to derive information from each source separately and only then select the result that appears more credible. Another way is simply to cut down on the number of different data sources - or just to ignore some of them. The underlying reason for this is that each extra data source raises internal K-complexity and attempting to take account of *all* such data when making a decision can be suboptimal and can cause a business to stay anchored on a landscape peak of low fitness.

Business planning and modelling

We have already described how businesses competing with one another adapt to each other's actions. A business that coevolves internally (like VISA) is no exception. All but the very smallest businesses have business plans which will forecast profit, cash, assets, sales growth and other measures of fitness from projections of 'external' market and market share and 'internal' costs such as cost of goods sold, fixed costs and the like. Projections of market share are closely bound up with what the competition is expected to do, and computer models for planning brand growth should take competitor activity into account - plus growth in the market itself. Errors in forecasting will be tracked, and the models together with the sources of external data that drive them will be continually refined. Since the models themselves will evolve, we can put them into a coevolutionary context. Suppose a model can be pictured to have N of our 'radio knobs' (decision genes). The way in which turning one knob affects the forecast accuracy (fitness) of the model then determines K. If forecast accuracy improves smoothly as knobs are turned, one by one, with little relationship between the settings of different knobs, then K is low. If adjusting any knob slightly has a jerky and unpredictable effect and if the turning of one knob affects the function of one or more other knobs, K is high. C is simply the number of data linkages *between* the models in competing businesses; examples would be trade list prices or shelf prices of major competing brands. S is the number of directly competing brands.

So, as businesses coevolve, business-planning models within competing businesses must also coevolve. Depending upon the values of K, C and S, these models may behave:

- in a very responsive but relatively orderly way - when K is roughly S x C. This corresponds to critical coevolution or something close to it.

- in a smooth but forever changing way - where K is relatively low. The potential forecast accuracy achievable is high - higher than in the preceding case, but the knobs need to be continually and substantially adjusted to keep up with changes.

- in a 'jerky' way - when K is relatively high, and it becomes very difficult to improve forecast accuracy by adjusting any one knob because the results of adjusting it cause the system to go haywire. The forecast accuracy achieved is lower than in the previous two cases.

An internally over-complex (high-K) model will find it difficult to track closely the behaviour of another business, particularly when relatively few factors such as the quickly changing sales of one or two competing products are being tracked.

The behaviour of business planning models should influence the timescale - the planning horizon - over which they plan. There is always a temptation to forecast too far into the future, and this is particularly dangerous if the model of our business includes the modelling of competitor activity as well as our own. The part of the model which describes *our* business will include some logic about the future decisions we will make - new brand launches, capital investment and so on which may change as the external business environment evolves in ways we did not predict. So these are themselves questionable forecasts. But when we include in the model some logic about what our competition will do, we are probably not party to what their internal plans are and can only forecast them based on how their business has behaved in the past. The forecast accuracy of the model may then be seriously at fault. This may not matter if the serious errors are several years hence, but if we then make major decisions about our short-term activities based on such highly inaccurate forecasts of the future, such decisions may be very poor ones. Furthermore, the model will then have to undergo continual change when it is adjusted to cater for perceived shortcomings in its short-term accuracy, and, in a high-K model, these will then significantly affect the forecast of the future. Such a repetitive cycle is to be avoided as it will cause the business to change direction continually.

Brand vs. brand

In mature markets, the top two or three brands slug it out year after year for the odd point or two of market share. This happens in most developed countries apart from Japan; in Japan, maintaining an *unchanged* share appears to be the aim. Brand managers live on a tightrope: they must develop new product forms (propositions) without disturbing the successful but stagnant current ones. Their worst nightmare is to upset the status quo by too much innovation and kill their current brand. But at least this might be predictable, whereas successful attack from a new and unexpected quarter by a competitor's innovative product comes out of the blue. The militant but cosy market relationship between the top few brands can result from each brand manager pursuing a development strategy which happens to be the only one which maintains market share (in the sense that any deviation from it loses share). This is a result of what theoretical biologist John Maynard Smith called an 'evolutionarily stable strategy'. But if the status quo is disturbed ever so slightly too far by an external jolt, perhaps the entry into the market of a new competing brand, then almost anything can happen including - in theory at least - sudden extinction of all of them. The current generation of biologists understands this but it is probably news to brand managers who think that the market is rational.

Warring brands need military strategies. Those with poor strategies become extinct. This implies that strategies themselves might evolve and are subject to natural selection. Maynard Smith's evolutionarily stable strategy is similar in operation to Kauffman's way of inducing an NKC model to self-organize. In both cases, a species of low fitness is replaced by one of higher fitness. This fitter species brings its own 'fitter' K with it and the fitter K thus spreads through the population. If K were a strategy, the evolutionary spreading of this better K would be like the spreading of successful strategies. One problem peculiar to the spreading of successful strategies is that the more successful a species becomes through executing some particularly cunning or otherwise successful strategy, the more it is likely to end up competing with copies of itself. An evolutionarily stable strategy is a strategy that, in any one species, is successful when competing with copies of itself! Species, like businesses, have problems of success.

Assume that at any one time the market shares of the top three competing brands in a market sector were static. One brand marketing strategy could be 'if market share drops appreciably, increase advertising until share recovers; if above normal share, avoid spending much'. If this strategy were adopted by all three brand managers, the result would be more or less static share for all three if other factors such as product improvements were ignored. Assume now that there were two other similar strategies: one of increasing trade promotional spend for retail price protection and the other of increasing consumer (sales price independent) promotional spend. If any one of these strategies *in isolation* were always more cost effective than the others, it would seem that it would become universally adopted. The other two strategies would become moribund. But one such strategy is not executed in isolation: it competes with itself or with one or both of the other two strategies. If all three brand managers fought an advertising war, advertising itself becomes an arms race. When advertising wars are fought between the largest well-advertised brands, the relative effectiveness of advertising in stealing share falls. Taken to extreme, it could even push up the cost of prime advertising slots on television. The three brand managers might independently decide to adopt a mixed policy of *alternately* using each of the three strategies in some fixed split, say 60% for an advertising strategy, 30% for trade promotions and 10% for consumer promotions. If each manager uses a different split, the one with the best split *relative to the others at that time* will get most value for money. The cost-effectiveness of a particular split depends upon what the other two competing splits are. If the other brand managers are alert, they will also adjust their splits and watch the result. After a time, the three splits may each home in on either the best common split (if one exists) or three splits that are each as cost-effective as the other two *when pitted against the other two*. Note again that the cost-effectiveness of one split is only relative to those splits adopted by the other two parties.

The principles of coevolution apply here also. What is happening is that as one brand manager changes the split in a direction believed to be more cost-effective, the fitness landscape of those splits currently adopted by the other two is deformed. Other ways to adopt a split strategy are possible. The most obvious one is not to use the strategies alternately in the proportions indicated by the split, but to use all three strategies at the same time and split the money spent on each. This percentage split is not necessarily the same split as was adopted when the strategies were used alternately.

All this is somewhat idealistic and relies on knowledge of the competition that is not currently possessed by brand managers. It is rare at the moment for brand development strategy to be based on what the competition will do. More often it simply assumes that the competing brands will try to maintain their brand positioning. But this will change. The rapid rise in product differentiation for niche marketing implies that it will no longer be safe to assume that the competing brands will look tomorrow as they look today. A new competing product form of this type has no history, so assumptions need to be made about what the intended positioning is and how the product will evolve.

Advertising too, which for major brands means television advertising, is not as simplistic as portrayed above. Advertising spend is typically split several ways. The first - usually most of the investment - is for maintaining warm feelings, credibility and the desired image of a brand (and perhaps its manufacturer) in the hearts of consumers over a timescale of many years. Advertisements for capital purchases such as cars, expensive consumer electronics and domestic ‘white goods’ also provide reassurance for existing owners that they made a sound choice that they hopefully will repeat in due course. The second type of advertising investment is for maintaining resilience of brands to repricing or to attack by competition. The third is shorter-term publicity for brand relaunches, new sub-brands and new variety ranges. The fourth is for very short-term promotions: special seasonal or regional prices, competitions and so on. Contrary to popular belief, the effectiveness of advertising is almost all long-term: a ‘promotional’ advertising campaign may well induce someone to *try* a new brand or variety; it might also induce someone to *try* an existing brand, but is unlikely to get someone to switch permanently. Users are the only people who are really aware of the minor differences between competing brands - mere advertising is not enough. And with the possible exception of advertisements for cars and the like which can be bought in many variants and with or without many options, minor differences are never advertised. The point of all this is that advertising spend is not, in general, directly interchangeable with trade or consumer promotion spend - they work over a different timescale. This makes the choice of split described above more complex: repeated low-key advertising to maintain credibility and brand image does not have a short-term effect on market share. Promotions apart, short-term advertising to emphasize product differentiation with the aim of getting retail customers to switch immediately and permanently rarely has any effect at all. And product differentiation is usually not sustainable: it may well be as easy for your largest competitor to manufacture the more successful variants from your test marketing campaign as it is for you. If this sounds like evolution and natural selection with warring species’ variants invading others’ niches, that is exactly what it is. Any genetic advantage taken by one competitor can equally be taken by another. But the successful ones, the ones that come to dominate the market, are those that are faster on their feet and get new propositions to market sooner. Competition may be copying your every move, but being *first* is a decided advantage providing you can avoid extinction while building a customer base for your new proposition. Surprisingly, competitive brands have a long history of trying to *match* each other rather than be different, although this could partly be a reflection of the war between consumer products such as soap powders that simply do not have much differentiation to play with in the first place.

Typically also, at least in large businesses with big advertising budgets, detailed control of advertising spend and the monitoring of advertising effectiveness does not lie with brand managers (or agencies) but with media buyers. These buyers treat advertising like any other purchase, albeit one whose value for money is difficult to pin down.

Finally, trade promotional spend in particular is typically short term and is usually the province of Sales rather than brand management. Brand manager input may be constrained to setting strategy on the relative trade price (or recommended consumer price) differences between countries in order to manage cross-border trading.

Each brand manager has tried to do the best for his brand. No altruistic cartel was needed. And yet all homed in eventually on a stable strategy for their brand. If a competitor breaks ranks, it will lose money. But this might be a good strategy for someone who wants to live on a knife-edge of perpetual change, where share is continually rising and plummeting. It could also be a good strategy for someone who deeply understands his competitive position. Such a person is prepared to take a short term loss in order that his brand and those of his competitors all end up, after much chaotic steal-and-steal-about, in another evolutionarily stable state *but one which has a larger share for his brand*. It is unlikely that anyone has tried this with a real brand by planning brand development in such an analytical way. The unpredictability of the outcome would deter anyone who could not afford to lose all, although taking similar but smaller risks intuitively to break a share deadlock is not unusual.

Economic equilibria

Traditional economics and much of current economics as well presupposes that decisions on how to act are made in a predictable way. It assumes that if we feed information to a decision point - which for an economist may be an individual, a business or some larger conglomeration - the resulting decision on what (or how much) to do can be predicted. This is the basis of the huge computer-based economic models of the last thirty years used by most developed countries and by business schools. The latter at least should have known better. The idea is simple: create enough equations which reflect how each relevant piece of the economy should behave, couple them together such that the outputs from some equations form the inputs to others, and let the model run forward in time for weeks, months or years. The implementation is impressive with hundreds or even thousands of equations linked in complex ways. The results are voluminous, command considerable respect and are mostly wrong, particularly where they try to predict how measures of the economy as a whole, growth for example, will behave. The reason is not hard to find. People and businesses do not behave in predictable ways and their behaviour is influenced by the behaviours of other (C-coupled) people and businesses as well as by the economists' conventional indicators such as current prices, exchange rates, growth and inflation. This means that the type of equations on which economic models are based simply do not work, particularly at a high level - the whole economy of a country for example - or for a sustained period of time. Models that *do* attempt to mimic the interactions between individuals or between businesses are comparatively new and few. And those who use them are generally aware of the limitations on their forecast accuracy. This is at least an improvement.

The NKC model applied to coevolution of businesses is a model of this new class at the lowest ('micro') economic level. Objects are free to respond to the behaviour of other C-coupled objects and W-coupled external disturbances in complex and unpredictable ways. We have, however, gone a stage further and proposed that the competing objects may, under certain conditions, seek the boundary between order and chaos themselves or be nudged in the right direction by enlightened management.

The conditions for an ideal economic equilibrium - where price and wage stability reigns and where demand for consumer goods for example was exactly balanced by the ability of manufacturers to supply them - has been at the heart of economic thinking for the last hundred years and more. If this equilibrium could be tilted in favour of full employment, so much the better. But this utopia has latterly received short shrift by the theoreticians who have shown that the conditions under which such a competitive equilibrium could exist in a free market are so unlikely as to render the goal impossible. There are several reasons. One is that a competitive equilibrium implies that there must be a futures market in everything anyone could want to buy, not just the usual commodities familiar to those who read the financial pages of daily newspapers. Another premise is that people's behaviour is predictable; but, as we have seen, people's decisions are affected by how other people think and act. What makes things even more difficult is that, whether acting for themselves (as stock market investors, for example) or as managers of businesses, people make assessments of the likely outcome of their potential decisions based on a complex and unquantifiable mix of data, information and knowledge. This mix consists of hard fact - current economic indicators such as the Dow-Jones average and its recent trends for example, and less tangible information about their own and others' good and bad experiences, others' opinions, the outlook for businesses in their market sector, the overall economic outlook, and so on. This is all shovelled into the melting pot to give an intangible 'expectation' about the future outcome of their impending decision. Most decisions follow rules of thumb which may be simple to enunciate and equally difficult to pin down precisely. And people try to beat the system - or at least their competitors - by reacting to trends in the light of past experience: if they expect bull market conditions, they buy in anticipation of one happening - and if enough people do this, a bull market is created as a self-fulfilling prophecy.

Fortunately, our evolutionarily stable strategies shared by individually competing businesses fighting brand wars are not subject to quite the same unlikely assumptions as are competitive equilibria in whole markets.

Brand management is very visible externally: brand growth and share are monitored closely by the competition. Changes in the marketing strategy of a big brand can make a major dent in profitability either way. Changes in organization structure are, sales-force apart, internal and less visible to outsiders. Company performance is less sensitive in the short term to such changes. Organization development might, however, be a key - perhaps *the* key - to long-term profitable growth. And since 1996 there has started to be real data available on the command structure of businesses that relate directly to the constraints on growth and to company size.

Perhaps businesses could be re-engineered to profit from this.

CHAPTER 6 - STABILITY, COHESION AND GROWTH

Sand in the works

Businesses can get stuck in a rut for many reasons. An internally over-complex (high-K) business may be stuck at a peak of low fitness in the foothills. It is a sitting target for the competition and very vulnerable. Competition may deform its landscape and allow it to escape, only for it to be driven from that market altogether and perhaps even to extinction in all its markets. Careful and planned injection of change may have the unexpected side effect of kicking the business over an intervening fitness valley and into the foothills of an adjacent fitness slope whose peak is higher than the peak on which the business was stuck. This change could, for example, be a reorganization linked perhaps to a successful small venture into a new and risky market. Deliberately fostering an organization culture where experimentation and the resulting inevitable failures are not penalized has a similar effect. The business then exchanges safety and slow decay for an immediate risk - the valley - plus the prospect of longer-term improvement. This deliberate injection of a controlled disturbance has a precedent: it is the annealing process used by a metal-smith. Metal is alternately heated - the crystalline structure is disturbed - and then cooled to allow the atoms to align themselves. As the metal becomes stronger (fitter) with a better internal structure, the amount of heat applied in any one heating/cooling cycle is progressively reduced. But whereas metal, a sword perhaps, can reach maximum strength when the internal structure is completely aligned, businesses using controlled disturbances in this way can never rest on their laurels because their fitness is not absolute but only relative to their competition.

Managing stability

There is a fine line between stability and stagnation. In Chapter 5 we examined the impact of landscape complexity on the abilities of an object to react quickly or for a sustained period of time to extraneous disturbances and to deformation of the landscape itself by C-coupled competitors. But there is another more subtle way to alter an object's reactivity. It has been assumed until now that the decisions which represent gene values and which try to move the object upward on its landscape are equally likely, i.e. that a 'yes' decision to a request is as likely as a 'no' decision. For example, if in my high-K business I need many signatures from my management at the next level above for an investment request, the need to obtain the signatures adds delay and thus reduces our department's ability to move swiftly. But if most of these signatures are routinely and unquestioningly rubber-stamped, there is at least some predictability in the process. Assume however that two objects, departments for example, were coevolving with a third, and that the third object would always react in a certain way if given the right deformation of its landscape via C-couplings from the other two. For example, given signatures from the first two objects, the third might initiate some capital investment. There is an underlying assumption that the signatures deform the third object's landscape such that capital investment by the third object makes it fitter. If the third object's landscape is such that it only requires the signature of either the first or second (or both), its reaction - to invest - becomes a 'preferred' reaction, one that will tend to happen more often than not. If, at the other extreme, the object required signatures not just from *both* of the first two but from many others besides, the decision to invest will occur less often and will not be a preferred reaction. With a preferred reaction, the combination of objects will attain a more stable and perhaps steady state if the 'yes' decision by the third object feeds back positively to the other objects such that they continue to 'ask' for further investment. The resulting state is certainly more stable than when the reaction is not a preferred one.

This seemingly complex situation can be summarized as follows. If an object is C-coupled to others and responds in a certain way if any one (or more) of the coupled objects 'asks' *and* the response reinforces the other object's decision to make the request, then the collection of coupled objects is relatively stable. Perhaps they even reach an evolutionarily stable state and the landscape deformations cease. To understand why preferred reactions lead to stability, we need to look at the extreme case where each object in a large collection of objects is C-coupled to all the others. Any disturbance reverberates continually around the collection and the result is unpredictable (and mathematically chaotic). The reverberation peters out more readily if one or both of the following occur:

- ❖ the amount of C-coupling is reduced. Each object is then only coupled to some fraction of the total number of other objects. This confines the scope of the disturbance that can only be transmitted via C-couplings.

- ❖ each object reacts in a preferred way most of the time in response to stimuli from one or more C-coupled objects. If an object is giving its preferred reaction in response to one of its stimuli, it will in effect ignore any input from other C-couplings that might say 'do the opposite'. This limits the scope of the disturbance if it is one of these other inputs.

Such considerations are important if we break up a business into many coevolving objects, create a high degree of C-coupling between objects, have low K-complexity within the objects, and possibly have a partial mesh organization structure instead of a pure hierarchy. Such a collection of coevolving objects sounds inherently unstable and adding (or superimposing) preferred reactions may be the only way to stabilize it.

Decisions, decisions

The values of a business's decision genes give rise to a particular state of that business. The state is changed by making decisions. A decision can be a 'yes/no' decision - make a new product, invest in new equipment and so on, or a continuous value: *how much* to invest. All that matters is that the value, in conjunction with the values of the other associated K genes within the object and the C genes in one or more associated objects, can be turned into a measure of fitness. In a large complex business, a small object would bother only how well it met its own targets and not about the financial state of the business as a whole.

Adaptability

It is advantageous if the operations of a factory or business are resilient to disturbances while at the same time are adaptive to desirable change and responsive enough to adapt quickly. But it is now widely held wisdom that a business which cannot itself evolve ever more quickly or adapt to external change cannot grow. And if it cannot grow it dies. But growth can be organic (from within) or by acquisition - joining with other businesses through take-over, merger or merely via continuing close symbiotic business relationships. Such a business group takes on a life of its own which is distinct from the individual businesses that make it up. This increase in complexity may seem a backward step, but it now appears possible that biological species might mutually evolve to the boundary between order and chaos and then remain at the boundary growing more complex all the while. Businesses need to catch up with biology: at the 1994 Bionomics Annual Conference, Dee Hock was damning about the inability of organizations to change at the same rate as the world outside, and their equal aptitude for stifling human ingenuity with hierarchical bureaucracy and 'scientific management'. He reminded them that "... the *nature* of our expertise became the creation and control of constants, uniformity and efficiency while the *need* has become the understanding and co-ordination of variability, complexity and effectiveness". In other words, business management was way out of step with the real world it served.

Re-birth of growth

Reducing complexity is only part of the story. A decomplexed (lowish-K) business that is poised and reactive may well improve its own internal processes but the *process of improvement itself also needs to evolve*. The difference is subtle. When a business is created from scratch, its processes are also largely created from scratch but are strongly influenced by the founders' previous experience and perhaps also by the structure and facilities of computer packages such as SAP R/3 chosen to run major parts of the business. When the business creates a new major subsidiary or division, both the internal processes of the subsidiary and its other less formal ways of working will very likely be modelled on those of its parent. When a business wants to create a new brand or sub-brand, the market, product and manufacturing process research will follow existing practice. Up to a point, sticking to established processes cuts risk and saves money but it eventually leads to stagnation and a dearth of new ideas. Going back to the drawing board every so often forces new ways of working to emerge naturally. It allows businesses to develop new structures and allows brands to develop new product concepts in ways that are beyond what is possible through growth outward from an existing base. But there is more to it than this. A big business is a complex hierarchy. Development through process improvement can occur at any level, at departmental level for example, like the slow evolution of cells in a body. But what is needed in addition is evolution *of the whole structure*. This can only happen if the structure itself is *repeatedly* rebuilt. Richard Dawkins summarized this need for rebuilding thus: “In order to put together a complex ... organ you need a complex developmental sequence. A complex developmental sequence has to have evolved from an earlier developmental sequence that was slightly less complex. There has to be an evolutionary progression of developmental sequences, each one in the series being a slight improvement on its predecessor”. Even if the business has a process improvement programme, the methodology used needs to evolve!

There is a close parallel with how our brains are believed to work. For example, when we are born, the connections within our brains have some prefabrication (‘hard-wiring’) that supports the basic structures that appear to underlie human languages. But our brains also have a great deal of plasticity that allows us to learn *specific* languages. Evolution and natural selection have given us the right balance between the two. As we get older, the plasticity wanes. For language acquisition this happens surprisingly early at the age of around seven. I could not now learn to speak a new foreign language with the idiomatic fluency and accent of a native. If doing so would for some reason help my genes survive, the only choice left open to me is to encourage my children to learn it while they are still young.

It is worth noting that designing and constructing anything *in sequence* is typically a slow process but one that needs little communication. Doing related things *in parallel* is typically faster but needs lots of coordination if the end result is to work; and coordination implies communication (see Chapter 7).

How complex are real businesses?

Until 1996 there was no reliable information on the internal complexity of real organizations. Since then, Gene Stanley, Luís Nunes Amaral and colleagues at the University of Boston have analyzed statistically the growths of all US manufacturing public companies from data which the companies themselves have to file with the US Securities and Exchange Commission. Conventional business-school wisdom on business organization and development accounts quite well for growth due to investment in new plant and machinery; alert businesses track one another as closely in this area as in product development. But there is a difference in timescale: a competitor's innovative product can destroy your market share quickly, but differences in manufacturing efficiency (and hence product cost and perhaps quality) take longer. You die anyway; failure to invest in plant just gives a more lingering death. Stanley and the other Boston 'econophysicists' pointed out that the traditional models of growth did not include any real measure of the organization infrastructure, the time needed to build it, or how the number of organization units such as departments changed as a business grew. And since business school analyses typically concentrated on business units - divisions or other major parts of large enterprises - that make a particular type of product, the overall behaviour of real multidivisional businesses tended to be ignored.

The Boston team has demonstrated that as the years passed and businesses grew, the relative numbers of businesses of each size (in deflated gross sales) remained steady. There were not, for example, relatively more of the larger firms than smaller ones in later years. This is counter-intuitive: it would be reasonable to expect to find more of the larger businesses as computing and further automation progressively took hold. In addition, they found that the ranges of growth of different businesses of a particular size became narrower as the businesses became larger. In other words larger businesses were more likely to be growing at a similar rate than smaller businesses. This was less of a surprise as big businesses are more likely to be diversified with growth in one sector being masked by a drop-off in another. But what was unexpected was that it also seemed to apply to the many different ways in which the sizes of businesses could be categorized. As well as gross sales, the rule applied to business sizes specified in terms of cost of goods sold, assets and number of employees. This suggested that whatever governed both the growth rate and the variation in growth rate was independent of the type of business or the technology of the business.

They went on to look at one thing that is common to all firms in different sectors: a hierarchical management structure. They asked what would be the variation in overall growth of a business whose constituent units obey only *some* of top management's directives and make the other decisions themselves. More generally, if a business were a multi-level hierarchy, they asked what the effect would be of only *some* of the decisions at one level being obeyed at the next level downward. They also asked how this could be related to the mysterious common pattern of variation in growth rate of real businesses that were in different market sectors.

Their results indicate that for real companies with between two and ten levels of hierarchical organization, around 80% of the decisions which would result in a variation in sales growth are cascaded downwards - which suggests close coupling between the levels and 'dictatorial' management. This may sound like a good recipe for a reactive business: top management says jump and everyone jumps. But it may also follow that individual business units are precluded from taking opportunities themselves without agreement from all upper levels and perhaps other units at the same level as well - the fly caught in the web again.

It is worth restating these conclusions to highlight how they can be exploited. Manufacturing businesses of a particular size are constrained to a range of growth rates, and the constraint grows tighter as the business grows. The constraint appears to be related to the tightness of the organization structure - the degree of top management autocracy versus divisional autonomy. This suggests that if we change how tightly the organization structure is coupled, we can influence the constraint on growth. Thus a firm struggling for growth might find things easier if, instead of (or as well as) investing in plant or new products, it flexes this degree of coupling.

We are now back on familiar ground. This coupling in a coevolving business is Kauffman's C-coupling. A model of the organization could be created with known degrees of C-coupling and K-complexity. For example, how many (K) other people have to authorize my major decisions? Such authorization is not necessarily rubber-stamping: they may all hold conflicting opinions and their departments could be materially affected by my decision. The model could then be tuned either manually or self-adaptively to the boundary between order and chaos and the new K and C noted.

The amount K and C have changed is an indication of the degree to which an organization change (to a lesser or greater degree of divisional autonomy) might allow the constraint on growth to be widened. There appears, unfortunately, to be no definitive data yet extant on the organizational tightness of those businesses whose growth is at the high end of the growth range for their size. If there were, it would be easy to predict the additional growth that could ensue from a reorganization on the lines described. The best we can do at the moment is to suggest that there are intuitive grounds for believing that firms that change their complexity such that they sit on the boundary between order and chaos have moved in the direction of loosening the constraint on growth.

The new managers

A coevolving business would be run very differently from the ones we are used to. In addition to monitoring performance and setting plans and objectives, senior management would focus on defining and continually redefining the 'right' number of objects and their delineation and couplings. Lower down the tree, the aim of an object manager becomes more difficult: to meet his targets *and* to readjust the internal K-complexity of his area and C-couplings to other areas in order to balance responsiveness with stability. He may not, of course, have all the freedom he wants to adjust the internal complexity and couplings, particularly the latter. This is what earlier we called 'manual tuning': the object manager is attempting to predict in advance which way to change K (and C if possible). He has an alternative: self-organization.

He could keep changing K for example, watch how his object's performance varies - perhaps production throughput drops or energy consumption goes up - and choose better values of K and possibly C in that way. In a manufacturing plant this would be heresy; it is perhaps less so in the more glamorous side of the business where test marketing is an accepted practice.

If (a big if...) all this could be made to work successfully, the organization or facility would, without senior management assistance, seek out the special critical point between order and chaos and stay there forever, poised but wobbling slightly. If objects are manually tuned, there is the difficulty of knowing in which direction to tune. If the objects are allowed to self-organize, occasional avalanches of change are inevitable as efficiency rises and then levels off. Perhaps senior management would see their custodial role as one of putting some brake on how far object managers may tune K and C. Perhaps a little sub optimality and sluggishness is a good idea: there is some evidence that self-organization in Nature itself is biased on the side of stability and caution.

Effect on people

Surprisingly, current trends in Personnel management - such as empowerment of the individual - are aligned more closely with this new type of 'decomplexed' organization. Emphasizing the role of a manager as a leader rather than as a 'do this, do that' autocrat in the old mould is all-important. Equally, the manager of an object area has a responsibility to create the environment in which his people work: to set and continually adjust the complexity internal to the area and, where possible, the coupling between the area and others.

A business of this nature with a hierarchy which gives a great deal of autonomy (low C) to its constituent divisions nevertheless needs to have widely-communicated and coherent directions in which it wants to move and long-term goals it wants to attain. These are the currently fashionable Mission and Vision statements. People may - *must*, in fact - have local objectives. These local objectives must be attainable by them without a (high-K) web of interference caused by a mandate that everything they want to undertake has to be approved by others. For example, because a well-run business gives individuals suitable amounts of freedom and does not treat employees as machines, these individuals also need some context to guide their decisions in cases when there are no local objectives to steer by. A Services Buyer may be given an objective of minimizing the unit cost of buying electricity. He can do this quite single-mindedly, given only the approximate volume needed. But knowing the context in which electricity is being used - whether using less or more is better for the business and what the alternatives to electricity are - creates challenging dialogues between Buying and those who use the electricity. The result might be to take a short-term contract and then switch to another form of energy. There is, however, a yawning gap between remote and often unquantifiable Mission and Vision statements and individuals' local and more measurable objectives. This gap can be partly bridged by the Balanced Scorecard type of objectives and measurements devised Robert Kaplan and David Norton that are easier to align with the Mission and Vision statements and with measures of quality, customer service and the like. They can be very different to conventional cost-orientated management accounting measures.

Employees are suppliers of their own services to their internal 'customers' and to external (real) customers. They are also customers of other employees and of suppliers to the business. It is thus possible to create *efficiency* objectives, such as the minimisation of the *unit* price of something, which are concrete and clear and can be achieved autonomously. These can be related to *effectiveness* objectives such as *how much* of something such as secretarial assistance or preventative maintenance engineering is best for the business. 'Less' is not necessarily better, and deciding on the right level might need a wide knowledge of the business. My efficiency objective may be related to your effectiveness objective. I do not need your approval to pursue my objective and you can pursue yours without reference to me except that the unit cost at which I buy may influence your decision on how much of the service should be provided. Working together without needing each other's approval to act *and* with related but not directly interacting efficiency and effectiveness objectives is the essence of good organization. I support you but I do not tread on your toes; you do not interfere in how I achieve my own objectives. There is an apparent contradiction between an individual's having the freedom to pursue his own objectives without interference from other areas and the need for employees to see the wider context in which they work - *why* they do what they do. But the freedom to take decisions without reference to others, or with reference just to a limited few, implies that the organization is well structured such that the results of my decision do not mess up other areas. It may *affect* other areas, but such areas in a well-structured business are carefully delineated so that the impact is similar to that of, for example, a customer making an order on a supplier. Such an order may or (occasionally) may not be welcome but the supplier has organized his business processes around reacting to incoming orders.

The principle that an individual or department can be a customer to someone in the business and a supplier to someone else is at the heart of Activity Based Costing. What is proposed above is an elaboration that gives people more freedom while enabling them to know what their roles in the business are. For those unfamiliar with it, Activity Based Costing imposes two views on any business. The first is a 'process' view in which product manufacture, for example, is broken down into the series of interlinked (supplier - customer) activities mentioned above. In this view, each activity has one or more cost drivers. These tend to be non-financial but as nearly as possible directly reflect what causes cost within the activity: 'number of mechanical operations' is a common driver. Each activity also has some performance measures such that the result of the cost drivers can be tracked: 'elapsed time per operation' is frequently employed as it is a useful indicator of the related people cost as well as the degree of flexibility available to switch production between products. The process view is independent of particular orders or particular products. The second view is the ‘cost assignment’ view in which the need to make one hundred pink widgets, for example, is turned into a series of activities, each of which may need resources (people, materials and so on) to complete. Where Activity Based Costing differs markedly from conventional management accounting is that it includes every significant activity - overhead ones such as customer service as well as the more obvious manufacturing direct cost activities. This means that it can apply to service industries as well as process or assembly manufacturing.

The significance of Activity Based Costing for an internally coevolving business is that it provides a ready-made fitness management framework. Each activity could be defined as an object but this is probably extreme. In practice, activities would be grouped into manageable clusters such that the clusters were as autonomous as possible. One good way to delineate cluster boundaries is to choose the points where one potential cluster can internally supply two or more other clusters or can be an internal customer to them. The cost drivers for each activity should still be local to that activity. The process view would then be a series of clusters of activities. Each cluster would be C-coupled to its predecessor and successor activities. Each cluster of activities then undertakes commissions (works orders) to use its resources to play its parochial part in the manufacture of the order for pink widgets. The process view allows the efficiency of each cluster to be monitored, more or less irrespective of what is actually being made.

Missions, goals and ethics

A large and diversified enterprise will of necessity have *different* purpose and goal statements for each of its divisions if these statements are not to become so vague and high level as to lose all value. But something else is needed to make the enterprise hang together. This is usually a single underlying concept or a short list of ethical principles. IBMers from the 1930s onward were reminded that, whatever else they did, they must Think - the name of the company magazine and Chairman Thomas J. Watson Senior’s favourite exhortation, a reminder of which lived on until recently on many of the older employees’ desks. Other large businesses, particularly those of American origin, have a short statement of business ethics displayed in the reception area of each building. For some this is sufficient lip service. Others take it more seriously. The ‘Five Principles of Mars’ for example are displayed in the local language in *every* Mars open-plan office and on *every* meeting room wall worldwide. Having such common principles widely published and understood gives a large low-C enterprise the cohesion it would otherwise lack.

One consequence of giving divisions, departments, teams and ultimately individuals more autonomy is that power and responsibility must be delegated as far down the tree as possible. Giving people responsibility means treating them responsibly and not as machines. This then implies giving them the same status as far as possible. Names such as ‘partner’ (House of Fraser, UK) and 'associate' (J.C. Penney, Wal-Mart, Mars and several others) have been used to emphasize this equality of status.

Management rule

Management by setting and tracking financial objectives can be taken too far and can easily squash creativity, particularly if done in a large centralist organization. Three well-known and rather simplistic examples will suffice.

Accountant Harold Geneen, who died in November 1997, ran the sprawling ITT empire through a coterie of browbeaten senior managers. They were put on trial and individually interrogated every month by Geneen in front of their peers to establish 'unshakeable facts': management by objectives certainly, but guaranteed to promote safe and conservative ideas. An enterprise built from some 350 individual acquisitions sounds an ideal candidate for delegated decision-making but that was not Geneen's approach.

Geneen's style was mirrored in the way US automotive giant General Motors was run after 1958. To find the origins of this change we need to look further back. In the recession after the First World War, GM was out of control financially. It ran out of cash and had piles of unsold inventory. At that time, GM was the epitome of decentralized management but without commensurate decentralized financial responsibility. There was no common vision of where GM should be going. But even if there had been such a vision, the motley collection of businesses that GM comprised was not even financially audited so there was no obvious way head office would have been able to track progress. When founder William Durant was forced out of the business in 1920, financial man Pierre Du Pont took the reins as Chairman and President. But he left it to Alfred P. Sloan, who had a sound business brain *and* an engineering background, to recommend a structure for the corporation that combined the best of decentralized management and central control. Sloan himself acknowledged the contradiction inherent in this remit. The outcome was (initially) two central committees - a Finance Committee to define financial policy and an Executive Committee that defined operational policy. The Chairman of the Board was to be Chief Financial Officer. Each constituent business had its own financial staff who reported both within the business and to head office. The President, who was to be both Chief Executive and Chief Operating Officer, had to be someone who understood manufacturing, and Sloan himself was the first appointee after Du Pont relinquished the role in 1923.

In 1958 Frederic Donner was appointed Chairman and hence Chief Financial Officer. *But he was also appointed Chief Executive*. Donner was an archetypal figure of the Geneen mould. He did not have a close affinity with the business but possessed considerable ability to understand, remember and relate its formidable array of financial and other business performance measures - as he liked to remind people. Like Geneen, control stopped with him and attempts by the manufacturing or sales factions to have a detailed say in how the business was run were ruthlessly quashed. This slowly but radically changed the way GM worked. It was out of kilter with Sloan's conception of a decentralized enterprise that had the minimal overriding common policies and controls needed to ensure cohesion of the many business units. Arguably, GM is still paying the penalty today. In its original form, GM could be defined as a set of businesses, each a coevolutionary object and each of which was low-K and almost zero-C except where market share stealing occurred. The predictable behaviour was anarchy when seriously (W-) perturbed (see Chapter 2) by external influences such as an economic downturn. It ended up as an enterprise where each business was subservient to head office: the enterprise itself became an object with a high K - the result of lots of overlapping committees with different views. The equally predictable behaviour was sluggish responsiveness to structural changes in the market and successful attacks by VW and Japanese manufacturers.

In the 1960s and 1970s, IBM was the epitome of corporate success fuelled by the rewards of colossal investment in its S/360 mainframe computer line in the mid-1960s. In 1980 Chairman Frank Cary decided that Apple's attractive Apple 11 personal computer equipped with the VisiCalc spreadsheet system was offering something too useful to ignore. He had been trying for several years to launch a similar machine but IBM's well-entrenched bureaucracy and high fixed costs ensured that the results were slow, too expensive and completely uncompetitive. IBM's credo was 'sell to corporate customers', 'sell via the IBM sales force', 'have standard IBM maintenance agreements' and 'build with IBM-manufactured parts' - the way it had always built and sold mainframes. Cary finally got his way by hiving off a small development group to a remote corner of IBM's empire - a converted warehouse in Boca Raton, Florida - reporting directly to himself and run by engineer and corporate maverick Don Estridge. The result was the IBM PC that appeared in August 1981 and which within one year delivered a $1billion revenue windfall. It also legitimized the corporate market for personal computers that until then had been confined to niche areas and hobbyists. The PC was followed by the PC XT in early 1983 and the equally successful PC AT in mid-1984 by which time PC revenues had hit $4billion. But all was not well within IBM. Frank Cary retired and was followed by a succession of uninspiring Chairmen who lacked Cary's intellect and who preached decentralization - "just look at the success of the PC" - while doing nothing whatever to implement it. The success of the independent upstart unsettled IBM. It shepherded the PC development group back into the corporate fold as a Division in 1983 and replaced Estridge in 1985 by Bill Lowe, a dyed-in-the-wool bureaucrat. At the same time, revenue growth from mainframes - the bedrock of IBM's profits - faltered as corporate clients made unflattering value-for-money comparisons between mainframes and the new world of easy-to-use PCs with colourful graphics, a snappy response and Lotus 1-2-3 which was to the PC what VisiCalc was to the Apple. Worse was to come. Whereas cloning of mainframe computers and related devices such as VDUs had long been tolerated by IBM because of its traditional stranglehold over corporate customers and the difficulty clone manufacturers had in keeping in step with IBM developments, the PC was easy to clone. Many were sold direct to users and to those small businesses that were not typical IBM customers.

IBM had sown the seeds of its own near-demise by:

- using an autonomous (low C) and unbureaucratic (low K) business unit to develop a product which had a freely available specification
- removing that autonomy just when the product was successful and had created a new market
- reintroducing the unit's internal bureaucracy

IBM might, in the past, have got away with this. But the success of the PC laid waste to its traditional mainframe market and revenues. The PC clone makers had much lower fixed costs than IBM and had the autonomy that IBM now denied its fledgling offshoot. They profited hugely at IBM's expense.

The PC was not IBM's only flirt with autonomous business units. Its typewriter and printer business in Lexington, Kentucky was similarly shackled with corporate red tape and was unable to foresee the demise of the typewriter or respond to the need for small laser printers to attach to PCs for word processing. It was sold off in a leveraged buyout in 1991, scrapped layers of management, delegated decision-making and became a byword for informality and productivity and anathema within IBM.

IBM itself staggered on accumulating problems which culminated in the 1990 profit of $6billion turning into a loss of $5billion in 1992 and the ousting of Chairman John Akers in January 1993. His replacement was Lou Gerstner, a former McKinsey senior partner and CEO of RJR Nabisco: someone definitely not in the IBM mould.

One of Gerstner's first acts was effectively to abolish IBM's Management Committee. The Management Committee was the arbitrator between the warring factions within IBM that emerged as a result of the decentralization enforced by Chairman Tom Watson Jr in the 1950s. Watson encouraged this sparring as a necessary antidote to corporate complacency and, like Frank Cary, deliberately placed 'awkward' individuals - those with bizarre ideas and who were not afraid to vent them - in sensitive places. Watson pushed decision making down the tree with the Management Committee as a court of last resort. But as the organization exploded from a mere 50,000 employees in the late 1950s, the business gradually coalesced into a mammoth bureaucracy. The Management Committee then came to be used to make increasingly low-level decisions. The stultifying hierarchy had reappeared.

At first sight, it may seem that person-to-person communication and data flow within pre-Gerstner IBM or the 'new' GM would be better than within a collection of freewheeling units. After all, communication of business performance upwards and operational direction downwards in GM, for example, was enforced by corporate fiat from the very top of the tree - the Fourteenth Floor sanctum of GM's Detroit HQ. But this is a half-truth at best. Information overload is one of the chief bugbears of present-day business managers. Restructuring the organization to the right level of autonomy also provides the opportunity to improve communication by getting rid of unnecessary and unwanted communication channels. This is not a guaranteed panacea for information overload but it does at least ensure that information flowing around the business is of potential use to the recipient. If I receive two hundred electronic mail messages as soon as I start work for the day, the dross masks the nuggets of gold. Not only have I wasted time reading, filtering and consigning the rubbish to my electronic trashcan, but much of what is left is pointless: forms which need my signature quite unnecessarily; documents which I must review, again quite unnecessarily; and so on. The plaint of the overburdened manager is ‘just let me know what I need to know’. He relies on his network of contacts for the rest.

A re-engineered business needs information to flow around it efficiently. Unnecessary communication channels must be killed off at source. It also needs to support the evolving networks of contacts, particularly the internal ones. As any successful manager well knows, grapevines are not static: they must be nourished and cultivated with quid pro quo exchanges of useful or at least interesting information or they wither and die.

CHAPTER 7 - COMMUNICATION

Poor reactions

Coevolution of business objects implies communication between the objects. Without this (i.e. if C were zero), there is no spur to keep the coevolution process going. Objects would freewheel to meet their own internal objectives and ignore collaboration or competition with other objects. So when an organization has been re-engineered to the right level of complexity, two further things are needed in order that it works as intended and can be exploited effectively by the business people who *are* the business in the sense that they make the decisions that drive coevolution. One requirement is that communication happens efficiently. If a C-coupling has an inbuilt delay before its impact on the recipient object is felt, the effect on coevolution is unpredictable and depends on delays in other C-couplings. A simple example will suffice. A ship's helmsman, who may be a real person or the automatic equivalent, tries to correct the small deviations from the ship's planned course. There is inevitably some delay between noticing a deviation and his correction to the course taking effect. A slight slewing off course caused by a gust of wind may be self-correcting. But if the helmsman's reactions are so slow that his correction is only taking effect when the ship is swinging back under its own accord, his correction will make things worse and the ship will start snaking badly and will eventually go broadside on. What is certain is that if businesses have more efficient C-couplings between their internal objects, they will be able to change direction faster. More efficient external C-couplings, links between competitors via market share for example, make the business more responsive to changes in the market.

User friendly

Those who use a computer for business have a common complaint that has not changed much in thirty years: they want a system with the flexibility to do exactly what they want. But there are many underlying problems that will frustrate this. The first is that the system is probably used by many others who want it to work for *them* in *their* personalized way. People are not just different in what they want: their requirements change as they get used to the system. A shared system that is very user-friendly with lots of menus and helpful directions will all too soon be found cumbersome by someone who uses it all day to enter customer orders or payable invoices. This implies that a shared system for a large group of users must either:

- force users into the straitjacket of doing things one way, perhaps some common denominator which satisfies no-one

or

- must be extremely customizable by the user. It should preferably adapt automatically to a user's pattern of usage by progressively streamlining the system functions used frequently but letting the less-frequently used ones retain their somewhat baroque but easier-to-use screen displays.

Systems that try to adapt to the user's apparent needs and level of skill are sometimes too clever. Even personal productivity packages that are essentially for non-shared use can be a nuisance in this respect. For example, the current generation of word-processor packages will try to interpret 5th Jan as a date and put the 'th' as a superscript. This is fine until I want to type 5TH JAN as a product code...

There is no perfect system, but a business that has been successfully re-engineered into coevolving objects might reasonably demand that its computer systems should complement the newly found departmental autonomy and personal flexibility. Giving everyone a state-of-the-art PC and powerful set of personal productivity packages for word-processing, spreadsheets and so on is not the answer. But it might *seem* as if it should be. Giving adequate flexibility to allow individual initiative to be deployed effectively without unnecessary constraints is, after all, the essence of creating an efficient coevolving business. Exactly *why* it is wrong now needs some elaboration.

Decisions vs. data

The taking of decisions drives the evolution of a coevolving business. But the backbone of the operational side of any business is not decisions but data. Information and knowledge are also important but become more relevant as we move in timescale from day-to-day operations to planning and strategy; we will look further at both of these in the next chapter. Aside from physical operations such as manufacture or distribution, the operations of a business are just the processes whose transactions convert one form of data to another - from customer order to works order to depot picking note to delivery advice to invoice to electronic payment. Processing is not entirely 'automatic': an order that is not prepaid will be validated against the customer's credit limit for example. Any transaction may well contain some yes/no logic of this nature. Things become murkier as the decision rules become more complex. Agreeing a bank loan or taking out car insurance might be an automatic process. But the myriad decision rules embedded in the underlying 'expert system' need straightforward hard data from the hopeful applicant such as his age *and* also usually pose questions whose meanings are not straightforward. "Have you been convicted of any criminal offence?" will usually provoke the response "Is a parking ticket a criminal offence?" and the question "Have you ever been involved in an accident and if so what?" provokes more confusion: "Does an accident thirty years ago count", "Yes I have, but I was a passenger/onlooker/the injured or innocent party/...". And so on. The precise and unambiguous meanings of some of the questions will be lengthy and probably shrouded in legal jargon. The poor applicant needs some help to interpret what is meant. Even the well-trained bank manager or insurance salesperson will need to be empowered to use a measure of discretion. The real world is a messy gray area where things are not easy to define unambiguously. Legal statutes are criticized for being unreadable even by lawyers, but those who draft legislation try to cover all eventualities. Society is a forever evolving and amorphous beast, and where there is benefit in exploiting apparent loopholes in legislation, then someone somewhere will do so. Tax consultants make a living at it and the Revenue make a (less well-paid) living plugging the newly exposed loopholes or getting the legislature to plug them.

The upshot of all this is that systems themselves need to change perpetually. This is probably not a result of bad systems design. Neither is it likely to be lack of foresight. As we have already seen, the business planning processes and the supporting planning models coevolve with those of the competition. Business processes will change initially as a result of re-engineering the organization to a point on the boundary between order and chaos, but they also change continually as the business coevolves internally and as it responds to changes in its markets. If business planning predicts an uncertain future and if the planning process itself and its computer models evolve, how much more should the operational computer systems which support the structure of the business evolve. After all, they need to cater not just for changes in business process but also for changes in the user clientele. Perhaps experienced users who prefer quick shortcuts to friendly menus move to other jobs or retire and are replaced with others with less experience. These newcomers will initially need the help and intuitive way of working which a good system will provide.

What tend to be more stable are the underlying data items. The business may change how it makes widgets, how its sales division sells them, and how it accounts for widget sales, but the widgets tend to carry on. Larger ones, green ones and military specification ones may be added to the order form and the small pink ones may be discontinued, but widgets are widgets. Since the 1970s there has been a heated and rather pointless debate among IT business systems designers about whether systems should be designed around data or around business processes. The debate is fatuous: they are two aspects ('duals') of the same thing. The process bigots create huge process maps of the business and then depict data flowing from process to process. The data bigots create definitions of data items, then create repositories - databases - for them, and finally superimpose business processes to convert data from one form to another. What really matters is agreeing definitions for data and for what the business processes do to the data. But the ways in which users manipulate data must, as we have seen, be malleable by the user to suit personal preference as long as the underlying ways of processing the data - company rules for allowing credit for example - are not modified.

Implementing an enterprise-wide system such as SAP AG's R/3 to handle both transactions and business planning is not necessarily at odds with the personal and departmental autonomy to make decisions that is enjoyed in a low-K business. If the system provides the flexibility to add new business processes without disrupting the flow of existing ones and the capability to enter into a dialogue with its users in the way *they* wish - whether slick and terse or menu-ridden and friendly - it will not stifle individual freedom. But it should not be criticised if, in spite of this, it is complex to use because the business processes are themselves necessarily complex. If a system is *logically* complex and not just complex because a systems designer overreached himself or because an influential user insisted that the system cater for every conceivable business eventuality, it is certainly worth asking if the complexity is actually profitable to the business. Sales managers for example have a tendency to put together elaborate ways of promoting their wares. They introduce strange promotional bonus systems that are difficult to explain, conflict with one another, and are even more difficult to implement in computer programs - which is a very good test of whether they can be defined precisely! The justifications are invariably that 'the competition also does it', that 'the retail trade expects it' or that 'sales will otherwise fall off long term' - all of which are usually unproven and unprovable. But if a process is really and necessarily logically complex, then there is a strict limit on how far the complexity can be masked from users without inhibiting the experienced user.

What do people want?

Data flows are merely the bedrock upon which decisions are founded. Anyone who makes decisions needs something more - a formal and informal network of contacts inside and outside the business.

Any multinational business has communication problems, if only those caused by time-zone differences. A manager within any business that has grown quickly will often hark back to the 'good old days' when the person he needed to talk to was in the same room or just down the hallway. Life soon ceases to be that simple, not just because the two parties are separated geographically but because people's ways of working have grown more individualistic. Businesspeople travel more and have more varied ways of organizing their time. With the greater pressures to meet their own objectives and with fewer supporting staff to help them, they often do not want to be interrupted merely when it is convenient for someone else to interrupt them. That creates a dilemma. As organization structures flatten and personal responsibility gets forever more onerous, there is a pressing need for everyone to be contactable at all times. But there is also an equally pressing need for people to be able to manage or even bar interruptions in order to finish some heads-down planning work, for example, or even to get some sleep. The present author has lost count of the times he has been woken at 3.30 a.m. EST in US hotels by 08:30 a.m. GMT calls from UK. Most were the result of a momentary lapse of commonsense by the caller but around a quarter were simply inconsiderate. People want technology to give them a 'virtual' equivalent of the single large office they once knew and in which communication was face to face. This *virtual office* is slowly becoming a reality but the technology is struggling to keep pace with the changes in how people work, the growing multinational nature of individual jobs and individual mobility.

Person to person

Person-to-person communication has much in common with telecommunications. We will now look at this similarity in some detail since person-to-person communication in business will soon become much more automated and time-critical. The main difference is that much of what goes on between people is implied and not stated. Say, for example, I walk to the podium to give a lecture. Hopefully, the audience quietens down and waits for me to begin. There is an implied *protocol* that I will speak first and that the audience will listen. I may later invite questions and I am then explicitly passing control to someone in the audience who will expect me to answer when they have finished. Not only is there a protocol agreed but there is an informal negotiation process to agree it: I offer a protocol and the audience accepts it. This sometimes fails at political rallies and stockholder meetings where there is something controversial on the agenda; the offered protocols break down because the audience does not want debate to be shackled by the speaker. Informal protocols sometimes fail also. When two people stride towards each other on the sidewalk, each tries to avoid the other by stepping out of the way. We look for small signals from the other party about which way *they* will move. If both parties step to the same side there is an embarrassing impasse, particularly as the natural inclination is for both to step the other way and repeat the process with more embarrassment ("Shall we dance?"). The impasse is often resolved by one party taking the initiative to stand stock-still. Where errors could be more serious, we have evolved more formal rules such as 'drive on the left'. When person-to-person protocols have been agreed, explicitly or implicitly, there is still a problem of language (*format*). Within a country or community this is usually implied. American-based multinational companies usually make English a de-facto standard for communication between units in different countries, whether for oral communication or electronic mail. The speed of communication is usually implied also: sensible and considerate people will slow down and use simpler words when talking their native language to a foreigner who cannot speak it as well as they do. The experts at this will be observed to modify their talking speed and choice of words on the fly as they discover how well or badly the other person understands and speaks to them in return. But in spite of all this we may still experience communication errors where what was spoken does not make sense. The person on the receiving end can either guess what was said from the context (a telecommunications engineer would call such successful guesses 'in-flight error correction') or ask for the words to be repeated. Typically though we ask for repetition from a convenient checkpoint - the last point in the conversation we clearly understood - and not from the beginning. The considerate speaker has to achieve the right balance of slowness and *redundancy* - the extra words or rephrasing inserted to allow the listener to guess the meaning of words or phrases he could not understand. This idea of an adaptive dialogue will be revisited in the next chapter when we look at how to tap the knowledge held by an expert or specialist. It is one of the fundamental differences between talking to a person and talking to a machine.

Technology

Communication is a compromise between the initiators who want it and those at the other end who may not - or at least not then. If you carry a mobile phone, I can ring you almost anywhere, but you can switch it off and have calls diverted to a voice mail system. In other words a potential recipient can choose whether or not to be interrupted. Some technologies give interactive communication; others give the recipient the choice of when to participate. There is no one right solution as the initiator would usually like the recipient to participate immediately. From the point of view of a coevolving business, though, responsiveness is important. The ability to defer or bar interruptions can have a significant impact if taken to an extreme. This will be only too familiar to anyone who has had the frustrating experience of trying to contact someone in a business whose culture is to forward all desktop phones to voicemail. Since this is usually done for internal as well as external calls, the impact on internal responsiveness can be profound.

Some popular person-to-person communication products are listed below with an indication of their ability to interrupt:

	Will Interrupt?	Examples of current products
Interactive		
Phone	Yes	
Videoconferencing	Yes	ViewStation (Polycom Inc.); SwiftSite (PictureTel Corp.)
Collaborative tools	Yes	NetMeeting (Microsoft Corp.)
Pager	Yes	
Delayed		
Electronic mail	No	Outlook and Exchange Server (Microsoft Corp.); Notes (Lotus Corp.); Internet service providers' generic mailbox services
Shared database or folder	No	Exchange Server; Notes; Internet service providers' Web servers such as Apache, and Web server products such as IIS (Microsoft Corp.) and Domino (Lotus Corp.) built primarily for use within organizations (on an In<u>tra</u>net)
Voicemail	No	
Fax	Yes	

These are, of course, in addition to conventional post and courier services.

The table is, however, an over-simplification and the boundaries between the areas are not clear-cut:

- ❖ electronic mail, voice mail and computer-based fax systems can be set up to page the recipient when a message arrives
- ❖ electronic mail systems can store, file, display and forward received faxes
- ❖ a variant of electronic mail known as 'instant messaging' provided by, for example, Lotus's SameTime messaging system, allows *interactive* mail communication between groups of users to be set up spontaneously. It is essentially a 'text' version of videoconferencing
- ❖ voicemail systems can store faxes for later retrieval by a fax machine
- ❖ PC products such as NetMeeting provide passable audioconferencing and videoconferencing as well as document collaboration, although audio and videoconferencing* need additional equipment to support conferences between more than two users

Then there is the Internet that can act as a transport mechanism for all the items in the table. As a digression, it is worth pointing out that the quality of voice and videoconferencing over the Internet is unpredictable and will remain so until the Internet service providers upgrade it to support guaranteed qualities of service with low upper limits on end-to-end transmission delay and variation in delay. The human ear - or rather the brain - has been honed by evolution to expect no discernible *variation* in the transmission speed of speech; such variation has only been introduced by electronics over the past fifty years. Some *consistent* delay is tolerable provided we cannot see the speaker, but poor ‘lip synch’ - the lack of synchronization between speech and the movements of the speaker’s mouth - is more unsettling.

The Internet is now an indispensable part of most businesses. Apart from its use in advertising and selling, it is fast becoming a way to give a business a single standard method of linking people together, both staff members and those in the outside world. A decomplexed coevolving business in which individuals and departments have an unusual amount of autonomy still has a great need to communicate, as we shall see in the next chapter.

For a business, the Internet consists of two things that are apt to get overlooked in the general euphoria surrounding its multimedia use in the home:

- a mechanism for shipping data packets from source to recipient. This part works a bit like a railway that will carry many different types of trucks and carriages provided they conform to the same track width and they keep within specified height and width limits

- separate and very different services which are provided by the Internet service providers themselves, by companies who find it worth their while to advertise or to accept commercial transactions such as orders, and by users who can provide their own services (often “my personal Web page” with photographs of family and friends). These different services need different programs running in the user’s PC in order to access them, although some popular services such as mail, interactive chatting and newsgroups are often provided from within a single package. (Newsgroups are typically bulletin boards where questions, answers and comments can be displayed for all to see, along with some indication of which answer was to what question)

Electronic mail and the Web are the two areas growing massively. Electronic commerce - the ordering of items by retail customers from a displayed catalogue, for example - uses the underlying facilities of both but is essentially a Web application with added security. The popularity of the Web stems from two appealing characteristics:

- it is visually attractive: pages can contain still pictures, video clips and voice for example
- index search facilities are available. These are provided by indexing ‘search engine’ companies who make money through advertising rather than through charging for enquiries. Most search engines, which are in essence automated versions of Web browsers, continually trawl the

huge number of Web sites looking for new material to add to their indices (although Web site owners can request a visit from a search engine). Interestingly, the doyen of indexing companies, Yahoo, does not trawl in this way: Web site owners ask for their material to be indexed in predefined categories.

Most large businesses have built their own internal ('intranet') versions of the Internet for use solely within the business in order to hold and display corporate information. These have not superseded data processing systems - yet - but can if necessary act as a means of displaying and interrogating corporate data. Many enterprise-wide transaction and planning systems such as SAP AG's R/3 can now be accessed with a Web browser. Some businesses discourage internal electronic mail being sent to more than one recipient. Instead they ask users to put the material on a Web site that has newsgroup-like facilities with suitable privacy restrictions. This allows users within a team to keep all the team correspondence in one place where useful material - including dialogues worth keeping - can be archived for posterity. This is one way to prevent valuable corporate information being lost but is not a popular one.

Person-to-person communication in business is normally thought of as something that happens on a timescale of a few hours or days. Electronic mail messages for example are delivered much faster than the recipients open and read them. Phone calls are returned at the caller's convenience. So what was the point of describing at some length the communication protocols we use and the impediments to immediate response? The answer is that the ways in which we communicate today will be automated very much further within the next few years. Those businesses that believe that straightforward electronic mail is the height of technology are in for a surprise. Rudimentary mechanisms are already provided in the popular electronic mail programs to allow the text of incoming messages to be analyzed, a reply sent and the incoming message filed in the right place - all automatically. The user merely sets up a series of rules to tell the system what to do. But these rule-based systems are a pale shadow of what is to come. Firstly, mail, voice and other multimedia sources of information will merge. Systems are already available that read a user's electronic mail to him over the phone and, less reliably, turn voice into electronic mail. More significantly, PCs will start to observe how their users work and then automatically generate their *own* rules of how to handle mail of any type. Responses - and responses to responses - will happen within seconds. Person-to-person protocols will emerge naturally and become embedded into systems that learn. The user will be interrupted by the system for things it cannot handle or is not allowed to handle. The user will become part of the system. He will be the slowest part but the one with abilities to recognize subtle patterns and formulate strategies. Automated personal communication will be of particular value in a critically coevolving business. Such a business will have been deliberately re-engineered with the aim of making it as responsive as possible without instability.

From the invention of the phone onwards, electronic communication has changed the need for people to be together in one place. Proximity now means part of a communications network. This is important when we consider the impact of people, systems or businesses upon each other. A few hundred years ago, this impact was mostly local - perhaps restricted to the same village. Now it is global.

Lessons from telecomms

It was shown earlier how person-to-person communication had much in common with the standards that have been developed for data communication. This is, however, only part of the story. The IT industry has been searching for thirty years for a Holy Grail: a way to build large systems that are:

- resilient to failure - if one part fails, other parts are unaffected. If this is logically impossible because the first part is a key component that feeds the rest, then the remaining parts must still work in a degraded but useful way. If one PC in an office fails, the remainder will continue to work. But if the underfloor cabling system to the PCs fails, what happens then? Worse, if the links from this office to other offices fail, how long can the office continue to function?

- widespread - with employees around the world using the same system. In the days when all computing - apart from perhaps manufacturing plant and process management - was done in a large data centre, the system may well have run on a single computer there. Nowadays, several copies might run on different computers, perhaps one in each country or even one in each office.

The centralist and distributed approaches each have problems:

- centralist approach:

  - if the central system fails or the links to it fail, then all users are cut off
  - if an erroneous change is made to the central system, then all users are immediately affected

- distributed approach:

  - if data - sales orders perhaps - are accumulated on systems which run in each office, how are they consolidated into regional or worldwide sales reports? Transmitting the data to a central point is relatively easy, but what happens if one of the office systems is out of commission, perhaps because of a computer failure, when the time comes to consolidate?

- when changes are made to the system itself, how are these changes distributed and implemented on each local computer? All at once? One by one? And if they are implemented one by one what happens if the consolidation is performed while half way through the implementation and the change affects the format of data collected?

Perhaps unsurprisingly, the quest for the Grail continues ever and anon. These problems are the same problems as business organizations themselves face but in a technical guise.

One area of IT, data communication, is particularly affected by such problems. Data communication networks are typically a spaghetti of high-capacity links. These are joined together by equipment which routes traffic from one link to another, manages errors, smoothes out overloads and reroutes traffic if a link fails. These routers need to know how to forward a packet of data from X to Y: should it go via links C, D and E or via F and G? Each router thus needs a map of the network. But the network never stands still. In a large network, links are added and modified daily. Worse, links experience temporary failures and this information needs to get to each router quickly in order that traffic is not sent over a route which contains a dead link: links C and D, for example, may be fine, but final leg E may be defunct. Life becomes even more difficult when it is realized that distributing this information uses the links themselves and it may take many minutes to be distributed from router to router over a worldwide network. Not infrequently, by the time the information has found its way to all routers, the dead link is back in service and a new piece of information indicating this fact is then distributed too. So a router which is many links removed from the point of failure is bombarded with contrary information from the several routers to which it itself is connected: link E is dead; link E is live again; no - link E is dead; wrong - it is live; and so on. Link E, which is the only element in the network that knows the true state of affairs, is clearly not in a position to manage the publicity of its decease and resurrection. A large network might grind to a standstill while the routers argue among themselves about exactly what had happened.

Worse is to come. Routers are special purpose computers that, like any other computers, run programs. These programs will be updated regularly by technical experts with corrections and enhancements. This immediately raises another dichotomy:

either

- each updated version must be able to communicate flawlessly with both older and newer versions

or

- all routers are updated at once

The problem is similar to that of a new word being introduced into a language. Either everyone is told about it all at once or it must be possible for a dialogue to take place that nevertheless makes sense to a listener who has not yet come across the new word.

Data communication equipment companies such as CISCO Systems that make routers for private business networks and for Internet service providers have wrestled with such problems for years. One way of making a network more resilient to problems is to split it into areas (sub-networks) that are autonomous. But the network is then of limited use to those who want worldwide communication.

The usual solution is a compromise: create freestanding areas and then link them together at one or two points on the boundaries that separate them. The routers in area A would contain a map of the links in area A alone. Any links in another area B would be invisible from within A. All that a router in A needs to know is that any packet of data addressed to a destination somewhere in B has to be forwarded to a special router on area A's boundary. This boundary router would then take responsibility for sending it to its opposite number in B that would be fully up to date with what routes in B led where.

Some communication of network information across the areas has to occur. If not, a router in A would not know which destinations lay in B. But information about what *links* lead where in B and which ones were currently operational stays confined to B. Routers in area A will discuss link availability with each other. Routers in B will do likewise. But this will not happen between a router in A and a router in B. Major failures in one area will have limited impact on another area. Both data and the information about link availability can flow uninterrupted around A even when B is struggling.

This sounds like a perfect solution. Why not create even more areas - like the creation of the progressively smaller and more numerous cells used by the latest generation of mobile phones? Firstly, the fewer the points of interconnection between A and B the greater the dependence on the availability of the boundary routers (and the links between them) that look after all communication between A and B. What we have gained in resilience *within* each area we have lost in the connections *between* areas. Secondly, we lose the ability to diagnose faults or sub-standard performance on a route from end to end - we can only look at each area individually. Thirdly, we lose the ability to select a globally optimal route: offices that lie a short distance apart but in different areas may have traffic between them taking a lengthy diversion up to a distant boundary router and back down the other side again. Lastly, it is more expensive.

In coevolution terms, the areas are the by now familiar *objects*. Network design follows exactly the same process as was outlined in Chapter 4. The links between boundary routers give the C-coupling between areas. The (average) number of links between routers in any one area gives K. The effects of a temporary technical problem - perhaps information about a link failure - which occurs in a high-K area reverberates around the whole area in an unpredictable way. Routers repeatedly receive and pass on contradictory and perhaps very erroneous information about route availability (a process delightfully known as 'flapping' with 'poisoned' broadcasts!). If the routers in an area are connected in a hierarchy or in the extreme case a simple low-K star with each link connected directly to the boundary router, this impact of network failures is confined. But now the system has become more vulnerable to a failure at the centre of the star. As in business (see Chapter 4 - 'How big should an object be'), managing a star network is easier than managing a mesh and this is not a matter of network size.

Lessons from IT

There is more to communication between objects than sending or receiving information and requests for action. Within a business, coevolving areas - departments for example - know of each other's existence: they are both documented somewhere in a corporate family tree. Communication *between* businesses is a different matter and the IT industry has developed ways to do this securely. Both parties need to be sure:

- of each other's identity ('authentication')
- that if one party requests something or commits to something, it is allowed to do so ('authorization')
- that any allowable commitment cannot be reneged upon ('non-repudiation')
- that any messages between the two are secure from being read by third parties ('encryption')
- that any such messages have not been tampered with on the way ('integrity')

Protocols are needed for each of these. The implication for any collection of coevolving objects is that if they cannot trust one another they need to adhere to a potentially complex set of standard protocols. This may limit their freedom to change quickly, particularly as the standards themselves are evolving (a protocol is needed just to negotiate which protocols are then to be used…).

Dissemination of data, information and knowledge via person-to-person communication is usually an ad hoc and unregulated process. There is another approach - a more formal process of collecting, indexing and displaying material gleaned from corporate experts. This is knowledge management.

CHAPTER 8 - KNOWLEDGE MANAGEMENT

Data sharing

We saw in the last chapter that some data needed to be shared between objects. An object manager did not have unfettered freedom to do what he liked with the data items within his object area unless they were purely and obviously local. Even though an employee works within an object area, other people need information about him. Corporate Personnel for example need it so that the employee's career path can be managed. It should be the rule rather than the exception that operational transaction data and the more static data such as personnel records should be shared between object areas even though one object area may be the custodian. One area might be the custodian of an entire collection of data, different areas might be custodians of different data items within the collection, or different areas might be custodians of different ranges of values of particular data items.

Data are not the only things that need to be shared, however. Information and knowledge - as well as data - need:

- to flow unprompted from their source to where they are needed

or

- their source to be known to those who need to tap it ad hoc.

But what is good for the source or initiator is not necessarily good for the recipient. This is why it was stressed in the previous chapter that the recipient of any communication may or may not want to co-operate. It may be inconvenient to take a phone call there and then or he may simply not want to talk to the caller. Incoming electronic mail messages may be sent to the electronic trashcan without further consideration. But perhaps more significantly for any modern business, the person on the receiving end may want either:

- to keep his data, information or knowledge to himself

or

- to play down the fact that he holds such useful snippets at all

or quite possibly both.

Some corporate cultures are all too strong at promoting and rewarding individual competence and responsibility. In cultures where team spirit is fostered, it is common for small teams to refuse to co-operate with each other lest a competing team is seen by senior management to gain the upper hand. The result can be the creation of numerous protective islands. This may appear very close to the spirit of coevolving objects - black boxes whose internal workings and data are hidden from others and whose only linkages are via C-couplings. But knowledge and information, like the backbone of corporate operational data, have to live somewhere in one or more of the coevolving objects. Like data, some are legitimately private to an object. The rest are shared and should be easily accessible to whoever in the business needs to access them. Corporate culture may also inhibit people *asking* questions. After all if you need to ask it is reasonable to assume that you do not know the answer. Consultants in particular are prone to this: customers employ them for their specialist expertise and their admitting to anyone, customer or colleague, that there is a gap in what they know just goes against the grain of their corporate culture. Anyone ignorant of one thing is perhaps ignorant of other things as well and the seeds of doubt about their personal competence are sown. Even worse, the question may be inadvertently naive or stupid and the last thing the consultant wants is to be shown up as intellectually shaky in front of his peers or - worse - his manager.

Some corporate cultures go further and positively reward those who hoard data, information or knowledge. If year-end bonuses are given for possessing unique knowledge, for example, and there are not corresponding rewards for sharing knowledge, everyone has a strong disincentive either to share what they know or to let others be aware of their unique expertise. For example, the traditional culture of global consultancy KPMG Peat Marwick LLP was to reward its consultants individually. These rewards were not just for the amount of business they brought in but also for how well known they were internally as possessors of unusual and valuable knowledge and experience and for the unique and personal ways they managed their customer clients. When KPMG Peat Marwick started to implement its own knowledge management initiative called K-World across four countries, it identified that a sea-change was needed in how its consultants were rewarded in order to flush out and disseminate the bedrock of valuable information hitherto safeguarded all too well by individuals.

Other business cultures reward those who by nature are part of the herd and do as they are told. The creative ones brimming with ideas and bursting to tell others and take new initiatives are penalized. They are categorized as distractions, troublemakers and upsetters of the status quo. But a business without a leavening of such individuals is wasting its time appointing a Chief Knowledge Officer, setting up an internal Web site on which corporate information will be collected and with all the other paraphernalia of knowledge management.

Information and data

Numerous books have been written over the past three years about knowledge management in response to legitimate corporate concern that the growing staff mobility and the demise of 'jobs for life' has led to a huge amount of extremely valuable understanding being lost of how the business actually worked. The corporate memory of the problems a business has experienced and how these were overcome fades rapidly as the staff involved move on. Many of these knowledge management books digress at length on minute differences between data, information, knowledge and something even more nebulous called wisdom. They totally ignore something more fundamental, however: that what we will refer to here as knowledge is not, in the form it is sent or stored, of full use until it is interpreted within the recipient's brain.

The difference between data and information is somewhat grey. A list of numbers extracted from a series of manufacturing process measurements or from the business's monthly financial accounts is regarded as data because it can be selected, sorted and analyzed by computer. My annual progress assessment written by my manager is, ideally, as objective as possible and will list in some detail what objectives I was given, how my performance against them would be measured and by whom, and how I actually performed against them. But the numeric performance and promotability ratings at the end will probably be the only things that are truly data. The remainder is factual information but not readily analyzable. Data and information are not confined to text and numbers. Some pictures can be usefully analyzed: fingerprints for example are comprehensively classified and fingerprint databases can be searched automatically. Web search engines will shortly offer ways to search the Internet for stored pictures in the way we currently search for stored text. If a data pattern can be derived from a multimedia source - picture, sound, text or numbers - then a computer can search for it. In such a case, the multimedia source is a combination of data and information and trying to make a precise distinction is pointless.

Knowledge is different. The distinction made earlier between knowledge and information or data is fundamental to how people can access it.

Knowledge

If I want to find out something which I can define precisely and know where it is held, the scheduled delivery date of an order for example, I might access it myself from my business's own computer system or ask someone else to do so. But once I step apart from the operational transactions and the recorded analyses of them such as our annual financial accounts, things become less certain. I might, for example, not know exactly what question to ask. I might need some further assistance to help me formulate the right question. I might even need a lengthy dialogue to firm up exactly what I want to know. What I would *really* like is:

- ❖ to find the person who knows most about what I think I want to know
- ❖ to have a dialogue with that person until I am clear what my question is

- to receive an answer which I understand without it's being clouded by other things I think are irrelevant

The answer may well be an *explanation* rather than a recital of information. There is a big difference between the two. An explanation implies that the explainer *understands* what he is explaining. This is probably why he was consulted in the first place rather than anyone else. 'Understanding' is itself a nebulous concept, but one we find very easy to pinpoint in ourselves and others - but especially in ourselves. We are only too clear when we do not understand something despite being able to give facts and forecasts about it to others, and our lack of understanding is openly apparent to an intelligent questioner. Equally, we can understand something when we are not aware in advance of understanding it. For our purposes, understanding and knowledge are the same thing and an explanation is the passing on of knowledge.

The implications from this are significant. Firstly, if I have a dialogue with a computer system, it will have a number of limitations. Present day 'expert' or 'help' systems will be unable to assist me interactively to refine my question intelligently. Secondly, it will give me much extraneous information that I need to sift. Thirdly, it is unlikely to pace itself according to my existing level of knowledge.

A human expert or specialist who is good at clarifying questions and then providing the answers is rare. He needs a number of attributes, not all of them readily found in one individual. He must be:

- *identifiable*. I need to know who knows what in order that I might contact the right person. Some form of Expert Yellow Pages is needed supplemented by an active grapevine which can point me in the right direction: "I don't know but I know someone who does"
- *motivated*. Corporate cultures which, as we have seen, champion the achievements of solo individuals or small teams, bond dealers for example, need to reward knowledge sharing as well as measurable individual achievement
- *adaptable*. I do not know what the expert knows. I may not even have the background to ask a sensible question without some assistance. I may not speak the same jargon or even the same language as the expert and our educational backgrounds and intelligence may be very different. It is clearly in my interests to accommodate myself to him as far as I can in the various ways described for person-to-person communication in the last chapter but he also needs to accommodate himself to me.
- *helpful*. I do not want to drown in information. I want the most succinct answer possible. I want a structured dialogue that homes in on what I want to know as fast as possible. An expert who gives all callers an overwhelming brain dump to demonstrate his prowess is not being helpful, although a brain dump is sometimes necessary to give the questioner a wider panoramic view of the subject in order to help him frame his question better.

- *knowledgeable*. In addition to his own specialized knowledge, the expert also needs to have his own network of contacts to call on if the dialogue wanders outside his own realm of expertise.

Knowledge management projects have a high profile at the moment. Nevertheless, all but a few will founder because:

- the culture of the business does not foster or reward knowledge sharers
- experts with something to impart have insufficient time allocated to do so
- others are not aware who knows what and rely on their own network of contacts to help
- there is a widespread misunderstanding that knowledge can be readily and comprehensively held in multimedia form on searchable Web servers. This is generally true for information but the storage of knowledge needs some qualification.

The study of knowledge, epistemology, dates back to the Ancient Greeks. Knowledge is something that exists inside people's heads. It can be passed on to other people through face-to-face dialogue, phone, the postal service, electronic mail or any of the other electronic aids described in the previous chapter. It can even be stored electronically. But what is transmitted or stored is not overt knowledge but is overt *information* with knowledge wrapped within it that needs unwrapping by the recipient's brain. This sounds a truly weird concept and is certainly at variance with the writings of the current crop of knowledge management gurus. But it is consistent with our equating 'knowledge' with 'understanding' since only the human brain is believed to be able to understand an explanation, to search for one or to pass it on to others - for the present at least. It is best illustrated by example. If I mention 'the sun' to you, you probably have a mental picture of a spherical blob in the sky that wanders about predictably during the day giving light, warmth and shadows. If I say that it is yellow, you have a concept of 'yellow' - assuming that you are not colour blind. But your concept of yellow is far removed from a formal definition of yellow as electromagnetic radiation of a particular frequency. Yellow means something to people but what it means no one currently knows. More precisely they 'know' but cannot express or define what is means. If I went further and described a beautiful sunset, we might both share roughly the same concept of 'beautiful'. The information that the sun is yellow and that a sunset was beautiful can be transmitted, stored and retrieved - you are, after all, reading these very words now, but no computer can as yet analyze what either concept means to the sender or recipient. The meanings are in some way encapsulated in the words but need a human brain to interpret them. And the meaning is more likely to be similar to both of us if our social and educational backgrounds and intellects are similar. They provide the 'context' within which information is turned into meaning and understanding. Three things are needed: something external (the sun, a painting…), a context (our respective backgrounds) and a brain to interpret the first within the context of the second. The context is, of course, part of the brain's wiring anyway.

The same is true for, for example, music. I can send you a CD of the works of J.S. Bach, one of my favourite composers, but why the tense and haunting Adagio from his Concerto for Violin and Strings in E Major is so appealing to both of us is something outside its digital representation on the CD. It is also beyond the analyses of the musicologists who can exhaustively dissect Bach's way of writing music without coming close to being able to express precisely what his music means to them - or to me.

Communicating business knowledge

All this sounds unproven, somewhat esoteric and light-years removed from the factual and analytical world of business. If it is true, however, it means that even if corporate culture actively promotes and rewards the sharing of knowledge, the present drive to create knowledge databases on internal corporate Web sites is doomed to failure, at least in the near future. They may capture useful *information*, much of which may be searched for in the way the police search their fingerprint archives, and they may store *encapsulated knowledge* waiting to be unfolded within the searcher's brain. The subtler and arguably more useful nuances need, however, the ad-hoc dialogue to refine each question and home in on the desired answer. Computers are currently poor at this as they are weak at detecting the more subtle patterns in the questions we ask. The implication for our coevolving business - or any forward-thinking business come to that - is unsurprising: the people who staff each object need the personal attributes we have described earlier. The object managers, who by definition have considerable autonomy to do whatever is needed to further the objectives set for *their* object area, must not rely completely on computer assistance to store the business knowledge from their area. They should instead concentrate on streamlining the ways people can communicate interactively with other people. This naturally includes making people aware of who is where in the organization, and perhaps even keeping formal tabs on those who have moved elsewhere but nevertheless have something useful to contribute if asked. It also includes the creation of expertise Yellow Pages so people both inside and outside each object can easily track down those they need to communicate with.

We have at this point described the characteristics of a coevolving business and the relative autonomy that is characteristic of low-K objects. We have also highlighted and explained an apparent exception to this autonomy: the need to share data, information and knowledge *between* objects and given some technical examples of ways to do so. We need now to bring things down to earth by looking at attempts to create free markets where coevolution could take place unhindered. In the course of this we will examine a real coevolving business that was set up recently from scratch using some of the ideas already described. We will contrast this with some of the more grandiose outcomes of Margaret Thatcher's pioneering attempt to privatize UK utility corporations, in the course of which they were either split up or subjected to real external competition.

CHAPTER 9 - THE FREE MARKETEERS

Markets - internal and overt

From the point of view of an end user or consumer there are two types of market. One is an 'overt' market in which he needs to make a choice of what to buy. The other is an internal market where he himself makes no choice but where there are competitive buyer and seller relationships further back in the supply chain. The Thatcher government that was running the UK in the 1970s and 1980s often seemed to have a one-track mind. Its aim was to create a mixture of overt and internal markets through privatization, disaggregation of utility companies into suppliers and distributors and, where it made sense, offering the public a choice of supplier. Some of the resulting businesses became very successful in their own right as well as putting cost pressure on their sector to become more efficient and to 'think customer' - a new and unsettling concept to the former featherbedded utilities. Others struggled. The US, for example, has led the way in showing the benefits of a free market and in raising the profile of customer service. But in no other country has market freedom been so thoroughly driven in such a short space of time as in the UK. What follows is a look at the creation of perhaps the most complex internal market - within the UK's National Health Service (NHS). The NHS is complex because the service objectives are very difficult to define, the organization is large and labyrinthine, the costs are huge and rising and almost everyone but the very rich are customers at some time or another. Patients are not in an informed position to fully understand the technicalities of the treatments they receive and are extraordinarily vulnerable to failures of the service. If the domestic electricity supply fails, the lights go out. If surgery fails, the patient may become a former patient.

The NHS was a fully-fledged dinosaur at birth, waiting for the Thatcher government to trigger its evolution into something more efficient. Other healthcare management services have not taken this approach and have been free-market creatures from their inception.

Healthcare

The costs of providing public healthcare are the bane of any government. Demand always seems to exceed supply for schemes such as Medicare that in the US supports the disabled and the over-65s and Medicaid that supports the poor. The archetype of such schemes is the UK's fifty-year-old National Health Service that is funded through general taxation. When set up by the post-war left-wing Labour government as 'free treatment for all', it was the epitome of vertical integration and either owned or controlled most of the country's medical facilities as well as paying the salaries of doctors and nurses. The government confidently predicted that, after an initial jump in the numbers treated, the resulting widespread improvement in the health of the population as a result of prompt treatment and preventative medicine would cause costs to level off. But this never occurred. The demand for treatment continued to rocket and was always much greater than any realistic level of funding could provide. The real cost per head of population has continued to rise inexorably since then, partly as consequence of the expensive technology that is now needed for surgery and intensive care. Previously fatal illnesses are now routinely cured - at a price. This pattern is repeated worldwide. Countries that have taxation-funded and privately-funded schemes running in parallel end up with two tiers of quality. The 'free' scheme usually has waiting lists and relatively second-rate accommodation. The expensive private healthcare schemes on the other hand are tailored to the better off for whom convenience and quality are well worth paying for. The state scheme is inevitably seen as a poor relation.

There are many ways in which national healthcare can be implemented. In Canada, Australia and the UK the state manages and pays for most of the service. In the US people are free to take out health insurance - or not as they wish - with the result that some 35 million Americans are not covered by any private or national healthcare scheme at all. In much of continental Europe, health insurance is compulsory - a scheme misleadingly called 'social insurance', although in Germany and the Netherlands people have the freedom to opt out of paying health insurance contributions to the state if they wish to use a private insurance scheme.

Switzerland, which has higher per capita income and higher standards of service than most, is stuck in the same cost spiral. But one group of enterprising Swiss doctors has created a mutual organization that is being remarkably effective in containing the real cost of medical treatment. Its self-adaptive organization structure was based from the outset on the evolving ecosystem approach popularized by Michael Rothschild and aimed to achieve a balance between:

- competition - which squeezes out the inefficiencies and the inefficient
- specialization - which reduces overlap in skills and equipment by promoting niche areas of expertise
- co-operation - which avoids the overhead costs of competition where competition is not obviously beneficial

The UK National Health Service also underwent a structural sea change to an internal market economy in the early 1990s, driven by Margaret (now Baroness) Thatcher. But a subsequent switch back to a mildly left-wing government has threatened to scupper some of the freedom to shop around given to doctors by Prime Minister Thatcher.

We will examine and contrast these two examples - the small flourishing Swiss healthcare consortium and the monumental lurches to the right and the left of the UK National Health Service. We will do so using the notions of fitness landscapes and measures of complexity, together with the requirements for efficient communication and information management outlined in the previous two chapters. Finally, we will look at some of the Thatcher government's other initiatives to liberalize the markets.

IGOMED

The IGOMED medical network is an association of some ninety Swiss medical practitioners. Some are general practitioners (family doctors), some are hospital-based generalists and some are specialists ('consultants'). The medical insurance funding for their work is provided by QualiMed. QualiMed is the brand name of a special high-value-for-money policy offered by the largest Swiss health insurer Helsana who also provide a full range of other more conventional policies including the mandatory basic health insurance. Within the first two years of full operation IGOMED managed to take over the custom of some 60% of Helsana's 10,000 policyholders in IGOMED's geographical area. IGOMED's aim was to reduce the total cost of patient care by providing a more efficient service without reducing its quality in any way. Indeed, because the new organization streamlined the internal communication processes between doctors, hospitals and pharmacists, the patients receive better attention.

The forerunner of IGOMED dates back to around 1986. A group of doctors in the Thun region of the Bern canton (state) banded together to establish advanced technical training and to improve their understanding of self-regulating public health organizations. They also wanted to sort out a long-standing dissatisfaction with the quality and degree of communication between prescribing doctors and the pharmacists who made up their prescriptions. In Switzerland, medicines prescribed outside hospital are supplied to patients in several ways. In around half the cantons and in rural areas, general practices dispense their own medicines. Patients generally like this as it saves time and it gives the doctor better documentation of exactly what each patient actually received. Pharmacists, on the other hand, naturally prefer the traditional system in which the patient takes a written prescription to the pharmacist. This has the disadvantage that the prescribing doctor receives no feedback about exactly what was prescribed unless, for example, the prescription is illegible or a prescribed drug is simply not available. Personal communication between doctor and pharmacist is infrequent and by phone. Most busy doctors would not, in any case, tolerate incessant phone calls from pharmacists.

It was clear to the group of doctors in Thun that some radical overhaul was needed if the growing cost of private medical insurance was to remain affordable. There was, however, also a potential business opportunity to work closely with a major health insurer in order to take over a major share of their business through cost efficiencies.

Cost drivers

Analyses of the cost structure of patient care showed that some 50% was attributable to in-patient (hospital stay) treatment, 15% to outpatient treatment by a hospital doctor or general practitioner and 11% to prescribed medicines. Dentists took a further 8% and the remaining 16% was absorbed by preventative medicine and other costs. The group decided to focus on the costs and incidence of in-patient treatment, on improving communication processes between all of the areas, on the processes for moving patients into and out of hospital and on the links to pharmacists.

Change was not easy. Traditionally, provision of hospital and other medical resources in Switzerland is a centrally planned and heavily regulated process with price controls, tariffs, quotas and subsidies. The group wanted to get away completely from this and build a organization where:

- the total cost of patient care was visible to all participants
- participants' decisions would minimize the *total* cost of patient care, not just the 20-25% or so they directly incurred
- the inefficient links in the chain of patient care, whether doctors, pharmacists, administrative staff or processes would *continually* be replaced by more efficient ones, or be automated, or (ideally) be simply removed. For example, it was shown that unnecessary and expensive in-patient treatment could often be replaced by a cheaper and better combination of in-patient and out-patient care.

These three measures to reduce cost were not just intended to generate once-off improvements but to give a *never-ending* stimulus to drive costs down irrespective of how the organization changed in the future. Doctors may join and leave the scheme, hospitals may add new facilities, the prescribing arrangement with pharmacists may change, but the cost-reduction drivers are always in evidence and pushing in the right direction. They need no central management to impose them and need little bureaucracy to manage them. Why? Health insurer Helsana maintains up-to-date details and analyses of the costs of patient care and it is always in *its* interests to provide these to participating doctors since this results in *its* policy claim costs being reduced. This in turn enables Helsana to grab a greater market share through keener pricing of QualiMed insurance premiums. In other words, the self-adaptive nature of the operation extends from IGOMED proper to those with whom it deals such as Helsana and also with ancillary agencies such as pharmacies. Patients have little knowledge themselves of the comparative costs of care with different doctors or of the costs of various forms of hospital treatment. But they have even less understanding of the value for money of the different ways they can be treated. They are, understandably, not in a frame of mind to shop around when sick. To them it is just another insurance claim.

Hospitals and pharmacies

With in-patient treatment costing half of the total cost of patient care, the group tried to identify those cases where extended in-patient treatment was justified and those cases where care could be better provided through more out-patient treatment or through a mixture of both. And since most patients would also prefer not to stay in hospital, everyone gained. In addition, studies highlighted that if combinations of in-patient and out-patient treatment were to be provided, then some quicker and more comprehensive form of communication was needed between hospital in-patient and out-patient departments and between both and the general practitioners.

Pharmacies too were not immune from the attentions of the group. Since prescription costs are part of insurance claims, each prescribing doctor has an incentive to fulfill prescriptions from the more cost-effective pharmacies, and insurer Helsana has an incentive to provide the cost breakdowns necessary to enable this to be achieved - natural selection drivers in action again.

There were naturally also some more familiar commercial advantages such as bulk buying. Conventionally, doctors buy alone or buy as part of a family-doctor practice partnership - an arrangement that is hardly cost efficient.

Communication and information management

As IGOMED grew, it quickly became apparent that the usual face-to-face discussions between doctors and the flow of casework paper would have to change. The group is now starting to discontinue as many as possible of the administrative case meetings and to move to a higher degree of both formal and informal contact made possible through new technology. Such automation of communication and documentation processes has the additional aims of removing unnecessary communication and administrative processes and of introducing a more reliable service with fewer administrative errors. One fundamental criterion for whatever replaced the traditional paperwork and doctor-to-doctor discussions was that both medical records and mail were secure from being accessed except by those who were authorized to do so. Lotus Notes was selected for both mail and documentation. Notes had several advantages for any application that stores and distributes sensitive textual and diagrammatic information:

- it allows a high degree of selectivity about who can see or change what
- it has a secure encryption system for both mail and stored data
- it has a secure authentication system where each 'end' of a communication link must prove its identity to the satisfaction of the other end. For example, when a doctor who dials in to the network identifies himself successfully to the central Notes system, the doctor's PC also challenges the central system to prove that it is who it says it is

These security features were built into Notes from the outset and not grafted on later when security became a hot topic.

Notes is also a very accessible system. Stored information can be accessed via the Internet or by a Web browser on a private intranet using Notes' Domino facility as well as from a normal Notes 'client'. (Domino is effectively Notes itself plus a Web 'server' which run together on one computer. There is an interactive link between the two that translates requests from the Web side to Notes or from Notes clients via Notes to the Web server - hence the two-sided 'domino'.)

This is the vision. In reality, IGOMED has some way yet to go in order to be an automated organization. There is still a lot of paper, fax and ordinary mail involved.

Communication between prescribing doctor and pharmacist has also been improved but there is lots of scope for further development. Instead of writing a prescription and expecting the patient to take it to a pharmacy, most IGOMED doctors currently fax prescriptions to preferred pharmacists. Medicines thus prescribed are then either sent by mail or courier service to the patient's home or sent to the doctor's practice for collection, as the patient wishes. The use of preferred pharmacists helps control costs, and the additional prescription fulfillment choices offered to the patients gives IGOMED a competitive advantage in customer service.

Benefits

Since the insurance company has statistics from other areas, it was relatively easy to define a control group in another region of the Bern canton in order to monitor relative as well as absolute cost movements per patient over a year. In 1997, IGOMED produced a remarkable 33% reduction in the unit cost of patient care relative to the control group. In absolute terms, costs went down repeatedly in all the key areas over the first two years of full operation:

	1996-7	1997-8
General practice	-25%	-26%
Out-patient services	-31%	-53%
In-patient services	-47%	-40%
Prescriptions	-	-10%
Home nursing	-	-43%

As is evident above, savings were not merely in costs within IGOMED itself: prescription costs were down by 10% in absolute terms in 1998. As patients get older, prescription costs per head not unexpectedly rise, but IGOMED's relative cost efficiencies show through here also:

doctor-dispensed medicines (in rural areas, for example)	100%
IGOMED plus associated pharmacies	115%
conventional Swiss organization	145%

The cost base chosen is that of doctors in rural areas of the Bern canton who dispense medicines themselves because of the lack of local pharmacies. The cost-efficiency of this latter scheme is also unexpected since there is no real brake on what such doctors prescribe. They could use the freedom as carte blanche to print money. But this simply does not happen.

One other surprising result of IGOMED's approach has been improved co-operation between prescribing doctors and pharmacists: both now have a shared objective of improving value for money. In addition, the preliminary fax links between them have improved the accuracy and timeliness of the prescribing process while still preserving their individual and unique strengths.

IGOMED's insurer Helsana is delighted at the success of the scheme and their customers appear to be happy also. Helsana's own 1998 statistics show that of those in the IGOMED geographic catchment area who moved their insurance from Helsana to another insurer, 90% had been registered with general practitioners who were outside IGOMED. Of those in the area who took out Helsana contracts, 80% then registered with an IGOMED practice.

Personal touch

Automation will not do away with case discussions between doctors. This has been found to be an important element in cementing the working relationships between general practitioners, hospital doctors and specialists. But the nature of the case meetings is changing as material starts to be filed on central Notes databases. Older-style presentations of medical cases using flip charts and overhead projector slides are giving way to instant access to current information from Notes that is displayed using PC projectors.

Patient files on Notes are far easier to locate as they stay in one place as the patient moves between doctors and hospital, and the transaction flow involved in house calls and hospital appointments is being streamlined. It will be quite some time before the majority of patient files are held electronically but when this is achieved it will lead to a more efficient and reliable service for the patients themselves.

Coevolution

IGOMED doctors have profited through an improved exchange of experiences. Problems can be shared and opportunities for advanced training identified. The aim is not, however, to make all doctors alike but, following Michael Rothschild's guidelines, to capitalize on the strengths of each in their particular niches. IGOMED has been a clear and unambiguous success but it would have worked differently if each participant's objective was purely local - to reduce the costs under his control only. Taken to its extreme, this would have led to hospital doctors discharging in-patients before they were truly ready; the result would have been an escalation in out-patient and home-care costs which might have outstripped the reduction in hospital costs. This is an example of the linked objectives discussed in Chapter 6 except that an individual doctor is enjoined always to reduce the *total* unit costs of patient care and not necessarily the ones directly under his control (and with merely an eye on the others). He might, for example, prescribe an expensive drug that may avoid higher costs of in-patient treatment later. For a complex case, he needs to be able to plot the various paths the patient could potentially take between general practitioner, hospital doctor, specialist, in-patient treatment, out-patient treatment, home nursing and so on. Medicine is not an exact science mainly because of the variability of patients and capriciousness of diseases, so our decision-making doctor must try to make the best decision that he believes will reduce total cost without sacrificing the quality of patient care. Sometimes he will get it wrong and the expensive drug he prescribed does not obviate in-patient hospitalization. But, overall, his clinical experience and advice from IGOMED colleagues will lead to many more cost-reduction successes than failures.

As IGOMED expands, there is a possibility that the total cost of patient care becomes so multi-faceted and remote from an individual doctor that it is impossible for him to make a true minimum-cost decision. He will then be forced into making decisions that are locally least cost at the expense of not being *totally* least cost. When this happens, IGOMED will need to strengthen the competitive element between such decision-makers. Doctor A may make a decision that is minimal cost for *him*, perhaps prescribing an inexpensive drug, which increases costs elsewhere. The organization then needs a 'quid pro quo' feedback mechanism that acts as a disincentive for him to do so when this would result in cost escalation in another area. In NKC terms, IGOMED then breaks down from being a large single low-K object coevolving with pharmacists and other external parties into several smaller objects, shared doctors' practices for example, which are C-coupled to other objects such as 'hospital in-patients'. This C-coupling then gives the quid pro quo feedback. In 'How big should an object be' in Chapter 4, it was stressed that it is not the *size* of an object that determines whether or not it is better for the object to be split up but its internal *complexity.* For IGOMED this means one of two things:

- that each doctor's decision on how a patient should be treated next must have a well-understood effect on the total cost of treatment. For example, the costs of each leg in the chain of a sequence of treatments, from general practitioner to in-patient treatment to subsequent home nursing, must be understandable by the doctor in order for him to make a totally least-cost decision. There must thus be a process engineered into IGOMED's structure which monitors the complexity of the chains of treatment and, ideally, keeps them at a level such that the total cost picture is understandable enough for a doctor to make treatment decisions which are *totally* optimal or

- ❖ the treatment - and hence cost - must be entirely within the doctor's control. This means either that doctors' practices must take on more partners in order to cover a fuller range of treatments or that individual doctors must themselves become specialists in many more areas

IGOMED rejects the latter course as it conflicts with its charter Bionomics principle that participants must make full use of their own niche specialities and must co-operate with others with different specialities.

Where next?

IGOMED is not resting on its laurels. There is still considerable duplication of information and too much unnecessary paperwork. It is still difficult for a doctor to search for salient facts in a morass of case notes, ECGs, laboratory test results, X-rays and photographs. But treatment guidelines are now starting to be documented and a new focus made on preventative medicine. The Notes network itself, which is supplied by telecommunications service provider Swisscom, is to be extended. IGOMED want to mechanize the process that links each prescribing doctor and pharmacist such that pharmacists can be paid automatically and to fully outsource this processing. And in keeping with the underlying free market approach, IGOMED wants the resulting computer system to be open in both the technical sense of following common and non-proprietary technical standards and such that prospective participants can join or leave the organization with the minimum of formality. It remains also to be decided how all the information about any one patient should be managed. It could, for example, be gathered into one physical collection - a single 'compound' Notes document, or it could consist of a series of links (Notes 'doclinks' or Web 'hypertext links') to the prime sources of information held elsewhere.

The acid test of whether the organization's self-adaptive structure is correct is whether total costs continue to be driven down *without deliberate management action* as new areas are added to IGOMED's scope. Performance in subsequent years suggests that this is correct.

The UK National Health Service (NHS)

In the last ten years, the organization of the NHS has changed more radically than at any time in its history, and it is in the process of changing yet again - this time back in the direction it came from. As is usual with British national institutions, precedents are mediaeval and arcane, but the first real funding aimed at the indigent sick did not appear until the 1600s. It was funded from *village-level* taxation and was managed locally. In the 1830s, the principle was refined with a succession of Poor Laws: England and Wales were divided into twenty-one districts with all district managers accountable centrally to a board of three commissioners. This revised management structure cut administration costs, and the commissioners themselves were more interested in reducing direct costs than in the quality of service provided. But it was not a notably altruistic age and life was very hard for some: the poor were poor, the sick were sick and both afflictions were ordained by God. In the late nineteenth century, private health insurance schemes ('sick clubs') become common in all industrial areas and the first major government-sponsored scheme was introduced 1911: a national insurance scheme for working men. The worker paid and the State contributed rather more, but the scheme gave nothing to the unemployed. Just before the Second World War, however, the British Medical Association (the BMA is the nearest thing to a doctors' trade union with four out of five doctors being members) finally produced a series of recommendations that had been in gestation for nearly ten years. Key points were that everyone should have a general practitioner of their choice, that specialist services should also be freely provided, and that the country's overall medical service - general practitioners, specialists, hospital facilities and so on should be *centrally planned and managed*. These recommendations were put somewhat on ice for the duration of the War although discussions continued between the state and the BMA, but were dusted off immediately after the war and implemented as the National Health Service in 1948. Doctors themselves were ambivalent about the NHS. In particular they resented the idea of being state-salaried. Professionals were traditionally their own masters. In addition, any bureaucratic impediment to treating a patient as they thought fit was contrary to medical ethics. But they were also concerned for their pockets: that their wealthy and most lucrative private patients would seek free treatment and cut off their prime source of revenue, some of which was used by the more caring members of the profession to subsidize treatment of the poor. The government was not, however, to be put off and after negotiation with BMA, doctors were paid a fixed fee plus so much per registered patient, and they received compensation for losing the right to buy or sell their medical practices.

From the outset, costs rocketed as a huge pent-up demand for treatment was unleashed. Three years after its introduction, a cap was placed on NHS spending and charges were introduced for medicines, spectacles, dentures and other requisites, but the services of a general practitioner, dentist or hospital were still free. General practitioners provided local surgery and call-out services under contract but were free to see patients privately as well. A general practitioner's patients had, however, to elect to be private or NHS patients - they could not be both with any one practitioner. Hospital doctors were salaried employees of one of the fourteen Regional hospital boards in England and Wales and of similar organizations in other parts of the UK. Those at the top of the tree with the rank of consultant could undertake a mix of NHS and private work. Then as now, private medical insurance schemes in the UK only indemnified patients for visits to consultants and even these visits must be sanctioned by the patient's general practitioner. Few people insure for private visits to their general practitioner although similar dental insurance is now commonplace. Finally, public health doctors were employed directly by local authorities.

This organization remained in place for some thirty-five years and worked - after a fashion. Successive governments were appalled by the cost of the service but did nothing significant to contain it. Left-wing governments wanted democracy and joint agreement instead of management. Right-wing governments were equally afraid of confrontation with doctors and, in particular, with consultants and the senior Royal Colleges who control standards of entry to the profession. 'Professional gentlemen' were the traditional heartland of right-wing support. Medicine in this respect was matched only by the even stuffier and more gentlemanly legal profession who "wore nineteenth-century clothes in eighteenth-century buildings and spoke a dead language [Latin]". As late as the 1960s, London's Royal Hospital of Saint Bartholomew founded in 1123 was still advertising for 'medically qualified gentlemen'…

There were two fundamental problems: there was no one actually in charge with the power to change things, and there were neither pressures nor incentives on doctors and other clinical staff to work in a cost-effective way. The system was positively counter-productive. For example, any well-known consultant with a long waiting list had little incentive to reduce it by working more efficiently. The daunting size of his waiting list acted as a spur for the better off - or more desperate - of the hopefuls to consult him as private patients. Neither was there anything to force hospitals to operate within their expense budgets. Hospital budgets were inflexible: the better hospitals which became noted for their expertise and success in treating particular illnesses received no additional funding for the patients referred from other hospitals' catchment areas. The NHS had a rigid mentality that extended to its customers. Commercial businesses could revamp and relocate with little public disapproval; nationalized manufacturing industries could also do so after some heart-searching and the counting of potential lost votes by the government. But the NHS, like other state-owned services, was constrained. Its customers would never accept it, the reactionary Civil Service mentality of the non-clinical administrative management only thought in terms of protecting their own budgets, and general practitioners had no incentive either way. There was no serious attempt to buy materials or bought-in services cost-effectively in spite of the fact that the growing private hospital sector did just that. Few people in the NHS had a clue what such services should cost. NHS hospitals also provided chargeable private facilities but the true costs of these tended to be masked by the larger NHS overheads. Then as now, the purely private hospitals concentrated on planned surgery rather than chronic illnesses or emergencies, and were seen by their customers whose subscriptions were often paid for by their employers as a way to jump the NHS waiting list. A well-regulated flow of discriminating customers, a strong profit motive and preferred status with particular private health insurers made a big difference.

Then, in 1983, came Margaret Thatcher. Her proposed changes were based on an astute study of medical services in the UK by Alain Enthoven of Stanford University that was published in 1985. It recommended an internal market in which specialist and hospital services would be bought from whoever gave best value. General practitioners - the larger practices at least - would have fixed cost-per-head budgets. Consultants also would be shifted from being paid for what they did on to a fixed cost-per-head system. There was a further control: patients needed referrals from general practitioners who were trusted not to refer patients for consultation and treatment unnecessarily.

In an annex to his report, Enthoven drew illuminating parallels with his experience of US health maintenance organizations (HMOs). These act as intermediaries to contract the services of specific groups of doctors and hospitals to groups of employees (via their employers) and sometimes direct to individuals. Unlike the UK health insurers at that time, these organizations had the power to improve the efficiency of what is provided by enabling the sharing of patient records between general practitioners and hospital specialists. This enabled major cost reductions through reducing by up to 40% the incidence and duration of in-patient stays - as confirmed by IGOMED's experience with QualiMed. The underlying principles of the US HMOs date back seventy years. The HMO Enthoven knew best, the venerable Kaiser Permanente, had its origins in the early 1930s when a small hospital set up to treat workers building the Los Angeles aqueduct ran into trouble getting paid by the workers' health insurance companies. What turned the hospital into the forerunner of all HMOs (and rescued the hospital) was that the insurance companies were persuaded to pay the hospital a fixed fee per head in advance. Workers also had a choice: five cents per day for work related problems and an additional five cents per day to cover the rest. This pay-in-advance per head principle was successfully replicated for Kaiser Industries' Grand Coulee dam construction workers and later, at the start of the Second War, in Kaiser Shipyards which had big wartime orders for building Liberty ships. The resulting organization is still with us: as the non-profit Kaiser Foundation which looks after health insurance and hospitals and the commercially-structured Permanente medical group which supplies practitioner services. HMOs do have some disadvantages. Unlike conventional health insurance, where there may be comparative freedom to choose a general practitioner, specialist or hospital, HMOs are something of a straitjacket: if the doctor you want to see is not a member of your HMO, you will foot the bill yourself. On the other hand, HMOs bring down the cost of premiums through better commercial deals and through efficiencies that either have to be engineered by an umbrella organization (the HMO) or built in to the suppliers' structure (like IGOMED). It is possible, as IGOMED and QualiMed together found out, to release greater economies if both parties use the same drivers to reduce cost while maintaining quality of service.

The Thatcher government's restructuring of the NHS started in 1984 when the first professional regional managers were appointed, only one of whom had clinical experience. This move was in response to a study commissioned by the government from four eminent businessmen whose report memorably damned the lack of accountability within the NHS: " ... if Florence Nightingale were carrying her lamp through the corridors of the NHS today she would almost certainly be searching for the people in charge...". A raft of management recruitment followed and the government was accused by both patients and doctors of spending money on administration instead of patient care. There was some truth in this: as managers were being appointed, hospitals were being shut or wards mothballed because of a lack of funding, although some were closed when managers found they were surplus to requirements. The absolute level of funding was also an issue: in 1988 the UK's health spending per head was only 44% of that spent by Switzerland and 58% of its nearest neighbour France, and the U.K. had only half as many surgeons as France. The UK employed fewer doctors per head of population that anyone else in the EU except Italy and spent a smaller fraction of its GDP on healthcare than any other European nations except Denmark and Greece. The NHS was creaking at the seams with underpaid and overworked staff, waiting lists to go on waiting lists, run-down hospital facilities and new high-technology ones which it could not afford to operate.

The 1984 changes created a centralized hierarchy of professional managers in place of the previous collegiate organization where consensus decisions were taken by medical staff. If this change disturbed the medical profession, the radically new structure proposed in 1989 rocked it to its very foundations. Within the new structure:

- ❖ the larger hospitals could opt out from central control, could have management boards appointed to run them and were encouraged to market their most capital-intensive diagnostic and clinical services to private hospitals.

- ❖ groups of general practitioners with 11,000 or more people on their registers (twice the national average) would be allowed to become budget holders. This limit was reduced substantially year by year as the new structure was rolled out. These practitioners had budget caps for hospital referrals and drugs, and drug budgets based on the cost of generic drugs rather than the more expensive brand-name equivalents. There was thus an obvious disincentive to prescribing highly efficacious but even more highly expensive drugs.

- ❖ hospitals were to be charged interest on the asset cost of buildings to give them an incentive to sell off redundant accommodation, and competition between hospitals was encouraged (with some central intervention rights to stop monopolies being exploited). This alone led to an explosion in the number of plant accountants employed by hospital management. It was exacerbated by the need for financial and commercial accountants to manage the new HMO-type service contracts. Overall administration costs went up from 5% to around 12% of the total NHS budget.

With these changes in place, the plan was that an internal market would emerge where the three types of buyers:

- general practitioner groups
- private practice
- local health authorities

would buy from three types of suppliers:

- independently-managed state hospitals ('NHS trusts')
- private hospitals
- those state hospitals which continued to be managed by the local health authorities

plus other service providers such as chiropodists, diagnostic laboratories and so on.

State hospitals would receive funds to cover their overheads and some essential services but would otherwise have to compete with each other for business from general practitioners.

A new government-led Policy Board set policy and an Executive Board was set up to implement it. The latter was, in particular, responsible for general practitioner services and had the sensible aim of implementing better links between them and hospitals to reduce the length of in-patient stays - as happened with the US HMOs. Performance-related pay was to be introduced for middle management. The self-governing hospitals would be allowed to set their own pay rates and a free-for-all fight between hospitals to attract and retain the best senior staff was indirectly encouraged. Hitherto, consultants for example were paid the same basic rate wherever they worked. Their total incomes might be somewhat different through additional merit pay and private practice but there was a traditional principle of equality that ensured that the best did not all gravitate to the more glamorous teaching hospitals in London. One final business practice was added: internal audits of both clinical and administrative quality and efficiency.

Funding was to be distributed based on both consumer demographics - age and health statistics - and on the relative costs of providing services in different areas. Emergency admissions apart, hospitals could charge extra for patients referred from other regions. There were some dangers and not all of them were foreseen. A super-league of hospitals could develop. The gap between the good and second-rate could widen, mirroring what was already underway in soccer clubs where the top few clubs had become extremely profitable and the rest struggled. There had always been a 'premier league' consisting of the famous London teaching hospitals that were well cushioned by comfortable endowments and were accustomed to receive a more than generous slice of the health funding cake, but this was at least under government control. The new free market could drive out of business those hospitals which were unpopular with larger general practitioner groups who could now spend their budget with whom they wished, a freedom which in 1993 was extended to buying services from dieticians, chiropodists and home nurses. The individual and the smaller general practitioner groups on the other hand, those without funds and budgets, would lose out as they had to send patients to whichever hospital the local health authority (in their new HMO) role had contracted with. General practitioners were even to be allowed to advertise in order to stimulate competition between them. This change was as symbolic as it was practical. Practising doctors had hitherto been required to shun personal publicity: newspaper columns written by doctors were published under pseudonyms, and radio or television appearances were simply by 'a doctor'.

The new organization was some three years in gestation and partial rollout, and was eventually fully launched in 1991. The HMO-type contractual arrangements between hospitals and the buyers - health authorities and the larger doctors' practices - were one of the first hurdles. Some contracts were 'cost-per-case', some were annual block sums of money, and some were block sums up to a certain number of cases and then cost-per-case on the excess. An interesting and unforeseen result of the need to undertake hospital audits emerged: hospitals that were self-governing were understandably very reluctant to share their detailed cost analyses with competing hospitals and, in particular, those hospitals that remained managed by the district health authorities. And much of the data needed to set budgets for general practitioners, such as their use of pathology services from hospital laboratories, was scant, impossible to analyze or simply unavailable.

Incentives were there in theory: general practitioners with budget-holding practices who saved money on drug prescribing could reinvest it, and even those who were not fund-holders were allowed to reinvest 50% of their notional saving.

Not all hospitals of a suitable size wanted to become self-governing. Differences emerged between the clinical staff who often preferred the status quo and the senior administrative staff who could see more power coming their way if they took the bait. Hospitals were not entirely free to decide either. The government exerted subtle funding pressure on those eligible hospitals that were apparently quite unable to appreciate fully the amazing advantages the government was offering them...

What went wrong?

It is simplistic to look for a single cause of the problems that ensued. Some of the upheaval was a natural result of creating a freer market and was not 'wrong', although the extent of the upheaval was not foreseen (and, as is the nature of free markets, could not be foreseen).

Change management:
The right-wing Thatcher government made the same mistake as Presidents Gorbachev and Yeltsin: they ignored Adam Smiths' warnings. Cultural change - in this case shifting the entrenched attitudes of senior hospital clinical staff - had to precede structural change, and that to liberate the efficiencies inherent in a free market they first needed to dismantle the administrative machinery set up to manage the old regime.

Free market side effects:
The government was so convinced it was right and that it understood the likely outcome of what was planned that there were no trials. The internal market was simply rolled out. It forgot, or perhaps no one had told it, that creating a free market has unpredictable side effects.

Systems and information:
We saw earlier the importance of good communication, processing systems, operational data and information to monitor progress. But there were no accurate billing systems for hospital services at the launch in 1991. This was not merely a computer systems problem: it was doubtful if the necessary systems could have been created given the uncertainty about how billing was to work. Proposals ranged from billing every case at cost, through to the US practice of billing on the average cost of a typical case of a particular illness.

Commercial objectives:
The essence of a free market is a running series of contracts between buyers and sellers. But for the launch of the restructured NHS, contracts and performance measures were largely drawn up and monitored by accountants and administrative staff with little assistance from clinical staff. This inevitably led to contracts that defined services of the wrong quality or which were clinically out of date and to measurement criteria that could be legitimately if deviously exploited by the hospitals. For example, criteria were set centrally (under both the Tory and successor Labour governments) for the length of waiting lists and for how long a patient would wait to see a member of staff after being admitted for accidents and emergencies. Hospitals reacted as 'good' free marketeers should by having a nurse see each patient as soon as he set foot through the door but then allowing the patient to wait perhaps for several hours before seeing a doctor who could sort him out. Setting measurement criteria in any field is like drafting tax laws: anything too general awaits for some defining precedent; anything too specific tries to include every eventuality and becomes unmanageably complex as well as creating loopholes which are exploited by the ingenious.

Financial:
Budgets were unbalanced. To reduce the length of in-patient stays, there needed to be additional funding for outpatient and local care, and the two sets of budgets needed to be in alignment or else patients could not be discharged

from hospital. Hospitals were only allowed to borrow money for capital spending subject to a government-imposed cap and were not even entirely free to reinvest elsewhere whatever they saved. Tariff anomalies surfaced immediately after the launch of the internal market in 1991. Hospitals' published tariffs showed wide variations in their costs for the same surgical treatment. Budgets for similar budget-holding general practices also showed wide variation, based as they were on historical spending of those practices. But many hospitals were simply unable to provide tariffs because they had insufficient time, money, staff and historic costs to prepare them.

The results

Start-up problems apart, it was expected that the scheme would make major shifts in how the larger general practices used hospitals and other services for which they held budgets. Surveys undertaken in 1997, when the scheme was well bedded down, showed otherwise. General practitioners simply did not make major changes in their choice of hospitals. Admissions to hospitals certainly increased but the polarization into super-league and second-best hospitals did not occur, at least not as a result of unlocking the market. Instead, doctors used their new power to demand changes in the services provided by their *existing* hospitals and did not shop around. Private hospitals also received more referrals paid for by general practice budgets, but the shift was minor. There were, inevitably, huge differences in the ways different budget-holding practices reacted.

Politics

With the change of government in 1997, it was widely expected that the Thatcher government scheme would immediately be dismantled. True Socialism is, after all, anti-capitalist and hence anti-free-market. The new Labour government was elected on a pledge to increase NHS funding in real terms by £1B. This never happened. The free market *had* reshaped the service - slowly - and instead of providing lots more money the new Labour government promised to cut administration and redirect the resulting savings into clinical healthcare. Part of the administration savings was to come from automation of manual processes. For example, by the end of 1999 the Labour government intended that all general practices that had a PC with dial-up facilities would be able to receive a proportion of hospital test results electronically, and by 2002 this was to be extended to all general practices. In 1998, the government announced that they would spend £1B over seven years to make *all* information in the NHS available electronically. Their target is that by 2005, electronic patient records will exist for every patient and will, in encrypted form, be accessible remotely and be moved around electronically. The aim, say the government is "integrated information systems for an integrated healthcare system".

The internal market has in theory been dismantled. General practitioners still hold budgets but in larger groups. All they have lost is their individual freedom to shop around as they chose - a freedom which as we have seen they chose in general to ignore. Sourcing of hospital facilities for general practitioners in any area of between 50,000 and 150,000 people is made by these groups of doctors with input from other local related healthcare organizations. This is, however, no different from what emerged naturally within the free market where groups of general practices within urban areas banded together to exert their collective influence on local hospitals. Neither was the introduction of private capital abolished. The ‘private finance initiative’ that enabled a hospital to be built through commercial funding and leased to the NHS is actively promoted. Such a hospital operates a bit like an HMO within an HMO and employs all the staff in the hospital except for the clinical staff. The changes forced on hospitals by the internal market were clearly too effective to ignore.

Coevolution

When the NHS was created, it was modelled on a combination of traditional UK Civil Service bureaucracy and Communism. Such an organization typically works effectively during wartime when a common cause overrides individual aspirations, but not otherwise. Detailed planning and control was done from the centre. In coevolution terms, a single large and apparently simple (low-K) object would represent this. Why *apparently* simple? Because - Civil Service style - it was in theory a hierarchy where geographic and other areas did not overlap. That was the theory. In practice, the areas were adjacent and disparities of funding between areas became all too apparent. Patients were not wedded to one funding region either and movement for specialist treatment such as radiotherapy crossed boundaries. Perhaps more relevant is that the NHS was under overwhelmingly strong government control. Parliament is composed of representatives from all areas of the country whose loyalties are compromised by the need to be elected by their constituents. These will almost all be NHS patients and want better services for their locality. But their parliamentary representatives also need to follow their political party’s line. The web of connections between the government and general practitioners, hospital clinicians, local healthcare administrators, the central NHS planners and the rest becomes complex. The hierarchy was actually a spaghetti-like mesh. In addition, those working within the NHS had no objectives that would make the NHS progressively more efficient (doing the same things cheaper) or effective (improving quality). This meant that unless the centre got it right all of the time, inefficiency and inadequate service would proliferate. Until the early 1990s, the NHS was a paper-based organization. Communication up and down the system from central planner to general practitioner was inevitably slow. In a static environment, this is not of great consequence. But since the NHS was under continual pressure with developments in clinical treatment, technology and particularly patient demographics - as well as government cost-cutting measures, the centrally run organization had decay built into its inception as it was unable to react quickly enough. A national postal service can run reasonably effectively (if not efficiently) as a hierarchy. Addressees have one address. The objectives are simple and measurable. Having a single national rate to deliver a letter and a single standard of service for (almost) all may not be equitable for the majority who live in urban areas but it makes costing the service and budgeting relatively simple. A national health service is very different. Any large mesh-connected object has high-K complexity. Any high-K object whose decision genes - all of them except those representing the central planners - have no 'knowledge' of the total fitness of the object (compare this with IGOMED) can, as we have seen earlier, only stay efficient if the central team has:

- ❖ swift communication up and down the organization - faster certainly that the rate at which the local demographics are changing and quick enough to respond to sporadic epidemics
- ❖ detailed and useful effectiveness measures (waiting list sizes and so on)
- ❖ an organization beneath them which will respond to mandates to change

The NHS planners had none of these.

There is an alternative: to split the mesh object into a set of coevolving objects. This was Margaret Thatcher's internal market. But why was this only partly successful? Some of the more practical reasons have already been detailed. But it was not an entirely free market anyway and was restricted by the political compromises needed to cushion some of the more painful rebalancing of resources which happens when any real-life market is freed. In theory, each general practice was an object. In practice, urban practices formed collectives to augment their individual impact via C-coupling to hospital management. One 'large' C-coupled link will distort the fitness landscape of a hospital - whether for lower cost or better service - more effectively than several smaller ones. These will distort it in a less co-ordinated way at different times. The formation of collectives implies that the practices had jointly decided (although not in these terms…) that the more effective impact of creating a larger object with 'larger' C-coupling outweighed the additional complexity (which is equivalent to a reduction in flexibility) that this would create between previously autonomous practices. It was not just a way to simplify administration and commercial negotiation. This exemplifies a point made in Chapter 4: having low K is not necessarily optimal; what matters is balancing K and C. As K rises when practices combine, C does not. There are only a limited number of local hospitals to contract with, a limited number of services whose cost and quality are to be negotiated and a limited number of general practices. There comes a point where the difficulty of reaching agreement within the collection of practices outweighs the positive effect of the joint commercial pressure on hospital services. The implication is that within urban areas where there are many practices and several hospitals to source services from in a confined area, practices would form substantial collectives. In rural areas collectives would be small or non-existent. This is exactly what has been observed in surveys.

Communication and common requirements also matter. Urban practices in close proximity face very similar problems: inadequate nutrition in poor inner-city areas, for example. Doctors from the different practices can meet easily to agree joint action with hospitals to address them. This again does not happen in rural areas.

Objects can combine into a larger object (or be considered as a larger single object) in two ways. One way is simply the result of splitting down a coevolving object into lower-level objects as described in Chapter 4. In this instance, the lower-level objects may be very different in nature and might even be entirely competitive. The objects may, on the other hand, be groups of similar objects that have banded together to gain more leverage over one or more other objects. The combination may be a total merger or simply a co-ordination of effort. The result is interesting in terms of coevolution. Assume that three budget-holding general practices in the NHS co-operate in order to buy services jointly to get better deals from their providers who are mostly the hospitals. From the point of view of their customers - the patients registered with them - they remain separate. Budget management may or may not be shared but since the practices are separate, budgets are still allocated separately by central government and each practice remains responsible for its portion.

Assume for the time being we treat the objects as separate rather than as a single larger object. Assume also that the practices previously operated entirely separately. Combining purchasing then adds C-coupling between each practice and the partner practices. However, it is also safe to assume that the hospitals that were used beforehand by each practice were more or less the same ones, although the practices may have used each to a different relative degree. For example, one practice may have referred 60% of its patients to local hospital A and 40% to local hospital B. One of the other practices might have done the converse. So, in reality, each practice was not independent but was linked indirectly via the common hospitals. Practice X which wanted a low price from hospital A for high-volume non-urgent ('cold') surgery might have been thwarted by practice Y which wanted a more responsive accident and emergency service. When practices co-operate, they add C-coupling between them until the benefits of additional leverage on hospitals is outweighed by disagreements between the practices. If they are all of such like mind that these C-couplings is always co-operative, then the couplings are having little effect: if practice X always thinks the same about service purchasing as practice Y, neither deforms each other's landscape. These C-couplings may continue to exist in the form of a joint purchasing committee or similar, but they are 'weak' couplings.

If we now view the group of practices as one larger object, this is why the added K-complexity does not spoil its responsiveness. The C-couplings to the preferred hospitals on the other hand are strengthened and are competitive.

When the practices decide on preferred suppliers, they may elect to give most business to one hospital. The C-couplings that formerly existed between the practices and the other hospitals become insignificant, while the C-couplings to the preferred hospital are strengthened. Each practice may now be giving more business to the preferred hospital, potentially acquiring volume discounts on its own, and the group of practices can, in effect, add the weight - landscape deformation impact - of all the C-couplings together. This is all common sense but it leads to the principle that in such a situation C-couplings are 'additive' as a minimum. If, when the practices were separate, practice X pushes hospital A to drop its costs for a particular surgical procedure by 10% and practice Y does the same *but not at the same time*, the hospital may find different ways to make the economies demanded by each practice. It has time to react to the first demand before responding to the second. Its C-coupling back to practice X may result, for example, in an increase in costs for practice X elsewhere in its budget, like the boxer in Chapter 1 riding a punch and coming forward again. But when practices combine their C-couplings, the result is similar to the effect on a boxer being hit by several punches *at the same time and in the same place*. Merely adding C-couplings together may well understate the resulting deformation of the recipient's landscape because the couplings now act in a coordinated way and make the same demands, volume discount for example, at the same time. The co-ordination comes via the C-couplings *between* the practices.

Other internal and overt market initiatives

The Thatcher government was besotted by the idea of a free market. The mere privatization of telecommunications, gas, electricity and water industries was a start but did not of itself create a buyers market. It did, however, exchange overriding government intervention for regulatory authorities whose main aim was to prevent the customers being ripped off until true competition supervened. Overt market competition in the telecommunications area was initiated by bringing in a second licensed competitor to British Telecom: Mercury. But since BT owned - and with the exception of domestic fibre-optic cable links still owns - the link between house and local exchange, competition in the domestic market had little impact. Mercury did, however, cream off some of the profitable business data and voice links but had little incentive to significantly underprice BT. Conventional telecommunications does not, however, sensibly split into suppliers and distributors or any other way, and BT currently has regulatory restraints on the added-value services it would love to provide. What created true competition was the explosion in mobile phone use and the government's willingness to license several competitors and simplify the planning restrictions on the siting of static aerials. Unlike the US that has a bewildering variety of US-specific technical standards for mobile phones, the UK and most other countries have concentrated on two versions of GSM - a digital phone standard that allows data and text messages as well as voice to be carried. Since there are many network service providers and GSM supports 'roaming' between them, users have a wide choice of alternative phones and competing service providers in most of the countries they are likely to visit, while still receiving a single consolidated bill from their 'home' service provider or other billing agent.

In contrast with BT, the splitting of the electricity and gas monopolies into privatized generators and distributors was clearly feasible. It created an overt market where services could be bought from one of several generators since there are single specifications for gas and for electricity. Electricity companies soon offered to supply gas and gas suppliers offered to supply electricity. The distributors of either could use their cross-country wayleaves to lay fibre-optic cable and offer cheap data communication services.

Splitting the railway network into infrastructure and service suppliers was less successful because any one journey may involve the services of two or more service providers. The competition is certainly there to a limited extent - each long distance route tends to be owned by one service supplier - but the customers are confused. In addition, there is a plethora of different services purchasable for any one leg of the journey: different seating; with or without free snacks; with or without validity for peak-time travel and so on - options which are different for each supplier. It is not at all obvious to the potential customer which selection of route segments and suppliers is optimal in cost and journey time. Choice is good, but customers are turned off by too much choice that needs to be made too frequently and without prior contemplation. It was noted earlier that IGOMED would have the same problem if competing sequences of treatments became too complex for a doctor to select the optimal one.

The Thatcher government also gave the BBC a nudge to bring itself out of the Stone Age. Since its creation, the BBC has been a notoriously reactionary institution (famously, BBC *radio* newsreaders were compelled to wear dinner jackets while reading the evening news). Since it was funded from television viewers' licence fees (and formerly radio licence fees), there was no need to follow commercial practices. There was also no need to compete with commercial television that is funded in the UK from commercial breaks within and between programmes and latterly from sponsored programmes (as opposed to 'messages from our sponsor' incorporated within programmes). During the 1990s, the then Director General (CEO) John Birt implemented a fundamental change: he split the programme producers from the resources needed to produce programmes. This 'producers choice' meant that internal suppliers - design, scenery, casting, photocopying, research services, libraries, archives and so on - set tariffs for their services. These tariffs were (or were supposed to be) the full price of the service including immediate overheads and intended to be broadly similar to average external market prices. Producers on constrained budgets were generally free to shop around, not just at the internal 'shops' but outside as well. Not surprisingly, many small outside suppliers could undercut the BBC's homegrown resources. But in many people's eyes, Birt went over the top. There comes a point where too many internal cost transfers for trivial amounts are counterproductive unless the internal administrative processes to make them happen are slick and efficient. What upset employees more was the overwhelming amount of external business consultancy, business practices and management jargon that suddenly appeared. This might have been justifiable if effectiveness, however measured, went up, but the aim seemed to be simply to remove waste and lower the cost of running the service. The danger of doing so - of concentrating on efficiency objectives and ignoring effectiveness objectives - has already been discussed in Chapter 6. But the BBC has a more difficult problem than commercial businesses, most utilities and even the health service. This problem is to define what its output should be. It would be easy to follow the example of commercial television and aim for a mass market and high programme audit ratings. It would also be relatively easy, if very expensive, to produce a run of high-quality costume dramas of which the BBC is justifiably proud. The right mix is elusive. So the BBC ended up in a bind that by now should be familiar: an internal market driving efficiency objectives only. There were no measurable objectives for the corporation as a whole with which any two groups of viewers and listeners would agree. 'Producers choice' was aimed at cost reduction. Effectiveness of any production was, as always, in the hands of the producer, but even with the massive importation of external management practices, no one could ever be clear whether the sum total of all this effectiveness met the corporation's objectives. Producing the right product mix in a factory is relatively easy if good cost accounting *and* product profitability systems are in place. The BBC tried to create the former as a by-product of the free internal market but never even got to first base on the latter.

Success and failure

What characterized the successful market liberalizations?

- having a country-wide standard product (the customer's choice can then be made on price; he only finds out about inaccurate billing when it is too late!)

- having a product with options and variants but whose facilities can be contemplated at leisure before purchase. Mobile phones for example have differences in the facilities available (roaming; multi-band support; data;...) but throughout Europe and elsewhere (apart from the US) technical standards (for GSM, for example) are the same.

- having effective competition. Mobile phone service supply is, for example, provided to customers through competing intermediaries. A phone of brand X can be bought from shop Y with billing provided by supplier Z and with calls made through several different network providers in many countries during the course of the billing period.

- easy supplier selection. In spite of the variety of ways a customer can contract for mobile phone service, he enters only two contracts: one to buy the phone and one to use a particular billing agent. The phones themselves may automatically offer as many as possible of the network providers that are on air in the vicinity of the phone and have cross-charging arrangements with the billing supplier. All locally available services say "use me" and the customer either selects a default or perhaps make a choice of one he knows has a strong signal. In other words, where customers need to make a complex choice (the phone purchase and service contracts), this must be done infrequently. But where frequent choices (a choice of network provider every call) are needed, the choice must be easy to make. The same is true for IGOMED's healthcare services. Although private health insurance of some kind is mandatory for most people in Switzerland, an intending customer can examine various insurance and managed healthcare proposals at leisure. When one is selected, the customer will stay with the insurer for a period of one or more years. But the customer needs little input into any decisions about individual treatments. This is the doctor's professional job. The customer - now a patient - merely needs to agree where and when to be treated.

It remains to describe how the business coevolution principles described earlier and their successful use by organizations such as IGOMED (and faltering use by the UK National Health Service) can by summarized into a plan of action. A free market between businesses or within a business is not necessarily a *critical* coevolving system as one or more parts may be stuck on hills of low fitness surrounded by deep valleys on a rugged landscape, or in an evolutionarily stable state, or both. A free market is to be desired. But within a free market, businesses want to be fitter than their competitors. It would be premature to write a detailed fitness plan for critically coevolving businesses, but we should at least be in a position to provide a health checklist.

CHAPTER 10 - IMPLEMENTATION

Checklist

This book does not provide a detailed pharmacopoeia for company doctors. Neither does it provide a proven step-by-step prescription for creating a coevolving business. Not enough is known yet on how to write either of these. What hopefully it *does* provide is insight into an entirely different way to structure organizations and to manage some of their activities such as advertising. The foregoing chapters have lots of meat that may be difficult to digest immediately and very little fat. Any summary of how to go about things - to restructure an organization for example - is thus bound to fall short. What follows is more of an informal checklist of things to refer back to in the preceding text.

1. Are you trying primarily to understand how your business competes with others or how to reallocate decision-making power and autonomy within your business? This is the starting point for deciding what coevolves with what. Do you want to understand better the attacks from competition through treating your *whole* business as an object or to give your business the right balance of agility, resilience and stability through treating the *several major parts* of the business as coevolving objects?

2. Define clearly what K and C mean in the context of your business. Remember that objects interact - disturb each other - through C-coupling. K-complexity merely indicates what happens to an object's fitness (does it go up or down - and how steeply) if it were to change in some way through making business decisions, for example.

3. Try to define values for K, C and S; you can always refine them later. Do not assume that there are single values for K and C that are optimal ('critical') throughout the organization. If you choose to let the system tune itself through self-organization, K will be selected and refined automatically. If each department of 100 employees were defined as an object, each decision-making point within the department (which could be an individual) could be associated with one or more decision genes. A senior manager might, for example, authorize capital investment *and* the introduction of new products. The extent of linkage between decision making points - how many people need to approve an investment proposal for example - is a starting value for K, and the average number of objects which are mutually coevolving is S.

4. Build a computer model with the right structure and right starting values of K and C. The hardest part is deciding what fitness values to assign to each of the N decision genes. Recall that each gene is coupled to K of the N other

genes and to C genes in each of the S coevolving objects. You could assign fitnesses randomly.

5. Decide whether you want to tune the organization manually or let it self-organize. For the modelling phase, self-organization is the best choice providing that the value of K is not very much larger that C x S.

6. Run the model until the objects attain stability - perhaps the result of evolutionarily stable strategies being found. Note the value of K. Repeat with different (random) fitness values for each decision-gene combination. With luck, the values of K at the point when the model reaches stability will be similar. This indicates how much internal complexity (average number of authorizations for a decision, for example) will need to be changed in order to bring the collection of objects that makes up the business to the boundary between order and chaos. Then disturb the model using a W parameter or change C slightly. See how much K changes in response and how long the system takes to settle down again. Ideally, K will oscillate up and down and then return to its equilibrium value.

7. Examine the data from the Boston econophysicists on company growth and growth variation.. This will indicate the likely reasonable limits on how much a company of a certain size is likely to be able to increase its growth. Let this be the justification for reallocating decision-making powers and autonomy.

8. Structure your real business into objects along the same lines as the model. Change your business's internal complexity to match the equilibrium (perhaps optimal) value of K. Redefine the objectives of the most senior managers of each object to include managing internal K-complexity and (where possible) the C-coupling to other objects. It is not unlikely that a little training will be needed ...

9. Track object performance using the process view of Activity Based Costing. Use coupled efficiency and effectiveness objectives for employees in each object. Make sure objectives are parochial to the object where possible.

10. Develop business-planning models that coevolve with the competition. Realistically, no one is sure yet exactly how to do this. You might be the first!

11. Manage advertising, trade and consumer promotional spend with the realization that their effectiveness - at least the short-term effectiveness - is relative to that of the competition and not absolute. This applies in principle to *everything* you have broadly in common with your competition including price, product quality and standards of customer service - unless, of course, you have a monopoly or are party to a price or distribution cartel.

12. Reassess the initial choice of object size (i.e. the choice of level in the hierarchy at which we define individual objects), re-model and re-implement.

13. Manage the effectiveness of (real) product test marketing by removing variants with low take-up quicker than new variants are injected into the market. Make sure that there is sufficient differentiation between variants to make customer preferences noticeable.

14. Speed up the real business planning cycle by critically examining sources of consumer and trade information and removing sources of excessive conflict (disagreements between data).

15. Speed up the way in which your business operates internally. A low-K business will make decisions quickly but the information that is needed to make the decisions still needs to get to the decision-making point and the resulting decision still needs to be disseminated. Common data, enterprise-wise operational systems with screen presentation and dialogues which can be tailored by the user, all-pervasive electronic mail and videoconferencing and adjuncts such as single corporate directories all tauten the person-to-person network. An organization with efficient decision-making but appalling internal communication can still be very unreactive.

16. Speed up the way a decision point can make a decision. In a low-K business the unnecessary barriers that will stop or delay a decision will have been removed. In a business with good communication all relevant information will have been made available. The decision-maker may still, however, need access to knowledge residing elsewhere in the business. If the decision-maker is a person, he might need to use someone who has 'been through it before' as a sounding board.

17. If things start to stagnate in the real business, with or without the objects' reaching an evolutionarily stable state, inject controlled disturbance. In modelling, this is achieved by introducing the W external disturbance parameter or by changing C. In the real business, it probably entails reassignment of managers or bringing in fresh blood.

18. If your business is not diversified, do you have well-understood Mission and Vision statements for the business as a whole. In other words, do your employees know what the business wants to do apart from produce good share dividends and grow market capitalization? If your business is large and diversified, do you have a statement of corporate principles or ethics that is widely disseminated?

These steps may by now look relatively straightforward, but building and experimenting with a non-trivial NKC model of this nature is surprisingly difficult. The structure of the model is relatively simple; the difficulties lie in the sheer size of the model, the time taken to run it, and the interpretation of what has taken place. An alternative is to bypass the computer modelling in steps 4 and 6 and to experiment on part of the real business using a target K of C x S as a guide. If the business is being decomplexed, the change manager will watch for the low-K signs of incipient criticality such as instability, over-swift reactions and avalanches of change. If on the other hand the business is being made more (higher-K) complex, and this is probably unusual, he will watch for greater sluggishness and inability to progress smoothly toward objectives without continual unexpected problems and unwanted side effects which are the results of additional and conflicting decision points.

A current example

A look at a running skirmish between the world's largest computer software business and a diffuse coevolving enemy illustrates what the future might hold. Microsoft Corporation is a huge centrally run business that dominates the personal computer software market. Threats to its dominance are met by a combination of marketing brilliance, guile and sheer size. It pays only lip service to common technical standards because the heartland of its revenue is proprietary software. The rise of Web services on the Internet caught Microsoft flat-footed and, for a time, Netscape Communications Corporation's browser package was the tool everyone used to access Web services. Microsoft's volte-face was stunningly and awesomely swift. Not only did it rapidly develop and give away a comparable browser but it went on to embed it in the PC's operating system as the standard way for anyone to access anything on the PC. Netscape, now part of America Online, protested at this abuse by Microsoft of its market position, and court cases in the US on this and other antitrust issues are continuing. The results are anyone's guess. Precedents are for a consent decree that severely limits competitive activities (such as the one IBM signed in 1955) or enforced breaking up as happened to Standard Oil and AT&T. Microsoft did in fact sign a minor consent decree in 1994 that imposed curbs on some of its business practices, but in 1997 the US Department of Justice took the view that Microsoft had deliberately violated the decree. Its 1998 antitrust suit was the result.

Splitting Microsoft into three parts - as has been proposed by the US Department of Justice - could backfire. The three resulting 'Baby Bills', named after Chairman Bill Gates and the 'Baby Bells' that resulted from splitting up AT&T in 1984, could end up growing even faster than the current parent. Lucent Technologies for example, which was de-merged from the original AT&T, is now worth some $4 billion more than the original telephone company.

In this instance, the Internet catalyzed Microsoft's reaction. Microsoft was not able to ignore the Internet much as it might have liked to: it was an alternative but lucrative source of business and one that was not going to go away. Microsoft's browser fought Netscape's in the open market, albeit with the US Department of Justice as referee. But how would Microsoft mount guerrilla warfare against a shadowy, coevolving and amorphous competitor - someone using the ubiquity of the Internet as a competitive weapon against it. It is a challenge a lot harder to fend off than the one faced by VISA and the other credit and debit card businesses from transaction processing over the Internet. At least these businesses can use the same technology as the potential competition.

We will shortly find out how Microsoft will react to such a threat. Microsoft's dominance lies in its almost total ownership of the market for personal computer operating systems with its Windows ME and big brother Windows 2000. Windows 2000 is also used on larger remote shared computers as well as PCs but has stronger competition in that more specialized market. Microsoft makes good money from them and for providing support. However, even a business of Microsoft's size and success has problems. Operating systems are large and very complex programs and many people are involved in their development and testing. Willing customers are roped in to undertake trials of pre-production versions and report back on problems. But Microsoft only sends out pre-production and production versions in a form which will run but cannot be 'read' or amended by a programmer; it keeps the readable and editable form under lock and key. This has the advantage that it retains absolute control over the product. Its coevolving competition had innocuous beginnings. In 1991, Linus Torvalds, then an undergraduate at the University of Helsinki, produced a PC version of UNIX, another operating system much older than Windows 2000, which unsurprisingly he called LINUX. He made the readable and editable version freely available on the Internet and encouraged other potential co-developers to contribute. Co-developers were free to market their own developments and even to charge for them. But there was one subtle rider: the readable and editable version of what was given away or sold must also be made available, not merely the version which 'runs'. Central control over only the small kernel (core) of LINUX was retained to avoid complete anarchy. In all other respects, LINUX took on a life of its own. The result has been remarkable. New versions with new specialist features supporting new computers are made available extraordinarily quickly. Problems can be diagnosed and fixed, often free, via the large number of developers on the Internet. In short, one coevolving entity - LINUX - has ridden on the back of another much vaster one - the Internet, and is using it in a way that is closely in line with the unwritten ethical principles of the Internet.

Microsoft may be a dinosaur but is certainly not one of the plodding varieties - Tyrannosaurus Rex is nearer the mark. But both died out while other smaller, swifter and more flexible ones evolved on different lines, took to the air and breakfast at my bird-table. "Businesses", Dee Hock once said, "die out not when they are defeated or suppressed, but when they become despairing and lose excitement and hope about the future". Perhaps today's organizations have got sidetracked by 'scientific management' and have lost sight of their real asset - human creativity nurtured by individual freedom. Perhaps they are already drifting slowly and blindly towards extinction.

But does it work?

In an interview by Peter Day for the BBC Radio 4 programme *In Business* in September 1997, VISA Executive VP Robert Miller had no doubt that the highly unusual decentralized and competitive structure of VISA was the key to its awesome growth. “We consider that to be one of our distinct advantages” he pointed out “... the fact that no-one really controls us. We wouldn’t have had the success we have had over the years if we did have one member holding a very large ... share in the company”. But Michael Lafferty, chairman of the International Cards Council trade association, was doubtful that mutual organizations like VISA can exist in competitive markets because of their difficulty in raising capital: “The reason you exist as a financial institution is to generate value for your shareholders. The same culture is going to be imposed by those banks and the organizations that supply them with services. And that’s why it’s inevitable that VISA, the mixed-up mutual, must give way to a commercial organization.” Not unnaturally, Dee Hock sees things differently. Day put it to him that VISA might just have been the lucky case where his ideas worked. Hock responded that it was absurd to think that the model was not universally applicable. “Evolution has been applying this way of thinking about organizational and structural creativity since the beginning of time. [And] ... what about the Internet ... [or] ... Alcoholics Anonymous?” Geoff Mulgan, then of the Demos think-tank, who advises UK Prime Minister Tony Blair on policy issues characterized to Day the survivors in business as those “... whose glue comes more from an ethos, a shared culture and set of commitments ... rather than through owning something like a building”.

And if it does...

Perhaps those businesses whose guiding principles are posted on every office wall should add one more. It is a maxim well-known to those who haunt cathedral bookshops: “Give me the serenity to accept the things I cannot change, the courage to change the things I can change, and the wisdom to know the difference”. In other words allowing individual initiative the freedom to be applied sensibly without waste is at the heart of the new organization. And those businesspeople with heads in the sand and struggling for growth who stolidly maintain “we would never run *our* Company like that” should ponder how much of it - or perhaps how little - would be left if their strongest competitor decomplexed or even ‘went mutual’.

ANNEX - THEORETICAL BACKGROUND

The underlying theoretical background covers many scientific disciplines. What follows is a personal guided tour through the maze, giving directions to the adjacent bibliography. It is not a long tour, but is intended to give enough jumping off points for anyone wanting to delve a little more deeply. The use of 'reference **nn**' in place of the conventional superscript notation is to help those who prefer to read the text on screen (the Acrobat-format document is optimized for printing rather than display).

Where did it all start?

Seven strands of research came together in the early 1990s.

1. Evolutionary modelling

The first is the study of self-organizing biological and chemical systems pioneered by Stuart Kauffman, variously of the University of Pennsylvania, the Santa Fe Institute and latterly Bios Group LP. His monumental 'The Origins of Order' (reference **24**) is probably the most cited publication in the whole field, with Part 1 (Chapters 1 - 6 inclusive) being most relevant. Kauffman's other well-known publication 'At Home in the Universe' (reference **8**) is a much shorter non-specialist work for those without a background in genetics or biology. It has useful additional material on the optimal size of an object that was covered briefly here in Chapter 4 (How big should an object be? - see also reference **43**). It also introduces a different method for disturbing coevolving systems into avalanche behaviour: invasion followed by extinction as opposed to use of the external environment W parameter in 'Origins' which was described here in Chapter 2. A species' niche is defined by which C-coupled neighbours it has. If this niche is invaded by another species that thrives better in that niche, then the less fit species is made extinct. There is, in this instance, a difference between invasion plus extinction and the 'development' of a species. In the extremal models and analogies described in Chapter 3 of which the Bak-Sneppen model is typical, there is no difference between extinction and development: an embankment leak selected for plugging is merely allocated a new leak size. It is immaterial whether this is regarded as the extinction of one leak and its replacement by another or as the development of the first leak. But to a biologist, the development of a new species from an existing one (speciation) is very different from invasion and extinction since speciation brings with it at least some of the characteristics of the parent species (although species cannot interbreed).

In both books, Kauffman describes another form of an NKC system called a Boolean network. A Boolean network is a linked collection of logic gates -AND, OR, and so on. The linkages can be in any arrangement whatsoever, although for convenience all multiple-input gates such as AND are assigned the same number of inputs (2, for example). Boolean networks can be made to play games against each other by being made to adjust their numbers of inputs, to rearrange their connections, or to 'canalize' their gate functions to prefer a dominant output (1 as opposed to 0, for example) all or most of the time. The aim of these adjustments is to maximize some measure of, for example, how *different* the '0 or 1' states of gate-outputs of different networks are from each other. Boolean networks and NKC systems are very closely related. In a Boolean network, each *site* (gate) is coevolving with the others to which it is connected. In an NKC system, each *whole NK system* is evolving with each other system to which it is C-coupled. A Boolean network is thus an NKC system with zero K and where connections between gates are C-couplings. Be aware however that most articles on Boolean networks use a 'K' for what we would use 'C'. In Chapter 6 we introduced the idea of 'preferred' reactions from C-coupled objects. This is a simplistic way to describe the behaviour of Boolean networks which are, for example, composed of 'canalizing' OR-gates with two inputs (as opposed to non-canalizing Exclusive-OR gates with many inputs), and also more complex ('homogeneous') gates whose output has a dominant value most of the time. Either of these can induce part or all of the network to freeze with one set of outputs (a 'forcing structure').

NKC models are not, in general, amenable to exact mathematical analysis. It is, however, possible to analyze one special case - the maximally rugged (very high K) landscape where every gene in an object is K-coupled to all the others within that object. The terrain on a maximally rugged landscape is unpredictable: mountains are not clustered together and neither are the plains. The topography - the 'height above sea level' - at any one point is uncorrelated with that at any other point. It can thus be treated as assigned at random. This assumption can be made when viewing things on a very large scale, an entire ecology for example, because any landscape correlations for particular species are spatially tiny. It is also mathematically convenient since correlations upset the mathematics. In two related papers in the same publication (references **37** and **34**), an NK system is approximated (in the first paper) by adaptive walks on such a random landscape. An adaptive walk is one where each new step is an adventure into the unknown: a step is only taken when fitness is improved. The NK system is then elaborated (in the second paper) into a competitive NKC system. Two further approximations are then taken in order that the NKC model is mathematically tractable. The first is that the number of different species is infinite. The second is that instead of merely choosing *at the outset* and at random the C genes to which any one gene is externally coupled, these C genes are re-chosen randomly *at each step*. The first is called a 'quenched' choice since the choice is frozen after selection; the second is called an 'annealed' choice since randomness is introduced afresh each time. Both terms have their origins in the metal-smith's art as described in Chapter 6. Analysis of this model shows that the NKC system cannot be driven to a state of *self*-organized criticality by the sort of 'select the extremal value' perturbation used in the Bak-Sneppen model: the inability to build correlations sees to that. But the second paper does prove the existence of the expected frozen, chaotic and borderline states in systems with manually tuned parameter settings. However, since the model also does not include extinctions and invasions, it is too unrepresentative of Kauffman's model or the way Kauffman proposes its self-organization that was described here in Chapter 3.

Hillclimbing on a fitness landscape is a subject in its own right, and many hillclimbing techniques were developed in the 1960s when computing power become available to researchers. Finding an efficient technique is important because many business and engineering problems as well as scientific ones can be formulated as 'optimize some criterion whose value is determined by the values of several variables'. For our purposes, the criterion is fitness and the variables are decision genes. Optimizing something while the variables can take on any values (for us this might be the amount of capital investment) is the easiest type of problem technically. More often, however, constraints exist on these values and the problem becomes much harder. These constraints may be upper or lower limits on the values of the variables (a cap on capital investment for example) or even the restriction that variables must take integer values (our genes can be thought to have two integer values - representing blue eyes or brown eyes for example). Reference **45** is a good if rather dated summary of the best-known and most effective techniques for solving unconstrained problems and some types of constrained ones as well. It might be thought that a 'greedy' approach in which steps are taken only in the *steepest* direction upward - changing any or all of the variables at the same time - is always the best. This is only true, however, on simple regular-shaped hills like Mount Fuji. Landscapes that are more rugged need techniques that adapt to the shape of the landscape. Greedy (steepest ascent) strategies on rugged landscapes follow strange zigzag paths in the foothills, and using them is a recipe for becoming marooned on the peaks of small hills. A mountaineer tackling an unknown climb will choose and amend his route as he goes along while keeping his eye on the peak. What appears to be the direction of overall steepest ascent might involve descending into valleys and re-emerging. It might also be found to be infeasible when an unanticipated bottomless ravine is encountered. Computer-based hillclimbing techniques are even more limited. They operate very much like a blindfolded climber who has no better way of knowing which direction is 'up' than taking a few steps in each of several directions and picking the best. He cannot see the peak or even the terrain immediately in front of him. Of the many techniques that exist for solving unconstrained hillclimbing problems, the present author has had considerable success with variants of Powell's method (reference **53**). This is as aimless as any of the others on very rugged landscapes but often 'climbs' quite competently in medium and low-K environments.

2. Thermodynamics

The second strand is the study of systems which are far from equilibrium and which is mainly associated with 1976 Nobel laureate Ilya Prigogine (reference **12** and **26**), doyen of the 'Brussels school' of thermodynamics. Prigogine showed that in any system that is far away from thermodynamic equilibrium, the laws of thermodynamics need reinterpreting. The importance of this is simply that self-organizing systems that live on the boundary between order and chaos are themselves far from equilibrium. Self-organization would at first glance seem to violate the second law of thermodynamics that says that in any closed system - one in equilibrium with its surroundings - disorder (entropy) always increases. But a self-organized system spends most of its time far from equilibrium, and the growth of order within the system is counterbalanced by *more* disorder elsewhere in its surroundings. Such a system takes in energy to retain its ordered existence and, since energy is conserved, dissipates it again as heat to its surroundings. (The value of Prigogine's contribution to this area is controversial. It seems equally likely that an earlier Nobel laureate Lars Onsanger developed the current basis of non-equilibrium thermodynamics.) Prigogine was also one of the first to study traffic flow theoretically at an aggregate level, like the flow of a fluid, in the late 1950s. Detailed modelling of traffic as a collection of individually interacting vehicles was not however feasible until adequate computing power became available.

Stuart Kauffman has described (reference **25**) his search for a general law, rather like an extra law of thermodynamics, governing the behaviour of the non-equilibrium systems in which he is particularly interested - the 'coevolutionary self-constructing communities of autonomous agents'.

3. Chaos, avalanche dynamics and universality

The third strand is the study of how waves and particles move in some medium when subjected to an external force. The mathematics of chaos itself and its occurrence in the natural world has been studied extensively over the last thirty years although it dates back a century to the French mathematician Henri Poincaré. The avalanches of change resulting from disturbing a system on the boundary between order and chaos behave mathematically in a similar way to the movement of a fluid being absorbed by a thick swab. As the fluid moves upward through the swab it percolates more quickly through the slightly more permeable parts than other areas. This in turn changes the flow of the remaining fluid that is adjacent to the more permeable parts. Paczuski et al (reference **51**) discusses the underlying behaviour of the Bak-Sneppen model (see below) and relates it both to one common variant of percolation (see reference **29**) - directed percolation - and to Reggeon field theory. The latter is a quantum theory devised by Russian physicist Vladimir Gribov to describe the behaviour of subatomic (quasi-) particles called Reggeons first introduced by Italian physicist Tullio Regge in 1959. Reggeon field theory has been shown (references **39** and **40**) to be a special form of statistical (Markov) process in which particles diffuse though space, annihilate other particles, are made extinct themselves, or create new particles. The relationship between this, percolation, and the Bak-Sneppen model is (very roughly) that the movement of one particle affects close neighbours; these in turn then affect *their* neighbours.

Bak's sandpile experiment and its subsequent analysis (references **33** and **2**) was the first of two landmarks in the study of self-organized criticality. The second was the introduction of extremal dynamics to this area by Sneppen (references **35, 38** and **57)** as a way to build up the spatially long-term correlations between objects. These correlations are necessary for avalanches and their 1/f power-law behaviour. There are fewer larger avalanches than small ones and the likelihood of any particular avalanche size is inversely proportional ('1/f') to that size. 'Size' here can be the duration of the avalanche or the number of sites involved (as in the number of embankment leak sites participating in any one avalanche described in Chapter 3).

Extremal dynamics ('select the largest or smallest') on its own does not necessarily lead to self-organized criticality; it is just natural selection. But if objects are C-coupled to neighbours such that when the extremal object is selected for replacement the neighbours are replaced also, it creates the opportunity for a cascade of replacements because the neighbours are C-coupled to *their* neighbours, and so on. The more neighbours to which an object is directly C-coupled - perhaps both immediate neighbours and *their* immediate neighbours, the earlier the barrier (see Chapter 3) is hit, the sooner avalanches happen, and the lower the fitness becomes at the barrier point. Bak and Sneppen introduced the very simple model (the Bak-Sneppen model in references **17** and **35**) that was described in Chapter 3.

Paczuski et al (reference **52**) contains by far the best and most detailed review of avalanche dynamics in both self-organized and manually tuned systems. She comments that whereas the Bak-Sneppen model achieves self-organized criticality through a deliberate policy of selecting the objects of extremal (lowest) fitness, Bak's sandpile model achieves self-organized criticality through selecting *by itself* the extremal (weakest) sites in the sandpile.

One aim of Paczuski's paper is to show the underlying similarity of the behaviour of many different physical systems at the point they attain self-organized criticality. Such systems act very differently when they are non-critical, but when they reach criticality they take on a new and surprisingly common behaviour. This is characterized by two things: self-similar (fractal) behaviour - in space or time - and universality. We have already met self-similar behaviour: avalanches are composed of sub-avalanches; each sub-avalanche is composed of other sub-avalanches. Each avalanche looks the same as any other at whatever level of detail. This applies both to *when* avalanches happen (temporal self-similarity) and to *where* they happen (spatial self-similarity). The reason that different systems suddenly behave in the same way when they reach criticality is that it is the point when each element of the system (each of Pooh's fir-cones; each embankment leak...) is in touch with its neighbours through C-coupling. More importantly, it is the point when *the chain of C-couplings suddenly extends right across the system* (the across-the-lake melting described in Chapter 3). When this happens, the differences between the systems are swamped by the similarity of behaviour caused by this chain of C-connected elements that is present in all the systems when criticality is reached. We encountered correlations earlier in this section when describing the topography of a landscape. The existence of a chain of C-connected elements stretching right across the system implies that the behaviours of all the C-connected elements in the chain are correlated: nudge one and the knock-on effect might disturb them all. Not all critical systems will behave in an *identical* way. Fir-cones on the lake are two-dimensional in that they can affect other cones in any direction (except up and down) whereas the leaky embankment is a (one-dimensional) straight line and neighbouring leaks are left- or right-hand neighbours only. Systems with the same dimension behave in the same way; systems with different dimensions merely follow the same general pattern of behaviour. A Rolls-Royce limousine and an Indy racing car do much the same thing but over different timescales and at different speeds whereas the behaviours of an Indy car and an aircraft (when aloft) are fundamentally different. Systems at criticality with the same dimensionality and other fundamental characteristics that together determine the neighbours to which each element is connected are said to belong to the same 'universality' class. Those with different dimensions belong to different universality classes. Universality classes are of great importance because once the behaviour of a class is known, by analysis or experiment or computer simulation, the same behaviour applies to all systems that belong to that class. The differences between members of a class when they are not critical vanish when criticality is reached. It is interesting to speculate whether this could be made to apply to businesses: whether the myriad individual differences between businesses would vanish as criticality approached. No one knows, perhaps because for a business the dimension and other fundamental characteristics are also unknown. It would be intriguing to have a universality class for businesses at criticality because the behaviour of one of them could predict the behaviour of them all!

Chapter 3's frozen lake is an example of what a physicist would call a 'phase transition'. When ice thaws or water boils, or when turbulence sets in as an aircraft wing moves faster, the change from one state to the other is quite sudden. Hitting the boundary between order and chaos with the ensuing cascades of avalanches is another example. Ice thawing is itself a change from order to less order, and the steam resulting from water boiling is highly disordered. Very recent work (see references **32** and **48** in the same journal edition) on better ways to solve intractable optimization problems of a particular type has shown that in many of these problems, the amount of computation needed to find a solution does not just get rapidly worse as the problem gets larger, but exhibits sudden jumps - phase transitions. A well-known example of this type of problem is the travelling salesman problem. Its aim is to find the least-distance itinerary for a salesman who must make several stops. The problem gets larger through adding stops. No quick computational methods exist for solving it. The implication is that there are levels of complexity in a system beyond which trying to manage it suddenly becomes *much* more difficult. Decomposing or otherwise simplifying it to stay just below one of these barriers makes predicting its performance much easier. Breaking up a large organization into autonomous smaller pieces as described in Chapter 4 under 'How big should an object be?' *may* thus not be a smooth process ("a bit more decentralization is a bit better"). There may, for example, be levels of decentralization that are markedly better then ones that are only slightly more complex. If this is true, and it is entirely unproven at present, the implications for organization design are far-reaching.

Finally, the self-organization of traffic into sporadic traffic jams as traffic volume builds up, with the consequent avalanches of congestion up and down the highway, is not the whole story. When traffic density increases on a multilane highway as peak time approaches, free-flowing traffic does not normally experience sudden traffic jams without passing through an intermediate synchronized state where all vehicles move at the same speed. Vehicles in *all* lanes then move in one synchronized block. When this happens, opportunities for overtaking become rare as there is no advantage to be gained in overtaking and there is usually no free space in an adjacent lane to do so anyway. As congestion rises further, synchronized flow breaks down into sporadic small jams and then into one big jam as all the smaller jams coalesce. These phenomena have been known for some time. But what *is* new is the discovery by Boris Kerner of DaimlerChrysler AG in Stuttgart (reference **44**) using data from the Frankfurt - Basel A5 autobahn that the changes from free flow to synchronized flow and from synchronized flow to jams are sudden phase transitions. Maximum traffic flow occurs just before the synchronized state suddenly breaks down into jams, and when such jams take place the traffic flow tends to self-organize and fight its way back into a synchronized state again. At the point where jams are just starting, traffic in different lanes tends to jam at different moments and the speed of traffic in the various lanes then becomes different. This creates opportunities for lane changing and tends to restore synchronized flow again. Bernardo Huberman of Xerox and Dirk Helbing of the University of Stuttgart (reference **42**) have shown using data from the Netherlands A9 highway that a leavening of trucks - say 2.5% of the total - increases the likelihood of synchronized flow. This is because the generally slower trucks regulate the speed of all the vehicles immediately behind them unless the latter can overtake. Thus as opportunities for overtaking become fewer, a small number of widely spaced trucks act as an involuntary mechanism to regulate the speed of all vehicles. The effectiveness of this regulation depends upon what type of overtaking is allowed. The US approach of permitting vehicles of all types to occupy any lane allows trucks to spread themselves across all lanes. The European approach of restricting trucks to the slower lanes and (in theory) permitting overtaking only by movement from slow lane to fast lane and back again does not assist the formation of synchronized flows and it reduces highway capacity by some 25%. Apart from allowing the road to be used more intensively, synchronized flow is relatively safe as overtaking only happens infrequently and vehicle speeds are similar. The transition from a series of sporadic small jams to one large jam is the point of self-organized criticality. In any one lane, each vehicle is affected by the behaviour of the vehicle in front and in turn affects the behaviour of the vehicle behind, and so on. Like the fir-cones in our frozen lake at the point of thawing, each vehicle can affect the behaviour of *all* other vehicles behind for the length of the jam. In addition, vehicles in different lanes influence each other through opportunistic lane changing.

These sudden changes from free flow to synchronized flow to sporadic small jams to a single large jam are all phase transitions and there are other phase transitions *within* a jam as congestion rises even further (see reference **41** and the further references therein). However, unlike Bak's sandpile where surplus sand is sloughed off in avalanches in order to maintain the critical slope of the pile, a highway has no way to get rid of excess vehicles forcibly and all vehicles eventually just grind to a standstill. Highways are however self-adaptive in the sense that as traffic flow slows and vehicles become more and more bunched, it gets harder and harder and eventually impossible for additional vehicles to join the highway.

4. Biology

The fourth strand is from a better understanding of the ecology and biology of evolving and interacting living species. The 'genotype' of N genes, K of which are coupled, has a fitness that is determined by its phenotype - the effect of the genes on the organism and its development. The effect of C-coupling *between* genotypes - the effect of one genotype on another - is the *extended phenotype* popularized by Richard Dawkins of Oxford University (references **5** and **23**). The war for dominance between species in NKC models is a battle between phenotypes but one commanded by the genes of the genotype. The deformation of the fitness landscape of one species by another *is* (part of) the extended phenotype of the latter. The extended phenotype potentially extends several hops across a chain of linked genotypes such as a food web.

The equilibrium resulting from brand competition, where each brand manager follows a fixed strategy simply because any other would be more costly, was analyzed nearly fifty years ago by American Nobel laureate John Nash (reference **49**). Nash showed that if each competitor adopts a single strategy, not necessarily the same one, then there always exists at least one collection of 'special' strategies, one strategy for each competitor. These special strategies are such that if each competitor adopts his allocated strategy on the assumption that his competitors adopt their own special strategies, then they are the best strategies for each competitor. 'Best' here means that if any of the competitors decides unilaterally to change strategy, that competitor is worse off. No competitor has any incentive to do so and thus relative stability reigns. 'Adopting a single strategy' in this context means that any competitor does not use a *mix* of strategies as was described earlier for brand competition. Similar results exist for the use of mixed strategies, however. But these collections of special strategies which result in 'Nash equilibria' have several problems: there may be lots of them, there is no route-map to find them apart from intelligent trial and error, and none of them may be particularly good for any one competitor. But they are stable.

It was left to English evolutionary biologist John Maynard Smith (reference **47**) to generalize the idea of Nash equilibria to coevolution. Maynard Smith equated the selection of a strategy by a competitor to the specification of a genotype within a species. If evolution between a collection of species has progressed to the point where *each* species becomes less fit if it changes one or more of its N genes or changes its internal K-complexity or external C-coupling (if it can), then the collection of genotypes has reached an evolutionarily stable state. There may be other collections of genotypes - other directions in which each species could evolve. But, apart from those that are also evolutionarily stable strategies, they are all unstable and will probably give way to a spate of rapid coevolution until the collection of genotypes reaches another evolutionarily stable state. There is, however, nothing to prevent any collection of species that has jointly adopted an evolutionarily stable strategy from going extinct! This suggests that some evolutionarily stable strategies may be better than others and it is in all the competitors' interests to find the best one. It is almost as if the species got together and between them forged a deal to change strategy all at once. Even more intriguingly, this sounds like natural selection at two levels: once at the individual (phenotype) level and once at the inter-species 'group of strategies' level. But the latter may be an illusion: it could well be a side effect of the normal mechanism of natural selection where genotypes with fitter phenotypes survive and those with poorer ones die out. The controversy about whether or not group selection exists was touched on in Chapter 2. If group selection does exist in some form, it may have a significant bearing on how collections of coevolving objects develop. Altruism - or lack of it - between a major customer and its dependent suppliers, dependent perhaps to the extent of being co-located on the same industrial site, might be understood and managed somewhat differently. This would also apply to the coevolving parts of a single business. But group selection remains an unproven and perhaps unnecessary concept.

It is worth emphasizing at this point (as was touched on in Chapter 5) that the description given hitherto about what natural selection operates on is a simplification. A species has been treated as if it has only one genotype - the particular values of its N genes - at any one time. This is untrue in reality. Each species consists of a large variety of genotypes that are in competition with those of their predators, their prey, and each other. This competition is dynamic in that new genotypes in each species are continually being created through reproduction while others are being killed off. The values of the N genes in the NKC models described should be treated as averages rather than single fixed values.

The concept of 'fitness' is also a little more complex. We exemplified fitness earlier as the ability of a rabbit to outrun or outwit a predatory fox through stronger legs, more stamina or whatever. Rabbits that were blessed in this way were more likely to survive and reproduce. Other rabbits were less fortunate. Two measures of fitness - success in evading predators and number of viable offspring left - were joined in the interests of simplicity. A better definition of fitness is simply that which tends to maximize the number of copies of an organism's genes. A gene 'wants' above all else to multiply faster and exist longer than its peers. If this can be achieved best by helping its peers, as happens with sterile sister worker bees, and not through having numerous offspring, then so be it. A grandmother has an obvious genetic reason for helping nurture her granddaughter. What is less obvious is that the mother's genes will be fitter if they *also* manage to influence grandmother to help in childrearing. The mother has the stronger vested interest genetically - the child is more closely related genetically to her than to her grandmother. If they can, genes manipulate whatever material they have at their disposal: the rabbit or rabbits they occupy, predators, prey, parents, siblings, offspring, and patches of nutritious grass (perhaps by fertilizing them).

We associated genes with, among other things, business decisions. Important and far-reaching decisions become strategies. Successful strategies are repeated and documented and have management books written about them. Unsuccessful ones are not repeated, at least not in quick succession in the same business. Managers who push unsuccessful strategies tend to become ex-managers, although they may surface again in other businesses where peddling the same nostrums may actually work because competitors or products are different. Strategies and ideas, like genotypes, evolve and propagate - or die out.

Propagation and replication are a result of communication, whether face-to-face, via electronic mail or through television. Natural selection culls those ideas that fall on deaf ears or stony ground. An idea develops when it takes up residence in a human brain before being passed on. Richard Dawkins (references **5** and **23**) coined the word *meme* for these non-genetic replicators. Memes can develop outside the brain also. My PC could be programmed to edit all my incoming electronic mail in some way and then forward it. It is arguable whether this constitutes development of the *ideas* within each incoming message - perhaps this depends upon the cleverness of the program and hence the programmer - but it is certainly development of the text itself.

It was mentioned earlier that where constraints on development exist - such as limits to an animal's muscle bulk or limits on how much a business can spend on advertising, constrained traits lead to mutual coexistence whereas unconstrained traits lead to constant coevolution. More precisely, it has been shown (reference **55**) that constrained traits give rise to evolutionarily stable strategies. The outcome is similar to that of evolution on a high-K landscape, and if the constraint is the result of an increase in fitness (market share) in one area (advertising) being counteracted by a decrease in another area (trade promotional spend), it is due to K-complexity.

The Eigen and Schuster fitness ('error') catastrophe invoked in Chapter 5 to define an economic upper limit to the number of test marketed product variants which could be concurrently on sale looks like an inviolable natural law. However, Eigen and Schuster also proposed an escape route. Earlier, we defined an object in this context as the collection of all the marketed variants of one product. If several of such evolving objects, each of which is experiencing an error catastrophe, are C-coupled in a circle, the overall coevolving system does not itself experience an error catastrophe. It is currently unknown whether there is a sensible business analogy of the C-coupled circle ('hypercycle') of objects: three or more test marketing campaigns being waged concurrently by competing businesses are unlikely to constitute a hypercycle since each business can gain or lose market share from *all* the others. The coupling is then a mesh rather than a circle.

5. Information technology

The fifth strand is information technology and, in particular, programming computer simulations of the real world. In Chapter 7, the 'duality' of data and processes within business systems design was touched on. To a computer program designer, something more is needed to ensure that the resulting programs are structured such that they are easy to maintain. The structured programming movement first latched on to this in the early 1970s and their aim was to design programs around the (physical) structure of the data files rather than around some artificial structure imposed by the programming languages being used. Quite separately, the relational database movement led by IBM's Ted Codd was a campaign for 'logical' data structures to mirror the real world and to avoid many of the inconsistencies that are imposed by database management and file systems in the interests of efficiency. Data items do, after all, have some innate clustering which is independent of how they will be held in a computer: my sex, nationality and date-of-birth have a strong affinity with the employee number which uniquely identifies *me* in some personnel system. Codd's elaboration of this natural clustering into his 'normal forms' (and later a set of rules) gave rise to relational database management systems which either handled data in simple tables or gave the programmer the illusion that it was doing so. In this way, they made the physical structure of a collection of data - the way it was stored in the computer - identical to the desirable logical structure as far as this could be tolerated without causing unacceptably bad performance. Structured programming and relational database principles were amalgamated in the late 1970s into a single systems design methodology. But there was something missing. It soon became clear that data and processes were not independent things merely imposed one on another (*which* was thought to be imposed on *which* was determined by whether you were a 'process' or 'data' bigot) but they were aspects of the same thing - an *object*. Readers who are part of the current generation of computer programmers will by now have found our coevolving ‘objects’ resembling the things they are used to manipulating in object-orientated programming languages where an object is a self-contained section of program (a process) with its own associated data. Such objects communicate with each other by passing messages using formal *formats* and *protocols* (see Chapter 7), but *how* they perform their functions is deliberately hidden from others. This ‘information hiding’ for computing objects was introduced by David Parnas in 1972 (reference **28**) as a way to protect an object from being tampered with by other objects or from suppositions being made on how it worked internally. These objects exist independently of others, hide internal information on *how* they do what they do from others, respond only to messages (changes in C-coupling), have standard ‘classes’ of object (VISA participants, field sales areas..) with ‘instances’ of each (a particular VISA financial institution or sales area) and so on. But perhaps this should not be surprising after all: object-orientated programming had its origin and first implementation in 1960s simulation languages such as SIMULA67 whose aim was to model the real world. To complete the picture, SIMULA came from the Norwegian Computing Centre in Oslo. Norway and Denmark have long been associated with investigations into the mathematics of system dynamics. The Universities of Oslo and Copenhagen have for many years had active research programmes on the use of fractals to describe the workings of natural processes such as fluid movement and earthquakes, and on the mathematical theory of percolation. Both areas involve self-organized criticality and consequent avalanches of change giving rise to evolutionary or ecological ‘punctuated equilibria’ that are periods of stability interspersed with sporadic bouts of rapid change or species extinction.

In Chapter 4 we equated decision points with organization groups *and* with individuals. For most purposes this is correct. But it should not be assumed that the *processes* used to arrive at decisions are the same in both; they depend upon the relative proportions of data-driven computing assistance (the business planning models of Chapter 5, for example) and knowledge- and information-driven human thought processes employed. To digress, Daniel Dennett (reference **6**) has proposed that thought processes are the result of our trying to graft a sequential ‘virtual machine’ on top of a naturally parallel-processing machine - the hardware of the brain. All current general-purpose computers work the other way around. They offer a virtual parallel-processing machine such as is provided by every computer operating system and, at a higher level, by IBM’s VM that allows more than one operating system to execute concurrently. (To Microsoft’s chagrin, there is now even a product from VMware Inc that allows several copies of Windows NT to run concurrently in virtual machines on top of LINUX - see Chapter 10). These multiple virtual machines run on essentially sequential real hardware, although most shared computers now offer genuine but very limited multiprocessing. If Dennett is right, it would explain why we are so much better than computers at pattern recognition and so poor at conventional sequential calculations. Mathematical ‘savants’ - calculating prodigies who are often autistic like

Dustin Hoffman's Rain Man portrayal or who are otherwise retarded - may conceivably be *directly* accessing their parallel processing mechanism. But they have somehow damaged the link between their sequential (virtual) and parallel (real) machines which would be needed to understand the implied nuances of language and facial expressions as well as facts and words; intriguing perhaps, but pseudo-science.

The idea that knowledge is encapsulated in information while being transmitted or stored has a remarkable precedent in theoretical physics. The late David Bohm of Birkbeck College, London first popularised in 1980 this idea of 'enfolding' (see reference **22**). He proposed that thought (consciousness) and the external world were one unbroken entity (which he called the *implicate* order) but that what we perceive is one or more outer wrappers (or 'projections') which he called the *explicate* order. This is precisely one stage removed from knowledge and information where knowledge is the implicate order and information is the explicate order.

6. Analyses of real businesses

Gene Stanley et al (references **30, 31, 36, 46, 56**) highlighted the similarity of the patterns of growth of different types of business with the universality concept of statistical physics. As has already been described, universality implies that the same underlying laws with roughly similar constants in their defining equations apply to very different types and sizes of things - in this case businesses. Stanley et al went on to 'scale' the empirical growth results by removing ('dividing out') the dependence of the results on business size. This gave curves for businesses of all sizes which, apart from near the upper and lower limits of growth, had collapsed on (merged with) one another, confirming the independence of the underlying growth and growth variation laws from anything connected with the firms' lines of business or their sizes. References **31** and **36** are in the same publication and are intended to be read as a pair. The first summarizes analyses of the performance of real companies; the second derives a hierarchical model of a business and relates this to data from the previous paper. These two together are an excellent starting point for those interested in the 'econophysics' of business behaviour rather than in the dynamics of financial markets.

7. Economics

Lastly, the Nash equilibrium and Maynard Smith's extension of it to the evolutionarily stable state had an earlier parallel in traditional economics. In the early 1900s, Italian economist and engineer Vilfredo Pareto introduced the idea of an equilibrium state of 'economic efficiency' for an economy. This state was defined such that any change to the strategies adopted by the participants, which include the State as a major player via taxation policy and controls over the free market, made someone worse off. Everyone had an incentive to preserve the status quo. This equilibrium, also called a Pareto optimum, was extended to apply to a competitive free market where demand, including employment, exactly met supply. The result, which was introduced in Chapter 5, was unsurprisingly called a *competitive equilibrium* and remains a cornerstone of modern economics despite having been shown in the early 1980s only to exist under some very unlikely conditions indeed (see page 224 in reference **50**). A competitive equilibrium is only a Pareto optimum in a cloud-cuckoo-land where there are no constraints - no monopolies, no de facto price cartels and no wage bargaining hammerlocks by trades unions. In addition, there must be either a population of consumers who make purchase decisions in unrealistic ways or a 'futures' market for everything. The last point may seem an odd intrusion, but the equilibrium being sought is not a state of stagnation with no change whatever, but a state where the inevitabilities of real life are included: insurance companies may have good years and bad years for example. So the 'equilibrium' is actually a series - really a continuum - of equilibria where the prices of goods and services continually adjust to meet demand exactly. So for someone to make a decision, he needs to be able to predict the future of the economy with certainly. If he cannot, he needs to know the prices of anything he might wish to buy or sell - including labour - in *all* the likely economic outcomes. If he cannot do either of these, he needs to have unlimited computing power to calculate them (see reference **54**). None of these obtain in the real word. The economy is made up of individuals acting on their own behalf or as managers of businesses. Their decisions affect the economy that in turn affects other individuals who are themselves making decisions. And influential individuals such as financial forecasters *directly* affect other individuals' decisions. As Radner (reference **54**) points out, "the introduction of information [to decision makers] about the behaviour of other decision makers introduces [side effects] among the [courses of action] available to them. A particular case of this results from the introduction of ... 'spot' markets. The presence of such markets enlarges the [courses of action] jointly available to the individual decision makers ... but also introduces the above-mentioned [side effects] ... Individual decision makers cannot, in principle, calculate their optimal [decision] without knowing something about the decision rules of other individuals". This is why the promotional strategies of competing brand managers (Chapter 5) have to be considered as a single ensemble: in competitive markets, a marketing strategy, like the products it seeks to push, is always competing with others. The economy is driven by this collection of decisions interacting in complex ways. Attempts to extrapolate from the behaviour of one hypothetical 'rational' individual acting in isolation lead nowhere, although government economic models are still based on this fatuous hope.

When we discussed in Chapter 5 the failure of economic modelling, it was mentioned that the equations used by the large-scale economic modellers were of the wrong *type*. More precisely, these tend to be collections of *linear* algebraic and differential equations which, although lengthy and complex, are readily and accurately solved though numerical simulation of the particular economic environment being modelled. The equations are generally insensitive to small disturbances and react to them in predictable ways. But the real world is not like that and the new more realistic models include non-linear equations and also programming logic (a yes/no decision that provides a step change, for example) that has the same effect. A non-linear model can be very susceptible to disturbances and to minor variations in initial conditions. The way in which it reacts to them can be unpredictable. The same disturbance injected at a different time during modelling - when the model is in a different state - can have a completely different outcome. Readers who have persevered thus far will find this no surprise at all: the NKC model of coevolving objects is a non-linear model. Objects deform other objects' landscapes in ways that depend upon the current states (gene values and K values) of *all* the C-coupled objects in the coupled collection.

These C-coupled objects are the active 'agents' in the economy.

BIBLIOGRAPHY

Non-specialist books

1. Axelrod, R. and Cohen, M.D. “Harnessing Complexity: Organizational Implications” (The Free Press 2000)
2. Bak, P. “How Nature works” (Oxford University Press 1997)
3. Capra, F. “The Web of Life” (HarperCollins 1996)
4. Coveney, P. and Highfield, R. “Frontiers of complexity” (Faber 1992)
5. Dawkins, R. “The Selfish Gene” (Oxford University Press 1976)
6. Dennett, D.C. “Consciousness explained” (Penguin Books 1993)
7. Hock, D.W. “Birth of the Chaordic Age” (Berrett-Koehler 1999)
8. Kauffman, S.A. “At home in the Universe” (Oxford University Press 1995)
9. Lewin, R. “Complexity - life at the edge of chaos” (Collier 1992)
10. Lissack, M.R., Roos, J. and Petzinger, T. “The Next Common Sense” (Nicholas Brealey Publishing 1999)
11. Lissack, M.R. and Gunz, H.P (eds.) “Managing Complexity in Organizations” (Quorum Books 2000)
12. Prigogine, I. and Stengers, I. “The End of Certainty” (The Free Press 1997)
13. Rothschild, M.L. “Bionomics: economy as ecosystem” (Owl Books 1992; first published by Henry Holt in 1990 as “Bionomics: the inevitability of capitalism”)
14. Smith, A. “The Wealth of Nations” Vol 1, Bk 4, Ch 2 - first published in 1776 (edited by Cannan, E. - Methuen 1961; note that the popular Penguin edition covers Books 1-3 only)
15. Waldrop, M.M. “Complexity” (Simon and Schuster 1992)

Non-specialist articles

16. Bak, P. and Chen, K. “Self-organised criticality” (Scientific American January 1991)
17. Bak, P., Flyvbjerg, H. and Sneppen, K. “Can we model Darwin” (New Scientist 12th March 1994)
18. Enthoven, A. “Reflections on the management of the National Health Service” (Occasional Paper 5, Nuffield Provincial Hospitals Trust 1985)
19. Kauffman, S.A. “Antichaos and adaptation” (Scientific American August 1991)

Specialist books

20. Auyang, S.Y. “Foundations of Complex-System Theories: In Economics, Evolutionary Biology and Statistical Physics” (Cambridge University Press 1998)
21. Axelrod, R. “The Complexity of Cooperation” (Princeton University Press 1997)

22. Bohm, D. "Wholeness and the implicate order" (Routledge and Kegan Paul 1980)
23. Dawkins, R. "The Extended Phenotype" (W.H. Freeman and Co. 1982; also available in an Oxford University Press 1983 paperback edition)
24. Kauffman, S.A. "The Origins of Order" (Oxford University Press 1993)
25. Kauffman, S.A. "Investigations" (Oxford University Press 2000)
26. Kondepudi, D.K. and Prigogine, I. "Modern thermodynamics" (Wiley 1998)
27. Mantegna, R.N. and Stanley, H.E. "An Introduction to Econophysics" (Cambridge University Press 1999)
28. Parnas, D. 'On the criteria to be used in decomposing systems into modules' article in Yourdon, E. (ed.) "Classics in Software Engineering" (Yourdon Press, New Jersey 1979)
29. Stauffer, D. and Aharony, A. "Introduction to percolation theory" (2nd ed. - Taylor and Francis 1994)

Specialist papers

30. Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Salinger, M.A. and Stanley, H.E. "Power law scaling for a system of interacting units with a complex internal structure" (Phys Rev Lett 80 No 7 16th February 1998)
31. Amaral, L.A.N., Buldyrev, S.V., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., Stanley, H.E. and Stanley, M.H.R. "Scaling behaviour in economics: 1. Empirical results for company growth" (J. Phys. I. France, 7 1997)
32. Anderson, P.W. "Computing: solving problems in a finite time" (Nature 400 8th July 1999)
33. Bak, P., Tang, C. and Wiesenfeld, K. "Self-organised criticality" (Phys Rev A. 38 No 1 July 1988)
34. Bak, P., Flyvbjerg, H. and Lautrup, B. "Coevolution in a rugged fitness landscape" (Phys Rev A 46 15th November 1992)
35. Bak, P. and Sneppen, K. "Punctuated equilibrium and criticality in a simple model of evolution" (Phys Rev Lett 71 No 24 13th December 1993)
36. Buldyrev, S.V., Amaral, L.A.N., Havlin, S., Leschhorn, H., Maass, P., Salinger, M.A., Stanley, H.E. and Stanley, M.H.R. "Scaling behaviour in economics: 11. Modelling of company growth" (J. Phys. I. France, 7 1997)
37. Flyvbjerg, H. and Lautrup, B. "Evolution in a rugged fitness landscape" (Phys Rev A 46 15th November 1992)
38. Flyvbjerg, H., Sneppen, K. and Bak, P. "Mean field theory for a simple model of evolution" (Phys Rev Lett 71 No 24 13th December 1993)
39. Grassberger, P. and Sundermeyer, K. "Reggeon field theory and Markov processes" (Phys. Lett. B 77 220 1978)
40. Grassberger, P. and de la Torre, A. "Reggeon Field Theory on a Lattice" (Ann. Phys. 122, 373-396 1979)
41. Gupta, H.S. and Ramaswamy, R. "Backbones of traffic jams" (J. Phys. A. Math. Gen 29 No 21 November 1996)
42. Helbing, D. and Huberman, B.A. "Coherent moving states in highway traffic" (Nature 396 24/31 December 1998)

43. Kauffman, S., Macready, W.G. and Dickinson, E. "Divide to co-ordinate: coevolutionary problem solving" (Santa Fe Institute 15th October 1994)
44. Kerner, B.S. and Rehborn, H "Experimental properties pf phase transitions in traffic flow" (Phys. Rev. Lett. 79 4030 1997)
45. Kowalik, J. and Osborne, M.R. "Methods for unconstrained optimisation problems" (Elsevier 1968)
46. Mantegna, R.N. and Stanley, E. "Scaling behaviour in the dynamics of an economic index" (Nature 376 6th July 1995)
47. Maynard Smith, J. "The theory of games and the evolution of animal conflicts" (Journal of Theoretical Biology 47 1974)
48. Monasson, R., Zecchina, R., Kirkpatrick, S., Selman, B. and Troyansky, L. "Determining computational complexity from characteristic 'phase transitions'" (Nature 400 8th July 1999)
49. Nash, J. "Non-cooperative games" (Ann. Math. 54 1951)
50. Newbery, D.M.G. and Stiglitz, J.E. "The choice of techniques and the optimality of market equilibrium with rational expectations" (Journal of Political Economy 90 No. 2 1982)
51. Paczuski, M., Maslov, S. and Bak, P. "Field theory for a model of self-organised criticality" (Europhys. Lett. 27 (2) 1994)
52. Paczuski, M., Maslov, S. and Bak, P. "Avalanche dynamics in evolution, growth and depinning models" (Phys Rev E January 1996)
53. Powell, M.J.D. "An efficient method for finding the minimum of a function of several variables without calculating derivatives" (Computer Journal 7 1964)
54. Radner, R. "Competitive equilibrium under uncertainty" (Econometrica 36 No. 1 January 1968)
55. Rosenzweig, M.L., Brown, J.S. and Vincent, T.L. "Red queens and ESS: the coevolution of evolutionary rates" (Evol. Ecol. 1 1987)
56. Stanley, M.H.R., Amaral, L.A.N., Buldyrev, S.V., Leschhorn, H., Mass, P., Salinger, M.A. and Stanley, H.E. "Scaling behaviour in the growth of companies" (Nature 379 1996)
57. Sneppen, K., Bak, P., Flyvbjerg, H. and Jensen, M.H. "Evolution as a self-organised critical phenomenon" (Proc. Natl Acad. Sci. USA Vol 92 May 1995)

Web sites

These are correct at the time of publication but, like most Web sites, are subject to frequent change.

WWW.BIONOMICS.ORG - Bionomics Institute
The Bionomics Institute was set up to study the application of biological and evolutionary principles to business and economics. It sees the economy as a "self-organizing chaotic information ecosystem". The Bionomics Institute is now less active and promotional activity in this area among others has been taken over by the Cato Institute. Bionomics founder Michael Rothschild

subsequently founded Maxager Technology Incorporated (WWW.MAXAGER.COM). Maxager develops and sells methodology and tools for analyzing and improving the performance of capital-intensive batch manufacturing businesses. The methodology is akin in some ways to a cut-down version of Activity Based Costing (see Chapter 6).

WWW.BIOSGROUP.COM - Bios Group LP
Stuart Kauffman's joint venture with the Ernst and Young Centre for Business Innovation to develop and market complexity tools and concepts.

WWW.CALRESCO.ORG - CALResCo Group
Contains a useful list of other relevant Web sites and documentation

WWW.CATO.ORG - Cato Institute
Cato's aim is to promote the benefits of "limited government, individual liberty and peace"

WWW.DECOMPLEXITY.COM - Decomplexity Associates
The 'home' of this book. The present author is a principal of Decomplexity Associates Ltd, a business development company

WWW.SANTAFE.EDU - Santa Fe Institute
See especially Stuart Kauffman's series of lectures "The nature of autonomous agents and the worlds they mutually create". Kauffman subsequently wrote a book based on these ideas (reference **25**)

INDEX